Available online at www.eccomasproceedia.org
Eccomas Proceedia SMART (2023) 1065-1075

ECCOMAS
Proceedia

X ECCOMAS Thematic Conference on Smart Structures and Materials
SMART 2023
D.A. Saravanos, A. Benjeddou, N. Chrysochoidis and T. Theodosiou (Eds)

REAL-TIME CONDITION MONITORING OF MECHANICAL SYSTEMS USING CNNS TRAINED BY MULTIBODY SIMULATIONS

JOSEF KOUTSOUPAKIS*, DIMITRIOS GIAGOPOULOS†

* Aristotle University of Thessaloniki
Department of Mechanical Engineering, Thessaloniki, Greece
e-mail: jkoutsoup@meng.auth.gr

† Aristotle University of Thessaloniki
Department of Mechanical Engineering, Thessaloniki, Greece
e-mail: dgiagopoulos@meng.auth.gr

Abstract. Vibration response analysis is perhaps the most common method of diagnosing the health state of a mechanical system and monitoring its condition. By analyzing a system's response and assessing certain quantities either in time or frequency domain, information on the system's condition can be collected and its maintenance can be scheduled accordingly. While conventionally performed by engineers monitoring certain features in a system's response, Condition Monitoring (CM) can be automated by the use of Artificial Intelligence (AI) methods, aimed at locating and identifying damages in a system and characterizing its health state. The obvious and significant shortcoming of this method is the need for large amounts of high-fidelity data which is necessary for training the AI algorithms and can at times be virtually impossible to acquire. To this end, a novel CM methodology is presented in this work, utilizing a Convolutional Neural Network (CNN) as a means of identifying the damages present in a system. To overcome the problem of data scarcity, optimal Multibody Dynamics (MBD) models are used as a means of mining high-fidelity training data for the network. The proposed framework is then validated by application on a real-life experimental system of an elevator door, where the trained CNN is proven capable of accurately identifying the damages present in the system and classifying its health state accordingly. The proposed methodology allows for real-time CM of virtually any mechanical system and can include any type of fault, given that the training data is procured by numerical simulations.

Key words: Condition Monitoring, Damage detection, Artificial Intelligence, Convolutional Neural Networks, Multibody model.

1 INTRODUCTION

Deep Learning (DL) algorithms have been extensively reviewed for Condition Monitoring (CM) and Structural Health Monitoring (SHM) on mechanical systems, with applications of Artificial Intelligence (AI) based CM/SHM being investigated as solutions to the damage detection problem in mechanical systems such as gearboxes and transmission systems [1-4], aircrafts [5], ground vehicles [6,7], bridges [8,9], elevators [10-13] and so on. Vibration based

Peer-review under responsibility of the organizing committee of SMART 2023.
doi: 10.7712/150123.9857.449373

methods are one of the most common practices for fault detection in mechanical systems [14-16] as accelerations measured from various locations of a structure can provide data on a system's health state. As is common however, faulty state data is usually diffcult to acquire due to the prohibiting costs of deliberately damaging a structure. Due to this, AI-based damage detection is usually limited to unsupervised cases, allowing only for anomaly detection or unlabeled clustering in a system [1,17-18]. To alleviate this problem of data scarcity, optimal Finite Element (FE) and Multibody Dynamics (MBD) models have been previously used to produce training data for damage detection both in the time and frequency domains [19,20].

In this work, data is generated by high-fidelity MBD simulations and used to train a CNN for damage classifciation on two damaged states of an experimental elevator door system. An optimal MBD model is first obtained by parameter updating and is then used to simulate the two fault cases. Data is then prodcuced by the optimal MBD models and is used to train a CM-CNN which is then validated on experimental measurements of the real set-up during normal operation as well as in the presence of faults. This novel framework may find applications both in laboratory and real-word tasks by providing an alternative to the acquisition of experimental training datasets via MBD models.

The rest of this paper is structured as follows: In Section 2, the theoretical formulation is explained. Section 3 presents the physical system examined in this work as well as its corresponding MBD model. Section 4 shows the structure of the CM-CNN and presents the training data generated by the MBD models. In Section 5 the CNN's prediction results are displayed, validating the network's performance on the experimental data. Finally, the conclusions are drawn in Section 6.

2 THEORETICAL FORMULATION

2.1 Contact force formulation

For a MBD system containing non-linear contact mechanics, the normal contact force between two contacting bodies can be estimated as:

$$F_n = K\delta^n + \chi\delta^n\dot{\delta} \tag{1}$$

where K is the contact stiffness, δ and $\dot{\delta}$ are the indentation depth and velocity respectively, n is the nonlinear exponent and χ is the hysteresis damping factor. This damping factor is what differentiates the models, as its formulation varies from one to another. The contact force models and corresponding damping factor formulations used in this work are presented on Table 1, where c_r denotes the coefficient of restitution and $\dot{\delta}^{(-)}$ is the indentation velocity at the start of each contact [21].

Table 1: Nonlinear contact normal force models and damping factors.

Gharib & Hurmuzlu	Herbert & McWhannel	Flores et al.	Lee & Wang
$\frac{1}{c_r}\frac{K}{\dot{\delta}^{(-)}}$	$\frac{6(1-c_r)}{(2c_r-1)^2+3}\frac{K}{\dot{\delta}^{(-)}}$	$\frac{8(1-c_r)}{5c_r}\frac{K}{\dot{\delta}^{(-)}}$	$\frac{3(1-c_r)}{4}\frac{K}{\dot{\delta}^{(-)}}$
Hunt & Crossley	**Lankarani & Nikravesh**	**Zhiuing & Qishao**	**Gonthier et al.**
$\frac{3(1-c_r)}{2}\frac{K}{\dot{\delta}^{(-)}}$	$\frac{3(1-c_r^2)}{4}\frac{K}{\dot{\delta}^{(-)}}$	$\frac{3(1-c_r^2)e^{2(1-c_r)}}{4}\frac{K}{\dot{\delta}^{(-)}}$	$\frac{1-c_r^2}{c_r}\frac{K}{\dot{\delta}^{(-)}}$

2.2 MBD model optimization

Aiming to acquire an accurate MBD model and mitigate the differences between the simulated and actual system responses, a model parameter vector θ_n containing the contact normal force stiffness K, nonlinear exponent n, coefficient of restitution c_r, contact normal force model index I, static and dynamic coefficients of friction μ_s, μ_d, flexible body damping ratio C and part mass scaling factor MF is updated using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [19-20]. The algorithm optimizes the model by updating the values of vector θ_n in an attempt to minimize an objective function denoted as $J(\theta_n)$:

$$J(\theta_n) = \sum_{i=1}^{M} \sqrt{\sum_{j=1}^{N} \frac{\left(\hat{y}_{ij}(\theta_n) - y_{ij}\right)^2}{N}} \tag{2}$$

Here, y_{ij}, $\hat{y}_{ij}(\theta_n)$ are the PSD estimates of the physical and MBD system's accelerations respectively, measured at accelerometer channels i, ranging up to M channels, over the frequency values j, ranging up to N Hertz.

2.3 Convolutional Neural Networks

CNNs have proven to be efficient damage detectors both in time and frequency domain [19-20]. CNNs receive input data which they convolve with a previously learned filter which is comprised of several matrices of weights, commonly know as kernels. The CNN's output f_k of length t is calculated by convolving the input signal in and applying a filter w_k on a neuron k as shown in the following equation:

$$f_k = w_k * in \text{ with } (w_k * in) = \sum_j in(j) w_k(t-j+1) \text{ (j takes all valid values)} \tag{3}$$

The output of the convolutional layer is then fed to the hidden layers of the nework where more processing is performed before resulting in the final output. The CNN is trained through Backpropagation (BP) on a labeled dataset, updating the weights and biases of each neuron based on the minimization of a selected cost function. In this work, the categorical cross-entropy between the predicted and true labels of the data is used and can be formulated as:

$$CE = -\sum_{i=1}^{N_d} Y_i \ln(\hat{Y}_i) \tag{4}$$

where, Y_i is the true label of the input data, $\hat{Y}_i$ is the network's prediction and N_d is the number of input datasets.

3 PHYSICAL SYSTEM AND MBD MODEL

3.1 Experimental system

The experimental set-up used in this work is an elevator door mounted on a steel test frame as shown on part [a] of Figure 1. In this work, two damaged states are examined for the elevator door system, both of which are shown on part [b] of Figure 1. The first damage case, denoted as Damage 1 (D1), examines the existence of debris in the bottom railway ridge hindering the doors' sliding motion and was indcued in the system by adding small bumps from a soft tape material to the bottom door rail. The second damage case, denoted as Damage 2 (D2), is that of a surface defect on the top right door wheel preventing it from rolling properly and was created by filing a small portion of the wheel's external surface, mimicking the wear caused by prolonged use. The two cases were selected in a manner allowing for detection of both severe (D1) and more subtle (D2) damages, proving the method's general applicability. Last, accelerometers A1 and A2 are placed as shown on the same figure in order to measure the system's vibration response at the accelerometers with a sampling rate of 2048 Hz.

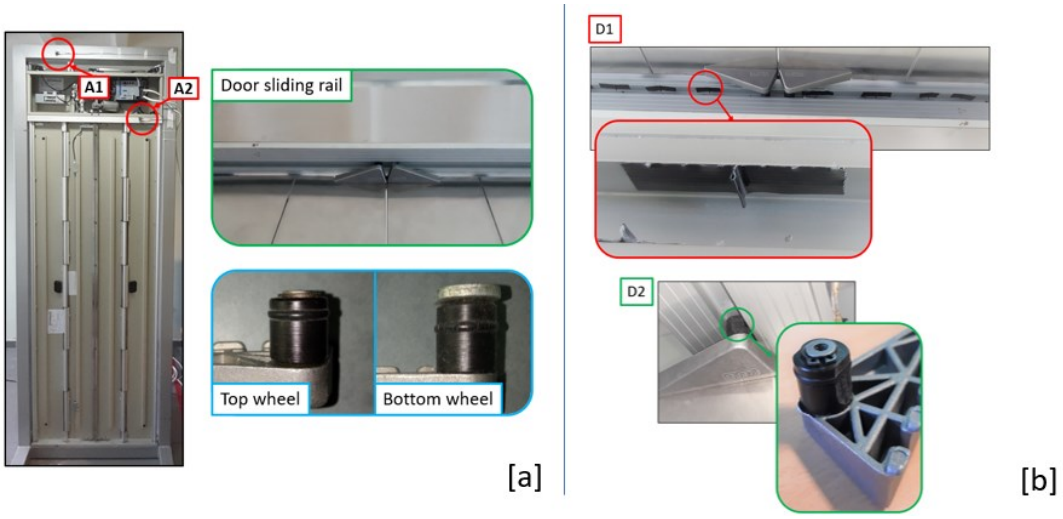

Figure 1: Elevator door experimental set-up and sensor configuration [a] and fault mechanisms [b].

3.2 MBD model

For the MBD model, the majority of the components were modeled as rigid bodies, excluding the test frame along with the door rails and the cylindrical wheels that they come into

contact with, which were considered flexible. The contact force formulation of Equation 1 is also integrated in the model to simulate the contact mechanics. The MBD model is shown in Figure 2, where the geometry of the contact mechanism of the wheel-rail system is also displayed. The present contact problem is treated as a sphere to an infinite flat surface contact. The two faults related to the wheel-rail mechanism were simulated by making changes to the model similarly to the damages induced on the experimental set-up.

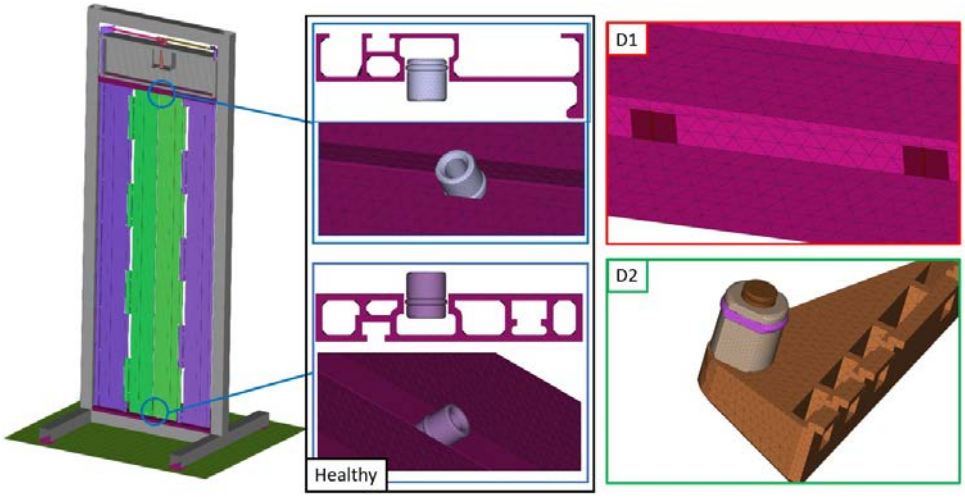

Figure 2: MBD model of the elevator door.

The optimal MBD model's parameters are shown in Table 2. Their values were estimated using the CMA-ES algorithm and a small number of initial healthy state measurements. The nominal value of each parameter is also shown to allow for comparison of the deviation between the nominal and optimal MBD models.

Table 2: MBD optimization parameters.

	Nominal	L/U Bounds	Optimal	%Deviation
$K\ [N/mm]$	1475	1300/2500	1863	26.3
$n\ [-]$	1.50	1.50/2.20	1.93	13.3
$c\_r\ [-]$	0.70	0.60/0.95	0.69	1.4
$I\ [-]$	1	1/9	6	-
$\mu_s\ [-]$	0.80	0.70/0.85	0.78	2.5
$\mu_d\ [-]$	0.76	0.60/0.80	0.77	1.3
$C\ [-]$	0.010	0.005/0.020	0.018	80
$MF\ [-]$	1	1/2	1	0

3.3 Frequency response comparison

Following the optimization process, the MBD model is expected to provide a good fit to the experimental data for the healthy as well as the damaged states, as they were built based on the

optimal healthy model. Comparison between the real and simulated responses is shown in Figures 3 & 4 for the healthy and damaged states respectively. It should be noted that while the damaged state responses are shown here, no such comparison was made during the model set-up as in practical applications damaged state data will usually be inexistent beforehand.

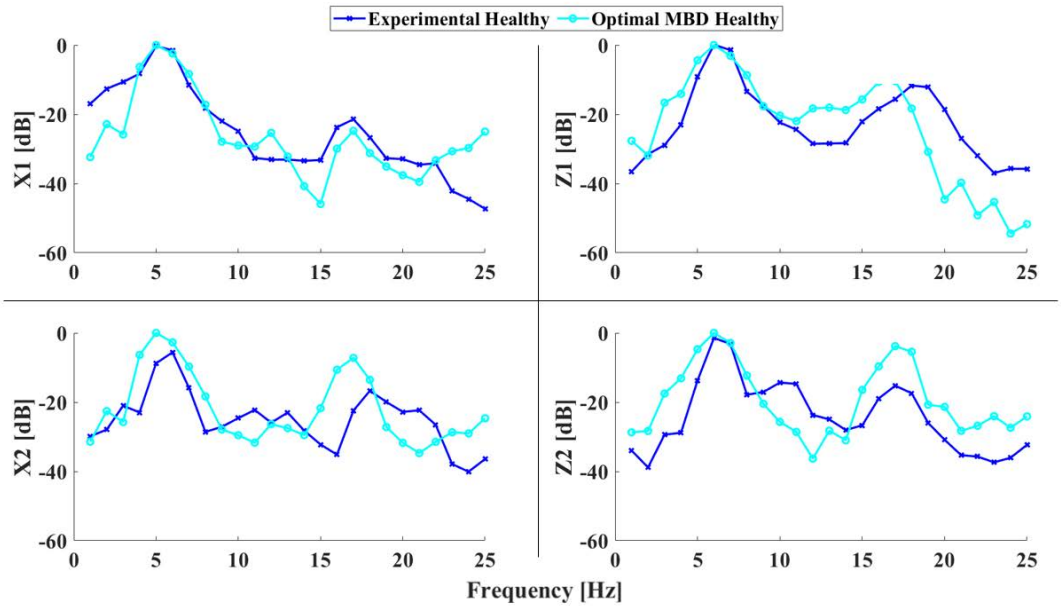

Figure 3: Healthy state Experimental and Optimal MBD PSD comparison for the transversal (X) and longitudinal (Z) axes of accelerometers A1 and A2.

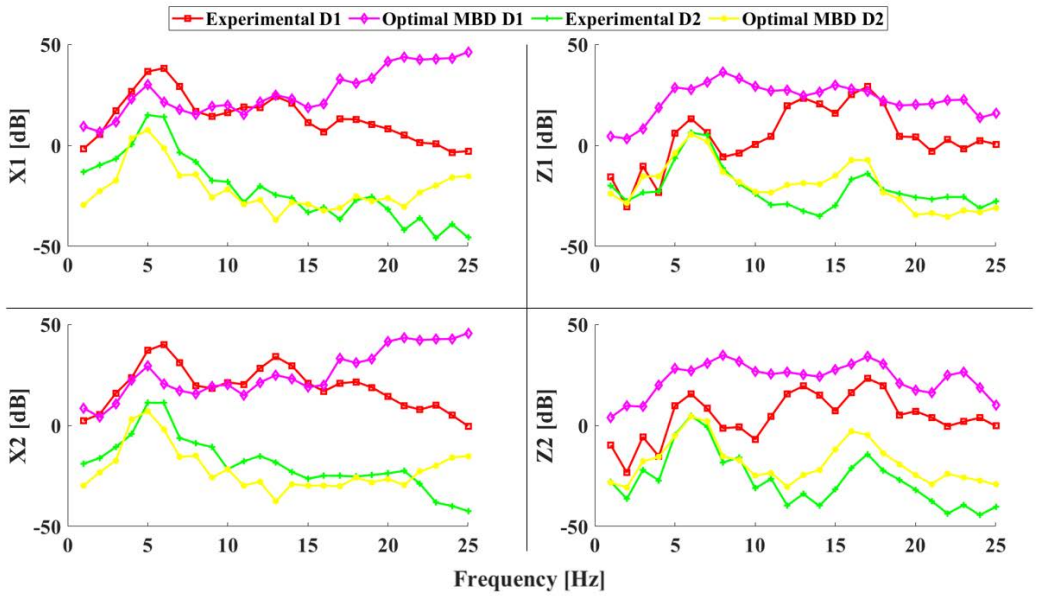

Figure 4: Damaged state Experimental and Optimal MBD PSD comparison for the transversal (X) and longitudinal (Z) axes of accelerometers A1 and A2.

4 TRAINING DATA AND CNN

4.1 Data generation

The training data for the CNN was produced by running repeated simulations for each health state and sampling the model's key parameters from a Gaussian distribution. To introduce the uncertainty that is inherent to the experimental system to the models, a variation of 1–10% was added to the optimal parameter values, creating the Gaussian sampling pool in the process. This data generation process has proven efficient in producing the high-quality data required for CM-CNN training [19,22,30]. Figure 5 shows the training data generated in comparison to the experimental test data. As shown, the training data is characterized by greater variance compared to the experimental ones, and as such, the network is expected to be able to generalize well, even to unknown datasets.

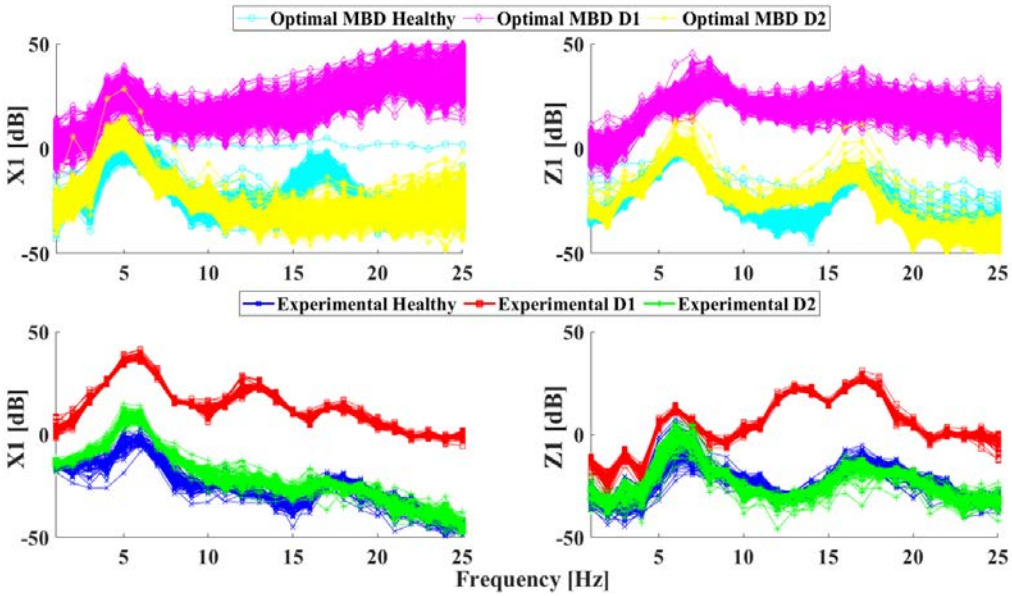

Figure 5: Optimal MBD training (Top) and Experimental test (Bottom) data comparison.

4.2 CNN

Previous work has indicated that a simple network configuration with a small number of trainable parameters leads to more robust and accurate predictions in this simulated training-experimental validation scheme [19,20,22]. As such, a shallow network configuration consisting of a small number of layers was selected in this work. The final CM-CNN structure is given in Table 3 along with the characteristics of every layer.

Out of 1800 training datasets, 70% is used for the training process and 30% is used for validation on simulated data. The Adam optimizer was used as it is one of the most popular gradient based optimizers, requiring next to no tuning while being highly efficient. The training was performed over 80 epochs, using batches of 100 samples, building an ensemble of 10 CNNs for an averaged decision in classifying each health state correctly.

Table 3: CM-CNN structure.

Layer	Activation	Units	Output	Parameters
Convolutional_1D	eLU	5	[-, 23, 5]	65
MaxPooling_1D	-	2	[-, 11, 5]	0
Flatten	-	-	[-, 55]	0
Dense	softmax	3	[-, 3]	168

5 EXPERIMENTAL VALIDATION

Figure 6 displays a bar chart of the predictions for 300 datasets, comprising of 100 datasets per health state. Assessing the certainty at which the network makes its classifications indicates that while the network is accurate in its predictions, its certainty levels are relatively low between the Healthy and D2 states, which is expected as D2 had much smaller an effect on the system's response. As shown in the figure, the D1 case is easily distinguished from the other two due to the response's larger amplitude as well as its different response shape corresponding to the different frequencies being excited.

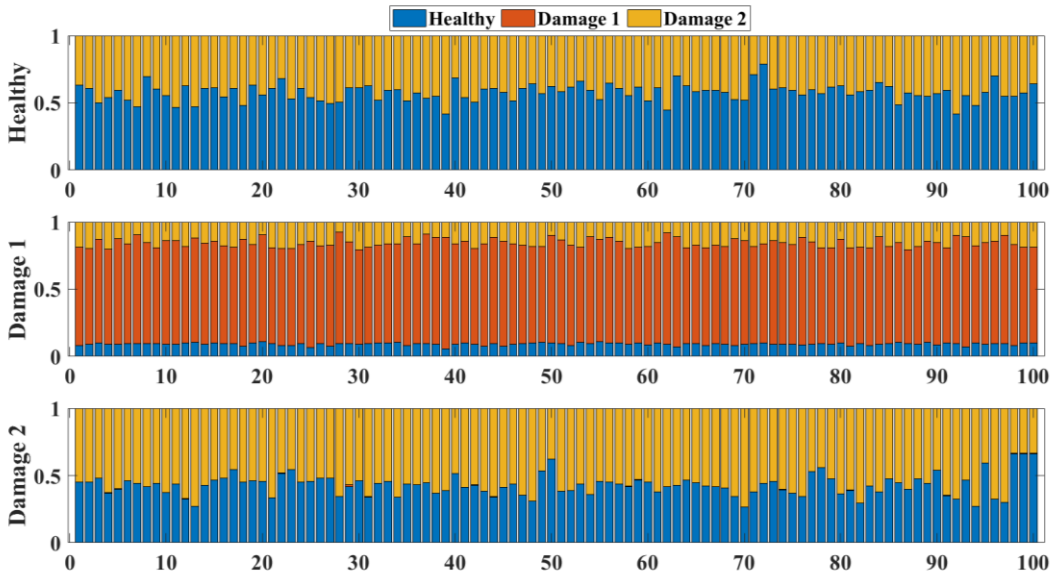

Figure 6: CM-CNN predictions on experimental data.

The predictions are also shown in the confusion matrix of Figure 7, providing a more compact display of the results. As shown, the CM-CNN makes accurate predictions with a total rate of 92.3 with the only misclassification error being the one between the Healthy and D2 states.

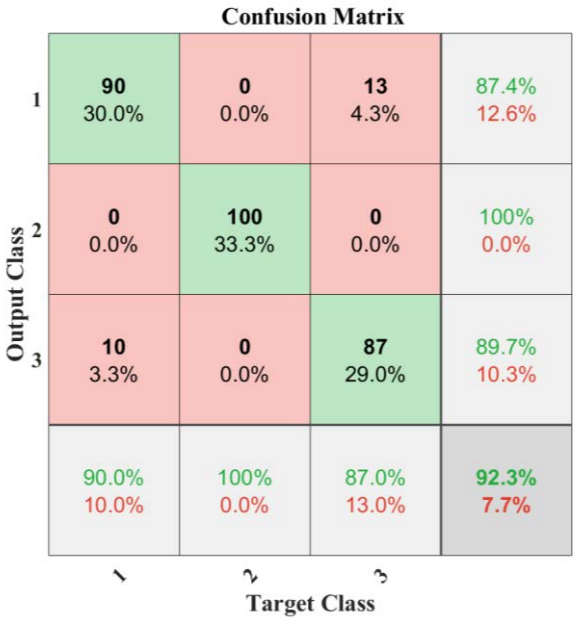

Figure 7: CM-CNN prediction confusion matrix.

6 CONCLUSIONS

A CM-CNN was presented for damage classification on mechanical systems, applied on an experimental elevator door system. The proposed methodology displays the potential of using an optimal MBD model to generate training data for an AI-based damage detector. The CM-CNN trained by simulation data alone resulted in accurate and robust classifications with an average accuracy of 92.3%. This high accuracy rate indicates that the training data produced by the MBD simulations were of high-fidelity, proving that using an optimal healthy state model to produce its damaged counterparts is possible. The CNN ensemble created was proven capable of generalizing to new and unknown datasets, distinguishing even damages of a more subtle nature and alleviating most of the concerns relative to overfitting. It was also validated that a relatively simple CNN with a simple structure and a small number of parameters can aid in this prediction generalization task, mitigating the errors emanating from the differences between simulation and measurement.

Use of a CNN ensemble as a damage detector allows for real-time damage detection and identification on a system, surpassing the time constraints of solving an inverse problem for parameter identification of a faulty system, while also eliminating the need for visual inspection of a system's response, which can be costly and requires experienced personel, especially in cases were multiple damages are present. The presented work hints to the potential of the proposed methodology and extends its application to industrial systems of complex nature.

REFERENCES

[1] Yongchao Zhu, Caichao Zhu, Jianjun Tan, Yong Tan, Lei Rao, Anomaly detection and condition monitoring of wind turbine gearbox based on LSTM-FS and transfer learning, Renewable Energy, Volume 189, Pages 90-103, 2022.
[2] S V V S Narayana Pichika, Ruchir Yadav, Sabareesh Geetha Rajasekharan, Hemanth Mithun Praveen, Vamsi Inturi, Optimal sensor placement for identifying multi-component failures in a wind turbine gearbox using integrated condition monitoring scheme, Applied Acoustics, Volume 187, 2022.
[3] K.N. Ravikumar, Akhilesh Yadav, Hemantha Kumar, K.V. Gangadharan, A.V. Narasimhadhan, Gearbox fault diagnosis based on Multi-Scale deep residual learning and stacked LSTM model, Measurement, Volume 186, 2021.
[4] Ronny Francis Ribeiro Junior, Isac Antônio dos Santos Areias, Mateus Mendes Campos, Carlos Eduardo Teixeira, Luiz Eduardo Borges da Silva, Guilherme Ferreira Gomes, Fault detection and diagnosis in electric motors using 1d convolutional neural networks with multi-channel vibration signals, Measurement, Volume 190, 2022.
[5] Thomas Bergmayr, Simon Höll, Christoph Kralovec, Martin Schagerl, Local residual random forest classifier for strain-based damage detection and localization in aerospace sandwich structures, Composite Structures, Volume 304, Part 1, 2023.
[6] Yunguang Ye, Bin Zhu, Ping Huang, Bo Peng, OORNet: A deep learning model for on-board condition monitoring and fault diagnosis of out-of-round wheels of high-speed trains, Measurement, Volume 199, 2022.
[7] Özgür Gültekin, Eyup Cinar, Kemal Özkan, Ahmet Yazıcı, Multisensory data fusion-based deep learning approach for fault diagnosis of an industrial autonomous transfer vehicle, Expert Systems with Applications, Volume 200, 2022.
[8] Emmanuel Akintunde, Saeed Eftekhar Azam, Ahmed Rageh, Daniel G. Linzell, Unsupervised Machine Learning for Robust Bridge Damage Detection: Full-Scale Experimental Validation, Engineering Structures, Volume 249, 2021.
[9] Jinpeng Feng, Kang Gao, Wei Gao, Yuchen Liao, Gang Wu, Machine learning-based bridge cable damage detection under stochastic effects of corrosion and fire, Engineering Structures, Volume 264, 2022.
[10] Mikel Canizo, Isaac Triguero, Angel Conde, Enrique Onieva, Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study, Neurocomputing, Volume 363, Pages 246-260, 2019.
[11] Xinyuan Huang, Zhiliang Liu, Xinyu Zhang, Jinlong Kang, Mian Zhang, Yongliang Guo, Surface damage detection for steel wire ropes using deep learning and computer vision techniques, Measurement, Volume 161, 2020.
[12] Niu, G., Lee, SS., Yang, BS. et al. Decision fusion system for fault diagnosis of elevator traction machine. J Mech Sci Technol, Volume 22, Pages 85–95, 2008.
[13] Piotr Gołuch, Janusz Kuchmister, Kazimierz Ćmielewski, Henryk Bryś, Multi-sensors measuring system for geodetic monitoring of elevator guide rails, Measurement, Volume 130, Pages 18-31, 2018.

[14] X.Y. Li, Y.H. Guan, S.S. Law, W. Zhao, Monitoring abnormal vibration and structural health conditions of an in-service structure from its SHM data, Journal of Sound and Vibration, Volume 537, 2022.
[15] Peng, Q.; Xu, P.; Yuan, H.; Ma, H.; Xue, J.; He, Z.; Li, S. Analysis of Vibration Monitoring Data of Flexible Suspension Lifting Structure Based on Time-Varying Theory, Sensors 2020, 20.
[16] C. Ruiz-Cárcel, V.H. Jaramillo, D. Mba, J.R. Ottewill, Y. Cao, Combination of process and vibration data for improved condition monitoring of industrial systems working under variable operating conditions, Mechanical Systems and Signal Processing, Volumes 66–67, Pages 699-714, 2016.
[17] Yun Ji Kim, Weonwoo Nam, Jongsoo Lee, Multiclass anomaly detection for unsupervised and semi-supervised data based on a combination of negative selection and clonal selection algorithms, Applied Soft Computing, Volume 122, 2022.
[18] Kilian Vos, Zhongxiao Peng, Christopher Jenkins, Md Rifat Shahriar, Pietro Borghesani, Wenyi Wang, Vibration-based anomaly detection using LSTM/SVM approaches, Mechanical Systems and Signal Processing, Volume 169, 2022.
[19] P. Seventekidis, D. Giagopoulos, A. Arailopoulos, O. Markogiannaki, Structural Health Monitoring using deep learning with optimal finite element model generated data, Mechanical Systems and Signal Processing, Volume 145, 2020.
[20] Josef Koutsoupakis, Panagiotis Seventekidis, Dimitrios Giagopoulos, Machine Learning Based Condition Monitoring for Gear Transmission Systems Using Data Generated by Optimal Multibody Dynamics Models, Mechanical Systems and Signal Processing, Submitted 2021 (Under Review).
[21] Margarida Machado, Pedro Moreira, Paulo Flores, Hamid M. Lankarani, Compliant contact force models in multibody dynamics: Evolution of the Hertz contact theory, Mechanism and Machine Theory, Volume 53, Pages 99-121, 2012.
[22] Panagiotis Seventekidis, Dimitrios Giagopoulos, Model error effects in supervised damage identification of structures with numerically trained classifiers, Mechanical Systems and Signal Processing, Volume 184, 2023.