

A Two-stage Method for Damage Detection in Z24 Bridge Based on K-nearest Neighbor and Artificial Neural Network

Viet Hai Hoang¹, Ngoc Long Nguyen¹, Tien Thanh Bui¹, Ngoc Hoa Tran^{1*}

¹ Department of Bridge Engineering and Underground Infrastructure, Faculty of Civil Engineering, University of Transport and Communications, No.3 Cau Giay Street, Hanoi 100000, Vietnam

* Corresponding author, e-mail: ngochoa@utc.edu.vn

Received: 18 November 2023, Accepted: 23 February 2024, Published online: 30 April 2024

Abstract

In this paper, we propose an effective approach for identifying damages in the Z24 bridge, a large-scale bridge in Switzerland. The dataset of the Z24 bridge is evaluated as a benchmark, reflecting the behavior of a real structure that has been used for numerous studies. However, most of the previous studies have only addressed the issues of updating the model or estimating the location and severity of damages. The core idea behind our proposed approach is to leverage the strengths of two effective Machine Learning (ML) algorithms: K-Nearest Neighbor (KNN) and Artificial Neural Network (ANN), to assess both the location and severity of damages in the Z24 bridge. First, we employ KNN, an unsupervised learning algorithm, for pinpointing the damage location. This strategy proves highly efficient, significantly reducing computation time by circumventing the need for a loss function during KNN training. By adopting this approach, KNN effectively mitigates the risk of encountering local minima in the ANN optimization process. Subsequently, we deploy ANN to determine the damage severity. When compared to previous studies on the Z24 bridge, our proposed method (KNN-ANN) exhibits promising results. Furthermore, our results illustrate that KNN-ANN consistently outperforms traditional ANN methodologies.

Keywords

damage detection, K-nearest neighbor, Artificial Neural Network

1 Introduction

Bridges are fundamental components of transportation infrastructure, and ensuring their safety is crucial for maintaining traffic flow and preventing accidents. However, bridges are susceptible to multiple environmental and operational variables such as vehicle loads, temperature fluctuations, aging, and so forth, which may lead to damage and degradation over time. Detecting damages early is essential to avert catastrophic failures and uphold bridge safety. Recently, diverse techniques, ranging from conventional visual inspections to sophisticated monitoring technologies, have been suggested for detecting damage in bridges [1–6].

Machine Learning (ML), a developing technology, involves the utilization of algorithms to analyze substantial amounts of data [7–9]. A subset of ML, the Artificial Neural Network (ANN), comprises a network of linked nodes or neurons that interact via weighted connections to process information [10]. ANNs have multiple benefits compared to conventional methods, including their ability to learn from complex and nonlinear data, their

resilience to noise and uncertainty, and their capability to generalize to new data [11].

ANNs have been effectively employed in various domains, including natural language processing, image processing, control systems, and Structural Health Monitoring (SHM). Nevertheless, one limitation of ANNs is their tendency to be trapped in local minima, potentially diminishing result accuracy. To address this issue, many researchers have suggested the use of optimization techniques to lessen the effects of local minima. For instance, Khatir et al. [12] presented a hybrid method combining Particle Swarm Optimization (PSO) and Teaching-Learning-Based Optimization for optimizing weight and bias parameters in ANNs. In a related study, Tran-Ngoc et al. [13] applied Cuckoo Search to find the best training parameters (weights and bias coefficients) for ANNs. This strategy was shown to be effective in preventing the network from getting stuck in local minima at the early stages of training.

Lin et al. [14] introduced an innovative approach for forecasting travel time on a signalized corridor,

incorporating exponential smoothing, ANN, and Bayes algorithms. The developed models exhibit proficiency in capturing travel time patterns, including non-recurring congestion, surpassing traditional methodologies. The obtained results demonstrate favorable performance, with error rates below the standard deviation threshold, substantiating the efficacy of the proposed methodology. In their study, Wang et al. [15] introduced a novel two-stage approach that integrated structural equation models and ANN to investigate the influence of live streaming on sales within the fresh food e-commerce sector amidst the COVID-19 pandemic. The results revealed that the proposed methodology achieved a notable 83.76% accuracy in forecasting live broadcast sales. These findings provide valuable insights into content optimization and user experience enhancement in this domain.

Chaki and Biswas [16] focused on utilizing ANN to investigate the application of ANN for the prediction and optimization of engine performance characteristics in a four-stroke diesel engine fueled by Nahar oil-based biodiesel. To accomplish this, a comprehensive full factorial experiment was conducted to evaluate the impact of engine revolutions per minute, engine load, and blend mixture on the engine's performance. The ANN-entropy-hybrid model was employed in the analysis, demonstrating high precision in its predictions with a maximum absolute percentage error below 3%. Sreekanth et al. [17] employed ANN to identify delamination in composites through the analysis of vibration signals. The research employed inverse methodologies, specifically Response Surface Methodology (RSM) and ANN, to estimate the size and location of delamination in beams with varying positions and sizes. Although optimization methods have partially solved the problem of local minima in ANNs, the process of searching for the optimal solution is time-consuming [18]. Therefore, it is necessary to come up with more practical solutions.

It is widely recognized that health monitoring studies of structures are primarily conducted using numerical models or laboratory experiments. Consequently, the results often fail to fully reflect the true nature of the structure, as the influence of noise phenomena affecting the obtained data is not adequately considered. The dataset from the Z24 bridge is a real-world dataset and considered a benchmark for SHM studies. In the initial studies, this data was mainly used for model updating purposes. Subsequently, the Z24 Bridge dataset has been employed for structural damage detection problems. However, most of these

studies only identify the damaged area and cannot accurately predict the extent of the damage. Although some recent studies have possibly properly identified the location and extent of damage, the computational cost is high. This poses a challenge when attempting to apply these methods to real-world problems [19].

In this paper, we propose a two-stage method for damage detection in the Z24 bridge using K-Nearest Neighbor (KNN) and ANN. The use of KNN with unsupervised learning capability is particularly effective in reducing computation time since no loss function or optimization problem needs to be performed during KNN training. This approach allows KNN to eliminate most of the local minima challenges typically faced by ANN. Subsequently, ANN is employed to determine the severity of the detected damage. Furthermore, to establish a baseline model for generating training data, a numerical model of the Z24 bridge is developed and updated based on measured results. After this update, the numerical and experimental models exhibit a higher degree of correspondence compared to previous studies [20, 21].

Beyond the introductory section, the paper is divided into three main parts. Section 2 offers an overview of the essential components: KNN, ANN, and the newly proposed KNN-ANN method. In Section 3, the effectiveness of the KNN-ANN approach is assessed using Z24 data. Finally, the paper concludes by summarizing the pivotal findings.

2 K-nearest neighbor - Artificial Neural Network (KNN-ANN)

In this section, a two-step damage identification method based on KNN and ANN will be proposed. KNN is a ML algorithm applied for classification and regression functions [22]. It is a non-parametric technique that involves identifying the K-closest data points to a specific query point and subsequently forecasting the output according to the class or value of those K-neighbors. KNN is considered a simple yet effective algorithm that doesn't require any training process. It operates by comparing the distances between the features of the query point and those of the training data points. The used distance metric can vary based on the problem, but the most common distance metrics are Euclidean distance as shown in Fig. 1 and Eq. (1):

$$\text{Euclidean}(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

where (x_1, y_1) , (x_2, y_2) are the coordinates of the points. Fig. 2 illustrates the classification of the Iris flower [23] and Seaborn using KNN [24].

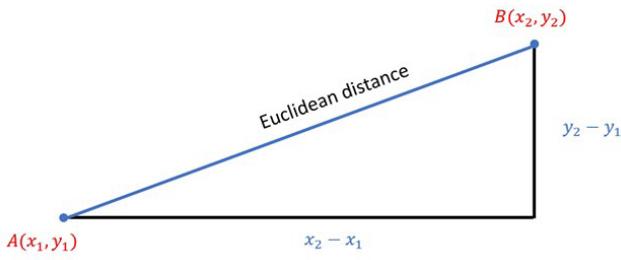
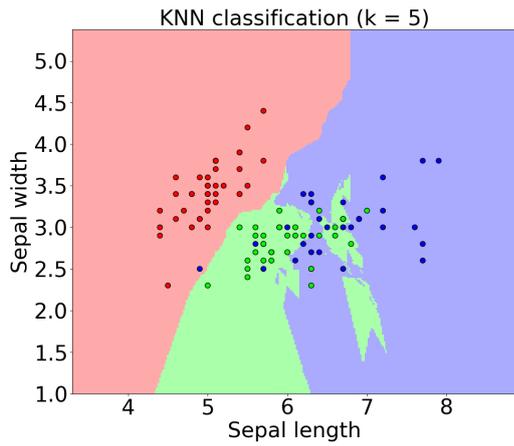
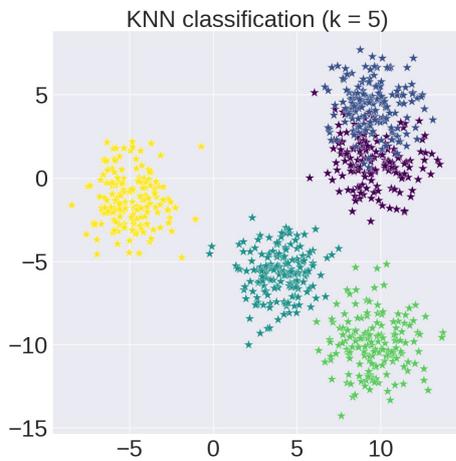


Fig. 1 Euclidean distance



(a)



(b)

Fig. 2 The classification uses KNN (a) Iris flower data, (b) Seaborn data

It is commonly acknowledged that KNN is very powerful when solving problems with the same properties of data [22]. With problems of structural damage localization, damage scenarios happening at elements have the same characteristics (data type). Therefore, KNN is suitable to identify damage locations of elements in the Z24 bridge. The classification problem can be approached using KNN, which determines the label of a new data point by considering its closest neighbors in the training set. Predictions for the test data can be made by means of either a majority voting or a weighted voting scheme, and then combining their contributions to derive the final label.

After the damage location is determined by using KNN, ANN is employed to identify damage level. A network includes one input layer, one hidden layer, and one output layer as Fig. 3 is applied.

The methodology for instructing the network involves utilizing a pair of equations, with the initial one being a summation function reliant on the training parameters (i.e., bias and weight) in conjunction with the input of the preceding layer as denoted by Eq. (2):

$$\varphi_{n_2}^{(1)} = \sum_{\tau_1, \tau_2}^{n_1, n_2} \mu_{n_1 n_2}^{(1)} \times \gamma_{n_1} + \varepsilon_{n_2}^{(1)}; \quad n_1 = (1: \tau_1); \quad n_2 = (1: \tau_2). \quad (2)$$

The value of $\varphi_{n_2}^{(1)}$ corresponds to the input of the hidden layer neuron at the n_2^{th} index, where $\mu_{n_1 n_2}^{(1)}$ and $\varepsilon_{n_2}^{(1)}$ represent the weight and bias coefficients, respectively, linking the n_1^{th} index neuron of the input layer to the n_2^{th} index neuron of the hidden layer. Here, τ_1 , and τ_2 signify the overall count of neurons present in the input layer and the hidden layer, respectively.

Subsequently, a sigmoidal function is employed to compute the resultant output of neurons in the hidden layer, which is denoted by $\theta_{n_2}^{(1)}$:

$$\theta_{n_2}^{(1)} = \frac{1}{1 + \exp^{-\varphi_{n_2}^{(1)}}}. \quad (3)$$

An analogous depiction of the process involved in transferring the neurons from the hidden layer to the output layer can be found in Eq. (4) and Eq. (5).

$$\varphi_{n_3}^{(2)} = \sum_{\tau_2, \tau_3}^{n_2, n_3} \mu_{n_2 n_3}^{(2)} \times \theta_{n_2}^{(1)} + \varepsilon_{n_3}^{(2)}; \quad n_3 = (1: \tau_3) \quad (4)$$

The quantities $\varphi_{n_3}^{(2)}$, $\mu_{n_2 n_3}^{(2)}$ and $\varepsilon_{n_3}^{(2)}$ fulfill similar functions as $\varphi_{n_2}^{(1)}$, $\mu_{n_1 n_2}^{(1)}$ and $\varepsilon_{n_2}^{(1)}$, respectively. Here, τ_3 correspond to the number of neurons in the output layer.

$$\theta_{n_3}^{(1)} = \frac{1}{1 + \exp^{-\varphi_{n_3}^{(1)}}}. \quad (5)$$

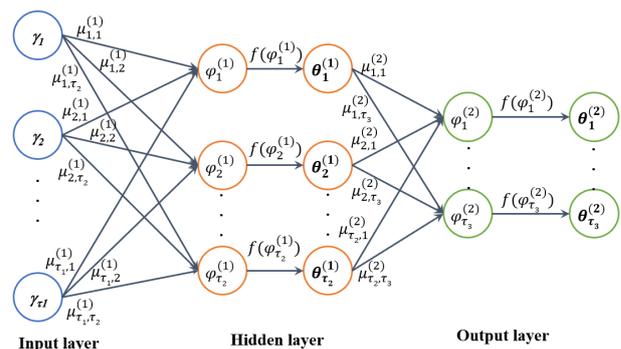


Fig. 3 ANN with one hidden layer

To achieve the global optimum in ML (the minimal values of the optimal function), commonly referred to as the loss function, is essential. However, accomplishing this task is a complex and challenging process owing to the intricate nature of the objective function's form. Typically, the prevailing approach involves commencing from a point that is near the solutions of the problems and subsequently employing iterative techniques to navigate towards the target points. This process employs backpropagation process.

In essence, the process of training a network involves the minimization of the disparity or deviation between the calculated output and the intended or desired output.

$$\Psi(\mu, \varepsilon) = \sum_{i=1}^{\tau_3} \frac{1}{2} \frac{(\theta_{n_3}^i - \bar{\theta}_{n_3}^i)^2}{\tau_3} \quad (6)$$

The training process involves obtaining new sets of parameters that connect the hidden layers and the output layer.

$$\begin{aligned} \mu_{n_2 n_3}^{2+} &= \mu_{n_2 n_3}^2 - \rho \times \frac{\partial \Psi(\mu)}{\partial \mu_{n_2 n_3}^2} \\ \varepsilon_{n_3}^{2+} &= \varepsilon_{n_3}^2 - \rho \times \frac{\partial \Psi(\varepsilon)}{\partial \varepsilon_{n_3}^2} \end{aligned} \quad (7)$$

Here, ρ represents the learning rate, $(\mu_{n_2 n_3}^2, \varepsilon_{n_3}^2)$ and $(\mu_{n_2 n_3}^{2+}, \varepsilon_{n_3}^{2+})$ denote the previous, and the new training parameter (weight and bias) connecting the hidden layer and the output layer, respectively. The $\frac{\partial \Psi(\mu)}{\partial \mu_{n_2 n_3}^2}$ represent the derivatives of the input of the output layer with respect to the weight and bias, respectively. Similarly, the training parameters for the input layer and the training layer are also redefined.

$$\begin{aligned} \mu_{n_1 n_2}^{1+} &= \mu_{n_1 n_2}^1 - \rho \times \frac{\partial \Psi(\mu)}{\partial \mu_{n_1 n_2}^1} \\ \varepsilon_{n_2}^{1+} &= \varepsilon_{n_2}^1 - \rho \times \frac{\partial \Psi(\varepsilon)}{\partial \varepsilon_{n_2}^1} \end{aligned} \quad (8)$$

3 Applications to the Z24 bridge

3.1 The description of the Z24 bridge

The Z24 overpass, situated in the west-central region of Switzerland, provided a vital linkage between the villages of Utzenstorf and Koppigen. Constructed of prestressed concrete, this structure featured an arrangement of 45 m, 75 m, and 45 m spans. The bridge was supported by a duo of abutments that were supported by concrete columns. The bridge was built in 1945; nevertheless, in 1998, it was dismantled since the need for a new bridge with a larger

expanse to accommodate the evolving railway infrastructure. The bridge's architectural features and dimensions are depicted in Fig. 4 [25].

Prior to its complete obliteration, the bridge was created various damage scenarios to assess the impact of diverse realistic damage cases on its dynamic response. A comprehensive account of the evolving damage scenarios can be found in [25].

3.2 A Finite Element Model (FEM)

A FEM of the bridge is created by using the Stabil toolbox, which was developed by [26] within the MATLAB environment. The model incorporates beam elements, with each node featuring six Degrees of Freedom (DOFs). The X -axis can be associated with the longitudinal direction, while the Y -axis can be linked to the vertical direction, and the Z -axis can be utilized to signify the horizontal one of the bridges. The main girder is modeled with a total of 94 beam elements, while the piers and abutments are simulated utilizing beam elements, as depicted in Fig. 5. The concrete is homogeneous, having an initial Young's modulus value of $E = 37.43$ GPa.

In order to acquire dynamic data including natural frequencies and mode shapes for the Z24 bridge, an experimental campaign was executed, with the specifics of the experiment outlined in [20]. Fig. 6 depicts mode shapes obtained from numerical model.

There is still a significant discrepancy in the natural frequency values between the calculated and the measured

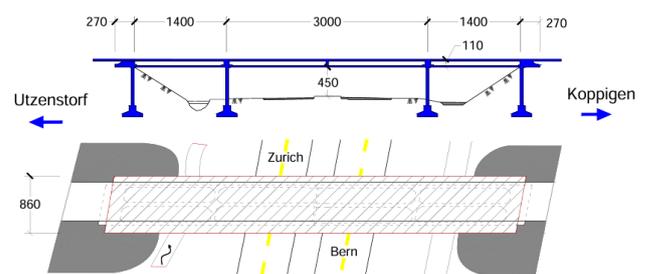


Fig. 4 The layout of the Z24 bridge

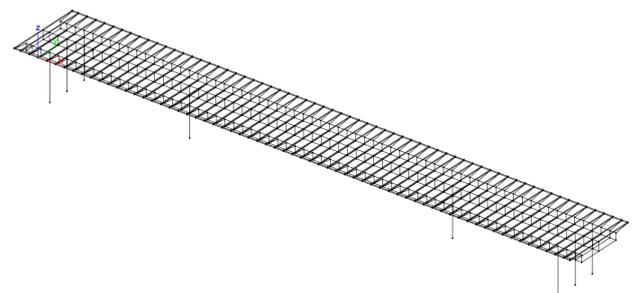


Fig. 5 The FEM of the bridge

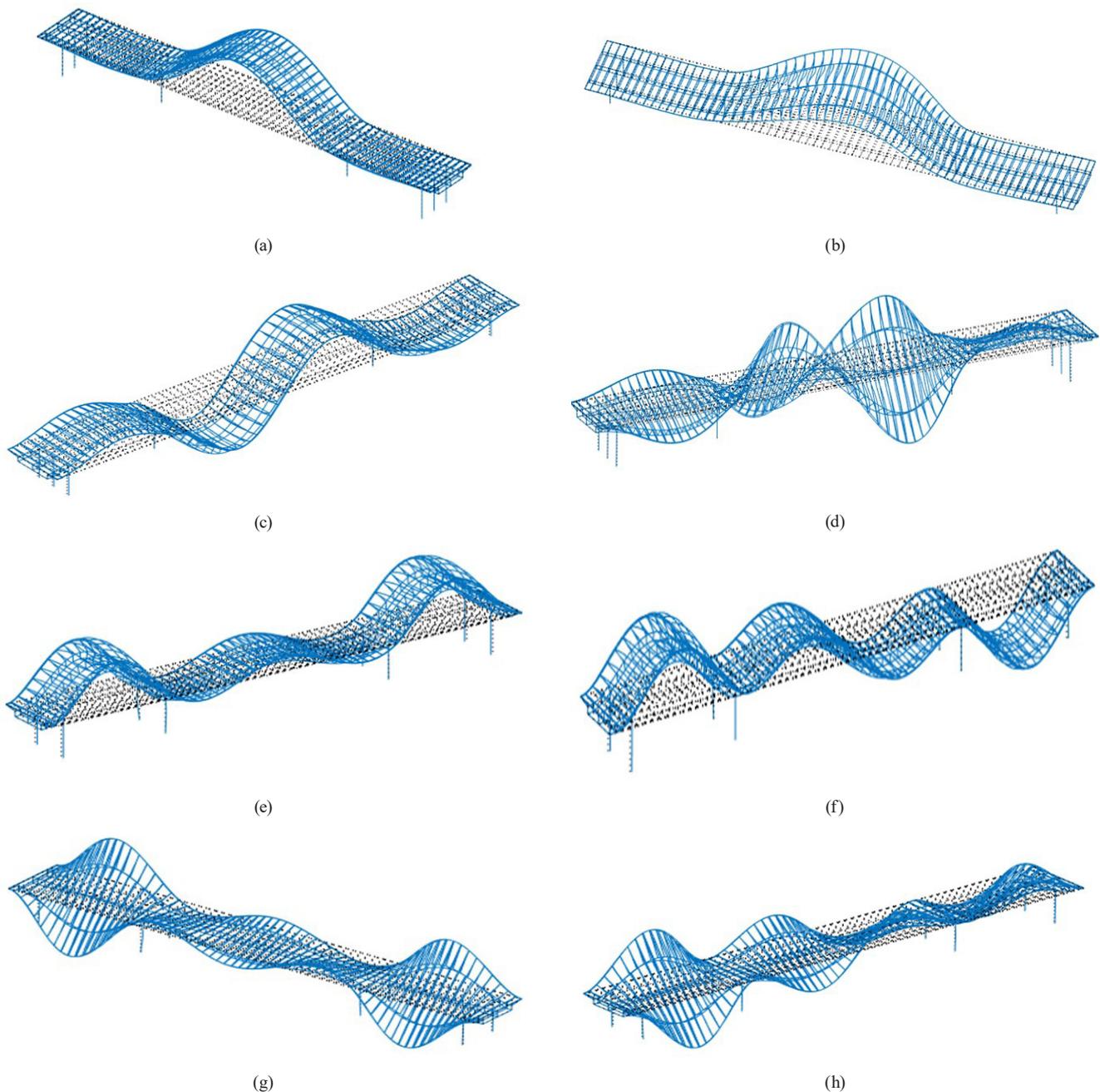


Fig. 6 Numerical modes ones: (a) Mode 1 – 3.96 Hz, (b) Mode 2 – 5.06 Hz, (c) Mode 3 – 9.85 Hz, (d) Mode 4 – 10.41 Hz, (e) Mode 5 – 12.94 Hz, (f) Mode 6 – 13.13 Hz, (g) Mode 7 – 18.40 Hz, (h) Mode 8 – 18.56 Hz

results in Table 1, ranging from 0.51% to 6.52%. These errors will be reduced by applying the model updating in Section 3.3.

3.3 Finite Element Model updating

In this section, the Z24 bridge model will be updated to generate training data for the network. PSO is employed to update the model. The parameters of PSO are employed as follows: c_1 and c_2 learning coefficients are assigned the value of 2; a population size of 50 is considered, and

the stopping criteria of the algorithm are either an error threshold of 10^{-5} between the computed and actual values or the number of iterations of 100. The findings obtained from the improved model are also contrasted with previous studies conducted on the Z24 bridge by other scholars.

From Table 2, the updated model presented in this paper is significantly superior to previous studies. In the study by Teughels and De Roeck [21] non-destructive, global damage identification technique, which is based on the fact that the modal parameters (eigenfrequencies

Table 1 Measured and numerical modes

Modes	Measured modes (Hz)	Numerical modes before model updating (Hz)
1	3.89	3.96 (1.76%)
2	4.93	5.06 (2.57%)
3	9.80	9.85 (0.51%)
4	10.34	10.41 (0.67%)
5	12.62	12.94 (2.47%)
6	13.34	13.13 (1.60%)
7	17.20	18.40 (6.52%)
8	19.28	18.56 (3.87%)
9	19.76	19.35 (2.11%)

Table 2 Measured and numerical modes

Modes	Measured modes (Hz) [20]	Numerical modes (Hz)-before model updating			
		Before	After [21]	After [20]	After – our model
1	3.89	3.96 (1.76%)	3.87 (0.25%)	3.86 (0.77%)	3.88 (0.25%)
2	4.93	5.06 (2.57%)	5.03 (1.98%)	5.04 (2.18%)	5.03 (1.98%)
3	9.80	9.85 (0.51%)	9.72 (0.82%)	9.72 (0.82%)	9.7 (1.03%)
4	10.34	10.41 (0.67%)	10.31 (0.29%)	10.33 (0.09%)	10.33 (0.09%)
5	12.62	12.94 (2.47%)	12.8 (1.40%)	12.74 (0.63%)	12.0 (5.16%)
6	13.34	13.13 (1.60%)	Missing	13.52 (1.33%)	12.7 (5.03%)
7	17.20	18.40 (6.52%)	Missing	Missing	18.26 (5.80%)
8	19.28	18.56 (3.87%)	Missing	21.55 (10.53%)	18.63 (3.49%)
9	19.76	19.35 (2.11%)	Missing	19.33 (2.24%)	20.04 (1.39%)

Table 3 Data used for KNN classification

f (Hz) The natural frequency of 3 modes of each sample	The number of samples (damages from 1%–50%)	Classes
$(f_1^{(1.1)}, f_2^{(1.1)}, f_3^{(1.1)}); \dots; (f_1^{(1.50)}, f_2^{(1.50)}, f_3^{(1.50)})$	50	1
$(f_1^{(2.1)}, f_2^{(2.1)}, f_3^{(2.1)}); \dots; (f_1^{(2.50)}, f_2^{(2.50)}, f_3^{(2.50)})$	50	2
$(f_1^{(3.1)}, f_2^{(3.1)}, f_3^{(3.1)}); \dots; (f_1^{(3.50)}, f_2^{(3.50)}, f_3^{(3.50)})$	50	3
.....
$(f_1^{(22.1)}, f_2^{(22.1)}, f_3^{(22.1)}); \dots; (f_1^{(22.50)}, f_2^{(22.50)}, f_3^{(22.50)})$	50	22
$(f_1^{(23.1)}, f_2^{(23.1)}, f_3^{(23.1)}); \dots; (f_1^{(23.50)}, f_2^{(23.50)}, f_3^{(23.50)})$	50	23
$(f_1^{(24.1)}, f_2^{(24.1)}, f_3^{(24.1)}); \dots; (f_1^{(24.50)}, f_2^{(24.50)}, f_3^{(24.50)})$	50	24
(3.69; 4.87; 9.18) [21]	1	24

and mode shapes, the relative error between measured and calculated results was quite large, and several modes, for example modes 6, 7, 8, and 9, were missing. In the study by Reynders et al. [20], which was conducted afterwards and improved the accuracy of the numerical model, the error between calculated and measured results is still large (10%), and mode 7 was still missing. After updating the model, the error between the calculation and measurement in this study is highly corresponded (the error ranges from 0.25% to 5.8% and is lower than previous studies.), indicating that the FEM constructed can be used to generate training data for KNN-ANN.

3.4 Damage identification

In this section, a two-step KNN-ANN based damage identification method will be proposed. First, the updated numerical model will be used to generate damaged cases. The extent of damage is determined based on the reduction in structural stiffness. For example, if the structure is undamaged, its stiffness is 100%. If the structure is damaged by 1%, the remaining stiffness would be 99% compared to the original stiffness. If the remaining stiffness is 80% of the initial stiffness, the level of damage would be 20%. In this work, the damage is generated by decreasing the stiffness of the elements within the range of 1% to 50%. Two datasets will be created for this purpose. The first dataset will be used for KNN, only labeling the damaged element positions for the datasets in Table 3 and Fig. 7. Since only 1/2 bridges are considered, elements 1 to 47 are considered.

Fig. 7 shows that the damage location is well-identified using KNN. For clearer visualization, only the data points belonging to classes 13 to 24 are shown. This is applied

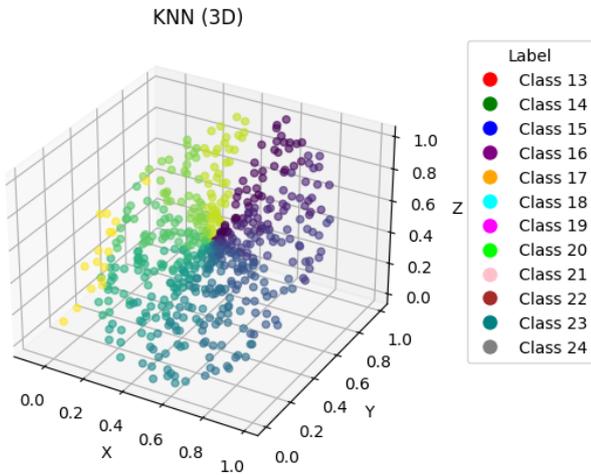


Fig. 7 Classification of damage location of the Z24 Bridge using KNN

to improve the clarity of the visualization, likely because showing all data points might be too cluttered or not informative. It suggests that the KNN algorithm is successful in determining where the damage has occurred.

After identifying the damaged element positions, the dataset will be used to identify the damage level by using ANN. To compare with the proposed method, the damage identification method using ANN will also be employed.

The network of both cases consists of three layers consisting of one input layer, one hidden layer, and one output layer (see Fig. 8 and Fig. 9). The data for training the

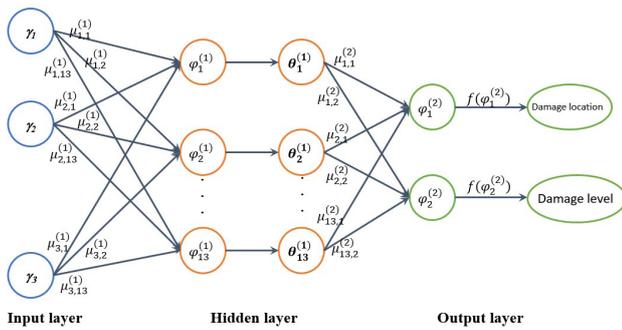


Fig. 8 ANN with 2 outputs

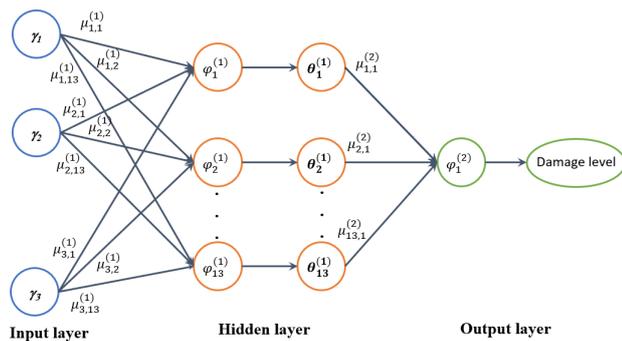


Fig. 9 ANN with 1 output

network will be the natural frequency of 3 first modes, whereas 13 neurons are utilized for the hidden layer. The datasets are randomly divided into training, testing, and validation sets, with a proportion of 70%, 15%, and 15%, respectively. Two metrics, namely regression (R) and Mean Squared Error (MSE), are utilized to evaluate the performance of the trained network. The Levenberg-Marquardt (LM) algorithm is employed for the network's training. The calculation of the sample size is determined by Eq. (9) considering the scenarios.

$$m = n_1 \times n_2 \quad (9)$$

The variable n_1 , represents the number of elements, which is assigned a value of 47. Similarly, the variable n_2 , signifies the number of damage scenarios and has a value of 50. In this study, a total of 2350 samples are utilized for training the network for ANN with 2 outputs, whereas, after using KNN to identify the damaged location, only 50 input data is used for (ANN with 1 output).

The diagram presented in Fig. 10 illustrates that both models demonstrate very good regression performance with very high R values (>0.99). This suggests that the models possess a robust ability to generate precise predictions and maintain strong generalization across the training, validation, and testing datasets. Notably, model in Fig. 10 (b), which has a single output, outperforms model in Fig. 10 (a), which has two outputs. This enhanced performance can be attributed to the strategic integration of the ANN and KNN algorithms, which appears to be more effective in this problem.

The chart depicted in Fig. 11 displays the discrepancies between the predicted value and the target value after the network was trained. The overall range of errors is divided into 20 separate columns, with the Y-axis denoting the proportion of samples from the dataset and the X-axis representing the variance between the computed outcome and the factual result. An error score of 0 signifies no difference between the two values. Most of the datasets are clustered closely to the 0-error column, implying a near-perfect match between the calculated and actual outputs. Fig. 10 and Fig. 11 also demonstrate that the proposed KNN-ANN method has higher accuracy compared to the conventional ANN method. Specifically, the discrepancy between the actual value and the predicted value for the KNN-ANN method is 0.03, whereas for the ANN method, it is 0.06.

Table 4 and Fig. 12 indicate that the proposed KNN-ANN method outperforms the method that only uses ANN for diagnosing the damage of Z24 bridge. Specifically, the

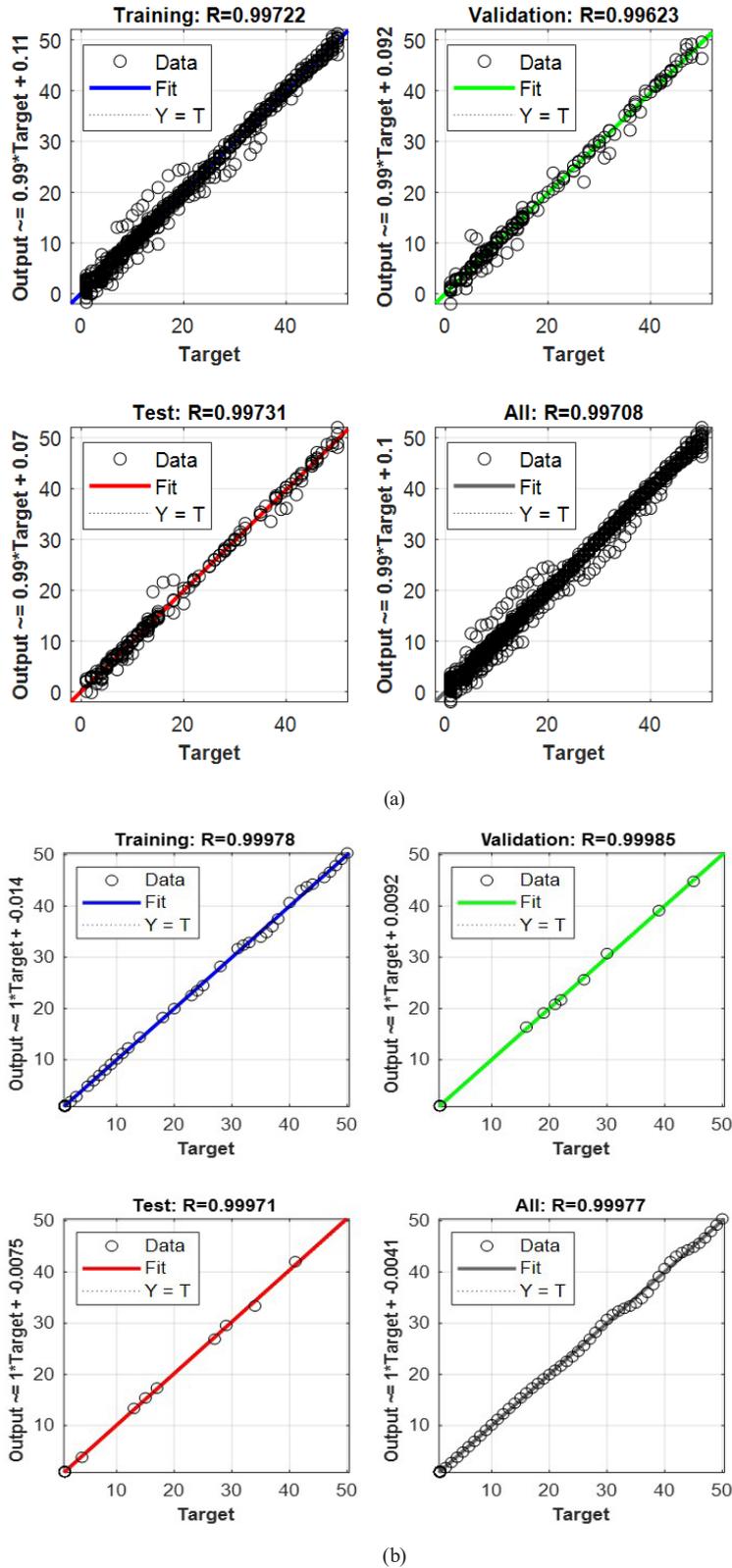
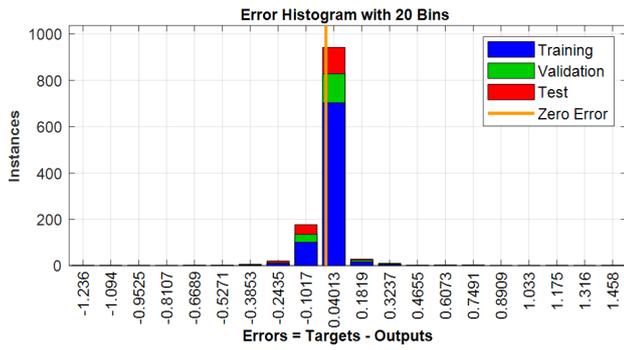


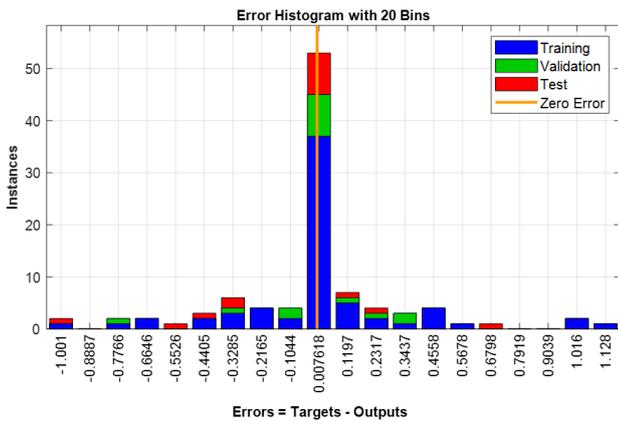
Fig. 10 Regression values (a) ANN with two outputs and (b) ANN with one output

MSE and R values computed by the KNN-ANN and ANN methods are (0.0048, 0.999) and (0.03, 0.997), respectively. The training time of the KNN-ANN method is also

shorter than that of ANN. The KNN-ANN method takes 43.21 s to train the network, while the ANN method takes 665.18 s for this process.



(a)



(b)

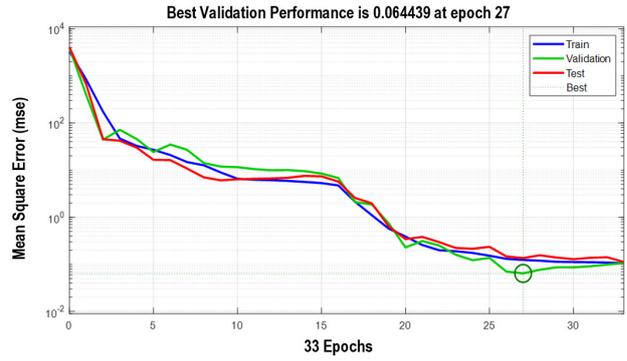
Fig. 11 Error histogram; (a) ANN with two outputs and (b) ANN with one output

Table 4 Performance evaluation metrics

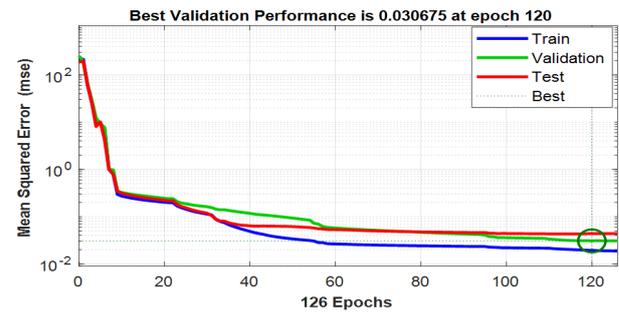
Methods	Mean Square Error (MSE)-Values	R-Values
ANN with two outputs	0.03	0.997
ANN with one output	0.0048	0.999

Fig. 13 illustrates the results of structural damage detection. The actual damage is created at the second pier with a stiffness reduction of 35%. These results have been presented in numerous previous studies [20, 21]. It can be observed that the proposed method accurately identifies the damaged area and provides a prediction of the damage level (32%) that closely aligns with the actual results (35%). In contrast, ANN incorrectly identifies both the damaged area and the extent of damage.

Specifically, ANN diagnoses damage occurring at element number 18 with a damage level of 26%. This demonstrates the advantage of combining the KNN and ANN algorithms in the task of structural damage identification. KNN is utilized to determine the damaged area, and then ANN is employed to assess the damage level accurately.



(a)



(b)

Fig. 12 Best validation performance (a) ANN with two outputs and (b) ANN with one output

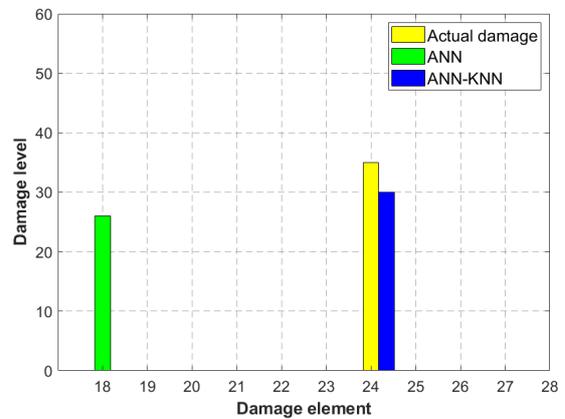


Fig. 13 Damage detection results at element 24 (on the top of the second pier) using ANN and KNN-ANN

4 Conclusions

The paper introduces an innovative two-stage damage detection method for the Z24 Bridge that combines KNN and ANN techniques to overcome the challenges faced by traditional damage detection methods. The KNN algorithm enhances the process by selecting salient features, while the ANN specializes in categorizing the extent of damage within the structure. This approach proves to be highly efficient in both identifying and pinpointing structural damage.

Empirical evidence from the study suggests that this method achieves remarkable accuracy in discerning structural damage. Initially, KNN refines the feature set for higher relevance, setting the stage for the subsequent analysis. Then, ANN takes over, applying these features to categorize data and assess the level of damage.

The effectiveness of the proposed technique is validated using data collected from the Z24 bridge. The results demonstrate that the proposed approach possibly accurately identify and localize damages in the bridge. The method also surpasses current damage detection techniques in both accuracy and computational efficiency. Notably, the Mean Squared Error (MSE) using the combined KNN-ANN method registers at a mere 0.0048, significantly lower than the 0.03 recorded when using ANN alone. Moreover, the KNN-ANN method provides a faster computation time, clocking in at 43.21 seconds, compared to ANN's 665.18 seconds.

References

- [1] Tran Anh, T., Hoang Viet, H., Do Anh, T., Tran Duc, N. "Effect of adhesion failure and temperature on the mechanical behavior of orthotropic steel bridge deck", *Transport and Communications Science Journal*, 73(1), pp. 52–60, 2022.
<https://doi.org/10.47869/tcsj.73.1.5>
- [2] Habashneh, M., Movahedi Rad, M. "An Investigation of the Recent Developments in Reliability-based Structural Topology Optimization", *Periodica Polytechnica Civil Engineering*, 67(3), pp. 765–774, 2023.
<https://doi.org/10.3311/PPci.22107>
- [3] Latifi Rostami, S. A., Li, M., Kolahdooz, A., Chung, H., Zhang, J. "Robust Topology Optimization of Continuum Structures under the Hybrid Uncertainties: A Comparative Study", *Periodica Polytechnica Civil Engineering*, 67(2), pp. 637–645, 2023.
<https://doi.org/10.3311/PPci.21562>
- [4] Pham, H.-A., Nguyen, B.-D. "Fuzzy Structural Analysis Using Improved Jaya-based Optimization Approach", *Periodica Polytechnica Civil Engineering*, 68(1), pp. 1–7, 2024.
<https://doi.org/10.3311/PPci.22818>
- [5] Rahimi, H., Esfandari, J., TahamouliRoudsari, M. "Experimental and Numerical Assessment of the Seismic Behavior of Non-uniform Slit Dampers and Bar Dampers in Moment Resisting Reinforced Concrete Frames", *Periodica Polytechnica Civil Engineering*, 68(1), pp. 314–324, 2024.
<https://doi.org/10.3311/PPci.22993>
- [6] Khatir, A., Capozucca, R., Khatir, S., Magagnini, E., Benaissa, B., Le Thanh, C., Abdel Wahab, M. "A new hybrid PSO-YUKI for double cracks identification in CFRP cantilever beam", *Composite Structures*, 311, 116803, 2023.
<https://doi.org/10.1016/j.compstruct.2023.116803>
- [7] Dang, H. V., Raza, M., Tran-Ngoc, H., Bui-Tien, T., Nguyen, H. X. "Connection stiffness reduction analysis in steel bridge via deep CNN and modal experimental data", *Structural Engineering and Mechanics*, 77(4), pp. 495–508, 2021.
<https://doi.org/10.12989/sem.2021.77.4.495>
- [8] Ho Viet, L., Trinh Thi, T., Ho Xuan, B. "Swarm intelligence-based technique to enhance performance of ANN in structural damage detection", *Transport and Communications Science Journal*, 73(1), pp. 1–15, 2022.
<https://doi.org/10.47869/tcsj.73.1.1>
- [9] Hoàng Việt, H., Đỗ Anh, T., Phạm Đức, T. "Utilizing artificial neural networks to anticipate early-age thermal parameters in concrete piers", *Transport and Communications Science Journal*, 74(4), pp. 445–455, 2023.
<https://doi.org/10.47869/tcsj.74.4.5>
- [10] Benaissa, B., Khatir, S., Jouini, M. S., Riahi, M. K. "Optimal Axial-Probe Design for Foucault-Current Tomography: A Global Optimization Approach Based on Linear Sampling Method", *Energies*, 16(5), 2448, 2023.
<https://doi.org/10.3390/en16052448>
- [11] Ghandourah, E., Khatir, S., Banoqitah, E. M., Alhawsawi, A. M., Benaissa, B., Wahab, M. A. "Enhanced ANN Predictive Model for Composite Pipes Subjected to Low-Velocity Impact Loads", *Buildings*, 13(4), 973, 2023.
<https://doi.org/10.3390/buildings13040973>
- [12] Khatir, S., Boutchicha, D., Le Thanh, C., Tran-Ngoc, H., Nguyen, T. N., Abdel-Wahab, M. "Improved ANN technique combined with Jaya algorithm for crack identification in plates using XIGA and experimental analysis", *Theoretical and Applied Fracture Mechanics*, 107, 102554, 2020.
<https://doi.org/10.1016/j.tafmec.2020.102554>
- [13] Tran-Ngoc, H., Khatir, S., Ho-Khac, H., De Roeck, G., Bui-Tien, T., Abdel Wahab, M. "Efficient Artificial neural networks based on a hybrid metaheuristic optimization algorithm for damage detection in laminated composite structures", *Composite Structures*, 262, 113339, 2021.
<https://doi.org/10.1016/j.compstruct.2020.113339>

Although the proposed method provided high accuracy in the problem of damage identification in the considered structures, it is necessary to extend the validation of the proposed methods with more diverse datasets from different structures to establish the robustness and generalizability of the approach.

In summation, this proposed method represents a practical and reliable strategy for damage detection and localization in large-scale infrastructural entities like bridges. Further research is encouraged to explore the applicability of this method in real-world scenarios and to expand its use across diverse structural types.

Acknowledgement

Ngoc Hoa Tran was funded by the Postdoctoral Scholarship Programme of Vingroup Innovation Foundation (VINIF), code VINIF.2023.STS.10.

- [14] Lin, W., Wei, H., Nian, D. "Integrated ANN-Bayes-based travel time prediction modeling for signalized corridors with probe data acquisition paradigm", *Expert Systems with Applications*, 209, 118319, 2022.
<https://doi.org/10.1016/j.eswa.2022.118319>
- [15] Wang, L., Li, X., Zhu, H., Zhao, Y. "Influencing factors of lives-tream selling of fresh food based on a push-pull model: A two-stage approach combining structural equation modeling (SEM) and artificial neural network (ANN)", *Expert Systems with Applications*, 212, 118799, 2023.
<https://doi.org/10.1016/j.eswa.2022.118799>
- [16] Chaki, S., Biswas, T. K. "An ANN-entropy-FA model for prediction and optimization of biodiesel-based engine performance", *Applied Soft Computing*, 133, 109929, 2023.
<https://doi.org/10.1016/j.asoc.2022.109929>
- [17] Sreekanth, T. G., Senthilkumar, M., Reddy, S. M. "Natural Frequency based delamination estimation in GFRP beams using RSM and ANN", *Frattura ed Integrità Strutturale (Fracture and Structural Integrity)*, 16(61), pp. 487–495, 2022.
<https://doi.org/10.3221/IGF-ESIS.61.32>
- [18] Bui-Tien, T., Bui-Ngoc, D., Nguyen-Tran, H., Nguyen-Ngoc, L., Tran-Ngoc, H., Tran-Viet, H. "Damage Detection in Structural Health Monitoring using Hybrid Convolution Neural Network and Recurrent Neural Network", *Frattura ed Integrità Strutturale (Fracture and Structural Integrity)*, 16(59), pp. 461–470, 2022.
<https://doi.org/10.3221/IGF-ESIS.59.30>
- [19] Tran-Ngoc, H., Khatir, S., De Roeck, G., Bui-Tien, T., Abdel Wahab, M. "An efficient artificial neural network for damage detection in bridges and beam-like structures by improving training parameters using cuckoo search algorithm", *Engineering Structures*, 199, 109637, 2019.
<https://doi.org/10.1016/j.engstruct.2019.109637>
- [20] Reynders, E., Teughels, A., De Roeck, G. "Finite element model updating and structural damage identification using OMAX data", *Mechanical Systems and Signal Processing*, 24(5), pp. 1306–1323, 2010.
<https://doi.org/10.1016/j.ymssp.2010.03.014>
- [21] Teughels, A., De Roeck, G. "Structural damage identification of the highway bridge Z24 by FE model updating", *Journal of Sound and Vibration*, 278(3), pp. 589–610, 2004.
<https://doi.org/10.1016/j.jsv.2003.10.041>
- [22] Zhang, Z. "Introduction to machine learning: k-nearest neighbors", *Annals of Translational Medicine*, 4(11), 218, 2016.
<https://doi.org/10.21037/atm.2016.03.37>
- [23] Gupta, T., Panda, S. P. "A comparison of K-means clustering algorithm and CLARA clustering algorithm on iris dataset", *International Journal of Engineering & Technology*, 7(4), pp. 4766–4768, 2018.
<https://doi.org/10.14419/ijet.v7i4.21472>
- [24] Waskom, M. L. "seaborn: statistical data visualization", *Journal of Open Source Software*, 6(60), 3021, 2021.
<https://doi.org/10.21105/joss.03021>
- [25] Brincker, R., Zhang, L., Andersen, P. "Modal identification of output-only systems using frequency domain decomposition", *Smart Materials and Structures*, 10(3), 441, 2001.
<https://doi.org/10.1088/0964-1726/10/3/303>
- [26] François, S., Schevenels, M., Dooms, D., Jansen, M., Wambacq, J., Lombaert, G., Degrande, G., De Roeck, G. "Stabil: An educational Matlab toolbox for static and dynamic structural analysis", *Computer Applications in Engineering Education*, 29(5), pp. 1372–1389, 2021.
<https://doi.org/10.1002/cae.22391>