Bobylev, Denis; Choudhury, Tuhin; Miettinen, Jesse; Viitala, Risto; Kurvinen, Emil; Sopanen, Jussi

Simulation-Based Transfer Learning for Support Stiffness Identification

Published in:
IEEE Access

DOI:
10.1109/ACCESS.2021.3108414

Published: 08/09/2021

Document Version
Publisher's PDF, also known as Version of record

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 21, 2021, accepted August 19, 2021, date of publication August 27, 2021, date of current version September 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3108414

Simulation-Based Transfer Learning for Support Stiffness Identification

DENIS BOBYLEV[1], TUHIN CHOUDHURY[1], JESSE O. MIETTINEN[2], RISTO VIITALA[2], EMIL KURVINEN[1], AND JUSSI SOPANEN[1], (Member, IEEE)

[1]Department of Mechanical Engineering, School of Energy Systems, LUT University, 53850 Lappeenranta, Finland
[2]Department of Mechanical Engineering, School of Engineering, Aalto University, 02150 Espoo, Finland

Corresponding author: Tuhin Choudhury (tuhin.choudhury@lut.fi)

This work was supported in part by the Academy of Finland under Grant 313676 and Grant 313675, and in part by Business Finland under Grant 1119/31/2018 and Grant 4329/31/2019.

ABSTRACT The support structures of a rotating machine affect its overall dynamic behavior. The stiffness of the support structures often differs at the actual sites compared to the test rigs, which leads to uncertain dynamics. In this research, a novel method is developed to identify the support stiffness for an in-situ machine using a simulation-data-driven, deep learning algorithm. In this transfer learning approach, a deep learning algorithm is trained with a simulation model and then tested with measured vibration of a rotor-bearing-support system. To validate the algorithm for multiple cases, an experimental test rig with variable horizontal support stiffness is used. The results from the deep learning algorithm are compared with Linear regression (LR), Artificial Neural Network (ANN), and Support vector regression (SVR) for benchmarking. The models are trained with filtered frequency domain response, and challenges due to measurement uncertainty are analyzed. With the proposed pre-processing steps of the frequency domain response and outlier elimination, the deep learning-based virtual sensor can predict the support stiffness with reasonable accuracy, where the limiting factor is the data quality and lack of excitation at critical speed frequencies.

INDEX TERMS Deep learning, machine learning, parameter estimation, physics-based simulation, support stiffness, transfer learning.

I. INTRODUCTION

The mounting of medium or large size rotating machines is a crucial step when preparing machines for operation. Machines have their own individual dynamic behavior related to their dimensions, manufacturing errors, support stiffness etc. The support stiffness has a large impact on the behavior of a machine after its delivery. Therefore, when estimating the dynamic behavior of a machine, a complete rotating system must be considered including, e.g., rotor, bearings and support structure. At worst, a poorly prepared mounting setup can lead to excessive vibration at the operating range, e.g., during on-site commissioning, causing the rejection of a machine or a rotor. Typically, a machine is put through acceptance tests by manufacturers, and then shipped to the customers. The final mounting position of the machine is connected to other structures and is part of a larger complex system. This creates design challenges from the mounting perspective, as typically the manufacturers have very stiff test beds. However, at the customer site the mounting structure can be relatively flexible, i.e., all machines are not located on desired or sufficiently stiff support. The change in support stiffness affects the critical speeds of the system and can thus lead to resonances occurring at operation speed range, which would normally not be the case when mounted on the stiff bed.

There are several traditional techniques for identifying support parameters. These include purely experimental methods such as the Experimental Modal Analysis (EMA), or a combination of measurements with numerical models and optimization using least squares, or extended Kalman filter to compare bearing forces or shaft displacement to parametrically identify the support properties [1]. Some studies have also investigated pedestal clearance as a support parameter using a dynamic rotor-bearing-pedestal model [2] and a stochastic model in Bayesian framework [3]. In recent years, intelligent diagnostics [4] including Machine Learning (ML) and Deep Learning (DL) models [5], have been introduced to analyze mechanical system states from vibration signals. These models have the advantage of automatic

The associate editor coordinating the review of this manuscript and approving it for publication was Alessandra Bertoldo.

state recognition and are often used for prognosis, diagnosis, virtual sensing, and parameter identification [6].

ML based methods for prognostic health monitoring aim to predict the development of the machine health state in the future, e.g., identifying of the remaining useful life (RUL) of bearings using Artificial Neural Network (ANN) [7] and support vector machine (SVM) [8]. For condition monitoring and bearing fault diagnosis, methods such as linear regression (LR) [9], ANN [10], [11], and SVM [12] are used in rotating machines.

Similarly, DL based methods have used feature extraction from frequency domain response to predict bearing degradation and RUL [13]. DL methods, such as convolutional neural network (CNN), have also been used for condition monitoring and bearing fault diagnosis, using either artificial vibration fault signal [14] or with real-time raw signals from motors [15]–[17]. For automatically extracting features from vibration data and classifying bearing faults, CNN is combined with methods such as SVM [18] or Long-Short-Term Memory (LSTM) structure [19], often utilizing frequency-domain signals [20]. Time-frequency domain analysis combined with vibration imaging and feature extraction with CNN, are also used for gearbox and rotor fault diagnosis [21]. LSTM based algorithms are also used in regression problems or as virtual sensing applications for identifying specific parameters such as turbine engine vibration [22] or rotation speed of a fixed shaft gearbox [23].

The distinguishing factor between ML and DL models is how the features are extracted from the vibration signals. ML algorithms often rely on features extracted from the vibration signals with signal processing techniques, and manually extracted features might be too sparse or biased towards a system. DL has a couple of advantages over ML. First, the DL algorithms learn automatically, to extract the relevant features from the data. Moreover, the feature extraction and system state recognition of the DL algorithm are optimized simultaneously. Second, DL algorithms can learn more complex non-linear relations between the input space and the diagnosis space.

In literature, it is common practice to use sensor data from a rotating machine to train a machine learning model which can be used for online condition monitoring [24]. In general, such experimental data driven applications require extensive historical data, in-service data with forceful, accelerated failures in a physical system for training the model to predict a fault [25]. Such experimental data driven, machine learning or deep learning methods have been used for detecting bearing related faults through various statistical features [24], [26], [27], [27], [28]. Cerrada *et al.* [29] comprehensively compared signal-based techniques with machine learning techniques for detecting bearing faults along with their severity.

However, as Sobie *et al.* [25] observed, such intelligent models trained with experimental data tend to become case dependent. Furthermore, the training procedure has limitations in the form of availability of in-service data and lack of generalization. Therefore, instead of using experimental data set for training, a simulation model or digital twin [30] is arguably an inexpensive tool to take various failure mode scenarios into account, or to analyze the dynamic behavior of a machine due to large scale parametric changes. Numerous researchers have used a combination of simulation and limited measured data in ‘model-based’ fault identification, estimation, and diagnostics of rotating machines [1]. However, there are limited studies which have used simulated data to teach machine learning algorithms for applications such as fault classification in a bearing [14], [25], internal combustion engine [31], [32], and electrical drives [33].

In conclusion, even though the existing literature demonstrates the successful use of DL in condition monitoring and diagnosis, most of these studies are based on classification problems, where the goal is to predict a categorical value, especially for bearing related faults. The existing models rarely focus on regression or parameter identification with a few exceptions such as [23] and [22]. There is a need for developing intelligent models to automatically identify support parameters and evaluate how the on-site support affects the behavior of a rotating machine. Another key unexplored area is the development of DL models trained by simulated data only.

The novelty of the study is to develop a DL model which is capable of identifying the support parameters based on the physics-based simulation model. In the study, a large rotor is studied with different support stiffnesses and the displacement at the middle of the rotor is captured at steady-state condition from the actual machine. In the simulations, the same dataset is generated. For the design of the CNN, the simulation dataset was used, and its performance was evaluated. The verification of the model was then conducted by comparing the labeled measured data (speed and stiffness) and it was then compared to the CNN model prediction. Speed was selected as a secondary parameter for verification, as it is typically easily available from other sources and it can be used to validate the CNN. A series of pre-processing steps are developed to prepare the simulation-based training data as well as the test data. The model is first tested with simulated response and then with measured response from an experimental setup and the results are benchmarked against models based on LR, ANN, and support vector regression (SVR). The model goodness was calculated with the mean average percentage error (MAPE), where the difference of the actual support stiffness and predicted was compared. The transfer learning concept, where valuable information is created with one model and then used to identify the state of another system, is the main idea in the manuscript [4]. The simulation-based data enables one to overcome the lack of labeled data from the measurements in order to train diagnosis models. The concept can be extended to different parameters of a rotor bearing support system to analyze their sensitivity to the dynamic response of the machine.

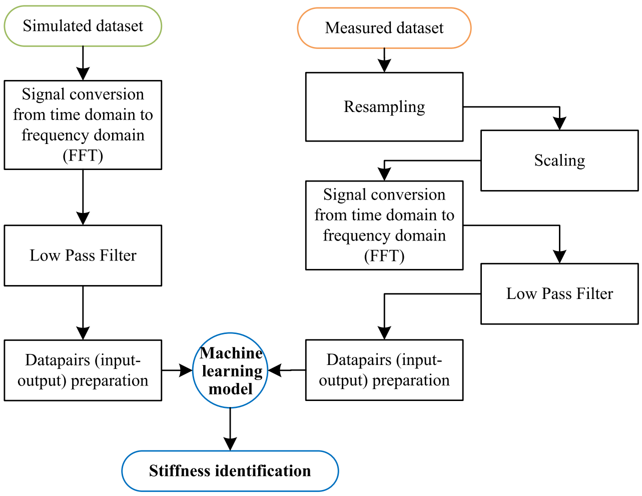


FIGURE 1. Pre-processing steps of training data (simulated) and testing data (measured).

II. TRANSFER LEARNING FROM SIMULATION TO REAL MACHINE DATA VIA NEURAL NETWORK

Deep learning is a subdivision of artificial intelligence (AI) and machine learning (ML), in which a distinctive feature is the use of ANNs, the operating principle whereby data processing operations are characterized by the extraction of high-level features from the raw data while applying data-driven decision making based on the raw data. The fact thereof determines the use of DL in fields such as object detection, speech recognition, natural language processing, etc. CNN is a DL algorithm, which is widely used in image and video processing. However, recent studies considered the use of this algorithm for one-dimensional sequential data gained by vibrational systems. This chapter describes the pre-processing steps and highlights the main theoretical aspects of CNN.

A. PRE-PROCESSING OF VIBRATION DATA

In general, DL models require a lot of data to make sure they are properly trained without overfitting or underfitting the training data. While a large training dataset is desirable, without appropriate pre-processing steps, the neural network will still not produce accurate prediction, i.e., the DL is blind to the source of data.

The datasets used for identifying the support stiffnesses are in the form of vibration signals. Since the developed method aims to use simulated response signal for the purpose of training, depending on the accuracy of the model, it can vastly differ from the measured response which could lead to an inaccurate estimation of system parameters. Therefore, figure 1 shows a few preliminary steps for creating uniformity in the system to make the comparison meaningful without over-fitting the data. For the simulated data used for training the model, the steps include data filtering and resizing. For the measured data which represent the test data, data filtering and re-sampling to match the simulated sampling rate are the preliminary steps. The strength and quality of the signals might differ in measurement and simulation and thus, scaling of the signals can be employed in cases where the dynamics related parameters, e.g., frequency domain response is observed. For training and testing the model using time domain response, both datasets can be prepared from this point to create data pairs. For feeding frequency domain response in the model, the signal conversion using Fast Fourier Transform (FFT) is required. A secondary filter to focus only on the frequency range of interest can be used to further focus and optimize the feed data.

B. CONVOLUTIONAL NEURAL NETWORK (CNN) MODEL

Convolutional Neural Networks (CNNs) belong to the class of neural networks primarily used for image data. They have demonstrated good performance in problems related to computer vision, while at the same time demonstrating excellent performance in tasks related to image classification, as well as being a part of hybrid models used for object localization, image captioning etc. [34].

This is made possible due to the operating principle based on the direct processing of raw data, instead of manually engineered features derived from the raw data. Typically, the CNN-based model is aimed at automatically extracting the features from the raw data that is useful, and closely related to the problem being solved. This is known as feature learning, and the CNNs can extract these useful features regardless of the way they appear in the data.

The fact that convolutional neural networks can extract features from raw input data makes them suitable for time-series data processing [14]. Sequential data can be treated as a 1D image from which a CNN model can treat and highlight the patterns within. This feature has had a major impact on the use of CNNs for time-series classification tasks such as fault diagnosis in the rotor bearing systems [19], or in classification of human activities based on raw accelerator sensor data [35].

CNNs are multi-stage neural networks that first extract features from raw data using the filters and then classifying them according to a certain parameter. The filter stage consists of several convolutional and pooling layers. At the next stage, fully connected layers are used for making the predictions based on the obtained features.

1) CONVOLUTIONAL LAYER

The convolutional layer, the local region of input signal with filter kernels, using activation function, produces the output features. For each filter the same kernel is applied, this is known as weight sharing and is used for local feature extraction from the input. The forward propagation to layer $l+1$ from layer l can be formulated as follows:

$$y_i^{l+1}(k) = W_i^l \cdot S^l(k) + b_i^l \tag{1}$$

where W^l refers to the weights of the i-th filter kernel in layer l, b_i^l is bias, $S^l(k)$ is the k-th local region in layer l, $y_i^{l+1}(k)$ is the input of k-th neuron in frame i of layer $l+1$, and $\cdot$ denotes convolution, respectively.

In the next step, the Rectified Linear Unit (ReLU) is used as the activation function of the layer to speed up the convergence of the model. It follows the next equation:

$$f(x) = \max\{0, y_i^l(k)\} \tag{2}$$

2) POOLING LAYER

After the convolutional layer, the pooling layer is typically applied. It is used to reduce the number of parameters by appending semantically similar features and downsampling its number in the model. Max-pooling layer is used to preserve the sharpness of the features and informative parts of the frequency spectrum. The layer finds the local max in the input features and helps to reduce the size of the vector, but at the same time the general behavior is kept intact. It can be expressed as

$$C_i^{l+1}(k) = \max_{(k-1)m+1\le t\le km}\{a_i^l(t)\} \tag{3}$$

where $a_i^l(t)$ is the value of the t-th neuron in the i-th frame of the layer 1, m represents the width of pooling region and $C_i^{l+1}(k)$ is the result of the operation, respectively.

C. DEEP LEARNING MODEL AND TRANSFER LEARNING MODEL

Typically, a CNN model consists of one or more convolutional layers followed by pooling layers that take the largest or average element from the rectified feature map. There can be a certain number of such pairs, depending on the requirements and input data. Then comes the flatten layer, which flattens the output of convolutional layers and feeds it into a fully connected layer, where the extracted features are analyzed and then the predictions are done based on them. Based on the loss and the optimizer, the model parameters are updated on and on until the desired results are achieved.

The number of parameters and output shape for each layer can be calculated using the next formulae:

$$\text{Output Shape} = H_i - H_f, L_i - L_f + 1, N_f \tag{4}$$
$$\text{No. of Parameters} = N_f \cdot (L_f \cdot H_f \cdot D_i + 1) \tag{5}$$

where, H_i stands for input height, H_f filter height, L_i input length, L_f filter length, D_i input depth, and N_f filter number.

In the study, the simulation data refers to displacements or vibration response obtained from a physics-based simulation model, which is verified by modal analysis of the actual machine. The verified model is used to analyze the different support configurations and used to train the CNN, therefore enabling transfer learning of the simulated output data to real world experimental data.

III. CASE STUDY, EXPERIMENTAL SETUP AND SIMULATION MODEL

To evaluate the performance of the CNN model over a range of values for support parameters, it had to be tested in a system where it was possible to continuously vary the support stiffness. This way, the dynamic behavior could be altered continuously, creating multiple different test runs for the CNN model using otherwise the same rotating system. Furthermore, although support parameters can be continuously altered in a straightforward manner in simulation models, a real experimental setup with adjustable support stiffness is required to incorporate experimental uncertainty in the test dataset and observe the prediction of the CNN model.

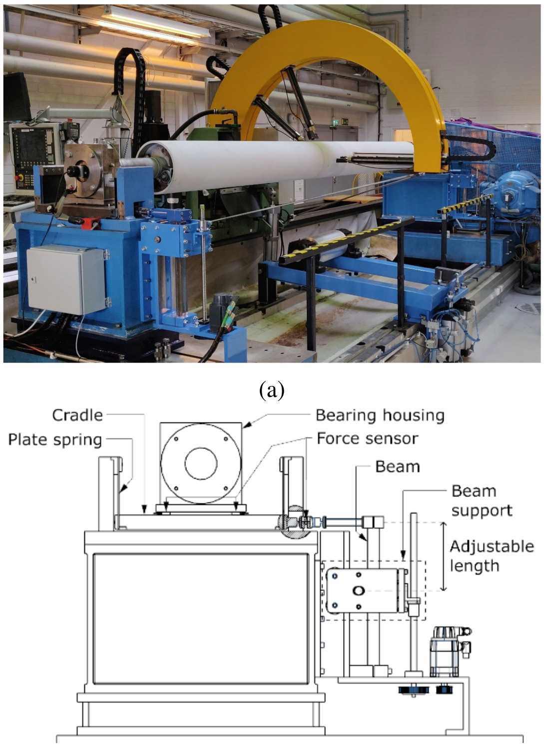


FIGURE 2. Experimental test rotor (a) Measurement setup (b) Mechanism for adjusting the support stiffness in horizontal direction.

For this purpose, the guiding roll of a paper machine, which has an additional mechanism for varying the horizontal support stiffness, is considered as a test case. The rotor used here consists of a 4 m long tube section and 0.5 m end shafts on either side. The overall rotor weight is 720 kg. The rotor is supported by two SKF 23124 CCK/W33 spherical roller bearings. The experimental setup, measurement procedure, the simulation model and its parameters are briefly described in the following sections. More details about the experimental setup can be found in previous research [36] while the details related to the simulation model are available in [37].

A. EXPERIMENTAL SETUP

The test setup has a possibility to vary the horizontal stiffness, and thus the measured dataset is similar to the data generated in the simulation. The adjustable stiffness in the test setup is implemented by using a similar structure as in balancing machines, where a rotor is supported by plate springs. Subsequently, the stiffness for the support is provided through an external beam. Changing the position of the beam support changes the stiffness of the beam end, which in turn alters

TABLE 1. Averaged stiffnesses in different horizontal stiffness adjuster (HSA) positions in the experimental setup [36]. Same stiffness values are used in the simulation model for data generation as well.

Position of HSA [mm]	Averaged stiffness in the driving end device [MN/m]	Averaged stiffness in the non-drive end device [MN/m]	Averaged stiffness of the system [MN/m]
0	17.86	18.77	18.32
30	12.77	12.88	12.83
60	9.18	9.55	9.37
90	6.79	7.76	7.28
120	5.60	5.90	5.75
150	4.55	4.56	4.56
180	3.64	3.68	3.67
210	3.11	3.12	3.12
240	2.74	2.61	2.68
270	2.35	2.36	2.36
300	2.01	2.06	2.04

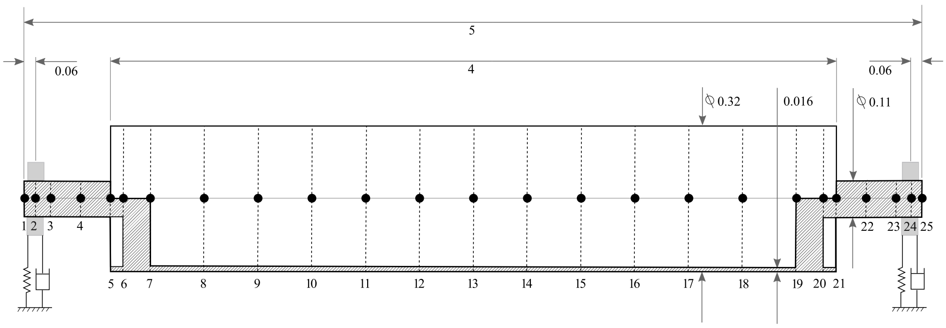


FIGURE 3. A sketch of the Paper machine's steel tube roll with FE discretization. All dimensions are in meters.

the stiffness of the horizontal support. The structure is built at both ends of the rotor. Figure 2(a) shows the outlook of the system and Figure 2(b) shows the mechanism for generating adjustable stiffness [36].

B. MEASURED DATASET

The measured data consists of 532 samples. All the samples are collected with the sampling rate of 1024 per revolution, and hence sampling rate was dependant on the rotating speed. 100 rounds of data were recorded per every speed/file. The varied parameters are the same as in the simulation dataset: speed and stiffness. The speed is varied between 4 Hz and 18 Hz with increments of 0.5 Hz, resulting in 27 subsets of files. The measurements were conducted at steady state condition. The stiffness was varied in ranges between 2.04 and 18.32 MN/m. Table 1 shows the final stiffness values of the devices, which are calculated as averages from 10 different measuring points. Finally, the stiffness values in drive end (DE) and non-drive end (NDE) were averaged and used for the purpose of this study. Further details about the measurement procedure are available in an earlier experimental study by Viitala *et al*. [36], and the measured data is available at [38].

C. SIMULATION MODEL

The main objective of the simulation model is to generate the required amount of data for training the CNN model in a computationally efficient manner, while representing the actual rotor bearing system accurately. To that end, a high-fidelity, rotor-bearing-support model of the test rotor is used. The model, which has been experimentally validated before [37], [39], uses 3D beam elements based on the Timoshenko beam theory for obtaining computationally fast and yet accurate dynamic response. Figure 3 shows the FE discretized sketch of the rotor along with its dimensions. The asymmetry of the tube section of the rotor is included by varying the thickness of the cylindrical section based on ultrasonic measurement similar to [37]. The thickness variation is implemented in the model by defining the thickness profiles of the cross sections along the length of the tube structure, thus affecting the area moment of inertia.

The spherical roller bearings are included in the modeling procedure by using a nonlinear bearing model [40], which considers the effect of clearance and bearing inner ring waviness. In the simulations, the nominal bearing clearance of 60 μm is used, and measured waviness profiles of measured sections are shown in Table 1. Waviness components from twice to six times per revolution were measured from the bearing inner ring for the non-drive end and drive end of the machine.

The supports are modelled using a concentrated parameter approach as mass-spring-damper elements, individually in horizontal and vertical directions. This simplification is possible because in this test rotor, there is almost no cross-coupling between supports [41], and it contributes to the overall computational efficiency of the model. Damping for support structures has 2% damping ratio for horizontal and 3% for the vertical direction. Modal damping ratios 1.5%

TABLE 2. The two datasets obtained from simulation model to be used as training data.

Dataset	Samples	Response node	Directions	Frequency range [Hz]	Speed [Hz]	Stiffness [MN/m]
Frequency domain	8711	13	Horizontal	0-25	4-18, 0.05 inc.*	2.04-18.32, 31 values

* inc. is short for increment.

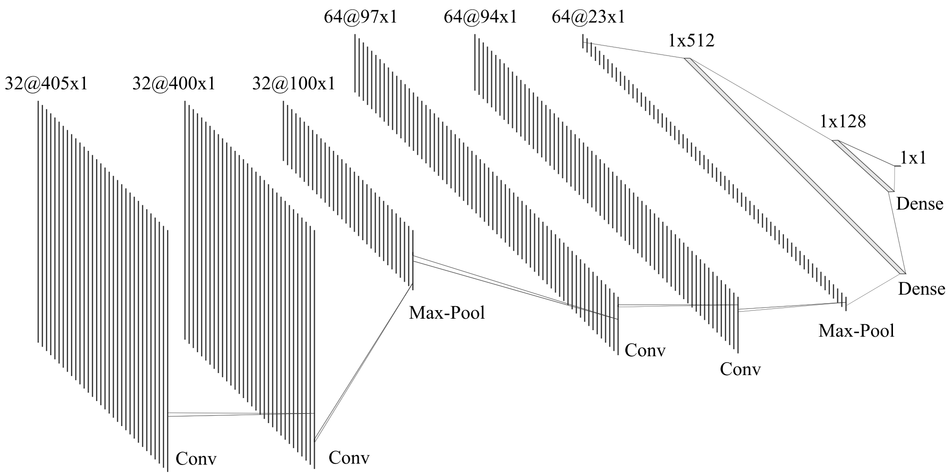


FIGURE 4. The deep learning model architecture. The layer configuration is shown as *a@b x c*, where *a* is the depth of the layer (number of filters), *b* is the height (length of the signal), and *c* is the width.

(1st), 2% (2nd), 2.5% (3rd) and 3.0% (4th to 6th) where the value in parenthesis corresponds to the flexible mode number. In the transient analysis, model reduction is applied and the number of retained modes is 16. Simulation runs are conducted for 9 seconds with a sampling rate of 2000 Hz (time-step of 0.0005s). The resulting responses are captured at the bearing locations and the centre of the rotor. The computational time is approximately 300 seconds per single simulation of 9 seconds. However, the computation time is dependent on the system parameters, such as parameters of nonlinear bearing model.

D. SIMULATION DATASET

The dataset contains the frequency-domain displacement at the centre of the rotor (Node 13) in the horizontal direction collected with the sampling frequency 2000 Hz, as shown in Table 2. Rotation speed and stiffness are the varied parameters. The simulation data in time domain format is available as part of a larger dataset in [42].

In this study the response of node 13 (center of roll) in the horizontal direction was used. For this purpose, the datasets have been preprocessed accordingly. Moreover, the dataset was also converted to frequency domain because it includes the relevant dynamic information and can be investigated thoroughly to identify the faulty parameter estimations root causes. In addition, it allows the use of more compact datasets. The proposed model was tested with the measured dataset which was also converted to frequency domain. The measured dataset is first resampled to match the simulated dataset sampling frequency of 2000 Hz. After the FFT conversion, only signal amplitudes of up to 25 Hz were used, as the critical speed varies from 9 Hz to 21 Hz, which determined the shape of the input equal to 410, which corresponded to the parameters used for Fast Fourier Transform (FFT).

Both measured and simulated data are pre-processed using the steps shown previously in figure 1. For simulated data the steps included data filtering and resizing. The measured data was also filtered and re-sampled using time synchronous averaging (TSA) from 1024 samples per revolution to 2000 Hz to match the simulation model with a higher sampling rate. After scaling and FFT conversion, a low pass filter of 0-25 Hz was used based on the machine's operational speed range to create optimized and focus pairs of datasets in the frequency domain.

E. CNN MODEL ARCHITECTURE

In this study, 1D CNN model is developed and used as a tool for feature extraction from the raw frequency domain vibration data obtained from measured and simulated datasets. Figure 4 shows the model architecture proposed in this research. The model consists of four paired convolutional layers, each pair is followed by a MaxPooling layer to decrease the number of parameters by four times. There are 32 filters in the first two convolutional layers and 64 in the second pair. The kernel size is 6 and 4, respectively. To proceed light features, the layers with the low numbers of filters are used, at the same time higher numbers result in the extraction of more complex features. The convolutional layers are followed by two simple fully connected layers. These layers are aimed at making the predictions based on the obtained features from the convolutional layers. The number of neurons in these layers is 512 and 128, respectively. The output layer has one neuron, because this type of problem belongs to regression and it predicts only one parameter as the output.

'ReLU' was used as an activation function and 'Adam' was chosen as an optimizer. The batch size was kept as default, which is equal to 32. By increasing the number of epochs, it is possible to increase the performance of the model, at the same time it is time consuming and overfitting is possible, so this value should be chosen wisely.

The proposed model was implemented and tested using the measured and simulated datasets. Frequency domain responses were used as the input for the proposed model, stiffness was the output variable. The model was trained for 100 epochs, Mean Squared Error (MSE) was used to calculate the loss in CNN model. This loss is calculated as

$$\mathrm{MSE} = \sum_{i=1}^{n} \frac{|y_i - y_i^p|^2}{n} \tag{6}$$

where, y_i is actual value, y_i^p is a predicted value, and n is the number of predictions.

IV. RESULTS

A. OPERATIONAL SPEED MAP

This section aims to evaluate the fidelity of the simulation model by comparing the simulated and the measured responses for a range of horizontal stiffness values. The responses are obtained over an operational speed range of 4-18 Hz when the horizontal stiffness is varied from 2.04 to 18.32 MN/m. Figure 5(a) and 5(b) show the horizontal and vertical responses from the simulated model whereas Figure 5(c) and 5(d) show the measured horizontal and vertical responses, respectively. The measured maximum speed was limited to avoid excessive vibration in the system. The plots capture the maximum amplitude with the speed and stiffness values from the dataset.

Firstly, due to the anisotropic support stiffness, the test rotor has different resonance frequencies in the horizontal and vertical directions. Since the test rotor operates in the low frequency range, for this particular test case, the subcritical vibrations are of key interest. The subcritical frequencies occur at the fractions of the natural frequency such as 1/2, 1/3, 1/4,..., times the natural frequency. Therefore, the 2^{nd} subcritical harmonic response occurs at half the natural frequency, the 3^{rd} harmonic resonance occur at $1/3^{rd}$ the natural frequency and so on. Therefore, in Figure 5, the first number denotes the natural frequency followed by the subcritical component as the second number. For example, 1-2, 1-3 represent the 2^{nd} and 3^{rd} subcritical harmonics of the 1^{st} natural frequency in individual direction. Figure 5 shows there are subharmonic excitation from multiple natural frequencies (1^{st}, 2^{nd} and 3^{rd} modes) in the horizontal direction in the operational speed range, whereas in the vertical direction only subcritical harmonics of the 1^{st} natural frequency are visible. The plots also show how the changes in horizontal support alters the horizontal subharmonic frequency peaks while the vertical response peaks remain unaffected both in simulation and measurement.

B. SIM TO SIM: TRAINING AND TESTING THE MODEL WITH SIMULATED DATA

For the initial testing purposes, the model was trained and tested using only simulation dataset. Raw vibrational signal was converted from time-domain to frequency-domain using FFT. Secondly, the frequency range was limited to 25 Hz, this was done due to the fact that this range contains all the necessary peaks, moreover, it allowed to keep the input length of the signal optimal for the proposed DL algorithm. The CNN performance is compared to three baseline models. First is an LR optimized with the ordinary least square method. Second is an ANN model with two layers. The first layer has 256 nodes and second and final has one node. ReLu was used as the activation function between the layers, and mean squared error (MSE) was used to compute the error. Adam was used to optimize the network. Third baseline model is SVR, which is based on LIBSVM implementation of ϵ-SVR [43]. Radial basis function kernel and parameters C=1.0 and epsilon 0.1 were used. Figure 6 depicts the predicted speed (a)-(d) with LR, ANN, SVR and CNN, and Figure 7 (a) to (d) stiffnesses, respectively. The blackline corresponds to the actual value and blue dots as predicted values from the dataset.

The LR, ANN and SVR, which are typically used for simpler problems were used as benchmark and justification for the need of DL. The mean average error (MAPE) for speed cases in LR is 14.4%, ANN 3%, SVR 7.6% and CNN 0.4% and for stiffness cases LR is 46.4%, ANN 6.9%, SVR 21.6% and CNN 0.9%.

For evaluation of CNN based model performance, the model with 4 CNN, 2 MaxPooling and 2 Fully Connected layers was built. The dataset with 8711 samples, where the rotational speed is varied in the range of 4-18 Hz and support stiffness 2.04-18.32 MN/m, was randomly split into train/test/validation batches. 70% (6098 samples) of the data was used for training and 30% (2613 samples) for testing the model. 20% (1220 samples) of the training data was used for validation. The validation data included broad distribution of the operating conditions. The CNN model validation was conducted by monitoring the validation accuracy and training accuracy. It was monitored that the validation error did not grow, and therefore preventing the overfitting. Ten trials were conducted. As a result, the average prediction error is 1.50%, and the standard deviation is 0.49.

C. SIM TO REAL: TRAINING THE MODEL WITH SIMULATED DATA AND TESTING WITH MEASURED DATA

Figure 8 demonstrates the predicted values against true values of random trial for predicted speed (a)-(d) with LR, ANN, SVR and CNN, and Figure 9 (a) to (d) stiffnesses, respectively. The blackline corresponds to the actual value and blue dots as predicted values from the dataset.

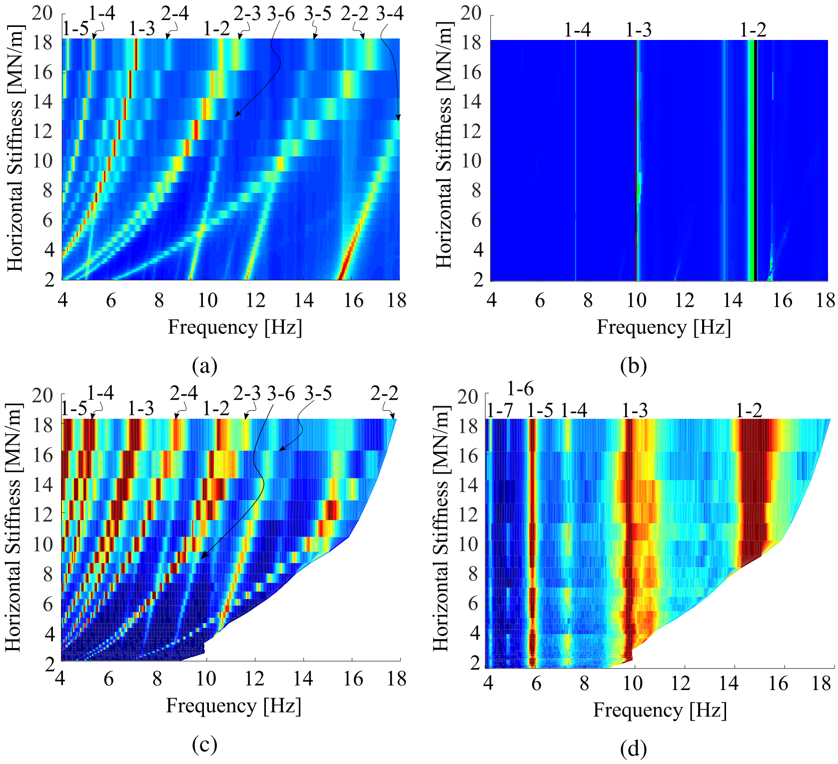


FIGURE 5. The operation speed map for the test rotor showing how the sub-critical harmonic frequencies for different bending modes are affected by the change in the horizontal stiffness. (a) and (b) are the horizontal and vertical response from the simulated model whereas (c) and (d) are the corresponding responses from the measured data, respectively.

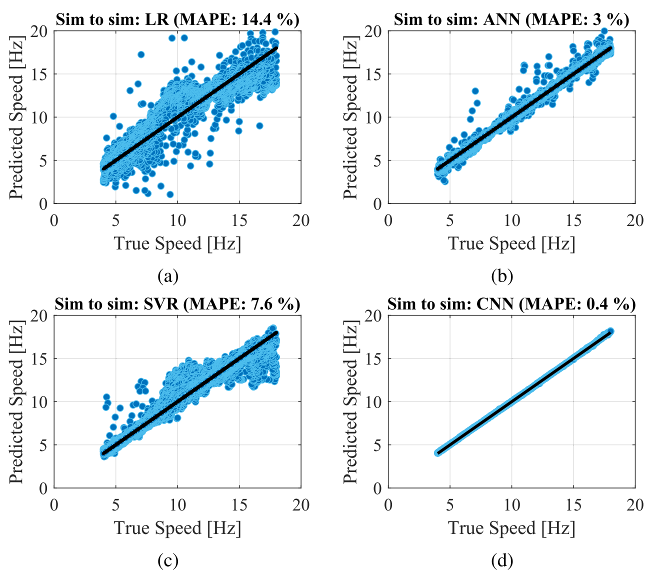


FIGURE 6. Speed prediction with (a) LR (b) ANN (c) SVR and (d) CNN, where the model is trained and tested with different segments of simulated data.

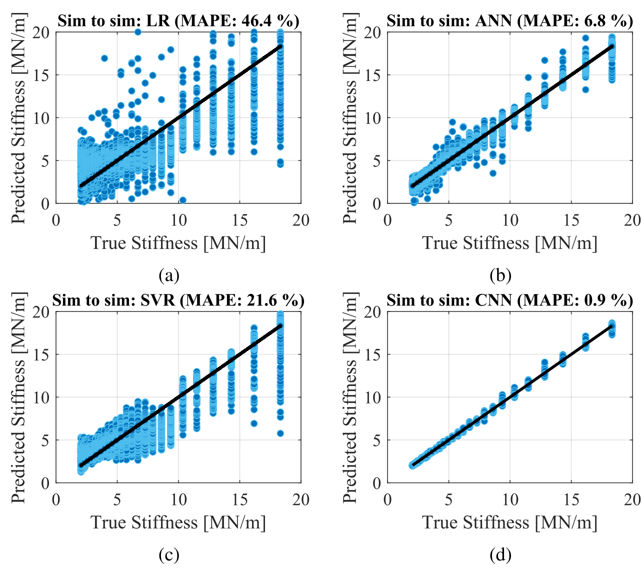


FIGURE 7. Stiffness prediction with (a) LR (b) ANN (c) SVR and (d) CNN, where the model is trained and tested with different segments of simulated data.

The LR, ANN and SVR, which are typically used for simpler problems, were used as benchmark and justification for the need of DL. The mean average error (MAPE) for speed cases in LR is 456%, ANN 71%, SVR 31%, and CNN 5% and for stiffness cases LR is 752%, ANN 81%, SVR 44%, and CNN 19%. In general, the proposed CNN model shows a good performance, although the deviation slightly increases with the increasing stiffness. This might be due to the coarse datapoints at high stiffness, and that the measured dataset does not include 1× response, due to resonance.

Ten trials were conducted. The proposed model predicts the stiffness based on the measured frequency-domain

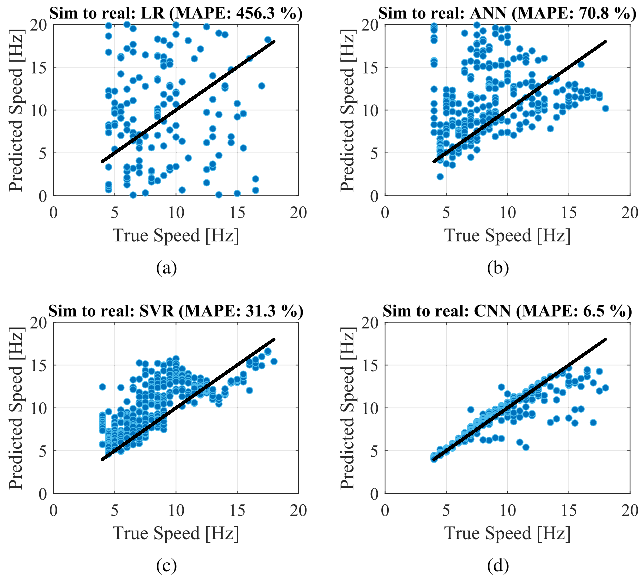


FIGURE 8. Speed prediction with (a) LR (b) ANN (c) SVR and (d) CNN, where the model is trained with simulated data and tested with measured data.

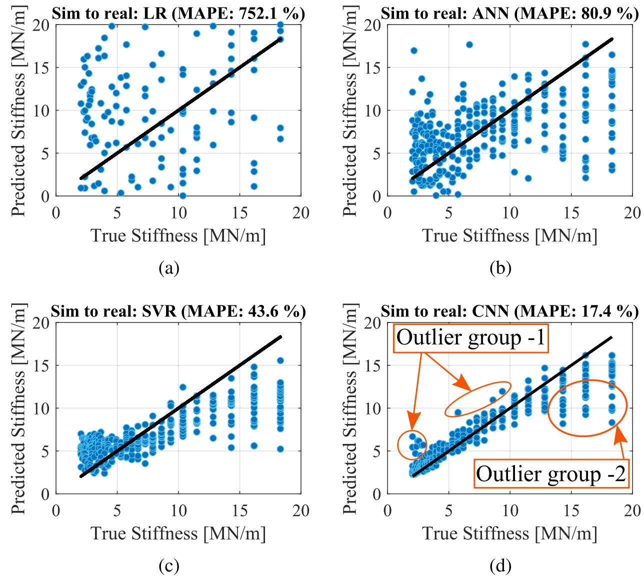


FIGURE 9. Stiffness prediction with (a) LR (b) ANN (c) SVR and (d) CNN, where the model is trained with simulated data and tested with measured data.

response with the average error of 18.64%, and standard deviation of 1.62 when the output value is stiffness and 4-5% when the output is speed. When speed is output, the predictions in general are good. Outliers are caused by samples where 1× component is not the highest component in the response, the speed is identified by the peak with the highest component, this logic is followed by the simulated data-set. In addition, the lack of excitation at critical speeds in steady state condition caused the outliers, as the critical speed information was weakly included in the signal. By excluding outliers, the predictions are closer to the real values.

V. DISCUSSION

Since the focus of the study was on the implementation of a DL model that can determine the stiffness of the support, it is clear that the model is very sensitive to input data and some degree of difference between the two datasets makes it almost impossible to identify the stiffness when the model is tested on a different dataset. It can be caused by non-optimal parameters of the DL model, sampling frequency, sampling method or input length. Furthermore, the neural network can only detect issues that are included in the training data. Therefore,

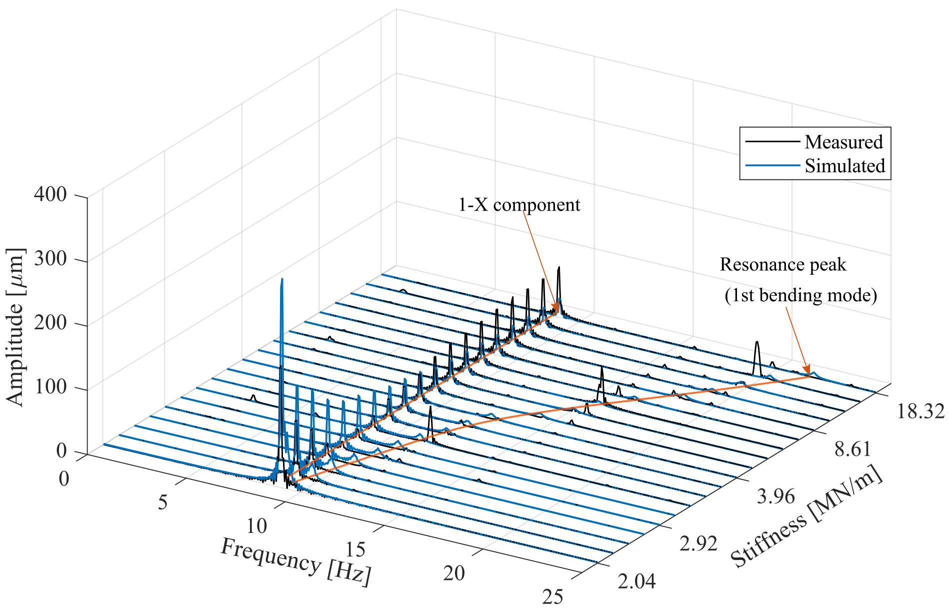


FIGURE 10. 9 Hz case, where the measured response has mechanical harmonics of speed overlapping with the resonance peaks.

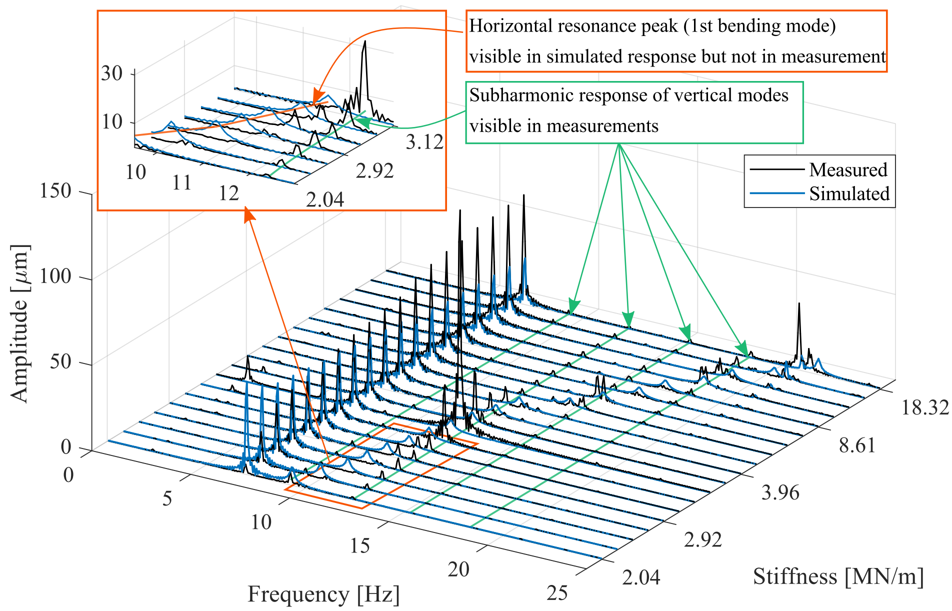


FIGURE 11. 7 Hz case, where the measured response at resonance frequencies are relatively low and not clearly visible.

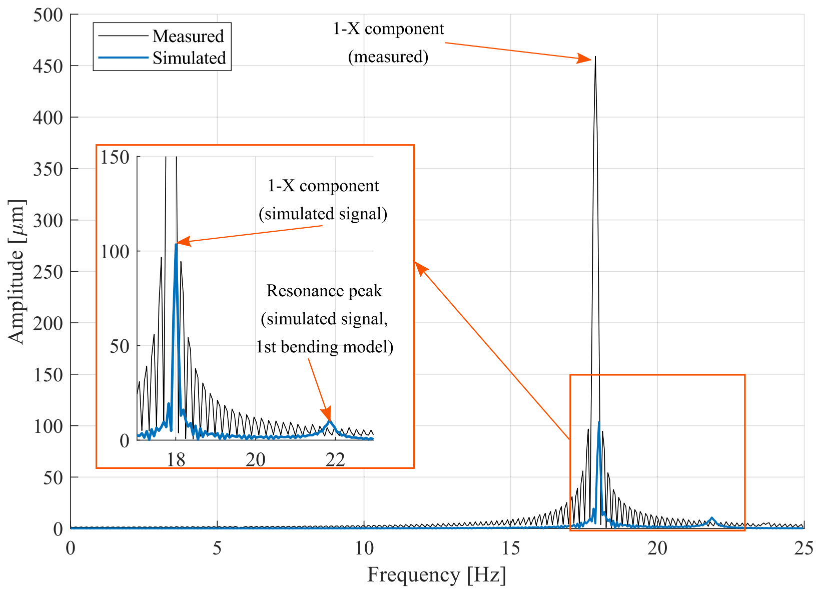


FIGURE 12. 18 Hz frequency domain response measured vs simulated.

a general solution would be to incorporate more complexities that influence the real machine vibrations into the simulation model to narrow the gap between the simulation and the real machine. That way, the training data replicates the real machine behavior more accurately which would lead to more accurate parameter predictions. Alternative methods such as graph theory [44], or multi-layer domain adaptation method [45] could be promising. In addition, adaptive batch normalization (AdaBN) [5] could be used to adjust the neural network to slightly different dataset distributions or Monte Carlo method could be used to generate large datasets fused from rotor systems with slightly varying system dynamics that could cover the real machine system dynamics.

In addition, the response in the frequency domain demonstrates more clearly, the dynamic features of the signal and simplifies the task associated with finding a suitable pattern in the sample, on which the model should concentrate on for stiffness estimation.

At the same time, the use of frequency domain response also has its own nuances caused by data imperfection. There are several outliers (see Figure 9 (d)) in the results, caused by phenomena of the measured data. The outlier group 1 in

Figure 9 (d) consists of samples with low stiffness. Some of the predictions correspond to the cases where the critical speed component coincides with the 1× component or some of the 1× harmonics, which makes it difficult for the model to identify the resonance component. This phenomenon can be seen in the example in Figure 10 in which case, the speed was set to 9 Hz and stiffness was varied from 2.04-18.32 MN/m. In this case, the resonance component of lower stiffness samples is close to 1× component or overlaps it, forcing the model to predict the stiffness based on the next peak in the response.

A second problem with predictions in the low stiffness, low speed range is that the resonance peak is poorly excited or almost invisible. In this case, the stiffness is determined by the next peak of the larger amplitude. Figure 11 shows an example where the speed is 7 Hz and the stiffness ranges from 2.04-18.32 MN/m. At lower stiffness, the resonance peak is low compared to adjacent peaks. The adjacent peaks of larger amplitude remain unchanged with variation in horizontal stiffness; which suggests that, these vibrations are probably transferred from the subharmonic resonance occurring in the vertical direction. The model then identifies stiffness based on some of the peaks along the green lines in Figure 11.

The outlier group 2 includes samples with higher stiffnesses. In this case, the resonance component is not visible from the response, as a result, the model then predicts the stiffness based on the 1× component. For clarity, Figure 12 shows an example of this situation, the speed is 18 Hz and the stiffness is 18.32 MN/m. As it can be seen in the measured case, the resonance peak is quite undetectable, while it can be visibly seen in the simulated case that was used to train the model. As a result, the model predicts the stiffness based on a single peak (1× component) visible in the measured response.

VI. CONCLUSION

In the study, the transfer learning from simulation to real machine was developed. The proposed deep learning 1D CNN model is able to extract the features from vibration data, and it estimates the varied parameter as the output, forming an end-to-end machine learning system. It was determined that the model is sensitive to input data, which entails the need for data pre-processing. The results from benchmarking algorithms, LR, ANN and SVR show reasonable predictions in sim-to-sim case, but the error percentage increased when the models are tested with real measurements. The CNN model, however, successfully predicts the output using frequency-domain data as input for both sim-to-sim and sim-to-real case. In sim-to-real, the average error for the stiffness is 19 % and the standard deviation 1.62, which can be considered reasonable values, as no additional excitations were used in the measurements to excite the system dynamics. The complexity of the model and the data sets justifies the requirement for the deep learning model, and is probably the reason why the CNN model performs better than the benchmark models for the test rotor system.

In future work, more attention should be paid to data quality and pre-processing. The complexity of the simulation model can be optimized while preserving the fault features to increase training efficiency. It is also possible to tune the performance of the proposed model by adjusting its hyper-parameters (number of layers, filters, filter size, etc.). Moreover, frequency-domain and time-domain data can be combined and tested with a 2D CNN model.

REFERENCES

[1] A. W. Lees, J. K. Sinha, and M. I. Friswell, "Model-based identification of rotating machines," *Mech. Syst. Signal Process.*, vol. 23, no. 6, pp. 1884–1893, Aug. 2009, doi: 10.1016/j.ymssp.2008.08.008.

[2] H. Cao, F. Shi, Y. Li, B. Li, and X. Chen, "Vibration and stability analysis of rotor-bearing-pedestal system due to clearance fit," *Mech. Syst. Signal Process.*, vol. 133, Nov. 2019, Art. no. 106275, doi: 10.1016/j.ymssp.2019.106275.

[3] Z. Taherkhani and H. Ahmadian, "Stochastic model updating of rotor support parameters using Bayesian approach," *Mech. Syst. Signal Process.*, vol. 158, Sep. 2021, Art. no. 107702, doi: 10.1016/j.ymssp.2021.107702.

[4] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mech. Syst. Signal Process.*, vol. 138, Apr. 2020, Art. no. 106587, doi: 10.1016/j.ymssp.2019.106587.

[5] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, p. 425, Feb. 2017, doi: 10.3390/s17020425.

[6] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019, doi: 10.1016/j.ymssp.2018.05.050.

[7] A. K. Mahamad, S. Saon, and T. Hiyama, "Predicting remaining useful life of rotating machinery based artificial neural network," *Comput. Math. Appl.*, vol. 60, no. 4, pp. 1078–1087, Aug. 2010, doi: 10.1016/j.camwa.2010.03.065.

[8] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, "Remaining useful life estimation based on nonlinear feature reduction and support vector regression," *Eng. Appl. Artif. Intell.*, vol. 26, no. 7, pp. 1751–1760, 2013, doi: 10.1016/j.engappai.2013.02.006.

[9] R. Zimroz, W. Bartelmus, T. Barszcz, and J. Urbanek, "Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings," *Mech. Syst. Signal Process.*, vol. 46, no. 1, pp. 16–27, May 2014, doi: 10.1016/j.ymssp.2013.09.010.

[10] K. Tarek, L. Abdelaziz, C. Zoubir, K. Kais, and N. Karim, "Optimized multi layer perceptron artificial neural network based fault diagnosis of induction motor using vibration signals," *Diagnostyka*, vol. 22, no. 1, pp. 65–74, Mar. 2021, doi: 10.29354/diag/133091.

[11] L. Hou, J. Zou, C. Du, and J. Zhang, "A fault diagnosis model of marine diesel engine cylinder based on modified genetic algorithm and multilayer perceptron," *Soft Comput.*, vol. 24, no. 10, pp. 7603–7613, May 2020, doi: 10.1007/s00500-019-04388-3.

[12] D. Goyal, A. Choudhary, B. S. Pabla, and S. S. Dhami, "Support vector machines based non-contact fault diagnosis system for bearings," *J. Intell. Manuf.*, vol. 31, no. 5, pp. 1275–1289, 2020, doi: 10.1007/s10845-019-01511-x.

[13] L. Ren, J. Cui, Y. Sun, and X. Cheng, "Multi-bearing remaining useful life collaborative prediction: A deep learning approach," *J. Manuf. Syst.*, vol. 43, pp. 248–256, Apr. 2017, doi: 10.1016/j.jmsy.2017.02.013.

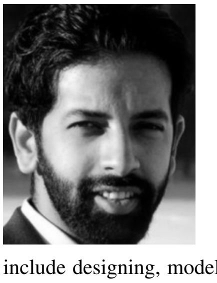
[14] M. Hemmer, A. Klausen, K. Huynh, K. G. Robbersmyr, and T. I. Waag, "Simulation-driven deep classification of bearing faults from raw vibration data," *Int. J. Prognostics Health Manage.*, vol. 10, p. 12, Mar. 2020, doi: 11250/2648526.

[15] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016, doi: 10.1109/TIE.2016.2582729.

[16] L. Eren, "Bearing fault detection by one-dimensional convolutional neural networks," *Math. Problems Eng.*, vol. 2017, pp. 1–9, Jul. 2017, doi: 10.1155/2017/8617315.

[17] M. Li, Q. Wei, H. Wang, and X. Zhang, "Research on fault diagnosis of time-domain vibration signal based on convolutional neural networks," *Syst. Sci. Control Eng.*, vol. 7, no. 3, pp. 73–81, Dec. 2019, doi: 10.1080/21642583.2019.1661311.
[18] W. Sun, R. Zhao, R. Yan, S. Shao, and X. Chen, "Convolutional discriminative feature learning for induction motor fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1350–1359, Jun. 2017, doi: 10.1109/TII.2017.2672988.
[19] H. Pan, X. He, S. Tang, and F. Meng, "An improved bearing fault diagnosis method using one-dimensional CNN and LSTM," *J. Mech. Eng.*, vol. 64, nos. 7–8, pp. 443–452, 2018, doi: 10.5545/sv-jme.2018.5249.
[20] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. Van de Walle, and S. Van Hoecke, "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib.*, vol. 377, pp. 331–345, Sep. 2016, doi: 10.1016/j.jsv.2016.05.027.
[21] H. Oh, J. H. Jung, B. C. Jeon, and B. D. Youn, "Scalable and unsupervised feature engineering using vibration-imaging and deep learning for rotor system diagnosis," *IEEE Trans. Ind. Electron.*, vol. 65, no. 4, pp. 3539–3549, Apr. 2018, doi: 10.1109/TIE.2017.2752151.
[22] A. ElSaid, F. El Jamiy, J. Higgins, B. Wild, and T. Desell, "Optimizing long short-term memory recurrent neural networks using ant colony optimization to predict turbine engine vibration," *Appl. Soft Comput.*, vol. 73, pp. 969–991, Dec. 2018, doi: 10.1016/j.asoc.2018.09.013.
[23] M. Rao, Q. Li, D. Wei, and M. J. Zuo, "A deep bi-directional long short-term memory model for automatic rotating speed extraction from raw vibration signals," *Measurement*, vol. 158, Jul. 2020, Art. no. 107719, doi: 10.1016/j.measurement.2020.107719.
[24] L. Eren, T. Ince, and S. Kiranyaz, "A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier," *J. Signal Process. Syst.*, vol. 91, no. 2, pp. 179–189, 2019, doi: 10.1007/s11265-018-1378-3.
[25] C. Sobie, C. Freitas, and M. Nicolai, "Simulation-driven machine learning: Bearing fault classification," *Mech. Syst. Signal Process.*, vol. 99, pp. 403–419, Jan. 2018, doi: 10.1016/j.ymssp.2017.06.025.
[26] M. Van, D. T. Hoang, and H. J. Kang, "Bearing fault diagnosis using a particle swarm optimization-least squares wavelet support vector machine classifier," *Sensors*, vol. 20, no. 12, p. 3422, Jun. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/12/3422
[27] Z. Chen, A. Mauricio, W. Li, and K. Gryllias, "A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks," *Mech. Syst. Signal Process.*, vol. 140, Jun. 2020, Art. no. 106683, doi: 10.1016/j.ymssp.2020.106683.
[28] D.-T. Hoang and H.-J. Kang, "A survey on deep learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, Mar. 2019, doi: 10.1016/j.neucom.2018.06.078.
[29] M. Cerrada, R.-V. Sánchez, C. Li, F. Pacheco, D. Cabrera, J. V. de Oliveira, and R. E. Vásquez, "A review on data-driven fault severity assessment in rolling bearings," *Mech. Syst. Signal Process.*, vol. 99, pp. 169–196, Jan. 2018, doi: 10.1016/j.ymssp.2017.06.012.
[30] J. Wang, L. Ye, R. X. Gao, C. Li, and L. Zhang, "Digital twin for rotating machinery fault diagnosis in smart manufacturing," *Int. J. Prod. Res.*, vol. 57, no. 12, pp. 3920–3934, 2019, doi: 10.1080/00207543.2018.1552032.
[31] J. Chen, R. Randall, B. Peeters, W. Desmet, and H. Van der Auweraer, "Artificial neural network based fault diagnosis of IC engines," *Key Eng. Mater.*, vol. 518, pp. 47–56, Jul. 2012, doi: 10.4028/www.scientific.net/KEM.518.47.
[32] J. Chen, B. Eng, and M. Eng. (2013). Internal combustion engine diagnostics using vibration simulation. School of Mechanical and Manufacturing Engineering and The University of New South Wales. [Online]. Available: http://unsworks.unsw.edu.au/fapi/datastream/unsworks:11591/SOURCE01?view=true
[33] Y. L. Murphey, M. A. Masrur, Z. Chen, and B. Zhang, "Model-based fault diagnosis in electric drives using machine learning," *IEEE/ASME Trans. Mechatronics*, vol. 11, no. 3, pp. 290–303, Jun. 2006, doi: 10.1109/TMECH.2006.875568.
[34] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6, doi: 10.1109/ICEngTechnol.2017.8308186.
[35] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Feb. 2017, pp. 131–134, doi: 10.1109/BIGCOMP.2017.7881728.
[36] R. Viitala and R. Viitala, "Method and device to investigate the behavior of large rotors under continuously adjustable foundation stiffness," *J. Vibroeng.*, vol. 22, no. 5, pp. 1037–1054, Aug. 2020, doi: 10.21595/jve.2020.21107.
[37] E. Kurvinen, R. Viitala, T. Choudhury, J. Heikkinen, and J. Sopanen, "Simulation of subcritical vibrations of a large flexible rotor with varying spherical roller bearing clearance and roundness profiles," *Machines*, vol. 8, no. 2, p. 28, May 2020, doi: 10.3390/machines8020028.
[38] R. Viitala, J. Miettinen, T. Tiainen, and R. Viitala, "Rotor & bearing vibration dataset," Mendeley Data, V1, 2020, doi: 10.17632/pdrxyfprfk.1.
[39] J. E. Heikkinen, B. Ghalamchi, R. Viitala, J. Sopanen, J. Juhanko, A. Mikkola, and P. Kuosmanen, "Vibration analysis of paper machine's asymmetric tube roll supported by spherical roller bearings," *Mech. Syst. Signal Process.*, vol. 104, pp. 688–704, May 2018, doi: 10.1016/j.ymssp.2017.11.030.
[40] B. Ghalamchi, J. Sopanen, and A. Mikkola, "Simple and versatile dynamic model of spherical roller bearing," *Int. J. Rotating Mach.*, vol. 2013, pp. 1–13, Jan. 2013, doi: 10.1155/2013/567542.
[41] K. L. Cavalca, P. F. Cavalcante, and E. P. Okabe, "An investigation on the influence of the supporting structure on the dynamics of the rotor system," *Mech. Syst. Signal Process.*, vol. 19, no. 1, pp. 157–174, Jan. 2005, doi: 10.1016/j.ymssp.2004.04.001.
[42] E. Kurvinen, D. Bobylev, and J. Sopanen, "Rotor vibration dataset," Mendeley Data, V1, 2021, doi: 10.17632/ny42v5cvt6.1.
[43] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, 2011.
[44] A. Sanchez-Gonzalez, J. Godwin, T. Pfaff, R. Ying, J. Leskovec, and P. Battaglia, "Learning to simulate complex physics with graph networks," *Proc. Mach. Learn. Res.*, vol. 119, Jul. 2020, pp. 8459–8468. [Online]. Available: http://proceedings.mlr.press/v119/sanchez-gonzalez20a/sanchez-gonzalez20a.pdf
[45] X. Li, W. Zhang, Q. Ding, and J.-Q. Sun, "Multi-layer domain adaptation method for rolling bearing fault diagnosis," *Signal Process.*, vol. 157, pp. 180–197, Apr. 2019, doi: 10.1016/j.sigpro.2018.12.005.

DENIS BOBYLEV was born, in 1994. He received the M.Sc. degree in mechatronic system design from LUT University, Lappeenranta, Finland, in 2019. He has served as a Research Assistant for LUT University, from 2019 to 2020. His research interest includes the implementation of artificial intelligence techniques in the field of rotating machines.



TUHIN CHOUDHURY received the B.Sc. (Tech.) degree in mechanical engineering from Sikkim Manipal University, India, in 2011, and the M.Sc. degree in mechatronic system design from LUT University, Lappeenranta, Finland, in 2018, where he is currently pursuing the doctor's degree with the Department of Mechanical Engineering. As a Design Engineer, he worked on the development of medical devices and diagnostic instruments, from 2011 to 2016. His current research interests include designing, modeling, and simulation of rotating machines, and the analysis of rotor behavior to understand the root cause of unwanted vibrations, specifically due to unbalance.

JESSE O. MIETTINEN was born in Vantaa, Finland, in 1993. He received the B.Sc. degree in mechanical engineering and the M.Sc. degree in mechanical engineering from Aalto University, Finland, in 2017 and 2020, respectively, where he is currently pursuing the Ph.D. degree in mechanical engineering. From 2019 to 2020, he was a Research Assistant with the Engineering Design Research Group, School of Engineering, Aalto University. His research interests include deep learning-based fault diagnosis and soft sensor applications for rotor systems.

RISTO VIITALA received the M.Sc. degree in mechanical engineering from the University of Oulu, in 2018. He completed his master's thesis on the radial bearing force measurement of large rotors, and now he has continued his research on rotor dynamics as a final year Ph.D. Student. His research interests include the effects of foundation stiffness and external load on rotor vibrations. His work can be applied widely in industries, where the large rotating machines are used.

EMIL KURVINEN was born, in 1988. He received the M.Sc. (Tech.) and D.Sc. (Tech.) degrees in mechanical engineering from Lappeenranta University of Technology (LUT), in 2012 and 2016, respectively. From 2014 to 2015, he visited the University of Virginia as a Fulbright Visiting Scholar researching active magnetic bearings. From 2016 to 2017, he has served as an Engineer of structural dynamics for FS Dynamics Finland Oy Ab. In 2017, he started as a Postdoctoral Researcher at LUT. He has a solid background in machine design, especially in the design, simulation, and analyzing of rotating machines. His research interests include rotating machines, especially high-speed machines, digital twins, and freeze crystallization.

JUSSI SOPANEN (Member, IEEE) was born in Enonkoski, Finland, in 1974. He received the M.Sc. degree in mechanical engineering and the D.Sc. (technology) degree from LUT University, Lappeenranta, Finland, in 1999 and 2004, respectively. He was a Researcher with the Department of Mechanical Engineering, LUT University, from 1999 to 2006. He was a Product Development Engineer of electric machine manufacturing with Rotatek Finland Ltd., from 2004 to 2005. From 2006 to 2012, he was a Principal Lecturer of mechanical engineering and the Research Manager with the Faculty of Technology, Saimaa University of Applied Sciences, Lappeenranta. He is currently serving as a Professor for the Machine Dynamics Laboratory, LUT University. His research interests include rotor dynamics, multibody dynamics, and the mechanical design of electrical machines.