Multiclass Damage Identification in a Full-scale Bridge using Optimally-tuned One-dimensional Convolutional Neural Network

Sandeep Sony[1], Sunanda Gamage[2], Ayan Sadhu[3], and Jagath Samarabandu[4]

[1]PhD (Corresponding Author), Department of Civil and Environmental Engineering, Western University, Canada. Email: ssony@uwo.ca

[2]PhD Student, Department of Electrical and Computer Engineering, Western University. Email: sgamage2@uwo.ca

[3]Assistant Professor, Department of Civil and Environmental Engineering, Western University, Canada. Email: asadhu@uwo.ca

[4]Professor, Department of Electrical and Computer Engineering, Western University. Email: jagath@uwo.ca

ABSTRACT

In this paper, a novel method is proposed based on a windowed one-dimensional convolutional neural network (1D CNN) for multiclass damage identification using vibration responses of a full-scale bridge. The measured data is first augmented by extracting samples of windows of raw acceleration time-series to alleviate the problem of a limited training dataset. 1D CNN is developed to classify the windowed time-series into multiple damage classes. The damage is quantified using the predicted class probabilities, and the damage is localized if the predicted class is equal to the assigned damage class, exceeding a threshold associated with majority voting. The proposed network is optimally tuned with respect to various hyper-parameters such as window size, random initialization of weights, etc., to achieve the best classification performance using a global 1D CNN model. The proposed method is validated using the Z24 bridge benchmark data for multiclass classification for two different damage scenarios, namely, pier settlement and rupture of tendons,

under the various extent of the damage. The damage identification is carried out on various bridge components to collectively identify the structural component with a damaged signature. The results show that the proposed windowed 1D CNN method achieves an accuracy of 97%, and performs well with different types of damage.

INTRODUCTION

In civil infrastructure, continuously increasing heavy traffic, unexpected natural calamities and human-made damages reduce their load-bearing capacity and service life. With ageing, the structures exhibit various damage signatures in several critical locations. In the absence of timely repair and maintenance, progressive damage leads to the collapse of structures. Despite the simplicity, the traditional manual inspection suffers difficulty while scanning the inaccessible areas in large structures such as bridges or tall buildings. Over the past few decades, structural health monitoring (SHM) has been a promising tool to supplement the knowledge of structural integrity over time. However, efficient diagnosis and prognosis of large-scale infrastructure require a reliable assessment of its damage under in-service conditions. SHM aims to provide suitable diagnostics and prognosis, and assist infrastructure owners and decision-makers in maximizing the safety, serviceability, and functionality of critical structures. An autonomous SHM will allow efficient and cost-effective disaster management and will lead to resilient infrastructure with faster recovery under natural disasters. In this paper, an autonomous multiclass damage identification method is proposed by utilizing artificial intelligence in the sequential data, such as vibration measurements.

Data-driven damage diagnosis is a critical component of infrastructure asset management (Piryonesi and El-Diraby 2019). Although there is a plethora of research on parametric methods based on time-frequency (TF) decomposition techniques (Staszewski and Robertson 2006, Sadhu et al. 2017, Almasri et al. 2020, Barbosh et al. 2020, Kankanamge et al. 2020, Sony and Sadhu 2020, Sony and Sadhu 2021), non-parametric methods (Nakamura et al. 1998, Wang and Ong 2015, Abdeljaber and Avci 2016) have shown significant promises in data-driven SHM methods. Parametric methods include extracting dynamic parameters such as modal parameters, while inferring the change in these parameters to detect any possible changes in the structures. On the other hand, non-parametric methods include extracting parameters that are estimated based on the computational models, where the parameters are mathematically derived in a statistical sense.

Structural damage identification can be considered as a pattern recognition-based non-parametric problem, which is divided into three stages, namely, data acquisition, feature extraction, and feature classification. With proliferation of various machine learning (ML) algorithms, the SHM community has prominently used various supervised learning algorithms (Hou and Xia 2020, Avci et al. 2021). In (Gardner et al. 2020), the authors explained the interface between nondestructive evaluation and machine-learning-based SHM for damage detection. In another study, Su et al. (2020) presented a critical review of field monitoring of high-rise structures. The study reviewed techniques for comfort assessment, seismic and wind effects, and environmental effect on monitoring of super-tall structures. Recently, the SHM community has explored both vibration and image data for structural damage identification and localization.

With advancements in artificial intelligence, image-based SHM has garnered as an inexpensive way to monitor large scale structures using Convolutional Neural Networks (CNNs). While image-based 2D CNN techniques remain a popular method for SHM (Kumar et al. 2019, Chang and Chi 2019, Gulgec et al. 2019, Sony et al. 2019), they involve significant complexity in obtaining a large amount of labelled data, pre-processing and classifying the images. As a solution, researchers have studied algorithms that directly operate on the sequential data such as vibration data. Guo et al. (2014) proposed sparse coding as a feature extraction method for unlabeled acceleration measurements obtained from wireless sensors. The damage classification was carried out using a CNN, and the results were compared with logistic regression and decision trees. A three-span bridge was considered to evaluate the efficacy of the proposed method, and it was shown that sparse coding-CNN based method outperforms other methods with an accuracy of 98%. Gulgec et al. (2017) conducted a simulation study on a steel gusset plate connection by varying the size and location of the damage. A CNN was used to classify damaged signals, and the proposed method achieved a testing error of 2% and showed robustness against environmental noise.

Fallahian et al. (2018) explored the applicability of dynamic features such as mode shapes, frequency response functions, and natural frequencies as damage indicators under varying temperatures. The authors used a combination of coupled sparse coding and deep neural network as an ensemble method for damage detection and localization. The proposed method was validated on a numerical truss bridge and experimental benchmark dataset. Bao et al. (2019) proposed a CNN-based anomaly detection using acceleration measurements by converting them into grayscale images. The authors used several anomaly parameters such as missing, minor, outlier, square, drift, and trend data points to train the datasets using a stacked autoencoder architecture. In another recent study, Shang et al. (2020) proposed deep convolutional denoising autoencoders for structural damage detection. The proposed method extracted damage features from field measurements under environmental noise. However, most of these methods are based on 2D CNN (Sony et al. 2021), which are primarily relied on images.

RECENT DEVELOPMENT OF 1D CNN-BASED SHM

Recently, 1D-CNN has shown promising results in capturing the temporal information and damage detection using vibration data. For example, Abdeljaber et al. (2017) introduced 1D CNN for real-time vibration-based damage detection. The authors trained 1D CNN on a vibration signal database obtained from a truss, named Qatar Grandstand, by damaging each joint and keeping the other joints undamaged. The proposed model was trained individually on each joint, and near-perfect classification accuracy was proposed. However, the proposed method was not tested for a multiclass damage scenario. Zhang et al. (2019) utilized the computational powers of 1D CNN to detect changes in structural parameters such as stiffness and mass. Three different structural components were used for data acquisition and model validation, namely, T-shaped steel beam, short and long steel girder bridge, and a mean classification accuracy of 98% was achieved. In another study, Ni et al. (2019) showed the applicability of 1D CNN with autoencoders for anomaly detection under data compression. The proposed algorithm was validated using a long-span suspension bridge with an accuracy of 97.53%.

A recent study by Azimi and Pekcan (2020) explored the concept of transfer learning in vibration measurements. The authors used a four-story IASC-ASCE SHM model for numerical training, and the proposed model was tested on experimental studies using IASC-ASCE SHM benchmark building and the Qatar University Grandstand Simulator with an accuracy of 90-100%. Recently, Sharma and Sen (2020) showed the applicability of 1D CNN for damage detection in the structural frames. Experimental validation was performed on a 2D-steel frame with different damage location and severity of the damage. The method was shown to identify different damage scenarios and the false-positive rate was also evaluated and found to be well within the acceptable limits. Furthermore, Liu et al. (2020) conducted a study by integrating traditional TF methods with the capability of neural networks. The authors used transmissibility function-based 1D CNN to effectively identify damage at the ASCE SHM benchmark structure. The proposed method was compared with the time-series and fast Fourier transform-based frequency-domain information, where the TF signals exhibited more significant damage-sensitive features.

Overall, 1D CNN exhibited superior performance over artificial neural networks (ANNs) in the context of computation efficiency and noise insensitivity for big data (Kiranyaz et al. 2019). Recently, Bao et al. (2020) evaluated a combination of finite element method (FE) and 1D CNN for localizing damage for a jacket-type offshore structure. However, the data was generated synthetically using a finite element model, which might not resemble the actual real-world data with operational and environmental noise contamination. In addition, the damage was induced artificially using the FE model that was clearly distinguishable from an undamaged structure that does not concur with the real-world data. In another recent study, Sarawgi et al. (2020) used 1D CNN for each joint in an experimental setup to identify hypothetical damage simulated using removal or addition of external braces. However, such individual CNN model per damaged node does not scale well to real-world structures with multiclass damages due to computational complexity. Moreover, the need for large datasets and the selection of the appropriate architecture of the network still remain a challenge.

The proposed research explores the existing challenges of multiclass damage identification in a full-scale bridge. Unlike the simulated data, the real-world data is limited and noise-contaminated, where multiclass damage localization becomes a significant challenge. In this paper, an enhanced windowed-1D CNN is explored for multiclass damage localization with varying damage severity under different damage scenarios. The issue of the limited dataset is solved by augmenting the data using windowing the acceleration measurements, and the classification results are improved using a novel majority voting approach of a global 1D CNN model. The effectiveness of the proposed 1D CNN model is evaluated using a systematic suite of hyperparameter tuning such as window size, random initialization of weights, and optimum learning rates to show the robustness of the selected optimal architecture of the 1D CNN network.

The paper is structured as follows. A brief introduction of the structural damage identification, its need, and a literature review based on 1D CNN techniques are presented first, followed by the theoretical background of the proposed algorithm. The capability of the proposed algorithm to identify multiclass damage, the importance of hyperparameter tuning and various metrics to show the damage parameters of the structures is explained later along with the key conclusions.

PROPOSED METHODOLOGY

Formulation of 1D CNN

Convolutional Neural Networks (CNNs) are a type of feedforward neural network model that is designed to approximate a function $y = f(x; \theta)$. For classification, the model maps an input x to a category (class) y. The parameters θ are learned to best fit a given training dataset by a gradient descent optimization algorithm (Goodfellow et al. 2017). The most common type of CNNs are 2D CNNs used in the field of computer vision for tasks such as image classification, where the inputs x are matrices (2D-shaped) representing images. 1D CNNs are a simpler variant of CNNs, where the inputs x are vectors (1D-shaped), typically representing a time-series. They are commonly used for tasks involving sequential signal processing such as speech recognition (Kiranyaz et al. 2019). Since the last few years, 1D CNNs became popular in SHM due to its computational simplicity in comparison to its parent family of 2D CNNs as it requires simple array application and a shallow network.

A typical 1D CNN architecture used in this study is shown in Fig. 1. It consists of an input layer (time-series), multiple alternating convolutional and pooling layers, and one or more fully connected layers at the end. An input time-series x presented to the input layer is transformed by the forward pass through the hidden layers and the output softmax layer produces the class label y. When the number of hidden layers is high, this architecture is referred to as a deep convolutional neural network. The convolutional layer is the core building block of a CNN. The parameters of each convolutional layer consist of a set of learnable kernels, which are defined by a kernel length (m). Convolutional layers have a reduced number of parameters in comparison to fully connected layers as a single kernel shares the weights for spatial locations in the input. The convolution process can be expressed as Eq. 1 (Goodfellow et al. 2017):

$$y(n) = f(x(n) \otimes h(m)) \tag{1}$$

where $x(n)$ is the input vector of length n and $h(m)$ is kernel of length m. The symbol $\otimes$ denotes the traditional 1D convolution between two signals as defined in Eq. 2,

$$x(n) \otimes h(m) = \sum_{k=0}^{n} x(k)h(n-k) \tag{2}$$

The function f is called an activation function, which is typically a non-linear transformation on the traditional 1D convolution. Non-linear activations enable the network to learn complex mappings between the input signal and the class labels. In this study, Rectified Linear Unit (ReLU) is used as the activation function, which effectively removes negative values from an activation map by setting them to zero. A pooling layer is added after the convolution layer to sub-sample the convolution output. The pooling operation reduces the dimensionality of a given mapping while highlighting the prominent feature and it also helps to reduce overfitting. Max pooling refers to selecting the maximum value in a window that slides over the input map. In Fig. 1, the max pooling layer has reduced the size of each convolution output size by a factor of two. For the output layer, the choice of activation function depends on the type of output. For classification problems, SoftMax activations are preferred. SoftMax function for a n-class problem (representing n probabilities of input belonging to each of n-classes) is shown in Eq. 3.

$$P(\text{class} = j|z) = \frac{e^{z_j}}{\sum_{k=1}^{n} e^{z_k}} \tag{3}$$

where z_j is the input to the softmax node j from the previous layer.

Multiclass damage detection using windowed-voted 1D CNN

A method based on a 1D CNN model is proposed to classify the vibration measurement into multiple damage classes (i.e., multiclass classification). First, each acceleration signal is scaled to fit a standard normal distribution. The scaling improves the convergence rate of models trained on the dataset and prevents any outlier from dominating the input (Ioffe and Szegedy 2015). In order to train a neural network to achieve high test accuracy, a large amount of training data is required. Due to the scarcity of vibration-based multiclass data for civil infrastructure, it is critical to augment the training dataset. In the proposed method, the dataset of raw acceleration signals is augmented by extracting windows of samples from the original signals, as shown in Fig. 2. Each extracted window is assigned the same damage level as the original time-series. A new dataset is formed by taking the extracted windows and their labels as the training instances. In addition to increasing the number of training instances, this windowing technique also reduces the data dimensionality (i.e., shorter input signals), which allows training machine learning models with less over-fitting.

The dataset is then split into training, validation and test sets. A 1D CNN model is trained on the dataset using a standard gradient descent optimizer. Hyper-parameters of the 1D CNN model include the number of layers and number of nodes in each layer, activation function and the kernel size in convolutional layers. Additionally, the length of the extracted windows is also a hyper-parameter. Finding the optimal hyper-parameters (also known as hyper-parameter tuning) is conducted using a random search over the parameter space and selecting the configuration that yields the high accuracy on the validation set (Bergstra and Bengio 2012).

In order to classify a new acceleration signal at test time, windows are extracted as before and fed into the trained 1D CNN model, which outputs a set of classification probabilities for each window. The predicted set of classification probabilities $P_p(y_c)$ for a full acceleration measurement is obtained by summing the class probabilities of all the window sequences in a single time-series. The class with the maximum probability is the predicted damage level classification of the series as shown in Fig. 2. It may be noted that this is equivalent to voting for the majority classification label of the individual window sequences to arrive at the prediction based on the entire time-series. Although it is possible for a well-trained model to misclassify some of the individual window sequences that comprise a time-series, the probability of misclassifying a majority of them is very low (Sony 2021). Therefore, the voting process improves the prediction accuracy and other evaluation metrics in the time-series. This result is empirically observed in the full-scale studies and discussed in model performance section. The proposed machine learning pipeline for the multiclass damage classification problem is shown in Fig. 3.

Performance criteria

In machine learning, a number of performance metrics are used to evaluate the efficacy of the computational model. A brief description of metrics used to evaluate classification models in the context of SHM is provided below. The confusion matrix is a tabulation of classifications made by a model, typically with the actual class in rows and predicted class in columns. There are various metrics that are derived from confusion matrix and are presented in Table 1. In the context of SHM and multiclass damage detection, ROC-AUC, Accuracy, FNR and F1 score are used to evaluate the performance of the proposed method. Accuracy is a primary performance metric used to evaluate the ability of a model to correctly classify the data samples into various class labels. Another important metric that has not been discussed in the literature is FNR. In the SHM context, it is critical that a damage detector have a low false negative rate, as a false negative corresponds to the potentially catastrophic case of a damaged signal being classified as an undamaged signal. Furthermore, a damage detector must have high values for accuracy, and F1 score. Additionally, two curves are used to evaluate the trade-off between performance metrics. The receiver operating characteristic (ROC) curve shows the trade-off between True Positive (TP) Rate (TPR) and False Positive (FP) Rate (FPR) as the decision threshold of the classifier varies. The precision-recall (PR) curve shows the trade-off between precision and recall as the decision threshold of the classifier varies. The area under the curves (AUC) for both ROC and PR is a summary metric that reflects the level of possible trade-off. Both curves are useful for an academic and practicing engineer to find a suitable decision threshold.

Algorithm for Component-level Damage Identification

Component level damage is identified for multi-class problems using Algorithm 1. The entire structure is modeled as one experiment rather than modeling each sensor separately as in (Abdeljaber et al. 2017) and prediction probabilities are acquired for each sensor location. This provides control over all the features for one experiment and prevents errors arising from multiple models working on different sensors. However, as there are multiple sensors covering the entire structure, different structural components are collectively used to localize damage using a limited number of sensors for each component. The damage is confirmed if the true predicted probability class is equal to allocated class label for all cumulative windowed series for each sensor location and $P_p(y_c)$ is greater than the threshold.

Algorithm 1: Multiclass damage identification

Input: A signal $x(t)$

Output: Prediction probabilities $P_p(y_c)$ for damage component level identification.

(a) The acceleration data is pre-processed into multiple windows of time-series and damage class-label is allocated to each windowed data.

(b) The structure is modeled as whole (i.e., a global CNN model) as compared to per sensor (Abdeljaber et al. 2017) for computational efficiency and ease of modeling.

(c) The windowed data is trained using 1D CNN with optimal parameters (e.g., window size, random initialization of weights etc.) and tested on a separate dataset.

(d) The probabilities of classification are obtained for each sensor of every windowed series and damage is confirmed if true predicted probability class of a sensor is equal to allocated class label, while an average of $P_p(y_c)$ for various structural components is used as threshold of damage.

(h) If the $P_p(y_c) \geq$ threshold, a localized damaged is confirmed.

FULL-SCALE STUDY

Details of the Z24 Bridge

Damage detection, where classification is more than two classes, is considered as a multiclass problem. In this study, two types of damage cases are considered, namely, rupture of tendons, and pier settlement of a full-scale bridge, namely, Z24 Bridge (Maeck and Roeck 2003). All the damage classes have multiple damage levels are used to evaluate the performance of the proposed method for multiclass damage detection. The bridge was located in the canton Bern near Solothurn, Switzerland. It was a classical post-tensioned concrete two-cell box-girder bridge with a main span of 30 m and two side spans of 14 m, as shown in Fig. 4. The bridge was demolished at the end of 1998 because a new railway adjacent to the highway required a new bridge with a larger side span. During the demolition, the bridge data was acquired using 15 accelerometers placed at different spans of the bridge, as shown in Fig. 5. The bridge was excited by two shakers, one at the mid-span of the bridge and another at a side-span. Because of the size of the bridge, response was measured in nine setups of up to 15 sensors each, with three accelerometers and the two force sensors common in all setups. The data was sampled at 100 Hz, and the data was made publicly available by the researchers at the Katholieke Universiteit Leuven (https://bwk.kuleuven.be/bwm/z24). The data was acquired by performing various progressive damage scenarios during the demolition period. For the brevity of this study, only two different damage scenarios are considered: rupture of tendons, and pier settlement. It may be noted that each damage scenario have different classes of damage, and they were chosen to evaluate the performance of the proposed method to classify various multiclass damage cases. For example, rupture of tendons have three levels, and pier settlements have four levels, and together they made a case of two separate damage classes. For detailed explanation of how the damages were induced in the bridge, the readers are suggested to refer (Roeck and Teughels 2004). The reference undamaged condition is considered as class-zero for all the cases and the other damages were assigned classes starting from 1 to n depending upon the level of damage, as shown in Table 2. For example, in the case of rupture of tendons, the damage was induced at first, rupture of two tendons, and second, rupture of four tendons, third, rupture of six tendons, thereby creating three classes of damages for rupture of tendons. Similarly, there are four classes for pier settlement. The rupture of tendon dataset contains 1,231 time-series (i.e., vibration signals) and the lowering of pier dataset contains 1,056 time-series. Each time-series contains 65,530 samples. Both datasets are class-balanced, and they are split into three sets of train-validation-test as 60%-10%-20% of the original suite of time-series.

Hyper-parameters of the 1D CNN model

A variety of hyper-parameters is incorporated to improve the robustness of the proposed model and avoid overfitting. For example, learning rate is evaluated for various scenarios to improve the accuracy while reducing the overfitting by empirically changing the gradient during back-propagation. In this study, first a search space of hyper-parameters is designed by considering a wide range of values for each hyper-parameter. In order to make the search feasible, extreme values (e.g., very small window sizes) that yielded poor performance or unstable training dynamics (training error does not decrease) were removed from the search space. A random search is performed on this hyper-parameter space, and the set of hyper-parameters that yielded highest accuracy is selected as the optimal configuration. The evaluation metrics based on the proposed method are described later with a comparison between window-voted and non-voted results. In this study, a range of hyper-parameters are selected first and tuned using random search algorithm to achieve a set of hyper-parameter that provides the optimal accuracy. The range of hyper-parameters used for 1D CNN is presented in Table 3. For example, window size is adopted within a range of 32-512 samples. Window size is the only external parameter and is decided by the user. Thus, a sensitivity analysis is performed to understand the behavior of the performance evaluation metrics (P_m) under different window sizes (w). Two different metrics, accuracy and FN, are used for sensitivity analysis as they represent overall accuracy of the model and false-negative alarm critical for civil infrastructure.

The optimal hyper-parameters of this dataset are obtained after tuning and are presented for all the models in Table 4. An analysis is performed to understand the effect of w versus P_m. The results are shown for various damage cases in Fig. 6. For example, Fig. 6 (a-b) shows that the optimal performance is achieved at w=256, with highest ROC and accuracy, and lowest false-negative. Although, the FNR remains consistent after w=512 and other metrics are at their peak, however, due to larger w, the data size reduces per damage class and it leads to over-fitting of the data.

Random initialization of weights

Deep learning algorithms are iterative and require the user to specify value of initial weights of neurons to initiate the iteration and its optimization. In practice, all weights in the model are randomly drawn from a Gaussian or uniform distribution (Goodfellow et al. 2016). However, the scale (low or high magnitude) has a large effect on both the outcome and optimization procedure. In this study, random initialization with early stopping criteria is used and Adam optimizer (Kingma and Ba, 2014) is used with dropout in each layer for regularization. After acquiring the optimal tuned parameters, a parametric study is conducted to understand variance in the metrics of 1D CNN model for random initialization of weights. The metrics used for evaluating random initialization of weights are ROC-AUC, accuracy, FNR, and F1 score and are shown in Table 5. It can be observed that for pier settlement, the mean (μ) of ROC-AUC is 0.97 with an accuracy of 0.85. The FNR is 0.15 and the standard deviation (σ) for all the trials is at its minimal of 1%. Similarly, for rupture of tendons, the ROC-AUC is 0.92 with an accuracy of 0.67 and FNR of 0.33 with minimal σ of 2%.

Effect of window size

The window size used to augment the data is an external parameter apart from other model parameters and it is critical to understand its effect on model performance. It can be observed that the best performance with a combination of maximum ROC-AUC and accuracy and minimum FNR is achieved at 256 samples per window. It is shown in Fig. 6 (a), ROC-AUC increases to 1.0 at 512, 800, 1024 samples per window, however, it leads to overfitting with increased FNR. A similar result can be observed from Fig. 6 (b) with optimal performance at 256 samples per window.

Model performance

The optimal parameters are first used to evaluate the performance of the proposed model on an entire series versus windowed and voted-windowed samples. The reason for comparison of the entire time-series, windowed and voted windows is to show the improved performance by voting the windowed samples. It should be noted that model performance on entire series results in very poor accuracy due to nonlinearity, and nonstationarity in the signal. It may be noted that micro-averaged is the average of area under the curve for all the classes. It can be observed from Figs. 7 (a) & (b) for pier settlement, the ROC-AUC of a full-series signal is merely 0.56 while PR-AUC is only 0.21. A similar observation can be deduced in case of rupture of tendon in Figs. 8 (a) & (b), where ROC-AUC and PR-AUC is 0.55 and 0.28, respectively. However, it is observed that voting on windowed dataset increases accuracy considerably and it exhibits in ROC-AUC and precision-recall (PR)-AUC curves, as presented in Figs. 7(c-f), and 8(c-f), respectively. The accuracy in terms of ROC-AUC and PR-AUC increased by 71% and 314% for pier settlement case, and by 58% and 153.5% in case of rupture of tendons. Moreover, as shown in Fig. 7, majority voting on samples has further improved the AUC for both ROC and PR. It can be observed that in case of pier settlement, there is meager increase on ROC-AUC, however, there is a considerable improvement in the area under the curve for PR. This behaviour can be attributed to a more localized damage in case of pier settlement. Similarly, as observed in Fig. 8, when the damage was considerably distributed in case of rupture of tendons, majority voting on windows highly increase the PR-AUC for rupture of tendons.

It can be observed that majority voting on windows of distributed damage signal increases the probability considerably by allocating the majority class and ignoring the non-prominent class along with augmenting the data samples per class. Another critical performance metric, FNR is used to evaluate the false negative alarm of damage in the proposed model. It can be noted that the FNR reduces as the entire time-series is windowed and further reduces with the majority voting. For example, in case of pier settlement, FNR reduces from 0.80 to 0.17 for the entire series versus majority voted augmented dataset. Similarly, FNR reduces from 0.71 to 0.34 for rupture of tendons, as shown in Table 6. The label 0, 1, and 2 are used to denote performance metrics of the entire time-series, windowed time-series and majority voted-windowed time series, respectively. The tuning process and the resulting performance improvement provide adequate justification to counter any potential errors caused by the windowing.

Component-level damage identification

Component level damage identification is performed using Algorithm 1, for two multiclass damage scenario, namely, pier-settlement and rupture of tendons. The sensor locations are identified first, then, three different structural components of the bridge are used to localize the damage caused by the pier settlement and rupture of tendons during the demolition period of the bridge. An undamaged pier (Utzenstorf), bridge deck, and damaged pier (koppigen) are used for representation of predicted probability (P_p) and infer damages in three components. The Koppigen pier is used for inducing the damage by lowering it in several increments starting with 20 mm, 40 mm, 80 mm, and moving to 95 mm at the last stage. Twelve different sensors are used to identify the location of damage, namely, 4 sensors (411, 421, 431, 441) on the undamaged pier (UDP), 4 sensors (216, 221, 226, 231) on the bridge deck (BD), and 4 sensors (511, 521, 531, 541) on the damaged pier (DP), as shown in Fig. 9.

The P_p is plotted against the sensor index and a dash-dotted average of P_p is shown as a representation of combined P_p for the corresponding structural component, as shown in Fig. 10 for 20 mm, and 40 mm and in Fig. 11 for 80 mm, and 95 mm lowering of pier, respectively. For example, Fig. 10 (a, b, c) represents P_p for undamaged pier (UDP), bridge deck (BD), and damaged pier (DP) for 20 mm lowering of piers. Similarly, Fig. 10 (d, e, f) is for 40 mm lowering of piers, respectively. It can be observed that, unlike in pier settlement of 40 mm, the proposed algorithm does not provide conclusive evidence of nominal damage of 20 mm. However, Fig. 11 (a, b, c) and (d, e, f) shows identification of damage for 80 mm and 95 mm, and it can be observed that the identification is clearly achieved through the proposed threshold where the P_p is highest for DP followed by BD which is affected by differential settlement of one of the piers.

It may be noted that, as the severity increases, the signals become more distinguishable and 1D CNN learns the classification more effectively. It can be observed from Fig. 12 that UDP shows lowest predicted probability due to its similarity to the response of the undamaged pier, however, both BD, and DP shows higher prediction accuracy. The reason for BD's highest probability is attributed to the surface area and larger affect of differential pier settlement of the entire bridge. In summary, it can be concluded that the proposed method performs well for damage levels of 40 mm, 85 mm, and 95 mm, however, does not perform well for a very low level of damage, as shown in case of pier settlement of 20 mm. Finally, Table 7 shows the confusion matrix for the multi-class pier settlement damage classification problem with five classes, starting from no damage (class-0) to 95 mm damage (class 5).

For rupture of tendons, the most affected area would be the bridge deck and the damage induced due to rupture of tendons will create a non-localized and distributed damage throughout the bridge deck in comparison to the damage in the bridge piers. The damage identification per sensors is avoided due to non-conclusive inference and a comparison between structural components of the bridge is provided in Fig. 13. It should be noted that rupture of tendons affects bridge deck and it is shown in Fig. 13, however, the proposed algorithm could not clearly show the effect of rupture of two and four tendons, while the rupture of six tendons proves to be the worse damage level scenario.

CONCLUSIONS

In this paper, a windowed-1D CNN is employed for multi-class damage detection using limited datasets. First, the limited dataset is augmented using windowing of the vibration data and the prediction accuracy is improved by a novel majority voting approach on windowed classes. It is observed that due to non-localization of sensors for data acquisition, damage identification for a minor level of damage (say, 20 mm of pier settlement) is a challenge to predict. However, the overall accuracy significantly improves with the increase in the severity of the damage (i.e., a pier settlement of 40-95 mm). The proposed algorithm is analyzed with a sensitivity analysis on window-size as the external parameter to the model as well as a parametric study to evaluate its sensitivity with random initialization of weights. The improvement in the accuracy is illustrated through a comparison between a single series dataset and windowed-voted time-series for ROC and precision-recall AUC. In this paper, it is demonstrated that a simple 1D CNN architecture with only one hidden layer is capable of classifying the time-series of vibration data into multi-class with high accuracy. There are still a few limitations of the proposed algorithm, which are reserved to be addressed in the future research.

- The obtained damage classification provided superior results; however, damage localization at the sensor level was not achieved. It may be noted that damage is identified with the exception of minor level of damage (e.g., 20 mm pier settlement). Future research is reserved to improve the proposed method and accommodate the minor level of damage by conducting an experimental simulation of a series of progressive damage cases with a wide range of minor damage.
- The proposed 1D CNN model based on sequential data is independent of any feature selection process and offers robust and accurate approaches to complex damage identification. However, like any other supervised technique, the 1D CNN also requires a significant amount of training data to classify and predict the damage, which can be considered as a limitation. However, with the recent advancement of remote and autonomous sensors and internet-of-things, long-term SHM technologies have shown significant promise to monitor critical infrastructure in smart cities. It is anticipated that the long-term monitoring will allow the SHM researchers and practitioners to acquire low-cost periodical data with multi-class health conditions of the structures, serving as the potential future training data for the deep learning techniques.

DATA AVAILABILITY STATEMENT

All of the data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

The authors would like to acknowledge funding provided to the first author through Government of Ontario, Canada as Ontario Graduate Scholarship (OGS), and Queen Elizabeth II Graduate Scholarship in Science and Technology (QEGS-II) for PhD studies. The authors also acknowledge NSERC for providing the financial support to conduct this research through the third and fourth author's NSERC Discovery grant.

REFERENCES

Abdeljaber, O., and Avci, O. (2016). Non-parametric structural damage detection algorithm for ambient vibration response: utilizing artificial neural networks and self-organizing maps. *Journal of Architectural Engineering*, ASCE, 22(2), 04016004.

Abdeljaber, O., Avci, O., Kiranyaz, S., Gabbouj, M., and Inman, D. J. (2017). Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *Journal of Sound and Vibration*, 388, 154-170.

Almasri, N., Sadhu, A., and Ray Chaudhuri, S. (2020). Toward Compressed Sensing of Structural Monitoring Data Using Discrete Cosine Transform, *Journal of Computing in Civil Engineering*, ASCE, 34(1).

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., Inman, D J. (2021). A review of vibration-based damage detection in civil engineering structures: from traditional methods to machine learning and deep learning based applications, *Mechanical Systems and Signal Processing*, 147, 107077.

Azimi, M., and Pekcan, G. (2020). Structural health monitoring using extremely compressed data through deep learning. *Computer-Aided Civil and Infrastructure Engineering*, 35(6), 597-614.

Bao, X., Fan, T., Shi, C., and Yang, G. (2020). One-dimentional convolutional neural network for damage detection of jacket-type offshore platforms. *Ocean Engineering*, 108293.

Bao, Y., Tang, Z., Li, H., and Zhang, Y. (2019). Computer vision and deep learning-based data anomaly detection method for structural health monitoring. *Structural Health Monitoring*, 18(2), 401-421.

Barbosh, M., Singh, P., and Sadhu, A. (2020). Emperical mode decomposition and its variants: a review with applications in structural health monitoring, *Smart Materials and Structures*, 29(9), 093001.

Bergstra, J., and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning*. 13, 281-305.

Chang, K., and Chi, S. (2019). Bridge clustering for systematic recognition of damage patterns on bridge elements, *Journal of Computing in Civil Engineering*, ASCE, 33(5), 04019028.

Fallahian, M., Khoshnoudian, F., and Meruane, V. (2018). Ensemble classification method for structural damage assessment under varying temperature. *Structural Health Monitoring*, 17(4), 747-762.

Gardner, P., Fuentes, R., Dervilis, N., Mineo, C., Pierce, S.G., Cross, E.J, and Worden, K. (2020). Machine learning at the interface of structural health monitoring and non-destructive evaluation, *Philosophical Transactions of Royal Society*, A378: 20190581.

Goodfellow, I., Bengio, Y., and Courville, A. (2017). Deep learning. Cambridge, MA: MIT Press.

Gulgec, N. S., Takáč, M., and Pakzad, S. N. (2017). Structural damage detection using convolutional neural networks. *Conference Proceedings of the Society for Experimental Mechanics Series*, 3 Part F2 (June), 331-337.

Gulgec, N. S., Takáč, M., and Pakzad, S. N. (2019). Convolutional Neural Network Approach for Robust Structural Damage Detection and Localization, *Journal of Computing in Civil Engineering*, ASCE, 33(3): 04019005.

Guo, J., Xie, X., Bie, R., and Sun, L. (2014). Structural health monitoring by using a sparse coding-based deep learning algorithm with wireless sensor networks. *Personal and Ubiquitous Computing*, 18(8), 1977-1987.

Hou, R., and Xia, Y. (2020). Review on the new development of vibration-based damage identification for civil engineering structures: 2010-2019, *Journal of Sound and Vibration*, 115741.

Ioffe, S., and Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167.

Kankanamge, Y., Hu, Y., and Shao, X. (2020). Application of wavelet transform in structural health monitoring, *Earthquake Engineering and Engineering Vibrations*, 19, 515-532.

Kingma, D.P., and Ba, J.L. (2014). ADAM: a method for stochastic optimization. *3rd International Conference for Learning Representations*, San Diego, arXiv:1412.6980v9.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., and Inman, D. J. (2019). 1D Convolutional Neural Networks and Applications: A Survey, 1-20. Retrieved from http://arxiv.org/abs/1905.03554.

Kumar, S.S., Wang, M., Abraham, D.M., Jahanshahi, M.R., Iseley, T., and Cheng, J.C.P. (2019). Deep learning based automated detection of sewer defects in CCTV videos, *Journal of Computing in Civil Engineering*, ASCE, 34(1), 04019047.

Liu, T., Xu, H., Ragulskis, M., Cao, M., and Ostachowicz, W. (2020). A data-driven damage identification framework based on transmissibility function datasets and one-dimensional convolutional neural networks: Verification on a structural health monitoring benchmark structure, *Sensors*, 20(4), 1–25.

Maeck, J., and Roeck De. G. (2003), "Description of Z24 benchmark," *Mechanical Systems and Signal Processing*, 17(1), 127-131.

Nakamura, M, Masri S. F., Chassiakos, A.G., and Caughey T. K. (1998). A method for non-parametric damage detection through the use of neural networks. *Earthquake Engineering and Structural Dynamics*, 27, 997-1010.

Ni, F. T., Zhang, J., and Noori, M. N. (2020). Deep learning for data anomaly detection and data compression of a long-span suspension bridge. *Computer-Aided Civil and Infrastructure Engineering*, 35(7), 685-700.

Piryonesi, S.M., and El-Diraby, T.E. (2019). Data analytics in asset management: cost-effective prediction of the pavement condition index. *Journal of Infrastructure Systems*, ASCE, 26(1), 04019036.

Roeck, De. G., and Teughels, A. (2004), "Structural damage identification of the highway bridge Z24 by FE model updating," *Journal of Sound and Vibration*, 278, 589-610.

Sarawgi, Y., Somani, S., Chhabra, A., and Dhiraj (2020). Nonparametric Vibration Based Damage Detection Technique for Structural Health Monitoring Using 1D CNN. *In book: Computer Vision and Image Processing*, 146-157.

Sadhu, A., Narasimhan, S., and Antoni, J. (2017). A review of output-only structural mode identification literature employing blind source separation methods. *Mechanical Systems and Signal Processing*, 94, 415-431.

Shang, Z., Sun, L., Xia, Y., and Zhang, W. (2020). Vibration-based damage detection for bridges by deep convolutional denoising autoencoder. *Structural Health Monitoring*.

Sharma, S., and Sen, S. (2020). One-dimensional convolutional neural network-based damage detection in structural joints. *Journal of Civil Structural Health Monitoring*, 0123456789.

Sony, S., Laventure, S., and Sadhu, A. (2019). A literature review of next-generation smart sensing technology in structural health monitoring. *Structural Control and Heath Monitoring*, 26(3), e2321.

Sony, S., and Sadhu, A. (2020). Synchrosqueezing transform-based identification of time-varying structural systems using multi-sensor data. *Journal of Sound and Vibration*, 486, 115576.

Sony, S., and Sadhu, A. (2021). Multivariate empirical mode decomposition–based structural damage localization using limited sensors. *Journal of Sound and Vibration*, 10775463211006965.

Sony, S., Dunphy, K., Sadhu, A., and Capretz, M. (2021). A Systematic Review of Convolutional Neural Network-based Structural Condition Assessment Techniques. *Engineering Structures*, 226, 111347.

Sony, S. (2021). Towards Multiclass Damage Detection and Localization using Limited Vibration Measurements, Western University, Canada, *Electronic Thesis and Dissertation Repository*. 7659.

Srinivasan, A. (1999). Note on the location of optimal classifiers in n-dimentional roc space. *Technical Report PRG-TR-2-99*, Oxford University Computing Laboratory, England.

Staszewski., W. J., and Robertson, A.N. (2006). Time-frequency and time–scale analyses for structural health monitoring, *Philosophical Transactions of Royal Society*, A365, 1851.

Su, J., Xia, Y., and Weng, S. (2020). Review on field monitoring of high-rise structures. *Structural Control and Health Monitoring*, 27(12), e2629.

Wang, V. Z., and Ong, K.C.G. (2015). Nonparamteric statistical formulations for structural health monitoring. *Computers and Structures*. 148, 63-74.

Zhang, Y., Miyamori, Y., Mikami, S., and Saito, T. (2019). Vibration-based structural state identification by a 1-dimensional convolutional neural network. *Computer-Aided Civil and Infrastructure Engineering*, 34(9), 822-839.

List of Tables

TABLE 1. Description of various performance metrics.

Metric	Formula	Remarks
ROC-AUC	Recall Vs FPR	Degree of separability between classes
Accuracy	$\frac{\text{TP+TN}}{\text{TP+FN+FP+TN}}$	Less useful for heavily imbalanced data
Precision	$\frac{\text{TP}}{\text{TP+FP}}$	Positive predicted value
Recall	$\frac{\text{TP}}{\text{TP+FN}}$	True positive rate or sensitivity
False Positive Rate (FPR)	$\frac{\text{FP}}{\text{TN+FP}}$	False alarm when there is no damage
False Negative Rate (FNR)	$\frac{\text{FN}}{\text{TP+FN}}$	No alarm for actual damage

TABLE 2. Description of the multiclass damage scenarios and the class labels.

Problem	Damage scenario	Class label
0	Undamaged	0
1	Rupture of 2 tendons	1
	Rupture of 4 tendons	2
	Rupture of 6 tendons	3
2	Lowering of pier, 20 mm	1
	Lowering of pier by 40 mm	2
	Lowering of pier by 80 mm	3
	Lowering of pier by 95 mm	4

TABLE 3. Hyper-parameters used in 1D CNN for tuning by random search algorithm.

Parameter	Values
Window size	32, 64, 128, 160, 256, 512
No. of hidden convolutional layers	1 - 6
No. of filters	1024, 512, 256, 128, 64, 32, 16
No. of fully connected layers	1 or 2 layers with 16 and 32 nodes
Learning rate	0.0003, 0.001, 0.01, 0.1
Batch size	32, 64, 256, 512
Kernel size	4, 8, 16, 32, 64

TABLE 4. Optimal configuration of the hyper-parameters of the selected 1D CNN.

Parameter	Values
Window size	256
No. of hidden convolutional layers	1
No. of filters	32
No. of fully connected layers	2 with 32 and 16 nodes, respectively
Learning rate	0.0003
Batch size	256
Kernel size	16

TABLE 5. Random initialization of weights.

Pier settlement				
Trial #	ROC-AUC	Accuracy	FNR	F1 score
1	0.98	0.85	0.15	0.85
2	0.97	0.85	0.15	0.85
3	0.98	0.86	0.14	0.86
4	0.97	0.83	0.17	0.83
5	0.98	0.86	0.14	0.86
μ	0.97	0.85	0.15	0.85
σ	0.00	0.01	0.01	0.01
Rupture of tendons				
1	0.92	0.69	0.31	0.69
2	0.93	0.68	0.32	0.68
3	0.90	0.66	0.34	0.66
4	0.91	0.65	0.35	0.65
5	0.92	0.66	0.34	0.66
μ	0.92	0.67	0.33	0.67
σ	0.01	0.02	0.02	0.02

TABLE 6. Training and testing performance of 1D CNN.

Lowering of pier								
Dataset	ROC	PR	A	P	R	FPR	FNR	F1 score
0	0.56	0.21	0.20	0.20	0.20	0.2	0.80	0.20
1	0.95	0.84	0.77	0.77	0.77	0.06	0.23	0.77
2	0.97	0.91	0.83	0.83	0.83	0.04	0.17	0.83
Rupture of tendons								
Dataset	ROC	PR	A	P	R	FPR	FNR	F1 score
0	0.55	0.28	0.29	0.29	0.29	0.23	0.71	0.29
1	0.87	0.71	0.59	0.59	0.59	0.14	0.41	0.59
2	0.92	0.82	0.66	0.66	0.66	0.11	0.34	0.66

TABLE 7. Confusion matrix for a pier settlement damage problem.

		Predicted Class				
		Pred-0	Pred-1	Pred-2	Pred-3	Pred-4
True Class	True-0	192	8	3	0	5
	True-1	56	137	16	0	3
	True-2	1	0	203	3	5
	True-3	0	0	24	183	5
	True-4	1	0	8	4	199

List of Figures

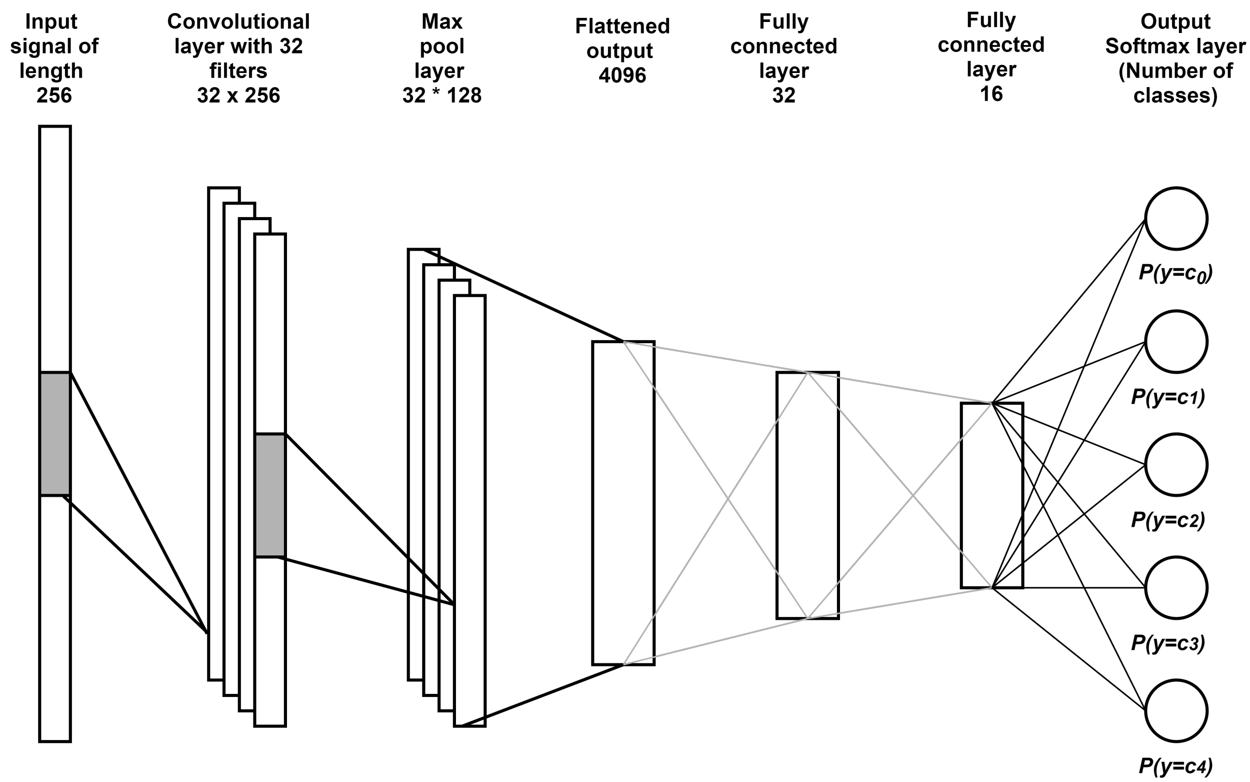


Fig. 1. The proposed 1D CNN architecture used in this study.

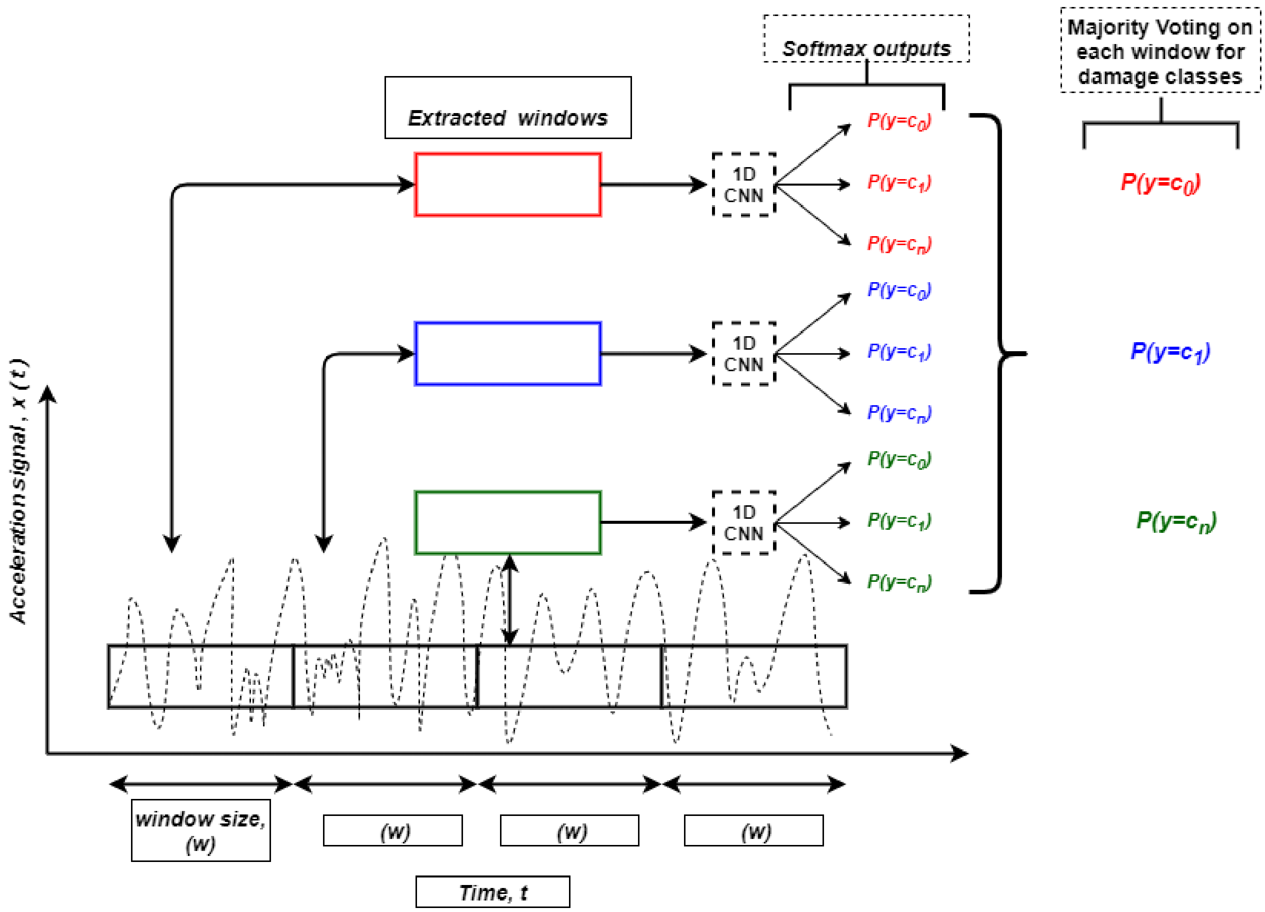


Fig. 2. Extracting data sequences of windows from the vibration data using 1D CNN architecture.

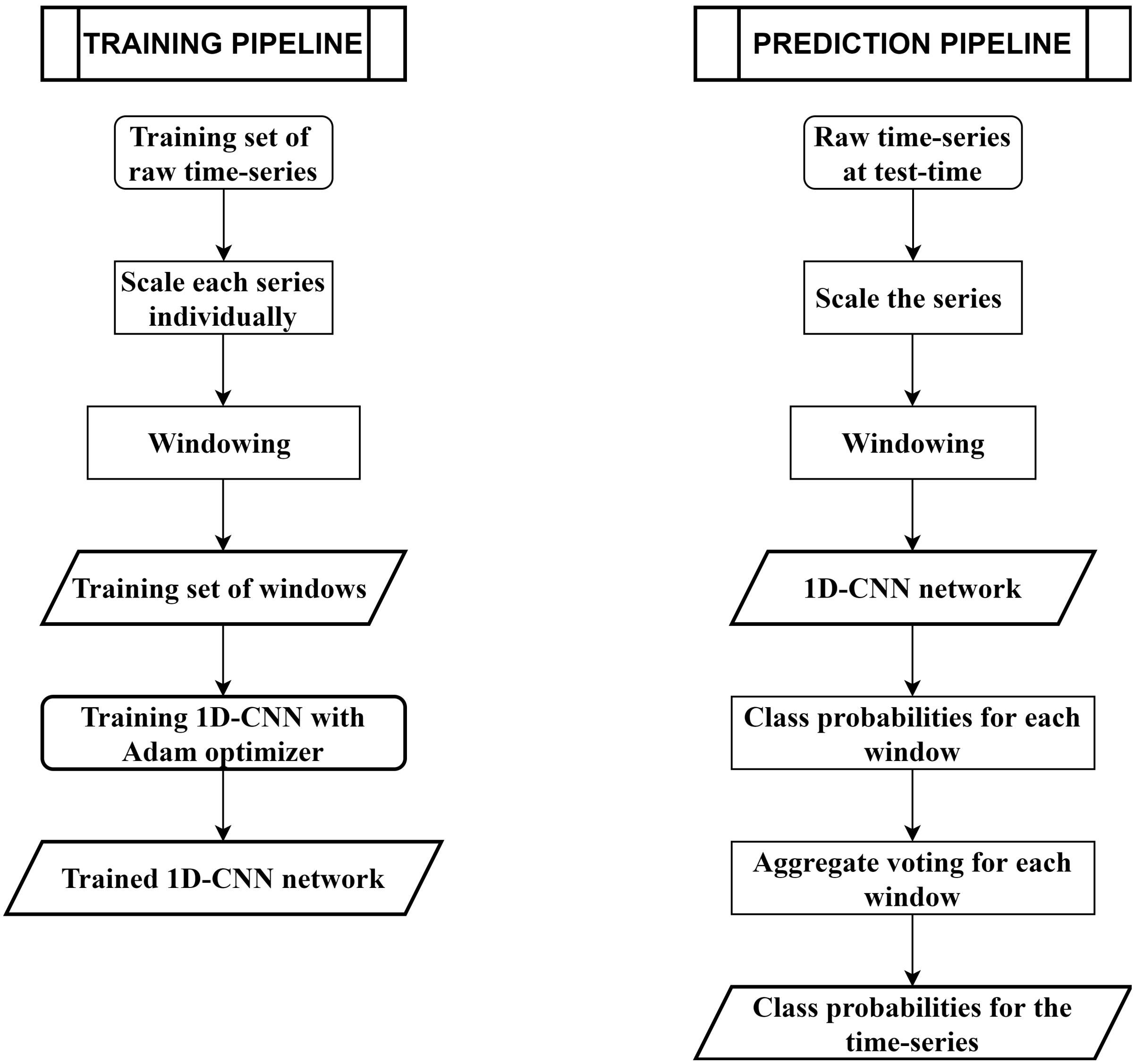


Fig. 3. Data pipelines for training the proposed 1D CNN network and obtaining predictions for a given time-series.

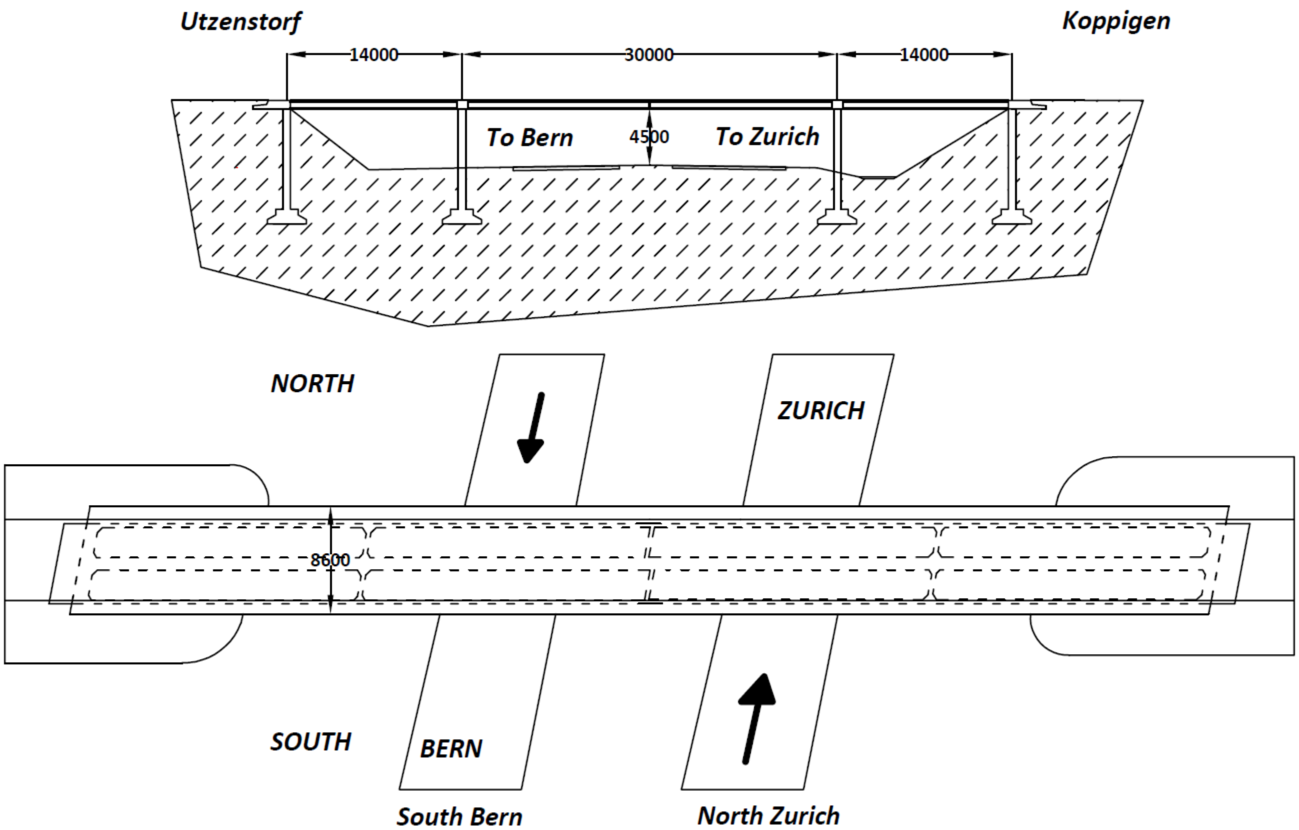


Fig. 4. Schematic of the Z24 bridge.

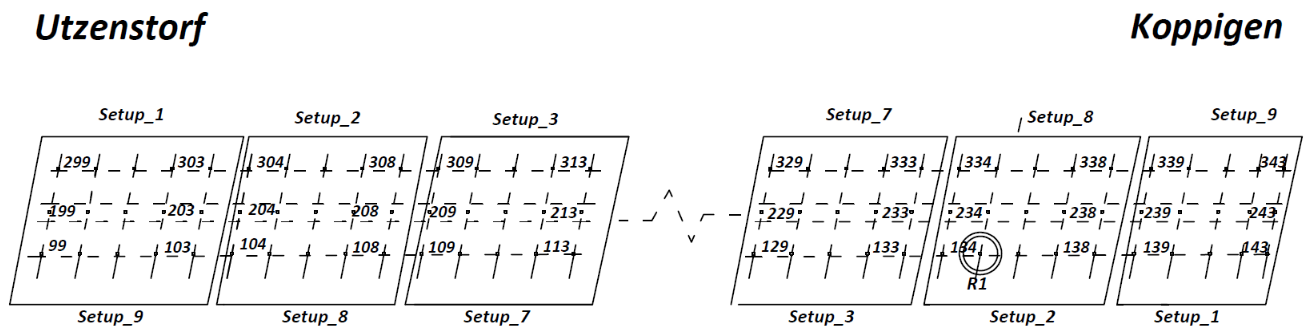


Fig. 5. Sensor placement for data acquisition in the Z24 bridge.

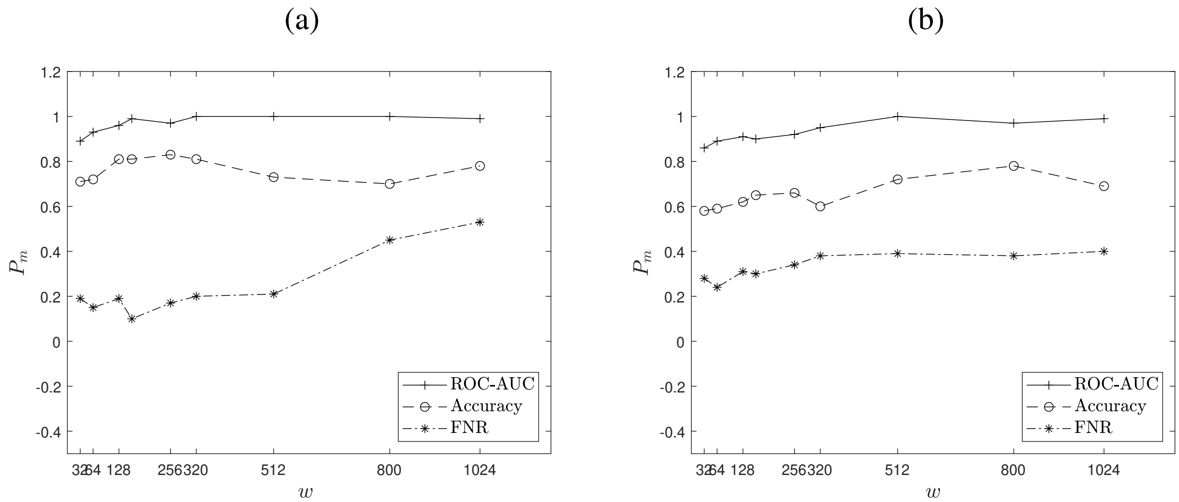


Fig. 6. Performance evaluation of 1D CNN based on window size for (a) pier settlement and (b) rupture of tendons.

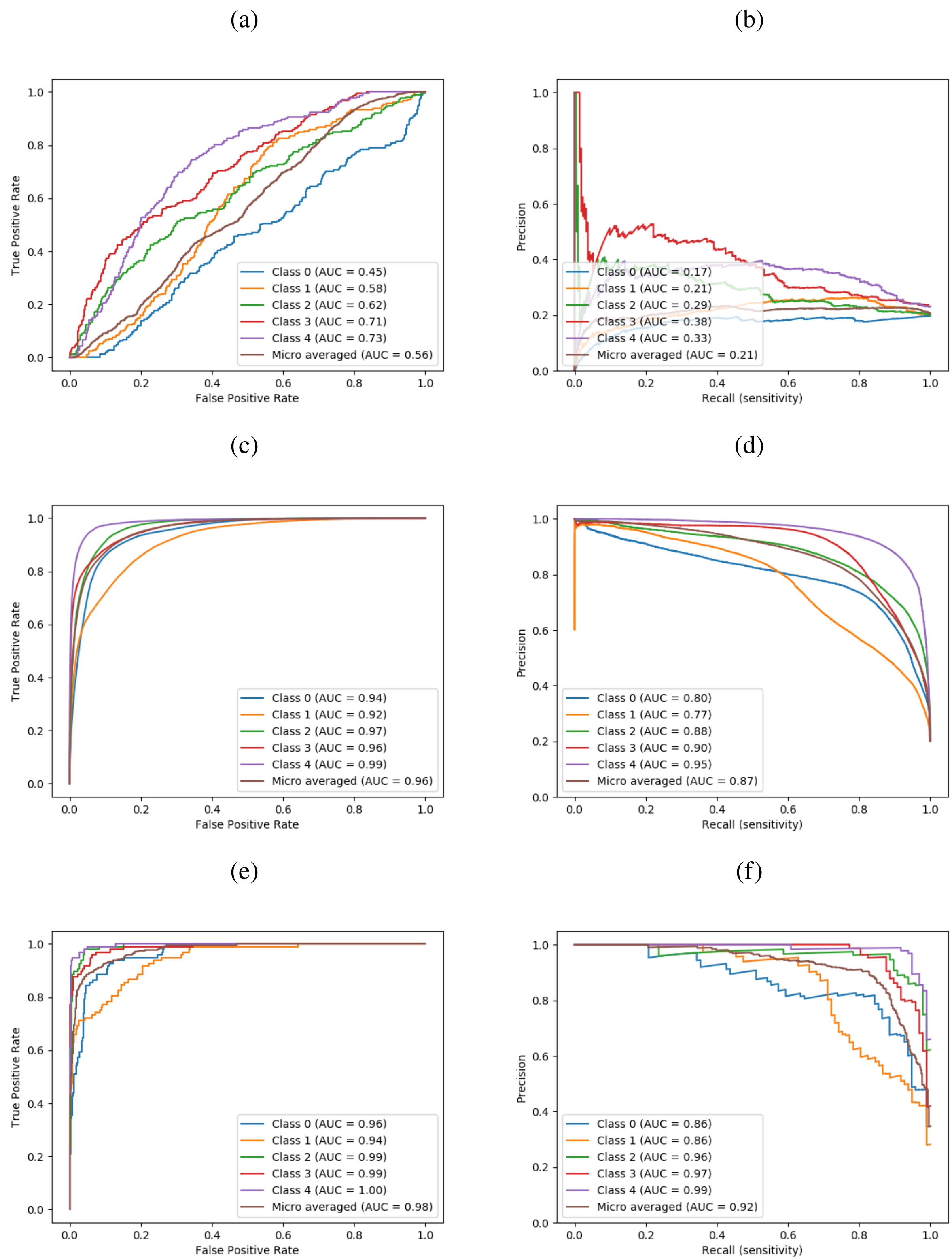
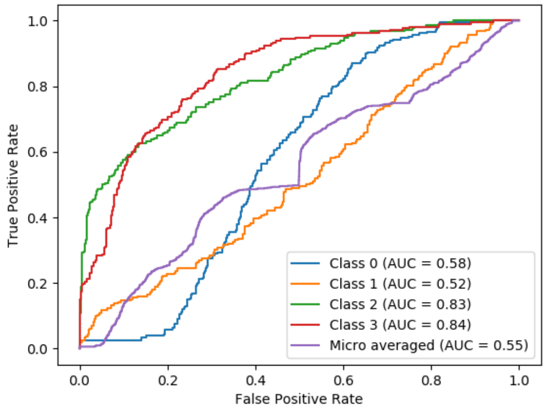
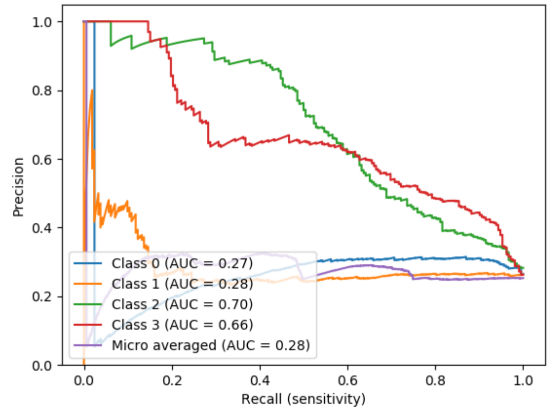


Fig. 7. Performance of 1DCNN by windowing of the data of pier settlement (a) full series ROC, (b) full series PR, (c) windowed ROC, (d) windowed PR, (e) windowed-voted ROC, (f) windowed-voted PR.

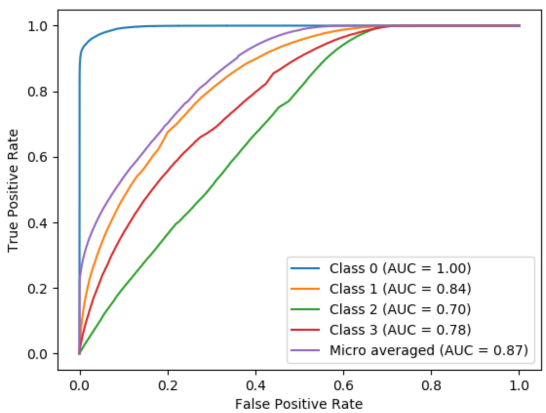
(a)



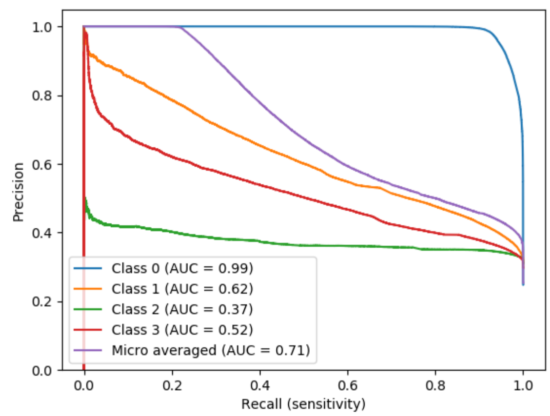
(b)



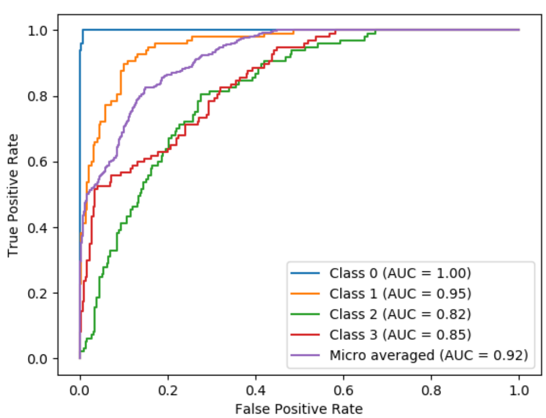
(c)



(d)



(e)



(f)

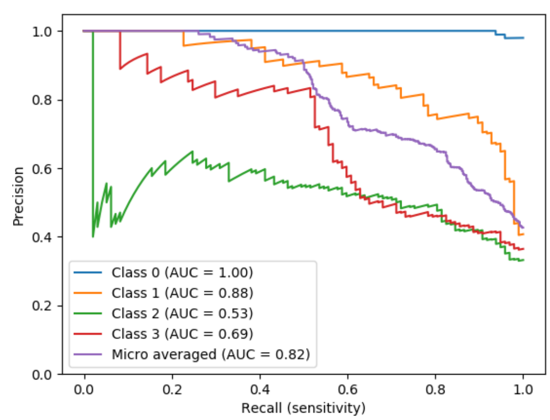


Fig. 8. Performance of 1D CNN by windowing of the data of rupture of tendons (a) Full series ROC, (b) Full series PR, (c) windowed ROC, (d) windowed PR, (e) windowed-voted ROC, (f) windowed-voted PR.

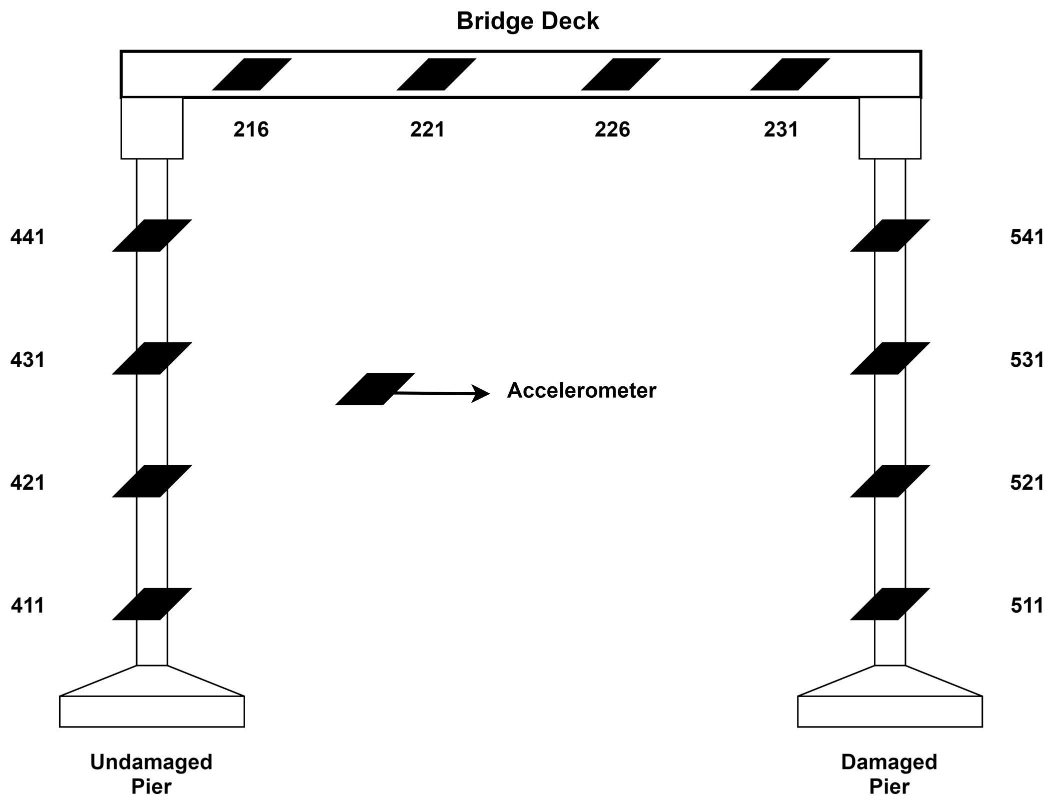


Fig. 9. Schematic showing the sensor location and their numbers used in the analysis.

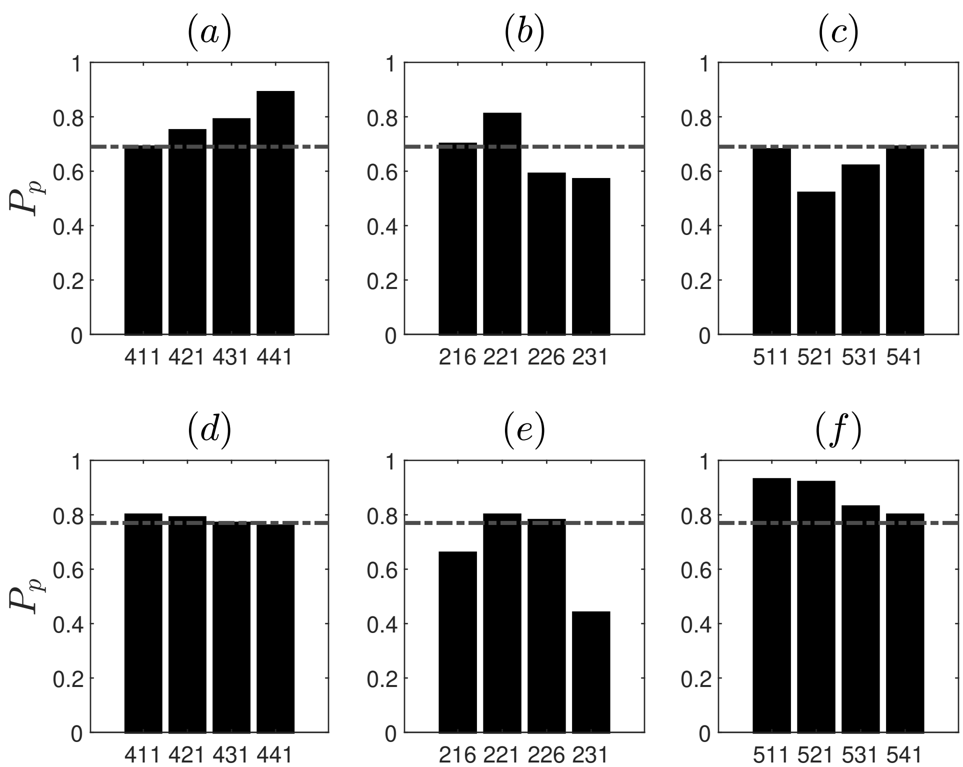


Fig. 10. Damage identification for pier settlement with two damage levels, (a, b, c): 20 mm and (d, e, f): 40 mm.

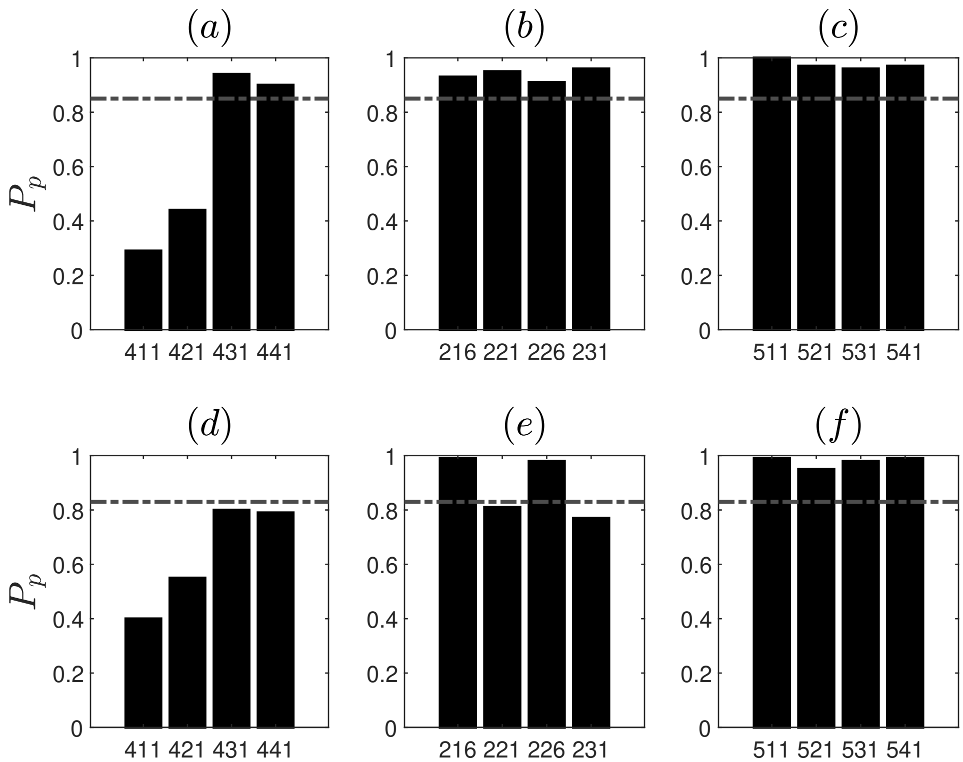


Fig. 11. Damage identification for pier settlement with two damage levels, (a, b, c): 80 mm and (d, e, f): 95 mm.

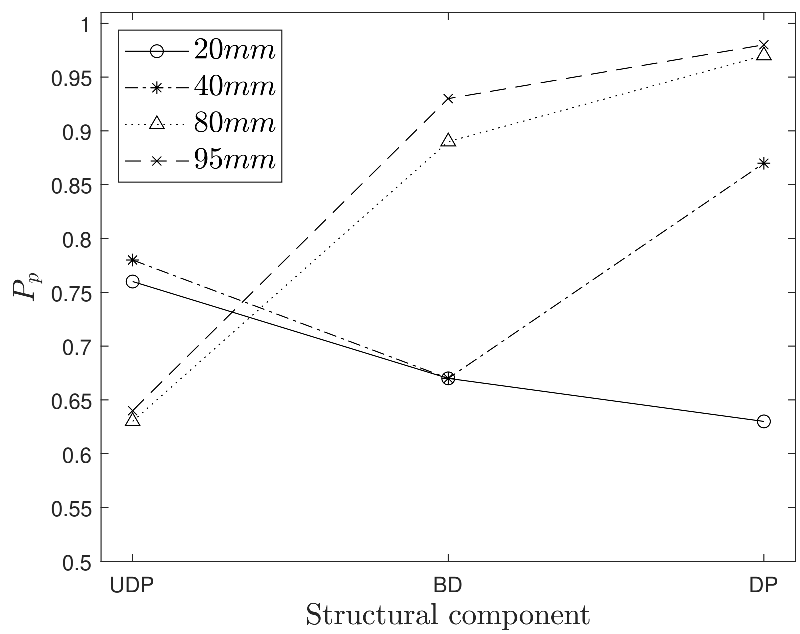


Fig. 12. Damage identification for the pier settlement.

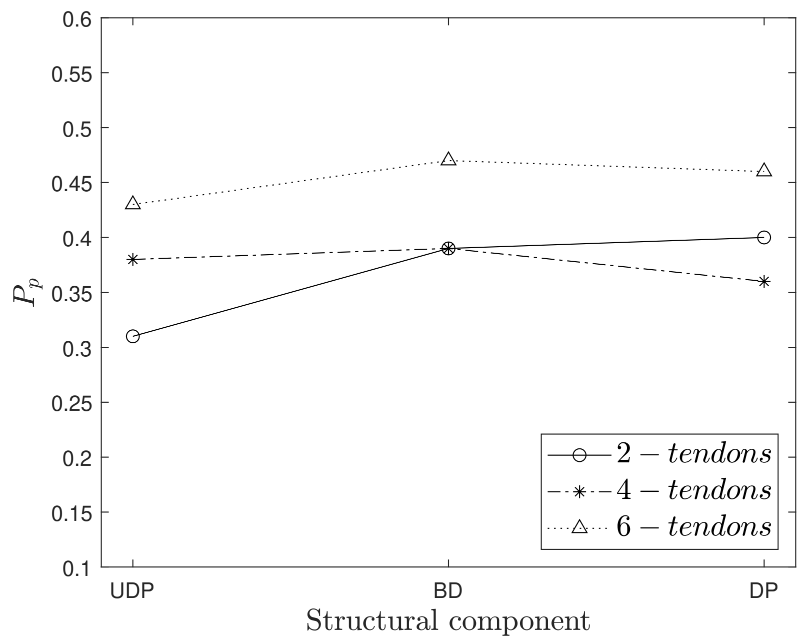


Fig. 13. Damage identification for the rupture of tendons.