Proceedings of the ASME 2021 16th International Manufacturing Science and Engineering Conference
MSEC2021
June 21-25, 2020, Virtual, Online

MSEC2021-XXXX

A DEEP NEURAL NETWORK MODEL FOR LEARNING RUNTIME FREQUENCY RESPONSE FUNCTION USING SENSOR MEASUREMENTS

Yongzhi Qu[1], Gregory W. Vogl[2], Zechao Wang[3]

[1]University of Minnesota Duluth, MN, US
[2]National Institute of Standards and Technology, Gaithersburg, MD, US
[3]Wuhan University of Technology, Wuhan, China

ABSTRACT

The frequency response function (FRF), defined as the ratio between the Fourier transform of the time-domain output and the Fourier transform of the time-domain input, is a common tool to analyze the relationships between inputs and outputs of a mechanical system. Learning the FRF for mechanical systems can facilitate system identification, condition-based health monitoring, and improve performance metrics, by providing an input-output model that describes the system dynamics. Existing FRF identification assumes there is a one-to-one mapping between each input frequency component and output frequency component. However, during dynamic operations, the FRF can present complex dependencies with frequency cross-correlations due to modulation effects, nonlinearities, and mechanical noise. Furthermore, existing FRFs assume linearity between input-output spectrums with varying mechanical loads, while in practice FRFs can depend on the operating conditions and show high nonlinearities. Outputs of existing neural networks are typically low-dimensional labels rather than real-time high-dimensional measurements. This paper proposes a vector regression method based on deep neural networks for the learning of runtime FRFs from measurement data under different operating conditions. More specifically, a neural network based on an encoder-decoder with a symmetric compression structure is proposed. The deep encoder-decoder network features simultaneous learning of the regression relationship between input and output embeddings, as well as a discriminative model for output spectrum classification under different operating conditions. The learning model is validated using experimental data from a high-pressure hydraulic test rig. The results show that the proposed model can learn the FRF between sensor measurements under different operating conditions with high accuracy and denoising capability. The learned FRF model provides an estimation for sensor measurements when a physical sensor is not feasible and can be used for operating condition recognition.

Keywords: Frequency response function, Encoder-decoder, Neural network, Deep learning

1. INTRODUCTION

Frequency response functions (FRFs), defined as the ratio between input and output spectra, describes the steady-state relationship between each possible sinusoidal input and the corresponding output under zero initial conditions for linear time-invariant (LTI) systems. Learning FRFs between system inputs and outputs is important for system identification and control. An accurate FRF can help with condition-based maintenance and system response prediction of mechanical systems as it models the system dynamics and can be used to monitor system parameters. FRFs are often estimated analytically or obtained experimentally. However, in certain dynamic operations, the FRF can be difficult to predict or obtain due to experimental difficulties, unsteady operating conditions, and mechanical noise. For example, the FRF is often measured between the dynamic force at the tool tip and the dynamic response measured on a stationary point on milling machine tools, such as the spindle headstock. Since the cutting tool is rotating, the transmission path between the tool tip and the spindle headstock involves a flexible component (the cutting tool) and a moving contact surface. The source excitation spectrum from the tool tip can be modulated by bearing ball pass frequencies and low-pass filtered by the flexible component along the transmission path. One frequency component can affect the harmonic frequency components at another measurement point due to mechanical modulation. Also, the structural resonance frequencies can serve as carrier frequencies that shift the original excitation frequency components to a new

frequency. Therefore, in practical systems, especially with relative movement involved between input and output points, the frequency components can affect one another.

Because the FRFs may be difficult to determine for complex systems with moving parts and varying operating conditions, this paper proposes to learn a ***runtime FRF*** from sensor measurements that would model the previously mentioned factors with a data-driven black box model. Such a method would have applications for manufacturing systems in which sensor-based systems are needed to decrease machine downtime and increase product quality and enrich the knowledge of complex manufacturing processes. For example, it is typically not possible to monitor real-time cutting forces during milling processes within machine tools. However, as a future application, by using runtime FRFs to relate force signals to on-machine sensor signals with known cutting force before cutting, on-machine sensors could be set up to monitor cutting forces during real-time machining for quality control purposes.

Note that the *generalized FRF* has been used to define FRFs for nonlinear systems under nonzero initial conditions, especially for nonlinear Volterra systems [2][3]. Generalized FRFs include the effects on the output response from previous inputs. The term of ***runtime FRF*** in this paper is ***different from*** existing definitions for *generalized FRF*. As argued in the above section, this work assumes that in general mechanical systems, one input frequency component may be related and can be used to calculate a different frequency component in the output spectrum. In the generalized runtime FRF, these cross-relationships among different frequencies can be learned.

It is generally understood that one wave frequency cannot affect another wave frequency. Hence, the system response at one frequency can only be determined by the same frequency component in the input spectrum. However, for certain mechanical and electrical systems, there are certain conditions needed for this constraint to hold in practical FRF measurements: 1.) The signals are not modulated internally between inputs and outputs; 2.) Sensors for data collection have wide-enough frequency range to capture the significant spectral amplitudes; and 3.) Nonlinearities in the system are insignificant and do not add extra frequency components to the output response. In practical systems, especially for assemblies containing multiple components, the above assumptions may not hold.

To account for all the previously mentioned factors, we propose to learn the ***generalized runtime FRF*** from sensor measurements with a data-driven black box model. The rest of the paper is organized as follows: Section 2 reviews related work for FRFs and vector regression; Section 3 introduces the basic background of encoder-decoder neural networks and presents the methodology; Section 4 provides the experimental setup for the case study; Section 5 presents the results and data analysis; and Section 6 concludes the paper.

2. RELATED WORKS

Existing research on FRF identification can be separated into several categories: physics-based FRF evaluation, experimental data-based FRF estimation, and hybrid methods for FRF identification. Theoretically, FRFs can be obtained analytically from the governing differential equations of the system [1]. Towards this direction, the effects of different factors on the FRF have been studied. For machine tools, the single-point FRF at the tool tip has been widely studied for the identification of chatter-free zones. The effects of rotational speed [5] and nonlinear behavior [6] have been considered for FRF identification. The effect of rotor rub-impact has also been modeled for the analysis of the FRF of rotor systems [7]. For nonlinear systems with fading memories, Volterra-based FRF models can be established when the system parameters are known using Volterra series [4]. However, Volterra models require the complete knowledge of the system parameters for the governing differential equations [3]. In applicable cases, system parameters can be measured offline or calculated using geometric and material parameters. However, for practical mechanical systems, the runtime boundary conditions and system parameters may change due to external loads and dynamic interactions between moving parts. Therefore, accurate estimation of system parameters are often infeasible.

With regards to experimental methods, impact hammer testing and sinusoidal function sweeping-based FRF identification are available in most commercial modal testing systems. Receptance coupling provides a hybrid approach to identify the FRF for assemblies. Several studies have tried to obtain FRFs for coupled mechanical structures with multiple components. For example, Schmitz and Duncan predicted the output frequency response of an assembly of nested components with common neutral axes using the receptance coupling approach [8].

There have been a couple of attempts towards runtime FRF determination. Kushnir proposed to estimate the runtime FRF for a lathe using the measured vibration spectrum at the top plate and spindle headstock [9]. The author took the average of the input and output spectrums under different spindle speeds, and then took the ratio between the averages to get the dynamic FRF. This ratio assumes no interaction among different frequency components in the input and output spectrums. Thenozhi and Tang used a radial basis neural network and support vector machine (SVM) to learn the frequency response function from simulation data [3]. The input space they used for the learning process was set as the system input amplitude and frequency, and learning outputs were set as the amplitudes of frequency response. They built a regression model between the learning inputs and outputs. While their input space is in the space of R^2 (i.e., two dimensions) and output space in R (i.e., one dimension), their learning model learned the output frequency response as $OFR = f(a, w)$, where a is the amplitude and w is the frequency, and OFR is the amplitude of the corresponding output frequency response. Similar to the model in [9], the model in [3] is equivalent to finding the average ratio between input and output spectrums. Once again, this ratio assumes no interaction among different frequency components and one common average model is obtained for all operating conditions.

Therefore, all existing approaches assume that the FRF function is strictly restricted to one input frequency component

affecting only the same frequency component in the output response. In this manner, the FRF is linear and cannot include nonlinearities under different operating conditions. The approach proposed in this paper is dramatically different from previous work, as this paper proposes a vector regression algorithm using neural networks in the whole frequency range, which allows the cross-relationships between different frequencies to be established. Since the model is nonlinear and includes interactions between different frequencies, it allows different ratios between input spectrums and output spectrums at different operating conditions. In other words, the model provides a regression relationship between the vector space of input spectrum and output spectrum. Also, this paper uses practical sensor measurements from mechanical systems with potentially low signal-to-noise ratios. It can provide a robust approach for real-time classification of operating conditions using a real-time input spectrum. As a general vector regression-based FRF identification approach, it can be applied to any system input and output spectrum measurements.

Next, related works in deep learning for vector regression will be reviewed. An encoder-decoder neural network is a special type of neural network that features dimension reduction during the encoding process and signal reconstruction during the decoding process. The most common application of encoder-decoder networks is an autoencoder, where the training input and training output are identical. An autoencoder is often used for feature extraction as an unsupervised algorithm for dimension reduction purposes and has seen many applications [10]. Supervised and semi-supervised autoencoders have also drawn much research attention for concurrent feature extraction and classification tasks [11]. Lei *et al*. proposed that the autoencoder can be used for regression purposes if the input and output are set with different time series as regularizers [12]. Inspired by Lei's work, this paper reveals an encoder-decoder structure for multivariate vector regression. Lei *et al*. used a supervised autoencoder for the classification task, and the encoding-and-decoding network serves as a regularizer. In this paper, we consider the vectors fed to the outputs of the decoder networks as underlying labels in vector form in the target space of regression. Therefore, the encoder-decoder network will simultaneously learn the mapping relationship as well as the regression tasks for different input-output groups.

An encoder-decoder structure has also been applied to sequence-to-sequence learning for speech recognition [13]. However, the structure used for sequence-to-sequence recognition took a recurrent neural network/long short-term memory (RNN/LSTM) structure to consider time-dependent relationships with a sequential memory mechanism other than an autoencoder-like structure, which does not focus on temporal dependency. While our model shares a similar common ground with sequence-to-sequence models, the proposed approach distinguishes itself from the RNN/LSTM models. The proposed encoder-decoder network aims to learn the intra-class and inter-class spatial relationship between R^N to R^N in vector space, while sequence-to-sequence models aim to model the dependency in temporal space inside the sequence vector and then perform regression tasks in the spatial space. Our model assumes a general regression relationship between two vectors without explicitly considering temporal dependency and therefore the application scenarios are different.

3. BACKGROUND AND PROPOSED METHODS

As illustrated in the Introduction section, we formulated the identification of the generalized runtime FRF as a vector regression problem. The regression can be understood in two levels. The first level is that each frequency component in the output spectrum will be determined by a regression model that takes a vector of the whole input spectrum. The second level is that the regression relationship is learned in the vector space for inputs-outputs pairs under different operating conditions, which can be understood as regression in Hilbert space or function space. Such a regression problem can be solved by a deep encoder-decoder neural network. Next, we introduce encoder-decoder networks and propose a customized model for FRF learning.

3.1 Background Introduction and Proposed Encoder-Decoder Network

An encoder-decoder is a type of neural network that first encodes the inputs to a low dimensional space, and then decodes the data back to high dimensional space. One example of encoder-decoder network is an autoencoder, where the inputs and outputs are the same, and it works as an unsupervised dimension reduction model. A more general form of an encoder-decoder network can take different data as inputs and outputs, and serve as a vector regression model.

In this paper, a partially-tied-weight deep encoder-decoder network for vector regression between the input and output spectrums is proposed. The structure of the deep network is shown in Figure 1. In the proposed structure, tied weights are set for the embedded space learning for both the input spectrum and output spectrum, and an independent output embedding layer is added to allow different coordinates in the embedding space for the inputs and outputs, as shown in the red box in Figure 2. Then, the input and output embedding layers are fully connected to learn the mapping relationship between the input and output embeddings. The number of nodes in each layer are: 200, 100, 50, 50, 100, and 200 with a symmetric layout. Layer 1 and Layer 2 share the same weights with Layer 5 and Layer 4, respectively.

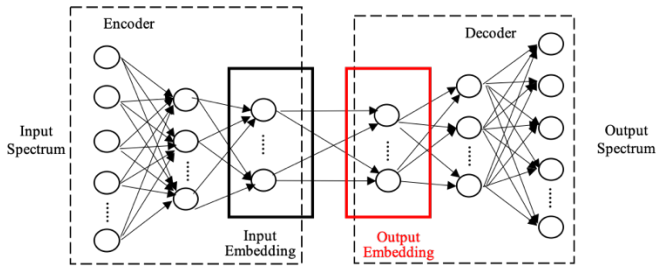

FIGURE 2. PROPOSED ENCODER-DECODER STRUCTURE FOR FRF LEARNING.

The whole network can be represented by the following equations:

$$X_{emb} = f_1(X) \quad (2)$$

$$Y_{emb} = f_2(X_{emb}) \quad (3)$$

$$Y = f_1^{-1}(Y_{emb}), or\ Y_{emb} = f_1(Y) \quad (4)$$

where X is the input of the neural network, Y is the output of the neural network, X_{emb} is the embedded coordinates for X, and Y_{emb} is the embedded coordinates for Y. The proposed encoder-decoder network includes a tied symmetric layer for the learning of input and output embeddings, as well as a fully connected layer from the input embedding to the output embedding, to enable learning of the mapping relationship between the two spectrums.

4. EXPERIMENTAL SETUP

4.1 Test Rig and Sensor Deployments

To validate the proposed FRF learning model, experiments were conducted on a high-pressure hydraulic pipeline system, as shown in Figure 3. For the given hydraulic system, the excitation force comes from the distributed dynamic pressure of the flowing liquid imposed on the inner surface of the pipeline. Measurement of such a distributed dynamic pressure would be physically infeasible. Therefore, in order to measure the excitation force on the pipeline, the dynamic strains on the outer surface of the pipeline were measured with fiber Bragg grating (FBG) sensors. The dynamic strains measured by FBG sensors serve as an indirect measurement of the dynamic pressure inside the pipeline. Using strain sensors for static pressure measurements is a common practice. Dynamic pressure measurements with strain gauges has been recently reported [14]. In this study, since the interested frequency range is about 0 Hz to 1000 Hz, which are the limits of the data collection capabilities, it can be assumed that the strain is linearly related to the dynamic pressure, which is the excitation to the pipeline. Therefore, the strain spectrum was considered to be the input spectrum of the hydraulic system and the FRF learning model, while the vibration spectrum was considered as the output spectrum.

In order to validate the proposed method for learning a generalized runtime FRF under different operating conditions, three sets of experiments were conducted under different operating pressures, namely, 4.62 MPa, 5.65 MPa, or 9.3 MPa. Vibration data and strain data over a section of straight pipeline were collected. A total of about 40 minutes of data were collected. Vibration data were collected using a tri-axial accelerometer and strain data were collected using FBG sensors. There are a total of five sensor measurements: X, Y, and Z accelerations, circumferential strain, and axial strain. For the vibration coordinates, the X axis points down to the ground, the Y axis points axially, and the Z axis points transversely, as shown in Figure 4. The sampling rate for vibration signals is 10240 Hz, and the sampling rate for FBG strains is 2000 Hz. While the sampling frequency for the FBG sensors (limited by the data acquisition system) falls a little short to cover 0 Hz to 1000 Hz, it would not significantly affect our validation process, due to the lower spectral amplitudes near 1000 Hz.

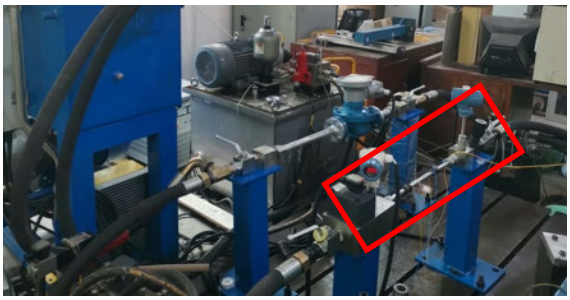

FIGURE 3. EXPERIMENTAL HYDRAULIC SYSTEM.

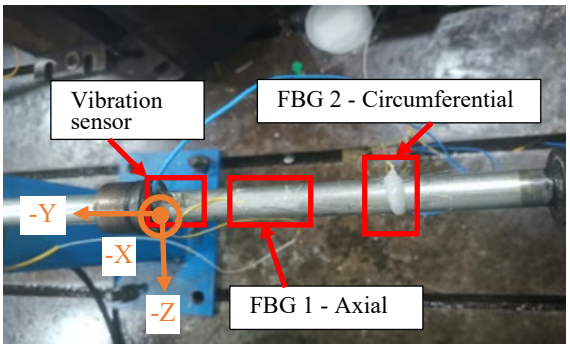

FIGURE 4. SENSOR LOCATIONS.

Based on experimental research under different pressures and quantitative calibration, it was found that the circumferential FBG is more sensitive to the pressure of the pipeline and presents a nearly linear relationship to the hydraulic pressure. Therefore, the FBG 2 sensor was selected for the FRF learning. For vibration, the Z-axis acceleration, which represents the transverse vibration in the horizonal plane and perpendicular to the axial direction of the pipeline, is chosen as the output corresponding to the FBG input.

4.2 Data Preparation

In order to provide rich enough data for FRF learning, each signal collected over about 40 minutes was segmented into multiple samples and then the Fast Fourier Transform (FFT) was used to yield the frequency spectrum of each segment. Since our vibration data and strain data were collected separately with different hardware and software systems asynchronously (misaligned by 10 s of seconds to a few minutes), in order to overcome the synchronization error and obtain more spectrums for training, the original data under the same operating condition was segmented with 50 percent overlap, which means adjacent data segments overlap by 50 percent with each other. Each data segment contains signals collected over 0.2 s, corresponding to 2024 data points for acceleration data and 400 data points for the FBG strain data, in the time domain. Both FFTs for acceleration and strain have a frequency resolution of 5 Hz, which provides 201 spectral data points for the frequency range of 0 Hz to

1000 Hz. The first FFT point, which is for 0 Hz, is removed, which yields 200 frequencies for learning purposes. A total of 23 254 pairs of sample segments for the training and test is obtained, where a strain spectrum and an acceleration spectrum form one pair of sample segments. The whole data sets are summarized in Table 1, in which each sample is a sample segment, a group of sequential data points, collected over 0.2 s.

TABLE 1. NUMBER OF SAMPLES FOR FRF LEARNING.

No. of samples	4.62 MPa	5.65 MPa	9.3 MPa	Total No.
FBG Strain	5242	12 011	6001	23 254
Vibration	5242	12 011	6001	23 254

During the model learning and testing process, about 85 percent of data was used for training and 15 percent was used for testing. Since the samples include overlap, which may help the testing processing if a similar sample with overlap has been seen in the training datasets, a random pool of datasets was not selected for training. Instead, for each of the operating conditions in Table 1, the first 85 percent of the data collected over time was used as training data and the latter 15 percent of data was used for testing. For example, for 4.62 MPa, the first 4500 samples out of the total 5242 samples were utilized for training. The rest of the samples served as testing datasets. A total of 20 000 samples from three different operating conditions were used for training and 3254 samples were used for testing.

4.3 Model Setting

This section provides the setting of our models for the deep encoder-decoder network. Since there exists a hybrid setting for the weight parameter, the total number of parameters for the network is 22 840. This is compared to 44 840 parameters if an un-tied encoder-decoder would have been used. Therefore, the partially-tied weights significantly reduced the number of parameters in the model. A batch optimization algorithm was adopted with a batch size of 100 samples, and the learning rate was set as 0.008. Twenty (20) epochs were performed to have the training process converge when the training loss and validation loss stopped decreasing. A quick overfit check with a 20% validation dataset over 50 epochs indicated that more epochs do not introduce overfitting since the accuracy on the validation set did not change, while the validation loss stayed relatively low after about 20 epochs.

The learning process was conducted on a MacBook Air computer with a 1.8 GHz Intel i5 processor and 8 GB of memory. On average, learning took about 11.8 s for 20 epochs, which is about 0.59 s for each epoch. Therefore, the training process is quite efficient.

5. RESULTS AND DISCUSSION

5.1 Results with the Proposed Model

In this section, the entire learning process is visualized, starting with the raw data and ending with the predicted output response from the learned FRF model. Figure 5 shows plots of the raw acceleration and FBG strain signals, and Figure 6 shows examples of FFTs of sample segments used for model learning. Note that many of the amplitudes in the strain spectrum are about 25% of the peak amplitude shown around 250 Hz. In contrast, the noise floor in the vibration spectrum is about 1% or less of its peak amplitude around 250 Hz; the vibration sensor has a much higher signal-to-noise ratio. Also, there is one dominant frequency component, which is suspected to be the first natural frequency of the pipeline structure that should capture the majority of vibrational displacements. The strain spectrum contains more noise and is flatter compared with the vibration spectrum. Yet, both spectrums share the same dominant frequency.

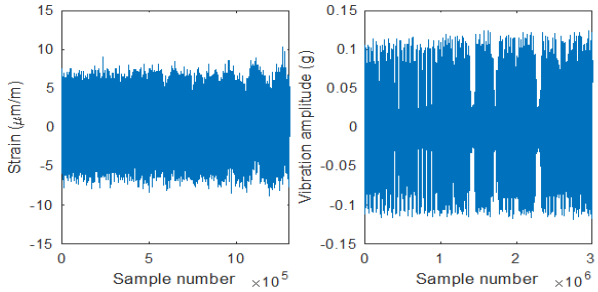

FIGURE 5. SAMPLE STRAIN AND VIBRATION SIGNALS.

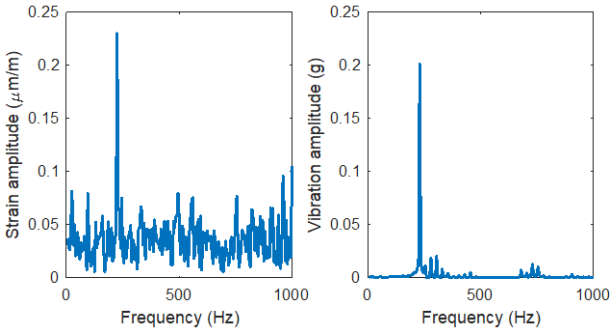

FIGURE 6. EXAMPLE FFTS OF STRAIN AND VIBRATION SIGNALS.

Without further processing of the data before use in the neural network, the peak frequency components would dominate over all other frequency components with regards to training errors. During the training process to map strain (the input) to vibration (the output), the low-amplitude spectral components would be neglected, since the error contributions from smaller amplitudes are much lower than those for higher amplitudes. Thus, the natural log of vibration FFT amplitude is set as the output for the neural network, which enables network learning to be influenced by all frequency content. A similar approach has been used in another method to allow all frequency content to influence the FRF-based solution [15]. Figure 7 shows the same sample spectrum as in Figure 6, except that the vibration spectrum is the natural log of its original spectrum. As one can see from Figure 7, the vibration spectrum is flattened by the log operation and shows more details for the lower-amplitude frequency components. The spectrums are then normalized to [0, 1] to eliminate scalings and were fed into the deep neural

network model. The normalization process rescales the data to a common range. Also, since Sigmoid is used as the activation function in this work, output of the neural networks needs to be normalized to [0,1].

Next, the learning results with the proposed encoder-decoder regression model are discussed. Figure 8 shows an example of the predicted outputted acceleration spectrum tested with the inputted strain spectrum for a sample segment. It can be seen that the predicted vibration spectrum closely follows the actual vibration spectrum corresponding to the testing data. However, the actual spectrum has more fluctuations, while the predicted spectrum is smoother. While it can be expected that the predicted spectrum presents an 'averaging' effect, it is surprising to find that output spectrums calculated for all test samples under the same operating condition are almost identical. This indicates that the trained deep encoder-decoder networks have excellent denoising capability. We are further interested in evaluating whether the predicted spectrums correctly reflect and separate with different operating conditions, which can help us evaluate the generalization capability of the trained deep neural network.

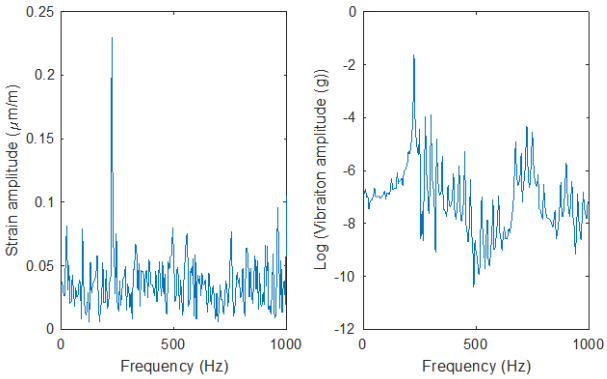

FIGURE 7. STRAIN SPECTRUM AND LOG-OF-VIBRATION SPECTRUM.

Since there are three different operating conditions, it is shown next whether the predicted output response correctly reflects the changing operating conditions, which represent the regression performance in the vector space. Figure 9 shows the predicted spectrums for three different operating conditions, a.k.a. 4.62 MPa, 5.65 MPa, and 9.3 MPa. Again, extensive examinations reveal that all predicted spectrums for one particular operating condition are almost identical even if different sample segments are used as the testing input. However, as Figure 9 shows, the learned generalized FRF captured the real difference between different operating conditions. Figure 9 shows that the predicted spectrum of 9.3 MPa clearly separates itself from those for 4.62 MPa and 5.65 MPa. The results for 4.62 MPa and 5.65 MPa can hardly be separated from each other. The reason is explained in that the data for 4.62 MPa and 5.65 MPa were collected from the same hydraulic system with a setting of 5 MPa. On the other hand, for 9.3 MPa, the control parameter was actually set to 10 MPa.

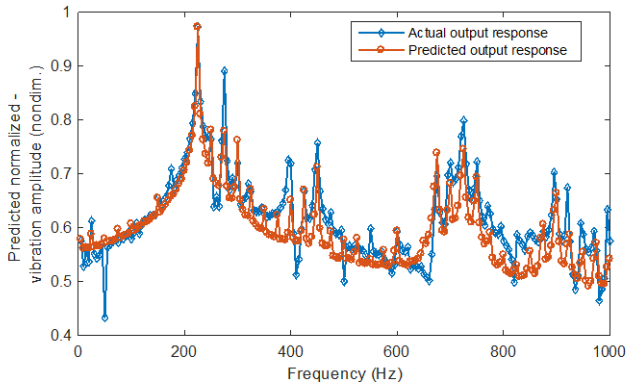

FIGURE 8. EXAMPLE OUTPUT SAMPLE SPECTRUM AND ITS PREDICTED SPECTRUM.

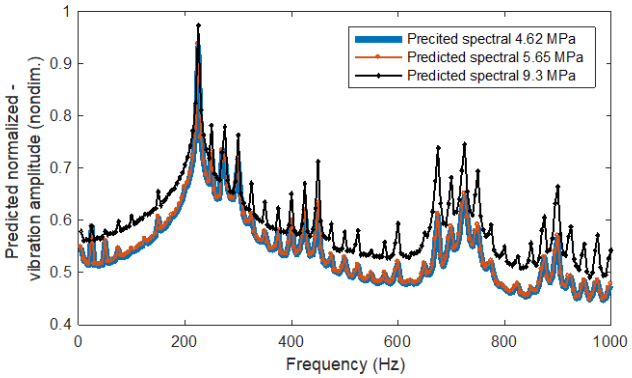

FIGURE 9. PREDICTED VIBRATION SPECTRUMS FOR THREE OPERATING CONDITIONS.

To further validate the results, the averaged spectrum for each of the three operating conditions are shown in Figure 10. The averaged spectrum is the average over all training and testing data. It is perhaps surprising to see that the learned spectrums approximate the average spectrums. Note that the amplitude of the averaged spectrum is slightly lowered due to the averaging effect over non-strictly aligned data, which is typical for averaging of FRFs with noise. Once again, the actual spectrums of 4.62 MPa and 5.65 MPa are very close to each other. Considering that the strain FRFs (the inputs) are extremely noisy with minor differences in the training targets fed to the model, the model cannot learn discriminative outputs for 4.62 MPa and 5.65 MPa. To better explain the averaging and denoising effects of the proposed model for FRF learning, examples of testing input spectrums and the corresponding output spectrums are plotted in Figure 11.

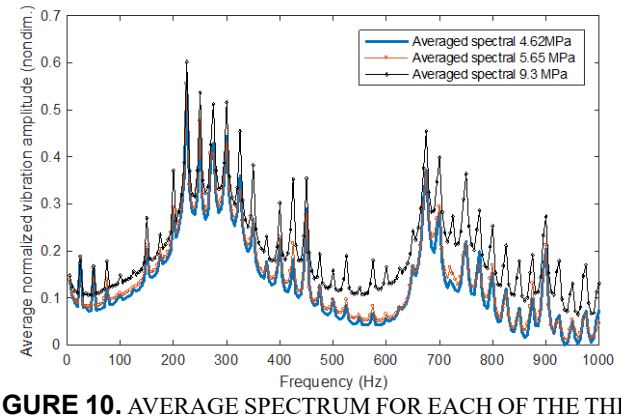

FIGURE 10. AVERAGE SPECTRUM FOR EACH OF THE THREE SEPARATE OPERATING CONDITIONS.

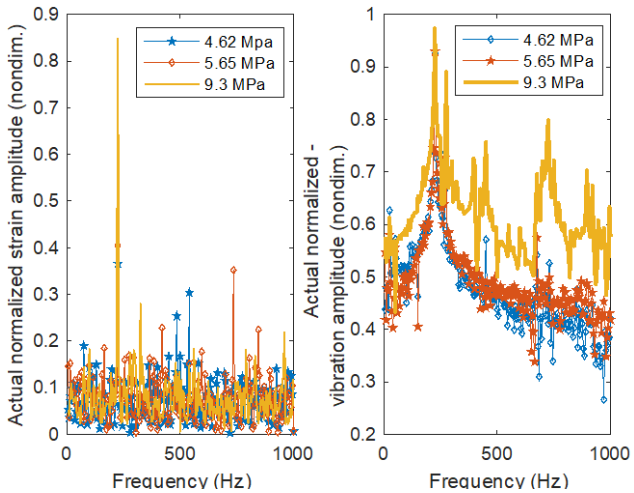

FIGURE 11. SPECTRUMS FOR AN ACTUAL SAMPLE USED FOR TESTING.

It can be seen from Figure 11 that the input spectrums are extremely noisy and the actual corresponding vibration spectrums are relatively noisy, as well. Furthermore, the vibration FFT amplitudes for 4.62 MPa are slightly lower than those for 5.65 MPa, and both are clearly separated from the FFT for 9.3 MPa. Note the differences of the average predicted spectrums in Figure 10 with the actual vibration spectrums in Figure 11; the noise-related fluctuations are filtered during the learning process, which justify the robustness of the proposed deep encoder-decoder network.

5.2 Comparison with Shallow Neural Networks

Next, the accuracy of the generalized runtime FRF is quantified. For various samples, the intra-group predictions under the same operating condition should be very close to each other, since they are sampled from the same process. In contrast, the inter-group predictions should be distinguishable in order to separate responses due to different operating conditions. Therefore, it is expected that the learned model could correctly predict discriminative results under different operating conditions.

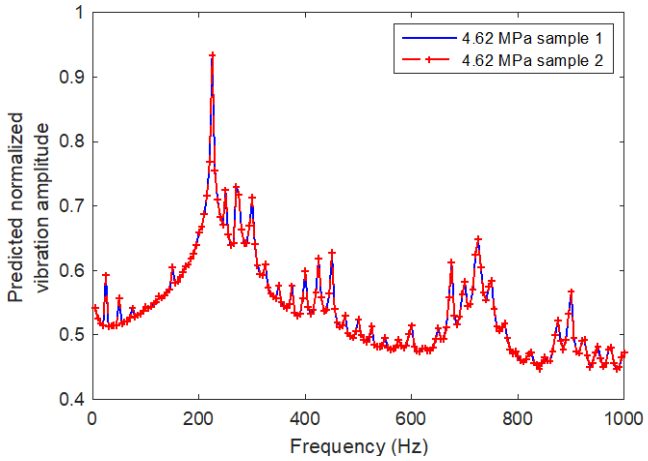

FIGURE 12. RESULTS COMPARISON FOR DIFFERENT TEST SAMPLES UNDER THE SAME OPERATING CONDITIONS FOR THE PROPOSED DEEP MODEL.

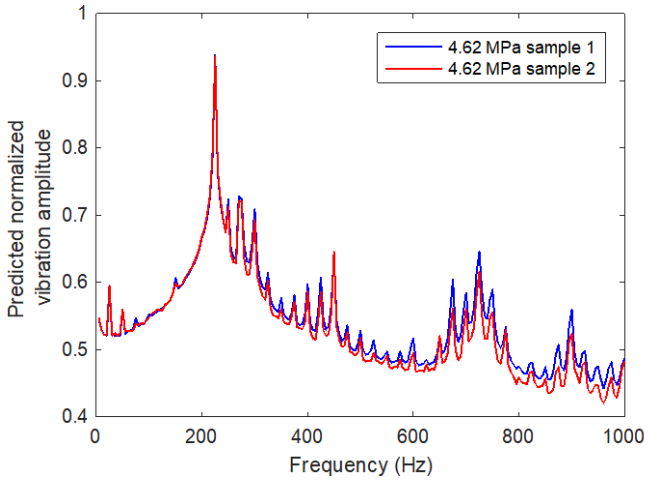

FIGURE 13. RESULTS COMPARISON FOR DIFFERENT TEST SAMPLES UNDER THE SAME OPERATING CONDITIONS FOR THE SHALLOW NEURAL NETWORK.

Figure 12 and 13 present the prediction results for the same two samples from the same operating conditions using the proposed deep neural network and a shallow neural network, respectively. It can be seen from Figure 12 and Figure 13 that the deep neural network performs much more robustly to produce almost identical results for the same process with noise filtering capability. However, the results from the shallow neural network present more fluctuations even if they come from the same process. For the prediction results of 9.3 MPa, the results from both the proposed deep model and the shallow network show fluctuations. However, the deep model still presents more robustness.

Also, Figure 9 shows that the deep neural network can predict clearly separable results for different operating processes. In order to quantify the performance of the proposed deep neural network and shallow neural network, we calculated and summarized the in-class fluctuations in terms of standard

deviation. Since the standard deviation (STD) is calculated over multiple vectors, the largest STD over all dimensions is used to indicate the fluctuation, as shown in Table 2. Similar results can be obtained for the average standard deviation as well.

TABLE 2. ROBUSTNESS OF PREDICTION INDICATED BY IN-CLASS STD.

In-class STD	Class – 1 (4.62 MPa)	Class – 2 (5.65 MPa)	Class - 3 (9.3 MPa)
Proposed deep NN	0.0083	0.0069	0.0238
Shallow NN	0.0200	0.0145	0.0310

Next, we show the separation capability of predicted results from the proposed deep and shallow neural networks. The estimated FRFs using trained neural networks serve as underlying labels to each operating condition. However, since there are no true labels for them, a k-mean clustering task is performed over both networks to see whether the predicted results are discriminative with regards to the operating process or not. Therefore, we can evaluate whether the predicted results can detect a pressure change in the pipeline.

The k-mean clustering results are shown for all conditions in Table 3, but because Class – 3 corresponds to the 9.3 MPa operating condition and has more noise in the test data, Table 4 shows the separation accuracy for only Class – 3. It can be seen in Table 3 and Table 4 that the proposed deep regression model based on encoder-decoder networks clearly outperforms the shallow neural network with both FRF prediction robustness and discriminative capability.

TABLE 3. GROUPING ACCURACY FROM CLUSTERING RESULTS (ALL).

Comparison of FRF prediction accuracy, All	Deep Encoder-decoder NN	Shallow NN
k-means clustering accuracy	98.59%	97.79%

TABLE 4. GROUPING ACCURACY FROM CLUSTERING RESULTS (9.3 MPA).

Comparison of FRF prediction accuracy, Class - 3	Deep Encoder-decoder NN	Shallow NN
k-means clustering accuracy	95.51%	92.91%

6. CONCLUSION

In this paper, identification of a generalized runtime FRF under multiple operating conditions was proposed as a supervised vector regression problem. The proposed model was tested for regression performance in the vector space to classify different inputs under changed operating conditions. A deep encoder-decoder-based model was proposed for the learning task. The model features partially-tied weights for the encoding and decoding process. An added layer for the output embedding was proposed to store different coordinates under the embedded space. The output embedding is fully connected with the input embeddings, which creates the mapping/regression relationship between them in the embedded space.

The model was validated using experimental measurement data from a hydraulic system, in which the relationship to be learned is between the vibration response (measured by an accelerometer) and the pressure (measured by an FBG strain sensor). The results showed that the proposed model can effectively learn the FRF function between the input and output spectrums under different operating conditions. The learned model also demonstrated excellent denoising capability, which was not shown by other FRF approaches. The presented work was for learning and classification of a system under very controlled conditions, and as such, is the first step of a longer process that will result in the generalized runtime FRF for prediction purposes.

NIST DISCLAIMER

Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States. Approved for public release; distribution is unlimited.

ACKNOWLEDGEMENTS

Yongzhi Qu acknowledges the Department of Commerce of the United States for partial support of this research under contract No. 70NANB20H175, and partial support from the University of Minnesota in Grant-in-Aid of Research, Artistry and Scholarship (GIA).

REFERENCES

[1] Liu MY, Wang ZC, Zhou ZD, et al. Vibration response of multi-span fluid-conveying pipe with multiple accessories under complex boundary conditions. Eur J Mech A-Solid 2018;72:41-56.

[2] Lang, Z.-Q., & Billings, S. (1996). Output frequency characteristics of nonlinear systems. International Journal of Control, 64(6), 1049–1067.

[3] S. Thenozhi, Y. Tang, Learning-based frequency response function estimation for nonlinear systems, Int. J. Syst. Sci. 49 (2018) 2287–2297.

[4] Xing Jian Jing, Zi Qiang Lang, Stephen A. Billings, Output frequency response function-based analysis for nonlinear Volterra systems, Mechanical Systems and Signal Processing, Volume 22, Issue 1, 2008.

[5] Niccolò Grossi, Lorenzo Sallese, Filippo Montevecchi, Antonio Scippa, Gianni Campatelli, Speed-varying Machine Tool Dynamics Identification Through Chatter Detection and Receptance Coupling, Procedia CIRP, Volume 55, 2016, pp. 77-82.

[6] A.S. Delgado, E. Ozturk, N. Sims, Analysis of Non-linear Machine Tool Dynamic Behavior, Procedia Engineering, Volume 63, 2013, pp. 761-770.

[7] Y. Liu, Y.L. Zhao, J.T. Li, H. Ma, Q. Yang, X.X. Yan, Application of weighted contribution rate of nonlinear output frequency response functions to rotor rub-impact, Mechanical Systems and Signal Processing, Volume 136, 2020.

[8] Tony L. Schmitz, G. Scott Duncan, Receptance coupling for dynamics prediction of assemblies with coincident neutral axes, Journal of Sound and Vibration, Volume 289, Issues 4–5, 2006, Pages 1045-1065.

[9] Kushnir, E. "Determination of Machine Tool Frequency Response Function During Cutting." Proceedings of the ASME 2004 International Mechanical Engineering Congress and Exposition. Pressure Vessels and Piping. Anaheim, California, USA. November 13–19, 2004. pp. 63-69. ASME.

[10] Qu, Y.; He, M.; Deutsch, J.; He, D. Detection of Pitting in Gears Using a Deep Sparse Autoencoder. Appl. Sci. 2017, 7, 515.

[11] Ranzato, M. and Szummer, M. (2008). Semi-supervised learning of compact document representations with deep networks. In Proceedings of the 25th International Conference on Machine Learning (ICML), pages 792–799.

[12] Lei Le, Andrew Patterson and Martha White, Supervised autoencoders: Improving generalization performance with unsupervised regularizers, 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

[13] Ilya Sutskever, Oriol Vinyals, Quoc V. Le, Sequence to Sequence Learning with Neural Networks, NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems, Cambridge, MA, USA.

[14] Fatima Garcia Castro, Olivier de Sagazan, Nathalie Coulon, Antoni Homs Corbera, Dario Fassini, Jeremy Cramer, France Le Bihan, μ-Si strain gauge array on flexible substrate for dynamic pressure measurement, Sensors and Actuators A: Physical, Volume 315, 2020.

[15] Gregory W. Vogl, M. Alkan Donmez, A defect-driven diagnostic method for machine tool spindles, CIRP Annals, Volume 64, Issue 1, 2015, pp. 377-380.