A Novel Orbit-based CNN Model for Automatic Fault Identification of Rotating Machines

Xiaomo Jiang\textsuperscript{1}, Shuhua Yang\textsuperscript{2}, Fumin Wang\textsuperscript{3}, Shengli Xu\textsuperscript{3}, and Xiaofang Wang\textsuperscript{3}

\textsuperscript{1}State Key Lab of Structural Analysis for Industrial Equipment, Department of Engineering Mechanics, Dalian University of Technology, Dalian 116024, China
\texttt{xiaomojiang2019@dlut.edu.cn}

\textsuperscript{2}Research Institute of Shenyang Blower Works Group, Shenyang 110869, China
\texttt{yshemail@sina.com}

\textsuperscript{3}School of Energy and Power Engineering, Dalian University of Technology, Dalian 116024, China
\texttt{wangfumin@mail.dlut.edu.cn}
\texttt{xusl@dlut.edu.cn}
\texttt{dlwxf@dlut.edu.cn}

ABSTRACT

Various faults in high-fidelity turbomachinery such as steam turbines and centrifugal compressors usually result in unplanned outage thus lowering the reliability and productivity while largely increasing the maintenance costs. Condition monitoring has been increasingly applied to provide early alerting on component faults by using the vibration signals. However, each type of fault in different types of rotating machines usually require an individual model to isolate the damage for accurate condition monitoring, which require costly computation efforts and resources due to the data uncertainties and modeling complexity. This paper presents a generic convolution neural network model for accurately automatic identification of various faults in common rotating machines by utilizing the shaft orbits generated from vibration signals, considering the high non-linearity and uncertainty of the sensed vibration signals. The sensor anomalies and environmental noise in the vibration signals are first addressed through Bayesian wavelet noise reduction filtering. Shaft orbit images are generated from the cleansed vibration data collected from different turbomachinery with four different fault modes. A multi-layer convolutional neural network model so-called OrbitNet is developed to classify and identify the shaft orbit images of each fault. Finally, the fault identification of rotating machinery is realized through the automated identification process. The proposed approach retains the fault information in the axis trajectory to the greatest extent, and can adeptly extract and accurately identify features of various faults. The effectiveness and feasibility of the proposed methodology is demonstrated by using the sensed vibration signals collected from real-world centrifugal compressors and steam turbines with different fault modes.

1. INTRODUCTION

Large turbomachines such as gas turbines, steam turbines and centrifugal compressors are the key industrial equipment in power plants, oil & gas, and petrochemistry. An unplanned breakdown due to component damage in these rotating machines may result in significant loss of properties and life. Many researchers from both academy and industries have attempted to enhance the safety, reliability, performance, availability and maintainability of these rotating machines thus improving their productivity and lowering their operation and maintenance cost. Real-time condition monitoring has become an increasingly important approach to pursue this purpose (e.g., Jardine et al. 2006). As the quick development of both turbomachines and information technology in the past decade, numerous data are continuously collected from the operation of the rotating machines. Recently image-driven methods have been proven to provide a powerful tool for damage identification and condition assessment of a large turbomachine, which is the focus of this paper.
In the past few years artificial intelligence methods have been exploited to utilize the shaft orbits for automatic fault identification and classification of turbomachinery (e.g., Carbajal-Hernández et al. 2016, Jeong et al. 2016, Wu et al. 2018, Khodja et al. 2019). Carbajal-Hernández et al. (2016) developed the Lernmatrix associative memory approach associated with orbital pattern recognition to classify the imbalance and misalignment faults of induction motors. Jeong et al. (2016) presented a convolutional neural network (CNN) based deep learning model to identify the fault of rotating machinery via orbit images. The images were preprocessed by re-orienting, rescaling, normalizing, and denoising to improve the identification accuracy. The laboratory testing data were used in the methodology demonstration. Wu et al. (2018) presented two shape description symbols from the orbits by using high functions and Fourier transform which are used in support vector machine (SVM) approach for fault detection of turbomachine. Khodja et al. (2019) also implemented the CNN deep learning model to classify the bearing faults of rotating machines based on the orbit images generated from the time-frequency features of the vibration signals. Noise impact on the identification accuracy was also investigated through the experimental simulation.

This study presents a new CNN deep learning model integrated with advanced signal processing for automatic fault identification of turbomachines by using the shaft orbits, considering the data imperfection. The raw vibration signals are first cleansed through waveform compensation and Bayesian wavelet noise reduction filtering. The clean signals are then used to construct the shaft orbits. A CNN model is established to identify the possible damage for turbomachine based on the obtained orbits. The proposed methodology is illustrated with the data collected from a real centrifugal compressor along with damage events. In the following sections of the article, the common fault modes represented by different orbit shapes are first introduced. The waveform compensation and wavelet denoising approaches are then explained. A convolution neural network model is developed to identify possible damage from the orbits. Data and events from centrifugal compressors are employed to demonstrate the feasibility and availability of the proposed methodology.

2. CONVOLUTION NEURAL NETWORK MODEL

The convolution neural network (CNN) provides a powerful tool for pattern recognition and image classification without the need of complicated feature extraction from images. It usually consists of three important layers: convolutional, subsampling or pooling, and fully connected. The convolutional layer intends to extract features from the images, the subsampling layer aims to reduce the dimensionality of the features through a fixed operator pooling in order to concentrate on the most important elements, while the fully connected layer rearranges the identified features to be a flattened form for prediction or recognition of a given image. Typically, a few of convolution and subsampling combinations are used in creating a CNN model to improve its accuracy. In comparison to the traditional hidden layers in a multilayer neural network model, the convolution layers associated with a fully connected layer have the properties of parameters reduction, sparsity of connection and capturing translation invariance, thus largely reducing the computing costs for parameters estimation and overfitting while providing enough accuracy. This study proposed a novel CNN deep learning model OrbitNet, specifically designed for orbits image-based fault identification of turbomachines. The model adeptly combines max-pooling and strides to increase the image compression efficiency while improving its learning skills. Figure 1 shows the architecture of the proposed OrbitNet model. The model consists of seven 2-D convolution layers with the filter size increasing from $F=16$ to $F=128$ at the power of 2. Of them,
the first, fifth and last layers perform the convolution and down-sampling (strides=2) at both x- and y- axes simultaneously, thus reducing 4 times of computation to be needed if a pooling layer is used conventionally at next layer. The kernel sizes of 5×5 in the first convolution layer and 3×3 in other layers are utilized to learn the larger spatial filters while reducing the volume size. The main features of this model are described below.

In the convolution neural network model, the convolutional layer is used to extract features from the images through filters and then a pooling layer is employed to down-sample the features. Let \( \{X_i, y_i\} \ (i = 1, 2, \ldots, N) \) denote a set of image training data, where \( X_i \) is the i-th feature map with the size of \( H \times W \times D \) in which \( H \times W \) is the height and width of the image and \( D \) is the number of channels for each image, for example, \( D = 1 \) for greyscale and \( D = 3 \) for colorful RGB (Red, Green and Blue) image. The variable \( y_i \in \{1, 2, \ldots, C\} \) denotes the corresponding damage type in this study in which \( C \) is the number of types. The objective of training the CNN model is to learn the filter weights and biases that minimize the classification error in the output layer. Suppose one convolution layer contains \( k \) filters, the \( m \)-th output of the convolutional layer can be represented by:

\[
X_i^{(m)} = f(W^{(m)} * X_i^{(m-1)} + b^{(m)}), \quad m = 1, 2, \ldots, M
\]

where the symbol * refers to the convolution operator, \( X_i^{(m)} \) is the i-th feature map at layer \( m \) for sample \( X_i \), the parameters \( W^{(m)} \) and \( b^{(m)} \) are the filter weights and biases of the \( m \)-th layer to be learnt, \( M \) is the number of layers in the CNN model, and \( f(.) \) is an element-wise nonlinear activation function. The ReLU (Krizhevsky et al. 2012) function is used in all convolution layers as it has demonstrated to provide powerful ability of modeling nonlinearity of the problem. It removes negative values from the activation map by setting them to zero thus effectively enhancing computational efficiency while maintaining sufficient modeling accuracy. The ReLU function has also been proven to be robust for gradient vanishing in parameter training (Bengio et al., 1994).

Instead of a stack of fully connected layers, a global average pooling is used in the end of the network by averaging each feature map. This layer is used to extract the deeper and more abstract features by reshaping the feature maps into an n-dimension vector, defined as:

\[
\bar{Y}_c = \sum_{i=1}^{i=C} g(W_cX_e + b_e)
\]

where \( X_c, Y_c, W_c \) and \( b_c \) are the input, output, weights, and bias of the fully connected layers, respectively, \( L_c \) is the number of the fully connected layers, and \( g(.) \) is the activation function for this layer, which is the softmax function used in this model to classify the images.

The multi-class categorical cross-entropy is employed as a loss function in this study to obtain the optimal model parameters. The L2 regularization is added to force the weights to be small non-zero values. The model is trained by minimizing the loss function, expressed as

\[
\min_{\theta} L(\theta) = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{C} y_{ij} \log(g_j(\theta, X_i)) + \lambda \sum_{l=1}^{L} \theta_l^2
\]

where \( \theta = \{W^{(1)}, \ldots, W^{(M)}, b^{(1)}, \ldots, b^{(M)}\} \) is the total parameters of the OrbitNet CNN model to be learnt, \( y_{ij} \) is the j-th actual element in \( Y \) for sample \( X_i \), \( g_j(\theta, X_i) \) is the j-th activation softmax function for sample \( X_i \), and \( \lambda \) is the regularization parameter to control the amount of punishing the weights. Note that the second item in Eq. (3) is the regularization of L2.

3. IMPLEMENTATION PROCEDURE

The procedure of implementing the proposed deep learning methodology for automatic detection of faults in different types of turbomachines mainly consists of nine steps as explained below.

1) Data extraction. In order to establish and test the generalized deep learning model, data from different types of turbomachines (e.g., gas turbines, steam turbines, and centrifugal compressors in this study) representing the health and damage status are extracted from various customers. The damage types can be also different such as misalignment, unbalance, rubbing, and oil whirl. The vibration history data is usually in the form of wave indicating the periodical rotating of the turbomachine.

2) Data processing. The possible outliers are removed from the raw data of each variable.

3) Data cleaning. Vibration signals are often contaminated with other undesirable noises which must be removed in order to extract the orbits for subsequent fault identification. Bayesian wavelet thresholding approach is employed in this study to clean the raw data. The denoising effectiveness is evaluated by using graphical comparison and signal-to-noise ratio (SNR).

4) Generating shaft orbits. The image of shaft orbits is created from the representative time series data in the x- and y-direction at each time interval.

5) Image compression. The yielded image can be compressed into a small-size for computing efficiency.

6) Data grouping. Data is divided into three groups for modeling purpose: training, validation and testing, in the ratio of 60%, 30% and 10%.

7) CNN modeling. A novel OrbitNet model is created from the convolution neural network for fault identification of turbomachines. The model is tuned by the cross-validation approach with two data sets.

8) Model testing. The trained model is tested with the
additional image data.

9) Model implementation. The established model may be applied for real-time monitoring and predictive maintenance of the turbomachine.

4. DEMONSTRATION EXAMPLE

4.1. Data Description

The effectiveness and feasibility of the proposed methodology is illustrated with the real-world data collected from four different customers with four various types of faults commonly observed on bearing axes, including oil whirl, imbalance, blade crack, and friction. The four customers contain two centrifugal compressors and two steam turbines to represent the machine diversity. The variety of fault modes, turbomachine types, and industries yields a big challenge of developing a generic methodology for various faults identification from different types of turbomachines.

4.2. Data Analysis

Vibration data collected from four eddy current sensors on each turbomachine are used in this study to generate shaft orbits as image features to validate the proposed methodology. The example time history data for one event with misalignment fault is shown in Figure 2 for the machine axis at the x-direction (Fig. 2a) and y-direction (Fig. 2c). The machine operation at the normal and fault status are also marked in the time series plots. The box plots of time series at the normal and four events status are embedded in Figure 2 for the x-direction (Fig. 2b) and y-direction (Fig. 2d), which graphically depict the five group of data at both directions. The box-plot is a descriptive statistic method commonly used to graphically represent five quartiles of the data, namely maximum (upper quartile) and minimum (lower quartile) excluding any outliers, first quartile (25th percentile or Q1), median (50th percentile or Q2), and third quartile (75th percentile or Q3). The distance between the upper and lower quartiles is calculated as the interquartile range (IQR). As shown in the box plots (Fig. 2b and 2d) the points beyond the 1.5 times IQR are usually identified as outliers and marked as whiskers.

The turbomachine is operated continuously at a high rotating speed, for example, up to 11900 RPM of the rotor. Only a portion of data points is selected to demonstrate the effectiveness of the proposed methodology in Fig. 2. Three observations are drawn from Fig. 2. First, the time series at the x-direction (Fig. 2a) shows the similar trend as that at the y-direction (Fig. 2b). Second, the amplitude of the data at the y-direction (Fig. 2b) appears to be larger than those at the x-direction (Fig.2a). Third and the last, obvious sparks are observed only in the unhealth conditions.

Figure 2 Data for centrifugal compressor with misalignment fault: a) time series of x-direction data, and b) time series of y-direction data

4.3. Data Denoising

The Bayesian wavelet multiresolution analysis is employed to remove the undesirable noise from each raw data at each scenario. The most popular Daubechies function (Daubechies 1993) is employed in the discrete wavelet packet multiresolution analysis. This wavelet function is selected because it has the orthogonal and compact characteristics thus effectively representing the local details in the vibration signals. Bayesian hypothesis testing is then able to more accurately identify possible noise from the decomposed wavelet coefficients. Given a raw signal, three-layer (J=3) discrete wavelet packet transform with 8 vanishing moments in the Daubechies wavelet (DB8) are found to provide enough accuracy in the denoising which produces eight decomposed coefficient series representing the details and approximations of a raw signal. The eight cleansed coefficients by Bayesian hypothesis testing are then reconstructed to form the denoised signal for shaft orbit generation.

The raw and denoised data can be visually compared in both time and frequency domains. As an example, Figure 3 shows the comparison of raw and denoised vibration data in the x-direction for one case without damage. Figure 3a shows only the raw time series as it can’t distinct the denoised graph from the raw one visually. However, the normalized histogram plots in Fig. 3b for the raw (filled) and denoised (solid line) time series demonstrate the obvious difference at two vibration ranges, [-5µm, -1.5µm] and [1.5µm, 5µm]. The significant difference is also observed on the Welch’s PSDs between the denoised data (Fig. 3d) and the raw data (Fig. 3c). The difference between denoised and raw signal
illustrate that the proposed Bayesian wavelet denoising approach provides an effective tool in removing the possible noise from the sensed data.

Figure 3 Comparison of raw and denoised data in the x-direction vibration for one case without damage: a) time series plot of raw data, b) histograms of raw data (filled) and denoised data (bold line), c) Welch PSD of raw data, and d) Welch PSD of denoised data

4.4. Shaft Orbits Generation

A shaft orbit graph is created as an image using a pair of 1024 data points (32 cycles of rotating) collected by the sensors installed at x- and y-directions of the turbomachinery. About 150 shaft orbits are created from denoised data points for each event. After image processing with rotating and shifting, totaling 1750 orbit images are obtained for fault identification modeling. As an example, Figure 4 shows the shaft orbit plots created from a pair of vibration time series collected from a centrifugal compressor with misalignment fault. For a comparison purpose, the two shaft orbits (Fig. 4b and 4d) obtained from the raw (Fig. 4a) and denoised (Fig. 4c) time series at the x- and y-direction are given to show the effect of denoising on the shaft orbits. The images are divided into three groups: model training, validation and testing at the number of 1050, 525 and 175 images, respectively.

4.5. Model Establishment

The obtained shaft orbit images with five types of status, normal, imbalance, misalignment, oil whirl, and rubbing, are used to establish the OrbitNet model. Table 1 shows the model structure (Fig. 1) and parameters to be estimated in this example. The total 180901 parameters need to be estimated in this model.

Cross-validation approach is employed in this example to train the CNN model. The training image data set is used to tune the model parameters while the validation image data set to monitor the model accuracy and prediction error. The model is selected when its validation accuracy matches with the training accuracy.

Figure 4 Shaft orbit plots of vibration data in a centrifugal compressor with misalignment fault: a) denoised vibration data at the normal status, b) shaft orbits from the denoised data at the normal status, c) denoised vibration data at the damage status, and d) shaft orbits from the denoised data at damage status.

Table 1 OrbitNet CNN model structure

| Model structure | Data structure | No of parameters |
|-----------------|----------------|-----------------|
| conv2d_1        | (None, 128, 128, 32) | 832             |
| conv2d_2        | (None, 120, 120, 32) | 25632           |
| dropout_1       | (None, 120, 120, 32) | 0               |
| MaxPool2        | (None, 60, 60, 32) | 0               |
| Normal_1        | (None, 60, 60, 32) | 128             |
| conv2d_3        | (None, 56, 56, 64) | 51264           |
| conv2d_4        | (None, 56, 56, 64) | 102464          |
| dropout_2       | (None, 52, 52, 64) | 0               |
| Aver_pool2d     | (None, 64) | 0               |
| Normal_2        | (None, 64) | 256             |
| dropout_3       | (None, 64) | 0               |
| dense           | (None, 5) | 325             |

Figure 5 shows the trends of model training and validation accuracy as well as the errors. The model reaches to the stable status with the training accuracy of 98.0% at the epoch of 50.
4.6. Model Testing

The 168 new image data with different fault modes and normal status is used to test the trained model. The obtained confusion matrix is shown in Fig. 6. It is observed that 3 oil whirl and 4 imbalance events are misclassified as normal in this example, resulting in the misclassification ratio of 4.0% (96.0% accuracy). This testing has demonstrated that the proposed model provides a high fault identification for multiple fault modes.

5. CONCLUSION

This paper presents a generic new convolution neural network model, OrbitNet, to automatic fault identification of multiple fault modes for turbomachines, using the shaft orbits generated from the vibration data. A general implementation procedure is proposed to integrate the advanced signal process with the CNN model. Shaft orbit images are generated from the sensed vibration data to represent the characteristic of four commonly observed fault modes including imbalance, misalignment, oil whirl and rubbing, and used as inputs to establish the OrbitNet model.

Currently the training image sets are not sufficiently large to tune up the model in order to cover all possible scenarios in practical applications. In future research the OrbitNet model will be improved and tested with more real-world application cases to demonstrate its generality and robustness.

REFERENCES

Bengio, Y., Simard, P. & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult, IEEE Transactions on Neural Network, Vol. 5, pp. 157–166, doi: 10.1109/72.279181.

Carbajal-Hernández, J.J., Sánchez-Fernández, L.P., Ignacio Hernández-Bautista, I., Medel-Juárez, J.J. & Sánchez-Pérez, L.A. (2016). Classification of unbalance and misalignment in inductionmotors using orbital analysis and associative memories, Neurocomputing, Vol. 175, pp. 838–850, doi: 10.1016/j.neucom.2015.06.094.

Jardine, A.K.S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mechanical Systems and Signal Processing, Vol. 20, pp.1483–1510, doi: 10.1016/j.ymssp.2005.09.012.

Jeong, H., Park, S., Woo, S. & Lee, S. (2016). Rotating machinery diagnostics using deep learning on orbit plot images, Procedia Manufacturing, Vol. 5, pp.1107-1118. doi: 10.1016/j.promfg.2016.08.083.

Khodja, A.Y., Guersi, N., Saadi, M.N. & Boutassetq, N. (2020). Rolling element bearing fault diagnosis for rotating machinery using vibration spectrum imaging and convolutional neural networks, The International Journal of Advanced Manufacturing Technology, Vol.106, pp. 1737-1751, doi:10.1007/s00170-019-04726-7.
Krizhevsky, A., Sutskever, I. & Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks, Communications of the ACM, Vol.60, pp. 1097–1105. doi:10.1145/3065386.

Wu, B., Feng, S., Sun, G., Xu, L. & Ai C. (2018). Identification method of shaft orbit in rotating machines based on accurate Fourier height functions descriptors. Shock and Vibration, Vol.2018, pp.1-10.

**Biographies**

**Xiaomo Jiang** was born in Lengshuijiang, Hunan, China. He received the M.S. and Ph.D. degrees in structural engineering from National University of Singapore, Singapore in 2000 and the Ohio State University, USA in 2005 respectively. He is a Professor in State Key Lab of Structural Analysis for Industrial Equipment and the Engineering Mechanics Department at Dalian University of Technology, Liaoning, China. He was a Technical Leader in monitoring, diagnostics, & prognostics of gas turbines and steam turbines at General Electric company, Atlanta GA, from 2008 to 2017 and a Postdoctoral Research Associate at Vanderbilt University, Nashville TN, from 2005 to 2007. His main research interests include Energy IoT, digital twin, data-driven reliability, performance and predictive analytics for turbomachinery such as gas turbines, steam turbines, and compressors. Dr. Jiang has co-authored 1 book, 5 book chapters, and more than 80 research articles and filed 10 patents in the cross-disciplinary fields of computer science, engineering, artificial intelligence and applied statistics. He has been invited to serve as a peer reviewer, editorial board, or associate editor for more than 40 different international journals and conferences.

**Shuhua Yang** was born in Chifeng City, Inner Mongolia Autonomous Region, China. He received the B.S. and M.S. degrees in engineering mechanics from Xi’an Jiaotong University, Shaanxi, China in 1994 and Fudan University, Shanghai, China in 2004, respectively. His main research interests include the strength design, vibration analysis and fault diagnosis of fluid machinery. Since 2004, he has been the chamber director, deputy director and director of R&D Department at Shenyang Blower Group Co., Ltd., president of Research Institute and deputy chief engineer of the company successively. He has published over 40 research articles on various technical journals and proceedings, like Journal of Computational Mechanics, Journal of Mechanical Engineering, Journal of Vibration Engineering, Journal of Mechanical Strength, AIAA and ASME Turbo Expo. Mr. Yang is the vice chairman of technical committee on Fans of Standardization Administration of China and the vice chairman of mechanical design branch of Liaoning Institute of Mechanical Engineering, as well as the editorial board member and editor-in-chief of the Chinese journal of turbomachinery.

**Fuming Wang** was born in Langfang City, Hebei, China. He received the M.S. degree in Energy and Power Department at Dalian University of Technology in 2020. His main research interests include signal processing, machine learning and deep learning with applications for turbomachinery damage detection.

**Shengli Xu** was born in Zhenjiang City, Jiangsu, China. He received the B.S. and Ph.D. degrees in engineering mechanics in 2002 and 2010, respectively, from Dalian University of Technology, Liaoning, China. He is an Associate Professor in the Energy and Power Department at Dalian University of Technology. His main academic and industry interests include structural numerical simulation method, structural optimization method, turbomachinery multidisciplinary optimization, turbomachinery health monitoring and engineering analysis and optimization software development. Dr. Xu is an ISSMO member and ASME member. He has published over 30 papers.

**Xiaofang Wang** was born in Liaoning, China. She is a professor of Energy and Power Department at Dalian University of Technology, Liaoning China. Her main research interests include design, optimization, aerodynamics analysis of turbomachinery such as centrifugal compressors, gas turbines, steam turbine, and pumps. Prof. Wang has completed over 60 research projects funded by various agencies like the National Natural Science Foundation of China and National Key Research and Development Program of China. She has authored 2 books and over 100 research articles, and has filed 42 patents in the multidisciplinary fields of turbomachinery. She has served as an editorial board or associate editor for 3 different technical journals, and has been the committee member of various scientific organizations.