Generalized Autoencoder-based Transfer Learning for Structural Damage Assessment

File Name: E6691.2021Spring.GADA.report.ll3097
Student: Lechen Li, ll3097
Columbia University

Abstract

The challenges of modeling for dynamical systems and the advanced machine learning methodologies have indicated a new direction for the structural damage assessment community. More and more data-driven approaches have been employed to extract damage sensitive features based on the monitored dynamic response of the structure, which have shown significant power and efficiency in structural-health-monitoring tasks. In this study, we implement a transfer-learning strategy by developing a generalized autoencoder based on a simulated dynamic structure. The autoencoder can then be transferred to learn the status of a real-bridge structure. Cepstral coefficients, generated from time-acceleration response, are employed to obtain effective damage sensitive features for the damage assessment. The main advantage of this approach is that the developed autoencoder can be well-trained by virtues of transfer learning which helps overcome the problem of the small data size under a real-life situation. The effectiveness and high accuracy of damage prediction have been validated through numerical simulations and experimental data.

1. Introduction

Over recent years, phenomenal advances in sensors and computer technologies have provided various promising structural-health-monitoring (SHM) techniques. Data obtained from sensors embedded in a structure can not only offer the advantage of reducing operational costs, but also guarantee a continuous and effective analysis of the structural condition. In terms of the techniques, vibration-based monitoring approaches allow engineers and researchers to collect data in real time, which makes it possible to assess damages of a structure online timely, and can even make predictions of its future status.

In terms of all possible structural characteristics and features (e.g., natural frequencies, mode shapes, stiffness coefficients, Autoregressive (AR) coefficients, etc.), the ones extracted from the structural response time history through digital signal processing or other data-driven approaches can be the most appealing due to their extraction computational efficiency without required high-level expertise from users. One type of such features are the power cepstrum coefficients, which was firstly introduced by Bogert et al. [1], where they worked on cepstrum domain to develop a method to detect an echo in a sound signal. Cepstrum-based features were later employed in structural damage assessment by Zhang et al. [2], where they generated mel-frequency cepstral coefficients (MFCCs) to characterize the bridge deck acoustic response to ultrasonic pulses for a study of the delamination of concrete bridge decks. After that, Balsamo et al. [3] employed the MFCCs of the vibration-based response of buildings and bridges as damage sensitive features, with a novelty detection strategy integrated with statistical analysis to conduct the damage assessment. In 2019, Civera et al. [4] developed an alternative of the original MFCCs, which is termed as the Teager-Kaiser Energy Cepstral Coefficients. Marcello et al. [5] conducted a vibration-based analysis and derived an analytical expression of the power cepstral coefficients of structural acceleration responses. They further proposed a damage assessment approach based on the power cepstral coefficients, principal component analysis, and statistical pattern recognition, and the corresponding numerical results showed appealing effectiveness of the proposed approach.

Recently, deep-learning-based approaches have shown powerful computational abilities and appealing performance in structural damage assessment by using monitored structural response data [6]. Abdeljaber et al. [7] developed a one-dimensional (1D) CNNs by using an adaptive strategy to detect loosened bolts at the joints of a steel frame. Their numerical results showed that the proposed CNN architecture can handle structural damage detection at a high-level accuracy. Overall, there have been many researchers employing supervised-learning strategies to train their neural networks to realize the damage assessment.

However, even though supervised strategies can produce accurate damage-assessment results, they commonly require numerous data from intact structures and various damaged structures for network training, which can be impractical and limited to be used for the damage assessment tasks of real-life infrastructures [8]. Therefore, many studies over recent years were conducted to investigate the approaches of unsupervised learning for the damage assessment. Rafiei et al. [9] employed several deep restricted Boltzmann machines (DRBMs) to estimate the both local and global conditions of a small-scale multi-story reinforced concrete building structure. From the automatic extraction of features from the vibration data of the building structure with DRBMs, their experimental results demonstrate the effectiveness of the proposed unsupervised method in the assessment of severe and near-collapse damaged cases. In this study, a generalized autoencoder (GAE) architecture is developed based on a transfer-learning strategy. The cepstral coefficients of the acceleration responses of both a simulated and a real-bridge structures are extracted and used as the inputs of the autoencoder, which can greatly improve the computational speed of the network and damage-or-not classification accuracy.

2. Summary of the Original Paper

2.1. Methodology of the Original Paper

For the original paper, they develop a classical deep auto-encoder (CAE) as an automatic extractor of damage-sensitive features from the acceleration responses of a simulated and a lab-based structures, with a one-class support vector machine (OC-SVM) as a damage detector. The flow chart of the original damage-assessment procedure is shown as Fig. 1, and the architecture of the developed autoencoder is visualized as Fig. 2.

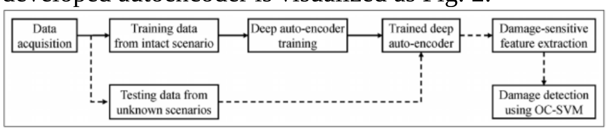

Figure 1. Schematic diagram of the unsupervised damage assessment method of the original paper.

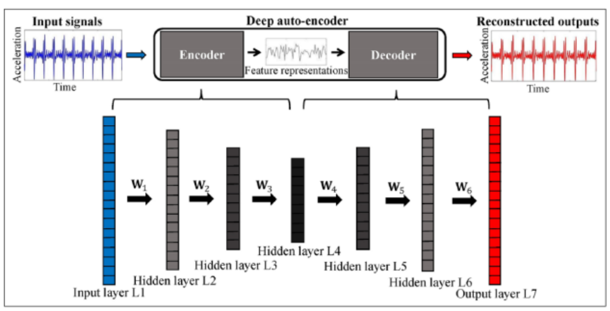

Figure 2. The architecture of the established autoencoder of the original paper.

2.2 Key Results of the Original Paper

After training and testing processes, the autoencoder of the original paper can provide the values of two damage sensitive features based on the acceleration response reconstruction, namely mean squared error (MSE) and original-to-reconstructed-signal ratio (ORSR), which are then employed as a two-dimensional feature space for the OC-SVM to handle the damage-or-not classification. The classification accuracies of the laboratory-scale steel bridge are shown in Table 1.

Table 1.
Classification results of the simulated system in the original paper. Damage detection accuracies calculated using an OC-SVM on a two-dimensional plane of features. IS represents undamaged scenarios, and DS1, DS2, DS3, and DS4 are four damaged scenarios.

Laboratory-scaled steel bridge	Tuned kernel parameter, γ	Damage detection accuracy				
		IS	DS1	DS2	DS3	DS4
Sensor #1	2	86.0%	100%	98.7%	96.5%	100%
Sensor #2	3	90.0%	93.3%	100%	100%	96.2%
Sensor #3	3	84.0%	86.7%	87.2%	100%	100%
Sensor #4	15	96.0%	71.1%	98.7%	92.9%	100%
Sensor #5	12	100%	94.4%	100%	100%	100%
Sensor #6	16	96.0%	41.1%	56.4%	100%	100%
Sensor #7	13	100%	21.1%	12.8%	92.9%	94.9%
Sensor #8	22	98.0%	11.1%	57.7%	100%	100%
Sensor #9	14	82.0%	27.8%	29.5%	100%	100%
Sensor #10	2	96.0%	100%	100%	100%	100%
Average (85.8%)		92.8%	64.7%	74.1%	98.2%	99.1%

OC-SVM: one-class support vector machine; IS: intact scenario; DS: damage scenario.

3. Method

In terms of our architecture, a GAE is set up for a pre-training over the cepstral coefficients of the simulated structure response data, to learn the fundamental vibration theory of structural dynamics. Then, the pre-trained GAE is converted into a CAE by replacing the compiled loss function by a mean squared error (MSE), and update the weights by fine-tuning through the cepstral coefficients of a new training set of the simulated structure or the real-bridge structure. After the fine-tuning process, the well-trained autoencoder will be used for damage sensitive feature extraction and the damage-or-not classification. In specific, this method section is organized as follows: Section 3.1 introduces the architecture of the developed GAE. Section 3.2 provides a summarized analytical expression of the cepstral coefficients. Section 3.3 illustrates the advantages and improvements of this study compared to the approach of the original paper.

3.1. Generalized Autoencoders

As shown in Fig. 3, the GAE differs from the CAE in two aspects, (i) the GAE learns a compressed representation y_i for an instance xi, and builds the relation between xi and other data $\{x_j, x_k, ...\}$ by using the compressed feature y_i to reconstruct each element in the set, not just xi itself. (ii) The GAE imposes a relational weight s_{ij} on the reconstruction error $\|x_j - x'_i\|^2$. Hence, the GAE captures the structure of the data space through minimizing the weighted reconstruction error [10].

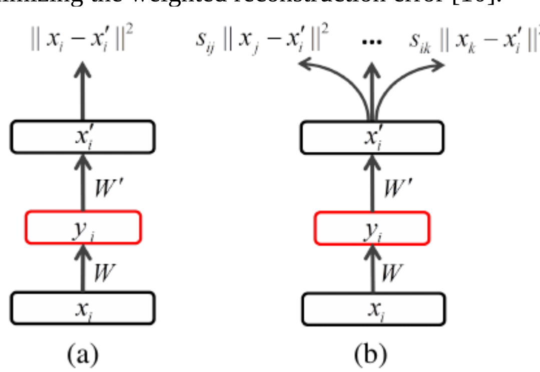

Figure 3. A generalized autoencoder. x_i involves the reconstruction of a set of instances $\{x_j, x_k, ...\}$. Each reconstruction error $s_{ij}\|x_j - \mathbf{x}_i\|^2$ measures a weighted distance between x_j and x_i.

There are several different types of the GAE, and we decided to use one of the proposed structures in the original paper, which is termed as GAE-LDA [10]. A modification implemented for the original loss function, which is defined as follows in this study:

$$L(\boldsymbol{W}, \boldsymbol{W}^{*}) = \sum_{i=1}^{N} \left\| \left(s \sum_{\substack{d=1; \\ d \neq j}}^{8} \boldsymbol{x}_{i(d)}\right) - D \circ \left(\boldsymbol{W}^{*}\left(E \circ \left(\boldsymbol{W}\, \boldsymbol{x}_{i(j)}\right)\right)\right) \right\|^{2} \quad (1)$$

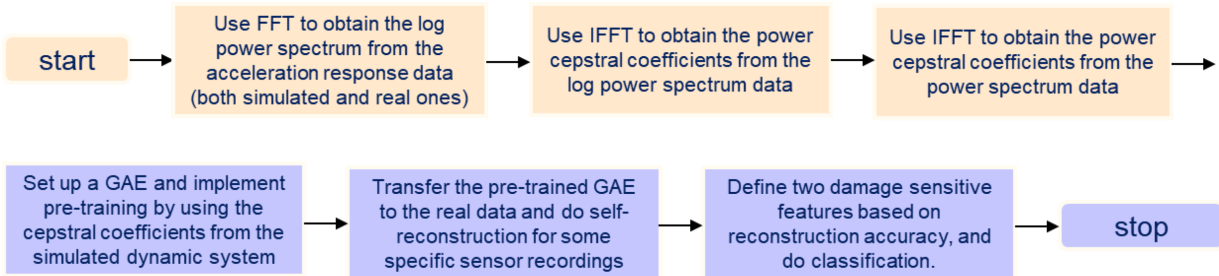

Figure 4. Block diagrams of the proposed methodology

3.2. Cepstral coefficients

In this study, we employed the same feature-extraction approach to generate the cepstral coefficients of the structural acceleration responses as introduced in [5]. The generated cepstral coefficients were then used as the inputs and outputs of the GAE for the signal reconstruction. The analytical expressions of the cepstral coefficients can be written as two parts as follows:

$$c_i(q) = \theta(q) + \gamma_i(q) \quad (2)$$

where i is the index of DOF, and the two parts can be expressed as follows:

$$\begin{cases} \theta(q) = \dfrac{1}{q} \displaystyle\sum_{l=1}^{N} 2e^{-\xi_l \omega_l T q} \cos\left(\omega_{l,d} T q\right) - 1 \\ \gamma_i(q) = -\dfrac{1}{q} \displaystyle\sum_{l=1}^{M} z_l^{(i)q} \end{cases} \quad (3)$$

where q is the index of cepstral coefficients.

3.3. Damage sensitive features

Two evaluation metrics are employed as the damage sensitive features in this study, namely coefficient of variation (CV), and standard-deviation ratio (STDR) between original and reconstructed signals. They are defined as follows:

$$CV = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\boldsymbol{x}_{i(j)} - \boldsymbol{x}_{i(j)}^{*}\right)^{2}}}{\overline{\boldsymbol{x}}_{(j)}} \quad (4)$$

$$STDR = \frac{1}{n}\sum_{i=1}^{n} \frac{std\left(\boldsymbol{x}_{i(j)}\right)}{std\left(\boldsymbol{x}_{i(j)}^{*}\right)} \quad (5)$$

3.4. Improvement of Method

In this study, first, cepstral domain signals, rather than acceleration response, are used as the features for autoencoders as they can provide a more robust representation of structural properties including natural frequencies and mode shapes. In addition, cepstral coefficients can significantly reduce the training times. Second, this project includes real-life structure dataset with more convincing damage assessment results. Third, a transfer-learning strategy by a GAE is implemented, which can obtain a more generalized model and solve data-limitation problems of structural health monitoring.

4. Implementation

4.1 Dataset

Our datasets include time-acceleration responses of a simulated lumped mass model of an 8 DOF shear-type system, and a real-life bridge. For the simulated 8 DOF system, as shown in Fig. 5 the mass and spring elements are numbered in ascending order from the ground constraint to the top. The system in its undamaged condition is characterized by horizontal springs of stiffness k_i = 25;000 N/m (i = 1, ..., 8), and each mass is equal to m_i = 1 kg (i = 1, ..., 8). The frame is supposed to have modal damping with a damping factor of ξ = 1% for each of the 8 vibration modes. Total 5 different structural conditions are considered in the analysis including 1 undamaged scenario and 4 different damaged scenarios (Table 2).

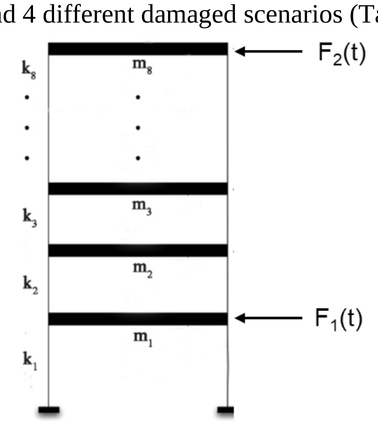

Figure 5. The simulated 8DOF system, and the excitation is applied at either DOF1 or DOF8, with the probabilities of 70% and 30% respectively.

Table 2.
Damaged and undamaged scenarios. k_i^0 = 25,000 N/m which represents the baseline stiffness value.

Scenario	Condition	Anomalies
1	Undamaged	$k_i = 1.0k_i^0$ for i = 1, 2, …,8
2	Damage#1	$k_i = 0.9k_i^0$ for i = 1
3	Damage#2	$k_i = 0.9k_i^0$ for i = 3
4	Damage#3	$k_i = 0.9k_i^0$ for i = 5
5	Damage#4	$k_i = 0.9k_i^0$ for i = 3, 7

The real-bridge dataset is from a case study of Z24 bridge, which was a post-tensioned concrete box girder bridge in Switzerland, with a main span of 30 m and two 14 m side spans [5]. A network of 16 accelerometers recorded structural accelerations at strategic locations on the bridge structure: for every hour, a total number of 65,536 samples (with a sampling time of 0.01 s) were recorded by each accelerometer, using an anti-aliasing filter with 30-Hz cutoff frequency.

For the modeling, we first simulated 800 realizations of the acceleration response of the 8DOF-system undamaged scenario, and then divided the 800 ones into two equal parts that 400 each. The two-part datasets are respectively termed as training set I and II. Then we generated the cepstral coefficients of the index 2-51 (q = 2, 3, …, 51) of each realization as the training-set I and II for the GAE. Next, another 200 realizations of the cepstral coefficients (q = 2, 3, …, 51) of each scenario (including the undamaged one) are generated in a same way, as the testing-set I. In terms of the real bridge, the acceleration recordings of 4 sensors (namely termed as Sensor 1, 2, 4, and 5) are considered in this study, which provides a training-set III that includes two undamaged scenarios with 310 realizations of each, and a testing-set II that includes one undamaged scenario with 100 realizations, and two damaged scenarios with 96 realizations of each.

4.2. Block Diagram and Algorithm

Fig. 4 is the block diagram of the overall damage assessment methodology, including the processes of the cepstral coefficient generation, setting up the GAE based on pre-training, post-training and testing, and the damage-or-not classification by the OC-SVM.

For the GAE architecture, we introduce a single-hidden-layer network, with the input and output are in 50 dimensions (same as the number of used cepstral coefficients), and the loss function is defined in eq. (3). Table 3 shows the algorithm of the training and testing procedures of the developed autoencoder: First, the cepstral coefficients of the training set I is used for pre-training of the GAE by 500 epochs. Second, the pretrained GAE is modified by changing its loss function into the mean squared error, and transfer the GAE to the training-set I or II for the post-training by 300 epochs. Therefore, during the post-training process, the GAE becomes a CAE that works on the signal reconstruction where the input is exactly the same as the output. Then, the fine-tuned autoencoder is tested by the testing set I and II, and then produce the results of the damage sensitive features (CV and STDR) of all training and testing sets for comparison. Finally, by employing the OC-SVM based on the obtained damaged sensitive features, we finally conduct the damage-or-not classification of the simulated 8DOF system or the Z24 bridge.

Table 3.

Algorithm 1: Generalized Autoencoder (GAE) Transfer Learning

Step 1: Prepare the training sets I, II, III, and testing sets I and II by generating the cepstral coefficients of each realization of the acceleration responses.

Step 2: Initialize a single-hidden-layer GAE, with the input and output are in 50 dimensions, and the hidden layer is in 10 dimensions, termed as G (W, b).

Step 3: Pretrain G (W, b) over the training set I (input: $x_{i(j)}$=[$c_{q(j)}(2)$, $c_{q(j)}(3)$, …, $c_{q(j)}(51)$] ; output: $x^*_{i(j)}$= [$s\sum_{d=1:d\neq j}^{8} c_q(d)(2)$, $s\sum_{d=1:d\neq j}^{8} c_q(d)(3)$, …, $s\sum_{d=1:d\neq j}^{8} c_q(d)(51)$]) by 300 epochs.

Step 4: Replace the loss function of the pretrained G (W, b) by traditional MSE (or no changes if using a pre-summation strategy), and do the fine-tuning for G (W, b) over the training set II or III (input & output: $x_{i(j)}$=[$c_{q(j)}(2)$, $c_{q(j)}(3)$, …, $c_{q(j)}(51)$]) by 300 epochs.

Step 5: Test the trained G (W, b) over the testing set I or II, and obtain the CV and STDR.

Step 6: Use the OC-SVM, based on the CV and STDR, to finish the damage-or-not classification.

Here is the link to the codes of this project:

5. Results

5.1. Project Results

As illustrated before, we produced the results of the damage-or-not classification of both the simulated 8DOF system and the Z24 bridge. In terms of the damage sensitive features (CV and ORSR) generated by the fine-tuned autoencoder, the corresponding results of the 8DOF system are shown in Figure 6, and the ones of the Z24 bridge are shown in Figure 8.

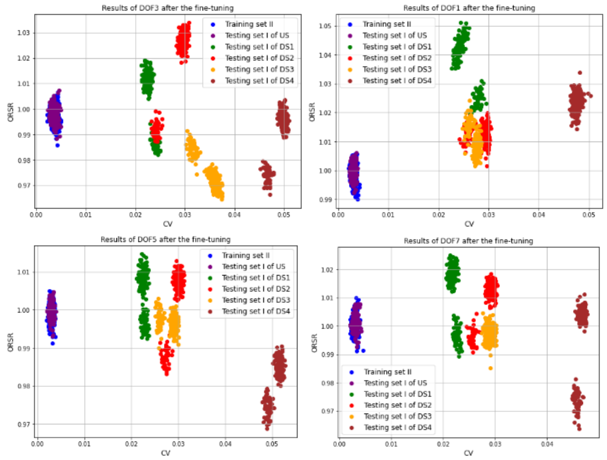

Figure 6. Damage-sensitive feature (CV and STDR) distributions on a 2D plane of features for each DOF of the simulated 8DOF system.

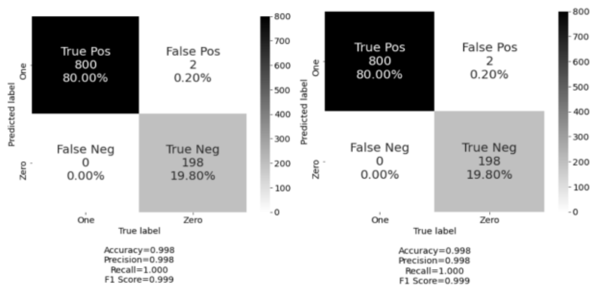

Figure 7. Confusion matrix of the classification results of the 8DOF system. (a) DOF#1 (b) DOF#7

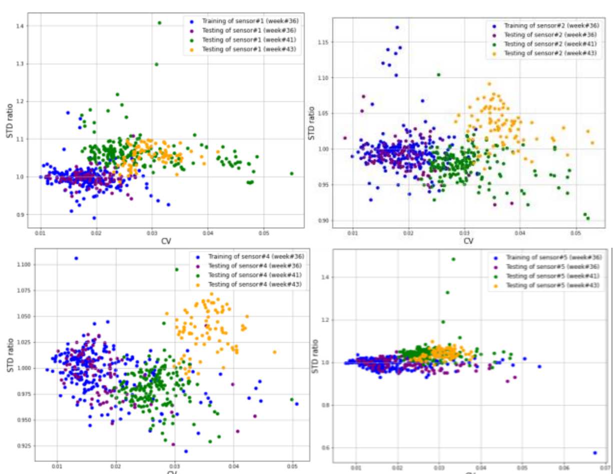

Figure 8. Damage-sensitive feature (CV and STDR) distributions on a 2D plane of features for each sensor of the Z24 bridge.

Then, the extracted CV and STDR were used as the input of the OC-SVM for classification. By tuning the kernel parameter γ, the best classification results of the testing sets of the simulated 8DOF system and Z24 bridge are shown in Fig. 7 and Fig. 9 of the confusion matrix.

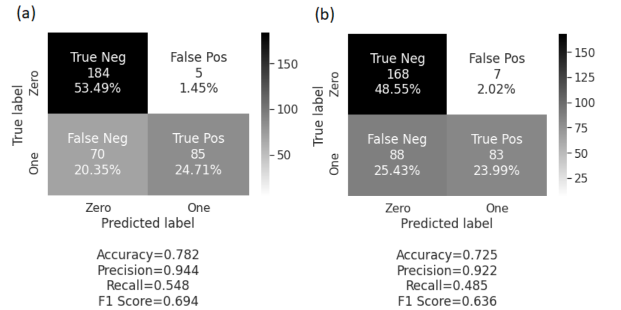

Figure 9. Confusion matrix of the classification results of the Z24 bridge. (a) Sensor # 2 (b) Sensor # 4

5.2. Comparison of the Results Between the Original Paper and the Study

In terms of the simulated structure, both the original paper and our study can produce a very high-level accuracy in the damage-or-not classification. However, there is a big difference in the used methods: In the original paper, time-acceleration-response data is used as the input and output of the autoencoder, which results in a time-consuming training process (approximate 30-40 minutes for each floor or sensor). For this study, due to the enhanced structural-property information obtained by the cepstral coefficients, a single-layer GAE is computational enough to handle the transfer learning, signal reconstruction, and the damage sensitive feature extraction within 10 minutes by a computer with an Intel Core i7-7700HQ CPU, 16 GB RAM, and a 64-bit operating system, which significantly accelerates the damage-assessment speed.

Regarding the Z24 bridge, it is a real-life bridge whose generated classification results are convincing and practical for the real-life applications of the structural health monitoring community. However, in the original paper, they only train their autoencoder by using the data from a simulated structure and laboratory-scale steel bridge, which may lead to a low-level generalized model for real-life damage assessment.

5.3. Discussion of Insights Gained

One meaningful finding of this study is that the developed GAE shows the potential abilities of removing the effects of various excitations applied on the dynamic system. As shown in section 4.1 and 5.1, for the simulated 8DOF system, the excitation is applied at either DOF1 or DOF8 which can generate two different distributions of the cepstral coefficients. However, after the fine-turning of the autoencoder, one can observe from Fig. 6 that the produced damaged sensitive features of the training sets (undamaged scenarios) gather into one single cluster, which means that the effects by different excitations have been removed. This greatly helps us further implement the abnormal detection for damage assessment by using the OC-SVM for the damaged scenarios.

As discussed in many literatures, a classical autoencoder whose input and output are exactly the same can only work on the limited representation-learning task, as its hidden outputs cannot be orthogonal to each other. By a gradient-descent-based algorithm, the weights of a classical autoencoder can only be closer and closer to a load-vector matrix (eigen-matrix) that can be obtained by singular-value-decomposition method or principal component analysis (PCA), but its column vectors can never be orthogonal to each other due to a redundant transformation in the hidden space by an inevitable matrix $\mathbf{T}$ [11]. The study in [5] has shown that the minor components generated by a PCA method can remove the different excitation effects, which can help the model concentrate on dynamic structural properties only without considering the effects brought by different monitoring locations. Therefore, the established GAE in this study can also provide a PCA-like results, which enables reconstruction of the cepstral coefficients and meanwhile, removes the adverse effects on damage assessment caused by different excitations of locations.

6. Conclusion

In this study, a generalized autoencoder (GAE) architecture is developed based on a transfer-learning strategy. The cepstral coefficients of the acceleration responses of both a simulated and a real-bridge structures are extracted and used as the inputs of the autoencoder, which can greatly improve the computational speed of the network and damage-or-not classification accuracy. Two limitations may exist for the proposed methodology: First, only one dataset of the Z24 bridge is considered in this study, and the developed autoencoder needs to be tested on more real-life dynamical systems. Second, the proposed method can only do damage-or-not classification in a global way, which means that we still need to further investigate the more advanced techniques for damage localization tasks.

7. Acknowledgement

The Z24 data have been collected within the Brite EuRam Programme BE-3157 SIMCES with a financial contribution from the European Commission. The data have been kindly provided by the project coordinator, the Structural Mechanics Section of KU Leuven. The authors gratefully acknowledge the Brite EuRam Programme BE-3157 SIMCES, the European Commission and the Structural Mechanics Section of KU Leuven for gathering and sharing the data of the Z24 bridge. This research was partially funded by the Center for Advanced Infrastructure and Transportation National UTC Consortium through the federal grant number DTRT13/G/UTC28.

References

[1] B. P. Bogert, "The quefrency alanysis of time series for echoes; cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking," *Time series analysis,* pp. 209-43, 1963.

[2] G. Zhang, R. S. Harichandran and P. Ramuhalli, "Application of noise cancelling and damage detection algorithms in NDE of concrete bridge decks using impact signals.," *Journal of Nondestructive Evaluation,* vol. 30, no. (4), pp. 259-72, 2011.

[3] . L. Balsamo, R. Betti and H. Beigi , "Structural damage detection using speaker recognition techniques," *In Proceedings of the 11th international conference on structural safety and reliability (ICOSSAR), Columbia University, New York, NY, USA*, 2013.

[4] M. Civera, M. Ferraris, R. Ceravolo, C. Surace and R. Betti, "The Teager-Kaiser Energy Cepstral Coefficients as an Effective Structural Health Monitoring Tool," *Applied Sciences,* vol. 9, no. (23), p. 5064, 2019.

[5] M. Morgantini, R. Betti and L. Balsamo, "Structural damage assessment through features in quefrency domain," *Mechanical Systems and Signal Processing*, no. 147, p. 107017, 2021.

[6] M. H. Rafiei, W. H. Khushefati, R. Demirboga and H. Adeli, "Supervised Deep Restricted Boltzmann Machine for Estimation of Concrete," *ACI Materials Journal,* vol. 114, no. (2), 2017.

[7] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. Journal of Sound and Vibration," *Journal of Sound and Vibration,* no. 388, pp. 154-170, 2017.

[8] Z. Wong and Y. J. Cha, "Unsupervised deep learning approach using a deep auto-encoder with a one-class support vector machine," *Structural Health Monitoring,* p. 1475921720934051, 2020.

[9] M. H. Rafiei and H. Adeli, "A novel unsupervised deep learning model for global and local health condition assessment of structures," *Engineering Structures,* no. 156, pp. 598-607, 2018.

[10] W. Wang, Y. Huang, Y. Wang and L. Wang, "Generalized autoencoder: A neural network framework for dimensionality reduction," *In proceedings of the IEEE conference on computer vision and pattern recognition workshops,* pp. 490-497, 2014.

[11] E. Plaut, "From principal subspaces to principal components with linear autoencoders," *arXiv preprint arXiv,* p. 1804.10253, 2018.

8. Appendix

8.1 Student info of the project

	UNI1
Lechen Li	ll3097