Robust Deep Learning-Based Diagnosis of Mixed Faults in Rotating Machinery

Siyuan Chen, Yuquan Meng, Haichuan Tang, Yin Tian, Niao He, and Chenhui Shao

***Abstract*—Fault diagnosis for rolling elements in rotating machinery persistently receives high research interest due to the said machinery's prevalence in a broad range of applications. State-of-the-art methods in such setups focus on effective identification of faults that usually involve a single component while rejecting noise from limited sources. This article studies the data-based diagnosis of mixed faults coming from multiple components with an emphasis on model robustness against a wide spectrum of external perturbation. A dataset is collected on a rotor and bearing system by varying the levels and types of faults in both the rotor and bearing, which results in 48 machine health conditions. A duplet classifier is developed by combining two 1-D convolutional neural networks (CNNs) that are responsible for the diagnosis of the rotor and bearing faults, respectively. Experimental results show that the proposed classifier can reliably identify the onset and nature of mixed faults. In addition, one-vs-all classifiers are built using the features generated by the developed 1-D CNNs as predictors to recognize previously unlearned fault types. The effectiveness of such classifiers is demonstrated using data collected from four new fault types. Finally, the robustness and ability to reject external perturbation of the duplet classification model are analyzed using kernel density estimation. The code for the proposed classifiers is available at https://github.com/siyuanc2/machine-fault-diag.**

***Index Terms*—Condition monitoring, convolutional neural network (CNN), deep learning, fault diagnosis, Kernel density estimation (KDE), model robustness, rotating machinery, rotor and bearing systems.**

I. Introduction

ROTATING machinery is an integral component in a broad range of applications, including motors, gearboxes, and generators. Ensuring optimal operating conditions for these components is essential for safety, efficiency, economy, and comfort considerations. Although sophisticated maintenance strategies exist for maintaining safe operation of many aforesaid applications, it is desirable to incorporate real-time, predictive monitoring methods to detect the onset, duration, possible cause of mechanical faults and perform countermeasures as fast as possible.

Detection of rotating machinery faults is most commonly performed by analyzing vibration signals [1], [2], as most mechanical faults, including eccentric rotor, bearing fault, and misalignment, result in an imbalance in the rotating element that materializes as shock response in the measured vibration signals. In many cases, the informative fault signal component may well be buried in the noise signal [3] and additional measures are required to extract useful information from the noisy source signal.

Popular signal processing techniques, ranging from time-domain statistical figures to frequency-domain methods like fast Fourier transform (FFT) [4], discrete wavelet transform [5], and higher-order spectrum [6], have been applied to the monitoring of mechanical faults in various rolling machinery with satisfactory performance. In many cases, the abovementioned methods are developed for a specific physical setup and do not give further insight into the root cause of the problem. During the last decade, there has been a rising interest in model-based fault detection techniques [7]. Maki and Loparo [8] applied feed-forward neural networks to a plant model and demonstrated the model's ability to operate on generic industrial process records. Frank and Köppen-Seliger [9] evaluated the features generated by neural networks using fuzzy logic. Bachschmid *et al.* [10] studied model-based diagnosis of multiple present faults by the means of least-square fitting in the frequency domain. Wang *et al.* [11] developed a novel variational mode decomposition method based on particle swarm optimization.

More recent works improve the performance of machinery fault diagnosis systems in the form of expert systems engineered with or without domain knowledge. Zhou *et al.* [12] examined a rule-based expert system for rotor machinery fault diagnosis that operates on a fault pattern library constructed with various hand-engineered features including common rotor fault patterns, root mean square values, peak-to-peak values, and FFT responses. Recent implementations of expert systems make use of Bayesian networks [13], support vector machines [14], random forest classifiers [15], and deep belief networks that operate on images converted from vibration signals [16] requiring less

Manuscript received December 6, 2019; revised April 7, 2020 and June 14, 2020; accepted June 26, 2020. Date of publication July 8, 2020; date of current version October 14, 2020. This work was supported by CRRC Corporation Limited. Recommended by Technical Editor H. (AISM FS SE) Ding and Senior Editor H. (AISM FS SE) Ding. *(Siyuan Chen and Yuquan Meng contributed equally to this work). (Corresponding author: Chenhui Shao.)*

Siyuan Chen, Yuquan Meng, and Chenhui Shao are with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61820 USA (e-mail: siyuanc2@illinois.edu; yuquanm2@illinois.edu; chshao@illinois.edu).

Haichuan Tang and Yin Tian are with the CRRC Academy, Beijing 100161, China (e-mail: thc@crrc.tech; ty@crrc.tech).

Niao He is with the Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61820 USA (e-mail: niaohe@illinois.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMECH.2020.3007441

domain knowledge specific to the physical setup. Sun *et al.* [17] introduced kernel sparse locally linear embedding that reduces data dimensionality by mapping the data to a low-dimensional manifold.

With the vast success of neural networks in tasks including image classification and semantic segmentation, many recent works focus on applying state-of-the-art neural network architectures on signal fault diagnosis [18]. Zhao *et al.* [19] applied a number of advanced neural network architectures including firefly artificial neural network, particle swarm optimization neural network, and genetic artificial neural network, and achieved diagnosis accuracy of more than 98% on rotor fault types. Additional works explored the effectiveness of generalized demodulation transform in adapting to varying speed of the rotating element [20]. Wang *et al.* [21] tackled the same problem by introducing multiscale analysis windows from sensor signals. Most recent works that employ neural networks for fault detections focus on the general applicability of data-driven models over a wide range of data or usage scenarios [22] and the ability of performing transfer learning based on existing models [23]. Others explore ensemble learning [24] and stacked sparse denoising autoencoders [25] for bearing fault detection, obtaining salient performance in their respective tasks.

At least two research gaps exist in the fault diagnosis of rotating machinery. First, the aforementioned methods only considered faults of one component, whereas in real applications it is possible that mixed faults in multiple components factor in the measured signals. Moreover, whereas many works introduced measures to passively reject noise and nonrelated factors such as varying rotational speeds [20], [21], none of the reviewed works actively evaluated the robustness of proposed methods by testing them against strategically introduced interference or perturbation signals. A robust model must reject a broader range of noise originated from external sources like signal acquisition process and varying operating conditions of the machinery, and it is not to be assumed that no noise or perturbation is deliberately introduced to disable the system. Such is especially true for data-based models that operate more or less in a black-box fashion, as it has been proven that most deep learning models are vulnerable to external attacks [26]. It is also possible to strategically engineer inputs to mislead the neural networks [26]. On this front, it is necessary to evaluate the effect of perturbation, at least those of a normally distributed nature, on the safety-critical fault diagnosis models.

This article seeks to explore these areas by introducing a 1-D CNN architecture for fault diagnosis under multiple mixed component faults on a rotor/bearing fault simulator setup. A duplet classification model is developed by combining two 1-D CNN models, each independently responsible for identifying faults from the bearing and the rotor. Experimental results suggest that the duplet classifier can reliably identify each known component fault cause with a combined accuracy of 95.93%. Using the features generated by CNN models as predictors, linear models are developed to distinguish unknown fault types from known ones. The effectiveness of these models is demonstrated experimentally. Additionally, we employ kernel density estimation to investigate the probability distribution of the vibration signals

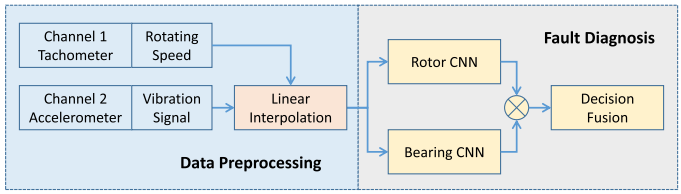


Fig. 1. Overall structure of the proposed rotor and bearing system fault identification model. Inference is performed by evaluating preprocessed data with the two independent fault diagnosis models before the decision is drawn.

and examine the robustness of the proposed model, concluding that the rotor fault detection model is more robust against Gaussian noise than its rotor fault detecting counterpart.

The rest of this article is organized as follows. In Section II, the proposed 1-D CNN architecture is introduced along with details on the collection and processing of the experimental data. In Section III, the model performance on the dataset with mixed faults is examined. In Section IV, we propose an approach to recognize unknown mechanical faults and recognize robustness of the proposed classifiers against external purturbation in the form of Gaussian noise. Finally, Section V concludes this article.

II. Methodology

Though CNN allows for the learning of fault patterns without extensive domain knowledge, examining the nature of the fault is beneficial for gaining insight and further improving the model performance. In this section, we first present the proposed architecture of the 1-D CNN that is used to analyze the vibration signals. The data acquisition process, experimental setup, and data preprocessing are then detailed. Fig. 1 illustrates the proposed processing flow.

A. 1-D CNN

As a nonsupervised feature selection method, deep learning has been widely applied in the field of mechanical fault diagnosis [27], [28]. In many cases, due to limited data availability and the characteristics of input signals, it is not feasible to apply deep networks and models frequently settle at local optima [19]. Inspired by demonstrated success of CNN in many image-related tasks, some recent efforts sought to operate CNN on vibration signals and FFT responses visualized as images [29]. However, this formulation might prevent the CNN to properly extract desired features from the input. For instance, the conversion of time-domain or frequency-domain profiles into 2-D images makes it necessary to divide a continuous signal into fixed-length segments so that the information can be represented as 2-D arrays. The 2-D filters then consider values across rows that are not adjacent to each other in time or frequency domain. As a result, the 2-D CNN might learn periodic features that are not present in the input signal, but are solely a result of reshaping signal to fit in input dimensions.

Instead of applying 2-D filters to 1-D signal of interest, we employed 1-D convolution layers for the proposed neural network architecture. As is suggested by the name, this operation

TABLE I
1-D CNN STRUCTURE FOR THE BEARING FAULT DIAGNOSIS

Layer name	Output shape
Input	$n \times 8 \times 512$
1-D Convolution	$n \times 64 \times 508$
1-D Convolution	$n \times 128 \times 506$
Flatten	$n \times 64768$
Dense	$n \times 128$
Dense	$n \times 64$
Output	$n \times 8$

TABLE II
1-D CNN STRUCTURE FOR THE ROTOR FAULT DIAGNOSIS

Layer Name	Output shape
Input	$n \times 8 \times 512$
1D Convolution	$n \times 64 \times 508$
1D Convolution	$n \times 128 \times 506$
1D Convolution	$n \times 256 \times 504$
Flatten	$n \times 129024$
Dense	$n \times 128$
Dense	$n \times 64$
Output	$n \times 6$

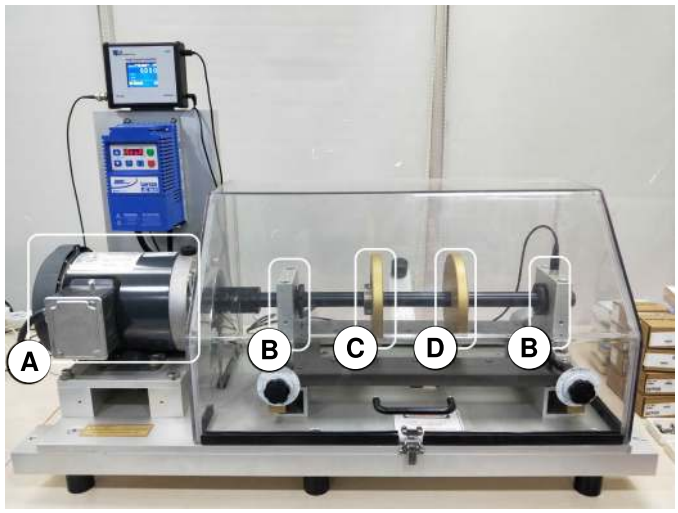


Fig. 2. MFS-LT MFS [32] used for data collection. Motor (A), bearings (B), rotor A (C), and rotor B (D) are labeled.

utilizes 1-D convolution kernels. Given an $n \times 1$ vector input x and an $m \times 1$ convolution kernel w, the ith element of the convolution output $(x * w)$ can be expressed as follows:

$$(x * w)_i = \sum_{j=1}^{m} w_j x_{(i-j+m/2)}. \tag{1}$$

Readers are referred to classic references for details on more commonly used CNN structures, including but are not limited to, ReLu activation layers [30] and batch normalization [31].

We hereby propose the use of 1-D CNN with filter dimensions 3×1 and 5×1. In this way, the 1-D convolution kernels will only operate on contiguous information in time or frequency domain. Considering that traces of fault in different components of the system may not necessarily be related to each other, we employ separate models each responsible to identify possible faults in one component. In this case, one model is dedicated to the detection of rotor faults and the other to bearing faults.

Based on the assumption that bearing fault state can be independent from that of the rotor's, joint fault diagnosis can be performed by simply bundling the two aforementioned models in parallel as a *duplet classifier*. Specifically, an m-class bearing state classifier and an n-class classifier can be combined to output in $m \times n$ result space with a look-up table. It is demonstrated below that in this layout, relatively simple two-layer or three-layer 1-D CNN architectures are capable of identifying the presence of various mechanical faults from the input signal. As a comparison, we have also trained a separate $m \times n$-class classifier. Details on the model architecture are presented in Tables I and II.

B. Data

A machine fault dataset was used to verify the effectiveness of the proposed CNN architectures and the duplet classifier. The dataset was collected on a specialized machinery fault simulator (MFS) manufactured by Spectra Quest, Inc. [32], as shown in Fig 2. The particular model used in this article, named MFS-LT, is equipped with two rotor disks and a shaft powered by an electric motor. The setup can perform controlled simulations on various mechanical fault conditions associated with rotor and bearing.

Unlike a few existing works such as [13] and [19] that study the identification of only rotor faults, we designed various combinations of both rotor and bearing faults, i.e., mixed faults, which resulted in a much more challenging diagnosis task. Details of the faults considered are provided below.

Rotor Faults: The MFS has two rotors where unbalanced rotor conditions can be simulated by attaching additional standard weights to the rotor plates. For simplicity, the left and right rotors are referred to as rotor A and rotor B, respectively. We simulated five fault conditions indicated by the location and number of additional weights installed to the two rotor elements, described as follows:

1) Normal condition: No additional weights were installed on either of the rotors. By including additional normal states to the collected data, it is desired to confirm that apart from classifying the various error states, the diagnosis model can also reliably distinguish any fault states from the normal state.
2) One weight on rotor A, one weight on rotor B, aligned (A1B1A): In this configuration, each rotor was fitted with an weight and the two weights were aligned to the same angular position.
3) One weight on rotor A, one weight on rotor B, opposite (A1B1O): Each rotor was fitted with a weight in opposite angular positions relative to the shaft.
4) Two weights on rotor A, adjacent (A2A): Two weights were installed on rotor A as close as possible. It is expected that this configuration will yield a similar vibration profile with configuration "A1B1A."
5) Two weights on rotor A, opposite (A2O): Two weights were installed on rotor A in positions symmetric about the shaft. It is expected that this configuration will yield a similar vibration profile with configuration "A1B1O"

and should generally result in lower magnitudes than the "A1B1A" and "A2A" configurations.

6) Three weights on rotor A, adjacent (A3A): Three weights were installed on rotor A as close to each other as possible. It is expected that this configuration will yield a similar vibration profile with configurations "A1B1A" and "A2A" but with higher magnitude.

Four additional fault states for the rotor are also simulated but are excluded from the following training process.

1) One weight on rotor A, large eccentricity (A1-L): Only one weight was installed on rotor A, with its location as far from the shaft as possible to create maximum unbalanced condition.
2) One weight on rotor A, small eccentricity (A1-S): Similar to "A1-L," only that the weight was installed close to the shaft for a less unbalanced condition.
3) One weight on rotor B, large eccentricity (B1-L).
4) One weight on rotor B, small eccentricity (B1-S).

These additional fault states are excluded from training data and are used exclusively for evaluating model robustness against unknown fault types, as later explained in Section IV.

Bearing Faults: The two rotors on the MFS are installed on a shaft by rigid connection, which is then mounted on the test bed with two bearings. The MFS provides a few bearings with mechanical faults, which help to simulate various bearing fault conditions along with rotor fault conditions. In this work, seven bearing conditions along with a normal bearing condition were simulated. Details of the simulated bearing conditions are provided as follows:

1) Normal condition: No mechanical fault associated with bearings is present. This state is to be used with various fault states of the rotor to verify the model's ability to identify the presence of multiple faulty components in the input signal.
2) Bearing ball wear: Four levels of wearing in the bearing's rolling element: slight, light, moderate, and severe, have been simulated on the test bed. The wearing levels are referred to as level 1–4 correspondingly.
3) Inner/outer race: Damage to the inner and outer race of the bearing.
4) Mixed: Combined bearing ball wear and damage to inner and outer race.

The abovementioned eight classes of bearing and six classes of rotor mechanical faults yield a combined 48-class dataset, with the detailed labeling scheme presented in Table III.

The dataset consists of two channels collected from sensors installed on the MFS test bed. Channel 1 is pulse width modulation (PWM) signal collected from a tachometer from which the rotating speed of the shaft can be determined. Channel 2 is collected from an accelerometer installed on the far end of the shaft to depict lateral vibrations.

All data were collected with a sampling rate of 1.28 kHz. A total of 82 h worth of data was collected, resulting in a dataset encompassing 48 health conditions, including one normal state and 47 mechanical fault combinations between rotor and bearing faults. The dataset were collected in equal portions of five speed setting: 12 Hz, 14 Hz, 16 Hz, 18 Hz, and 20 Hz to account for variations resulted from different rotational speeds of the shaft.

TABLE III
LABELING SCHEME FOR THE 48-CLASS DATASET

Class	Rotor	0	1	2	3	4	5
Bearing	Condition	Normal	A1B1A	A1B1O	A2A	A2O	A3A
0	Normal	0	8	16	24	32	40
1	Ball wear 1	1	9	17	25	33	41
2	Ball wear 2	2	10	18	26	34	42
3	Ball wear 3	3	11	19	27	35	43
4	Ball wear 4	4	12	20	28	36	44
5	Inner race	5	13	21	29	37	45
6	Outer race	6	14	22	30	38	46
7	Mixed	7	15	23	31	39	47

Although the dataset was collected in five levels of rotating speed, the data were processed into a single rotating speed, 10 Hz, by linear interpolation. This conversion is meant to adapt the model to different operating conditions and has been empirically proved to improve the model's performance in Section III.

C. Unknown Fault Classification

In real life, the fault types of a rotor and bearing system can easily exceed the aforementioned 47 types. To identify faults unknown to the dataset's 48 classes, an unknown fault classifier based on the duplet 1-D CNNs will be developed in Section IV. In essence, if all the unknown faults can be grouped into a single class known as the unknown fault, the problem can be formulated as developing a classifier for 49 fault condition classes. However, it is impossible to naïvely train a neural network for the newly added class, as unknown fault conditions will not appear in a dataset. A more realistic way is to develop a classifier that determines whether the input data corresponds to a fault condition known to us. To achieve this, 48 one-vs-all classifiers are proposed in Section IV. Such one-vs-all discriminant can be formulated as follows:

$$f_i(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \mathbf{X}_i \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $\mathbf{x}$ is a signal to be classified; $\mathbf{X}_i$ is a collection of possible signals from a known machine condition i (or class i in general terms).

Then unknown fault classifier is shown as follows:

$$F(\mathbf{x}) = \vee_{i=1}^{48} f_i(\mathbf{x}). \tag{3}$$

Such a classifier outputs 1 if the input signal belongs to a known class and 0 otherwise.

III. EXPERIMENTAL RESULTS

In this section, we present the performance of the proposed 1-D CNN architecture on the custom dataset. To demonstrate the effectiveness of data interpolation and separated training of bearing and rotor fault classifiers, we trained both separated and 48-class models and compare their performance on the test set. The best models for bearing and rotor fault diagnosis were then assembled to form a duplet classifier, while decision fusion is made by the look-up table provided in Table III. Additionally,

TABLE IV
BEST MODEL PERFORMANCE ON EACH VARIETY OF THE DATASET

Dataset type	Architecture	Training accuracy	Testing accuracy
(a) 6-class rotor with interpolation	3-layer 1D CNN	99.0%	**97.0%**
(b) 6-class rotor without interpolation	3-layer 1D CNN	99.63%	68.61%
(c) 8-class bearing with interpolation	2-layer 1D CNN	99.0%	**98.90%**
(d) 8-class bearing without interpolation	2-layer 1D CNN	99.99%	96.88%
(e) Duplet classifier	(a) and (c) in parallel	-	95.93%
(f) 48-class with interpolation	2-layer 1D CNN	99.38%	83.72%
(g) 48-class without interpolation	2-layer 1D CNN	99.38%	49.72%

TABLE V
CONFUSION MATRIX OF THE BEARING CNN MODEL

		Predicted label						
True label	1839	0	0	0	0	0	0	0
	0	1745	10	5	0	0	1	0
	0	0	1778	0	0	0	0	0
	0	10	31	1750	35	0	1	0
	21	1	0	81	1638	0	0	0
	0	0	0	0	0	1832	0	0
	1	2	0	0	0	0	1860	0
	0	0	0	0	0	0	0	1863

TABLE VI
CONFUSION MATRIX OF THE ROTOR CNN MODEL

	Predicted label					
True label	2350	1	15	0	30	0
	3	2281	1	47	2	11
	16	2	2394	4	51	0
	1	57	1	2336	1	3
	34	1	34	0	2368	1
	0	6	0	3	0	2452

the performance of the models on both interpolated and noninterpolated data are compared. The results are presented in Table IV. There models were trained using an 80-10-10 train-validation-test split.

Overall, one can conclude that the 1-D CNN architecture yields desirable performance on the dataset as it can classify the rotor and bearing fault states with relatively simple two-layer and three-layer structures. With carefully chosen hyperparameters including dropout rate, learning rate decay, and proper weight initialization, the rotor and bearing fault condition models achieve testing accuracies of up to 97.0% and 98.90%, respectively. Combining these models, and the bearing-rotor duplet classification model yields 95.93% accuracy on the test dataset, which consistently outperforms the 48-class model trained for comparison. One can also conclude that converting the rotating machinery data into uniform operating speed is a beneficial measure, as interpolating accelerometer data according to the rotational speed results in consistent performance improvements across all model types shown in Table IV.

Failure Cases Analysis: Though both the bearing and rotor CNN models in the duplet achieved high accuracy on the test set, fail cases are highly concentrated in certain fault types and prompts future improvement. Here, we examine the frequent fail cases faced by the duplet model performance of the duplet classifier on the test set is shown by confusion matrices presented in Tables V and VI.

TABLE VII
FAILURE CASES THAT OCCURRED FOR MORE THAN TEN TIMES IN ROTOR CONDITION TEST DATASET

True		Predicted		Occurrence
Label	State	Label	State	
0	Normal	4	A2O	30
1	A1B1A	3	A2A	47
1	A1B1A	5	A3A	11
2	A1B1O	0	Normal	16
2	A1B1O	4	A2O	51
3	A2A	1	A1B1A	57
4	A2O	0	Normal	34
4	A2O	2	A1B1O	34

The accuracy of the bearing fault classifier reached 98.9%, almost all the test data were classified correctly. For the rotor fault classifier, however, the errors are particularly concentrated in a certain number of positions. Summarizing all errors that occur for more than ten times yields Table VII .

Further inspection of Table VII suggests that the rotor CNN model is most prune to the following two kinds of errors.

1) Confusion between states "Normal," "A2O," and "A1B1O;"
2) Confusion between states "A1B1A," "A2A," and "A3A."

We believe that the performance of the unbalanced rotor state classifier may have been impacted by insufficient information from the signal. Namely, the accelerometer was installed at the end of the shaft and recorded only lateral acceleration only. In this case, the vibration signal for "Normal," "A2O," and "A1B1O" states may appear analogous, because in all three cases, the additional weights, if any, were installed in a symmetric manner across from the shaft, whereas in the cases of "A1B1A," "A2A," and "A3A" all the additional weights were arranged in a highly unbalanced manner. The performance of the rotor CNN classifier could be improved if inputs from additional sensors strategically placed throughout the test bed were available.

IV. DISCUSSION

It has been demonstrated in former sections that the proposed duplet classifier achieves salient performance for the task of rotor and bearing fault diagnosis. However, the proposed model is trained on a 48-class dataset, which effectively limits it to operating on these known fault types, despite that it is possible to encounter new fault types that were not presented in the training dataset in real-world applications. Moreover, it is not verified that the model will survive deliberately introduced external

noise and perturbation. This section focuses on these topics and discusses the possibility of detecting unknown faults and rejection to normally distributed noise.

A. Robustness Against Unknown Fault Types

A model that can identify previously unseen faults is highly desirable. Here, we integrate the 1-D CNNs with principal component analysis (PCA) in an effort to recognize unknown fault types.

1) Problem Formulation and Proposed Method: In order to separate known and unknown fault types, we set out to train one-vs-all classifiers f_i for each known fault types i for $i = 0, \ldots, 47$, as shown by (2). The following procedure is applied to train a one-vs-all classifier for class i.

1) Train the duplet classifier to classify all known fault types.
2) Feed data $\mathbf{x}$, which belongs to class i, to the neural network, and obtain the output from intermediate layers as an embedded vector $\mathbf{y}_b = G_b(\mathbf{x})$ where G_b is effectively the output of the second last convolutional layer of the bearing fault CNN model. $\mathbf{y}_b$ has a length of 64 which matches with the output size of the corresponding layer.
3) Perform PCA transformation and acquire the threshold $h_{b,i}^{(j)}$, which defines a class-distinguishing hyperplane, from PC scores with nearly zero PC coefficients (here, they are the 5th PC to the 64th PC), i.e.,

$$|P_{b,i}(\mathbf{y}_b)|^{(j)} \leq h_{b,i}^{(j)}, j = 5, \ldots, 64 \tag{4}$$

where $P_{b,i}$ is the PCA transformation trained for class i, $h_{b,i}^{(j)}$ is the threshold such that 97% of jth PC score in class i lies within a distance of $h_{b,i}^{(j)}$.

4) Repeat (2)–(3) for rotor CNN and acquire $P_{r,i}$ and $h_{r,i}^{(j)}$.

In this fashion, the trained network is considered as a means of feature extraction function g that nonlinearly transforms the data from data space to a refined feature space. Given the fact that the data in same class should have similar governing hyperparameters, the PCA model effectively reveals those hyperparameters and rank them in order. PC with a smaller coefficient means the feature has less variance, hence a better approximation of hyperparameters. To account for input noise and other possible perturbations, the threshold $\vec{h_i}$ has been introduced, shown in Fig. 3. In this article, f_i is taken as

$$f_i(\mathbf{x}) = \mathbf{1}_{|(P_{l,i} \circ G_l)(\mathbf{x})|^{(j)} \leq h_{l,i}^{(j)}, j=5,\ldots,64; l=b,r} \tag{5}$$

where g gives the output of the second last convolutional layer of the bearing/rotor fault CNN models. Due to the nature of convolution kernels, a CNN is intrinsically invariant under input translation. Therefore, f_i is guaranteed to be invariant under translation of input data in time. Repeat the procedure for 48 classes, namely, $i = 0, \ldots, 47$, and we have acquired 48 one-vs-all classifiers that will distinguish each of the 48 classes with others. If all 48 one-vs-all classifiers output 0, the input is considered to belong to an unknown class.

2) Prediction Algorithm: The following algorithm is proposed to generate a label for an input data:

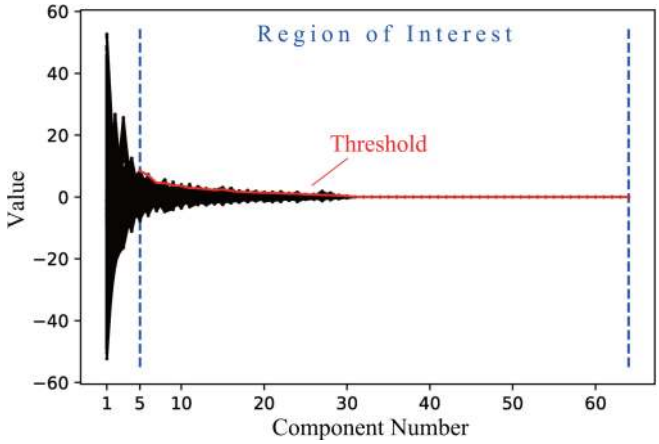


Fig. 3. PC scores of "Normal" state data by the PC transformation trained from the same state. Threshold is determined by the PCs with small variance.

TABLE VIII
CLASSIFICATION ACCURACY OF PROPOSED ALGORITHM

Input data	Prediction label	Accuracy
A1-L	Unknown fault	95%
A1-S	Unknown fault	92%
B1-L	Unknown fault	81%
B1-S	Unknown fault	85%
All unknown types	Unknown fault	88%
All known types	Corresponding type	93%

1) Feed the data into each one-vs-all classifier f_i and obtain their decisions (0 or 1).
2) Examine the output from all 48 classifiers:
   a) If at least one classifier outputs true (i.e., 1), then:
      i) If classifier labelled "Normal" outputs false (0), then feed the original data into the duplet CNN model. Output the label with highest probability excluding "Normal" class.
      ii) Otherwise, feed the original data into the duplet CNN model. Output the label with highest probability.
   b) If none of the classifiers outputs true, then identify the data as belonging to "Unknown Fault".

Note that above algorithm will output one of the 48 known state labels if the input data are recognized to belong to a known fault condition. The lower bound of predicting the correct known machine condition is given by

$$\mathbf{Pr}(\hat{c}_f = k|k) \geq 0.97^2 \times \mathbf{Pr}(\hat{c}_{\text{CNN}} = k|k) \tag{6}$$

where k is a class label of input data for one of the known machine conditions, $\hat{c}_{\text{CNN}}$ is the predicted label by the duplet classifier, $\hat{c}_f$ is the predicted label by the proposed algorithm and 0.97 is the value chosen to determine the thresholds.

3) Performance Evaluation: To evaluate the performance of the proposed hyperplane classifier, four data classes representing new fault types are introduced with names "A1-L" "A1-S," "B1-L," and "B1-S", with detailed definitions available in Section II.B. These four the 48 classes that were used to train the original 1-D CNN models. Table VIII shows the classification

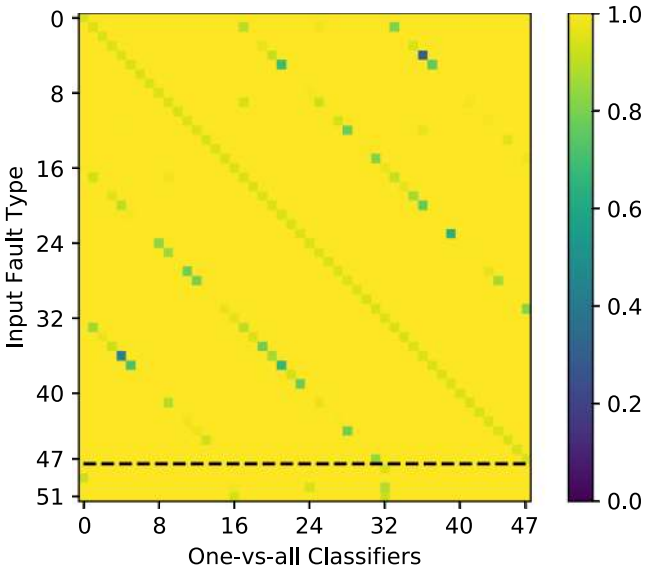


Fig. 4. Performance of one-vs-all classifiers on all fault types. *X*-axis represents 48 one-vs-all classifiers trained for each of the known classes. *Y*-axis represents the fault type to be judged, including 48 known machine conditions and 4 unknown fault types.

accuracy of proposed algorithm against unknown data classes and overall accuracy. The classifiers can identify unknown classes with an overall accuracy of 88%, which is highly desirable and indicative that the proposed models are robust against unknown fault types.

The classification accuracy of the one-vs-all classifiers and proposed prediction algorithm is further visualized in Fig. 4 and Table VIII. One can observe that while the collection of one-vs-all classifiers is robust against unknown fault types, it achieves slightly lower performance than the duplet 1-D CNN model on known machine conditions. There is a tradeoff between recognizing unknown classes and maintaining high performance for known machine conditions. However, equation (6) indicates that the accuracy of the proposed model on known faults is ensured by the duplet CNN model's accuracy and the predetermined threshold and will not be significantly deteriorated by the introduction of unknown fault detection.

In Fig. 4, diagonal entries of the confusion matrix represent the accuracy of each one-vs-all classifier against their respective known class. In this case, each entry has a value of 0.97^2, which is dictated by the threshold predetermined during the training process. At the same time, mis-classifications by individual one-vs-all classifiers (off-diagonal entries) for known classes do not have a decisive detrimental effect on the overall accuracy. In fact, a lower accuracy of one-vs-all classifiers will increase the overall accuracy, since for any known fault input, this will make it more likely to be fed into the highly accurate duplet CNN model that generates high-quality labels for known fault types. On the other hand, all one-vs-all classifiers demonstrate relatively high accuracy against unknown fault types, which leads to high detection rate for unknown classes.

One may notice that a few areas of improvement exist for the aforementioned collection of hyperplane classifiers. Firstly, since the backbone bearing and rotor fault detection CNNs provide embedding from data to feature space, the performance of these CNNs greatly influence the performance of one-vs-all classifiers. In this regard, improvements to the performance of the backbone CNN are always desirable. Secondly, because there is no guarantee that the learned embedding will convert data of same classes into a high dimensional representation, replacing the PCA method with a higher order, more expressive multivariate polynomial that is aided by random sample consensus for outlier rejection may bring significant improvement to model performance.

B. Robustness Against Gaussian Noise

In this section, the robustness of the proposed 1-D CNN classifiers against external noise and perturbation is assessed by deliberately adding different levels of Gaussian noise to the input signal, followed by analysis of the model robustness with kernel density estimation (KDE).

Table IX summarizes the performance of the 1-D CNN models after adding different levels of Gaussian noise to the raw signals. Here, we denote the mean and standard deviation of raw data as μ_0, σ_0, and the mean and standard deviation of the added noise as μ_ϵ, σ_ϵ, respectively. Under relatively small noise levels, e.g., $\sigma_\epsilon/\sigma_0 = 0.0001$, the bearing classifier retains a high accuracy (greater than 90%), while the accuracy of the rotor classifier declines dramatically to between 45% and 50%. Under $\sigma_\epsilon/\sigma_0 = 2$ or high noise levels, the bearing classifier maintains an accuracy of around 77%, while the rotor classifier achieves only approximately 20% accuracy.

To explain these observations, the input data are characterized using KDE. KDE is a statistical technique that calculates a probability density function (PDF) with derivatives of all orders, where a kernel function is employed to interpolate the distribution density [33]. A typical estimator is expressed as

$$\tilde{f} = \frac{1}{nb}\sum_{i=1}^{n} K\left(\frac{x - x_i}{b}\right) \tag{7}$$

where b is predetermined bandwidth, n is the total number of data points, x_i is a data point, and K is the kernel function, which can be uniform, Gaussian, Epanechnikov function, etc. A preliminary analysis reveals that the temporal correlation in the signals is weak. Therefore, we loosen the assumption and assume that the signals are independent and identically distributed random variables.

Fig. 5 shows the estimated PDFs of different fault types, where the normal distribution is obtained based on the mean and variance of the corresponding class. It can be concluded that the distributions of rotor fault data can be well approximated by normal distributions. Furthermore, it is revealed that rotor faults, including A1B1A and A1B1O, do not alter the distribution significantly, while bearing faults have a significant impact on the distribution. Specifically, a higher bearing wear level leads to increase in variance and deviation from the normal distribution. Bearing inner race fault data demonstrates a distribution significantly different from the normal distribution. The same

TABLE IX
CLASSIFICATION ACCURACY OF CNN CLASSIFIERS AFTER INTRODUCING GAUSSIAN NOISE TO RAW SIGNALS

$\left(\frac{\lvert\mu_0-\mu_\epsilon\rvert}{\mu_0}, \frac{\sigma_\epsilon}{\sigma_0}\right)$	(0.01, 0.0001)	(0.01, 0.0001)	(0.01, 0.0001)	(1, 2)	(1, 2)
Rotor classifier	46.41%	48.37%	45.75%	21.57%	24.83%
Bearing classifier	96.08%	92.81%	93.46%	77.78%	77.78%

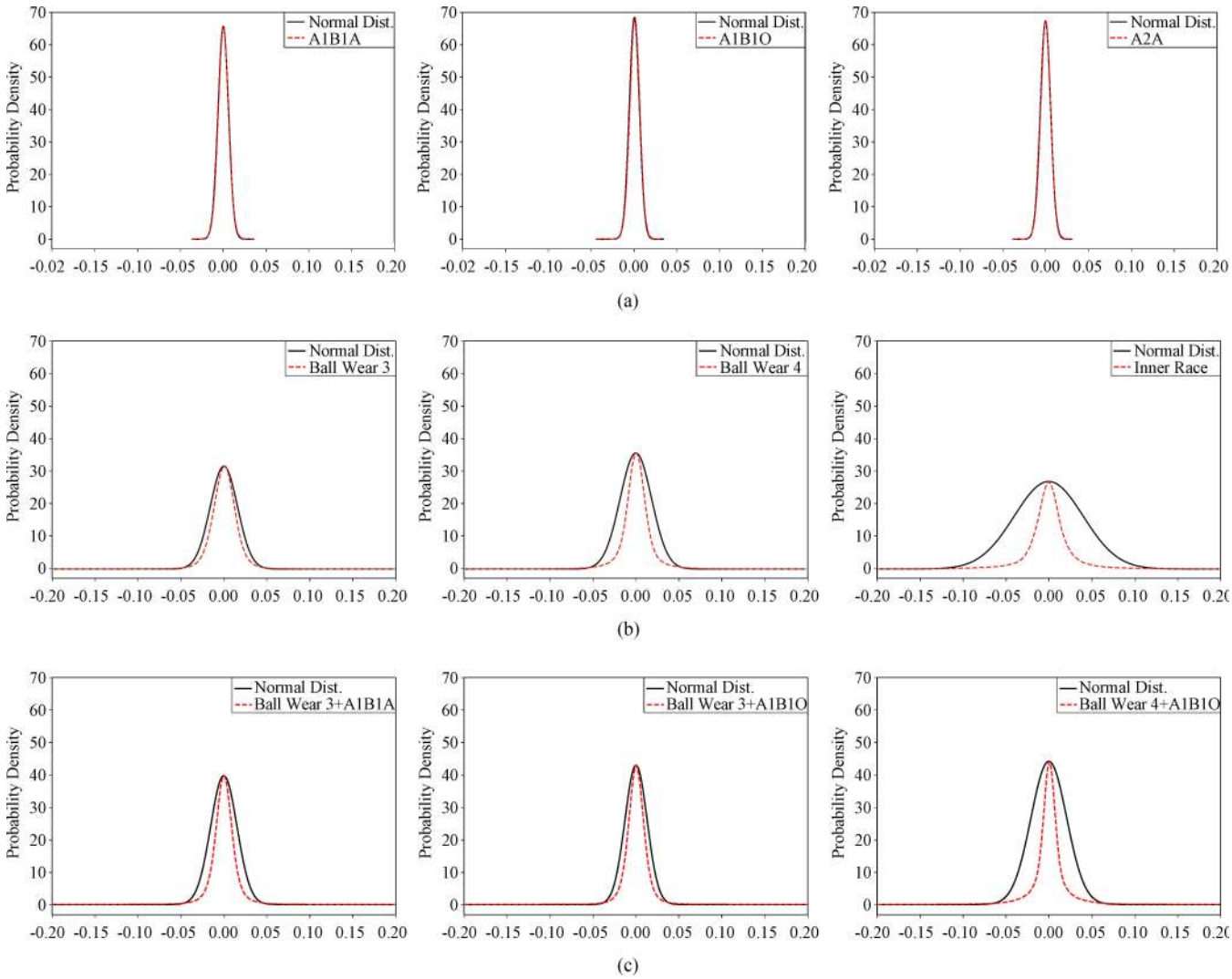


Fig. 5. Comparison of PDFs that are estimated using KDE and normal distributions. (a) Rotor faults. (b) Bearing faults. (c) Mixed faults.

phenomenon occurs in mixed bearing fault type, where the difference in distributions can be attributed to the bearing wear.

Since the bearing fault classes can be differentiated by ignoring the correlation between data points in different time, it is reasonable to conclude the bearing fault classes are easier to identify than rotor fault classes. By adding the Gaussian noise, the difference between bearing fault classes is still significant as seen in the corresponding PDF. Therefore, any classifier that successfully captures the characteristics of signals from bearing faults should be resilient to added Gaussian noise. However, for rotor fault types, the distributions of the signals are largely similar. To effectively identify each class, the classifier must learn features that are more specific to the dataset, namely, the correlation between data points must be accounted for. When high levels of Gaussian noise is introduced to the input signals, it may very likely alter the learned features, and consequently impose a negative impact on the classification performance. In conclusion, the rotor classifier is vulnerable to Gaussian noise.

V. CONCLUSION

A deep-learning-based approach was developed to diagnose mixed faults from multiple components in rotating machinery. A duplet classifier was established by assembling two 1-D CNN

models that are responsible for classifying rotor and bearing fault types, respectively. Experimental results show that the proposed classifier can reliably identify the presence of 48 machine health conditions that are combinations of eight bearing fault states and six rotor fault states. Based on the developed 1-D CNN models, one-vs-all classifiers were built to detect new, unlearned fault types, and their effectiveness was experimentally demonstrated. In addition, it was found that the bearing classifier is robust to Gaussian noise but the rotor classifier is very sensitive to Gaussian noise. A KDE-based analysis suggested that this is due to the nature of the signal patterns. These findings indicated that it is important to account for signal noises, which may be introduced either unintentionally from the environment or deliberately as a form of adversarial attacks, in the development of safety-critical fault diagnosis methods.

References

[1] Y. Zhang and R. Randall, "Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram," *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1509–1517, 2009.

[2] R. Li and Z. Sun, "A data-driven approach to detect mechanical faults in wind turbine gearbox," in *Proc. ASME 12th Int. Manuf. Sci. Eng. Conf. Collocated JSME/ASME 6th Int. Conf. Mater. Process*, 2017. [Online]. Available: https://doi.org/10.1115/MSEC2017-2736

[3] R. B. Randall and J. Antoni, "Rolling element bearing diagnosticsa tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, 2011.

[4] R. Walker, S. Perinpanayagam, and I. K. Jennions, "Rotordynamic faults: Recent advances in diagnosis and prognosis," *Int. J. Rotating Mach.*, vol. 2013, 2013. [Online]. Available: https://doi.org/10.1155/2013/856865

[5] I. S. Cade, P. S. Keogh, and M. N. Sahinkaya, "Fault identification in rotor/magnetic bearing systems using discrete time wavelet coefficients," *IEEE/ASME Trans. Mechatronics*, vol. 10, no. 6, pp. 648–657, Dec. 2005.

[6] B. Liang, S. Iwnicki, and Y. Zhao, "Application of power spectrum, cepstrum, higher order spectrum and neural network analyses for induction motor fault diagnosis," *Mech. Syst. Signal Process.*, vol. 39, no. 1–2, pp. 342–360, 2013.

[7] R. Isermann, "Model-based fault-detection and diagnosis–status and applications," *Annu. Rev. Control*, vol. 29, no. 1, pp. 71–85, 2005.

[8] Y. Maki and K. A. Loparo, "A neural-network approach to fault detection and diagnosis in industrial processes," *IEEE Trans. Control Syst. Technol.*, vol. 5, no. 6, pp. 529–541, Nov. 1997.

[9] P. M. Frank and B. Köppen-Seliger, "Fuzzy logic and neural network applications to fault diagnosis," *Int. J. Approx. Reasoning*, vol. 16, no. 1, pp. 67–88, 1997.

[10] N. Bachschmid, P. Pennacchi, and A. Vania, "Identification of multiple faults in rotor systems," *J. Sound Vib.*, vol. 254, no. 2, pp. 327–366, 2002.

[11] X.-B. Wang, Z.-X. Yang, and X.-A. Yan, "Novel particle swarm optimization-based variational mode decomposition method for the fault diagnosis of complex rotating machinery," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 1, pp. 68–79, Feb. 2018.

[12] J.-H. Zhou, L. Wee, and Z.-W. Zhong, "A knowledge base system for rotary equipment fault detection and diagnosis," in *Proc. IEEE 11th Int. Conf. Control Automat. Robot. Vision*, 2010, pp. 1335–1340.

[13] Z. Cai, S. Sun, S. Si, and W. Zhang, "Rotor fault diagnosis for machinery fault simulator under varied loads," in *Proc. IEEE Annu. Rel. Maintainability Symp.*, 2013, pp. 1–6.

[14] Z. Yin and J. Hou, "Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes," *Neurocomputing*, vol. 174, pp. 643–650, 2016.

[15] Q. Yao, J. Wang, L. Yang, H. Su, and G. Zhang, "A fault diagnosis method of engine rotor based on random forests," in *Proc. IEEE Int. Conf. Prognostics Health Manage.*, 2016, pp. 1–4.

[16] 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.

[17] C. Sun, P. Wang, R. Yan, R. X. Gao, and X. Chen, "Machine health monitoring based on locally linear embedding with kernel sparse representation for neighborhood optimization," *Mech. Syst. Signal Process.*, vol. 114, pp. 25–34, 2019.

[18] M. Xia, T. Li, L. Xu, L. Liu, and C. W. De Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 1, pp. 101–110, Feb. 2018.

[19] Y. Zhao, Z. H. Guo, and J. M. Yan, "Vibration signal analysis and fault diagnosis of bogies of the high-speed train based on deep neural networks," *J. Vibroeng.*, vol. 19, no. 4, pp. 2456–2474, 2017.

[20] D. Zhao, T. Wang, R. X. Gao, and F. Chu, "Signal optimization based generalized demodulation transform for rolling bearing nonstationary fault characteristic extraction," *Mech. Syst. Signal Process.*, vol. 134, 2019, Art. no. 106297.

[21] J. Wang, P. Fu, L. Zhang, R. X. Gao, and R. Zhao, "Multilevel information fusion for induction motor fault diagnosis," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 5, pp. 2139–2150, Oct. 2019.

[22] Y. Yuan *et al.*, "A general end-to-end diagnosis framework for manufacturing systems," *Nat. Sci. Rev.*, vol. 7, no. 2, pp. 418–429, 2020.

[23] J. Li, R. Huang, G. He, S. Wang, G. Li, and W. Li, "A deep adversarial transfer learning network for machinery emerging fault detection," *IEEE Sensors J.*, vol. 20, no. 15, pp. 8413–8422, Aug. 2020.

[24] L. Han, C. Yu, C. Liu, Y. Qin, and S. Cui, "Fault diagnosis of rolling bearings in rail train based on exponential smoothing predictive segmentation and improved ensemble learning algorithm," *Appl. Sci.*, vol. 9, no. 15, 2019, Art. no. 3143.

[25] J. Dai, J. Tang, F. Shao, S. Huang, and Y. Wang, "Fault diagnosis of rolling bearing based on multiscale intrinsic mode function permutation entropy and a stacked sparse denoising autoencoder," *Appl. Sci.*, vol. 9, no. 13, 2019, Art. no. 2743.

[26] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *Proc. 3rd Int. Conf. Learn. Represent.*, 2015.

[27] M. Gan, C. Wang, and C. Zhu, "Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings," *Mech. Syst. Signal Process.*, vol. 72, pp. 92–104, 2016.

[28] H. Shao, H. Jiang, H. Zhao, and F. Wang, "A novel deep autoencoder feature learning method for rotating machinery fault diagnosis," *Mech. Syst. Signal Process.*, vol. 95, pp. 187–204, 2017.

[29] S. Shao, R. Yan, Y. Lu, P. Wang, and R. Gao, "Dcnn-based multi-signal induction motor fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2658–2669, Jun. 2020.

[30] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 807–814.

[31] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Machine Learn.*, 2015, vol. 1, pp. 448–456.

[32] "Spectraquest inc.: Machinery fault simulator - lite," [Online]. Available: https://spectraquest.com/simulators/details/mfs-lt/, Accessed on: Jul. 16, 2020.

[33] V. A. Epanechnikov, "Non-parametric estimation of a multivariate probability density," *Theory Probab. Appl.*, vol. 14, no. 1, pp. 153–158, 1969. Accessed on: Jul. 26, 2020.

Siyuan Chen received the B.S. degree in mechanical engineering from Pennsylvania State University, University Park, PA, USA, in 2017. He is currently working toward the Ph.D. degree in mechanical engineering with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

His research interests include object tracking, sensor integration, and data analysis for scalable manufacturing processes.

Yuquan Meng is currently working toward the Ph.D. degree in mechanical engineering with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

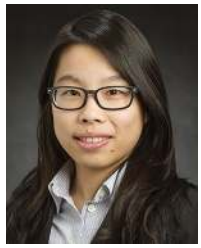
His current research interests include the simulation and optimization of manufacturing processes.

Haichuan Tang received the B.E. degree and M.S. degree in engineer diploma, in 2012, and the Ph.D. degree in electrical engineering from Southwest Jiaotong University, Chengdu, China, and also the M.S. degree from Ecole Centrale de Nantes, Nantes, France.

He is currently an Associate Researcher with the CRRC Academy, Beijing, China. He was a Visiting Scholar with the Rail Transportation and Engineering Center (RailTEC), University of Illinois at Urbana-Champaign, where he investigated train operations and its optimization. His research interests include intelligent transportation, machine learning, and Big Data analytics in health management.

Yin Tian received the B.E., M.S., and Ph.D. degrees in transportation engineering from Beijing Jiaotong University, Beijing, China.

He is currently a Senior Researcher with the CRRC Academy, Beijing, China. As a Project Manager, he has led the research and development of the CRRC rail prognostics and health management (PHM) system, intelligent wireless sensor networks for rail system, artificial intelligence-based behavior analytic system, etc. His research interests focus on intelligent products and their applications in transportation systems.



Niao He received the Ph.D. degree in operations research from Georgia Institute of Technology, Atlanta, GA, USA in 2015 and the B.S. degree in mathematics from the University of Science and Technology of China, Hefei, China, in 2010.

He is currently an Assistant Professor with the Department of Industrial and Enterprise Systems Engineering and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL, USA. Her research interests are in large-scale optimization, machine learning, and reinforcement learning.

Dr. He is also a recipient of the Best Paper Award at AISTATS 2016, the NSF CISE Research Initiation Initiative (CRII) Award, NCSA Faculty Fellow, and the CAS Fellow.

Chenhui Shao received the B.E. degree in automation from the University of Science and Technology of China, Hefei, China, in 2009, and the M.S.E. degree in industrial and operations engineering, the M.A. degree in statistics, and the Ph.D. degree in mechanical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2013, 2013, and 2016, respectively.

He is currently an Assistant Professor with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA. His research interests include smart manufacturing, machine learning, statistics, big data analytics in manufacturing, and materials joining.

Dr. Shao is a recipient of the National Science Foundation CAREER Award, SME Barbara M. Fossum Outstanding Young Manufacturing Engineer Award, SME 30 Under 30 Award, and multiple best paper awards.