Early Fault Detection of Machine Tools Based on Deep Learning and Dynamic Identification

Bo Luo, Haoting Wang, Hongqi Liu, Bin Li, and Fangyu Peng

***Abstract*—In modern digital manufacturing, nearly 79.6% of the downtime of a machine tool is caused by its mechanical failures. Predictive maintenance (PdM) is a useful way to minimize the machine downtime and the associated costs. One of the challenges with PdM is early fault detection under time-varying operational conditions, which means mining sensitive fault features from condition signals in long-term running. However, fault features are often weakened and disturbed by the time-varying harmonics and noise during a machining process. Existing analysis methods of these complex and diverse data are inefficient and time-consuming. This paper proposes a novel method for early fault detection under time-varying conditions. In this study, a deep learning model is constructed to automatically select the impulse responses from the vibration signals in long-term running of 288 days. Then, dynamic properties are identified from the selected impulse responses to detect the early mechanical fault under time-varying conditions. Compared to traditional methods, the experimental results in this paper have proved that our method was not affected by time-varying conditions and showed considerable potential for early fault detection in manufacturing.**

***Index Terms*—Computer numerical control (CNC) machine tool, deep learning, dynamic identification, fault diagnosis, feature extraction.**

Nomenclature

ANN	Artificial neural network.
BPNN	Back propagation neural network.
CNC	Computer numerical control.
CNN	Convolutional neural network.
DAE	Deep auto-encoder.
DBN	Deep belief network.
EMA	Exponential moving average.
FD	Frequency domain.
PdM	Predictive maintenance.
PCA	Principle component analysis.
SVM	Support vector machine.
SAE	Sparse auto-encoder.
TD	Time domain.
WD	Wavelet domain.

Manuscript received October 30, 2017; revised January 21, 2018 and January 23, 2018; accepted January 31, 2018. Date of publication February 19, 2018; date of current version August 31, 2018. This work was support-ed by the National Natural Science Foundation of China under Grant 51705174, Grant 51775210, and Grant 51625502. *(Corresponding author: Bo Luo.)*

B. Luo, H. Liu, B. Li, and F. Peng are with the School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: hglobo@163.com; liuhongqi328@163.com; libin999@hust.edu.cn; pengfy@hust.edu.cn).

H. Wang is with the Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061 USA (e-mail: haoting@vt.edu).

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

Digital Object Identifier 10.1109/TIE.2018.2807414

I. Introduction

CNC machine tools play an important role in modern digital manufacturing. They are used for shaping or machining metal or other rigid materials, usually by cutting, boring, grinding, shearing, or other forms of deformation. Due to long-term running, the abrasions and fatigues of its mechanical component are inevitable, resulting in poor product quality and low efficiency. In modern industry, nearly 79.6% of the downtime of a machine tool is caused by these mechanical failures [1]. Thus, early fault detection is highly desired to increase the product quality, expedite the progress of unscheduled maintenance, and shorten the machine downtime [2].

To address this issue, various signals, such as vibration analysis [3], current signature analysis [4], acoustic emission [5], stray flux analysis [6], etc., have been studied for condition monitoring. Among them, vibration analysis has been widely used and has shown satisfactory performance [7]. Generally, feature extraction is the key to early fault detection. The commonly used feature extraction methods can be mainly classified into three categories: the TD analysis [8], FD analysis [9], and TD–FD analysis [10].

However, most of the existing methods, which are based on a stationary assumption (i.e., constant speed and load), are only suitable for fault detection under constant operating conditions [11]. In actual machining processes, the working conditions are affected by many factors, such as different cutting tools, different workpieces, different cutting parameters, etc., [12]. Especially in long-term running, the machining load and speed vary frequently over time, which results in nonstationary signals, thus time-varying fault features. Fig. 1 illustrates a vibration signal collected from a cylinder machining process, consisting essentially of four operations: surfacing, side milling, drilling, and boring. Spectra I, II, III, and IV indicate the Fourier transforms of the vibrations in surfacing, side milling, drilling, and boring, respectively. As shown in the figure, under four conditions, the vibration characteristics in both TD and FD have changed significantly.

Usually, fluctuating cutting loads tend to cause amplitude variation of the vibration, while fluctuating cutting speeds induce frequency modulation. Such variations make conventional

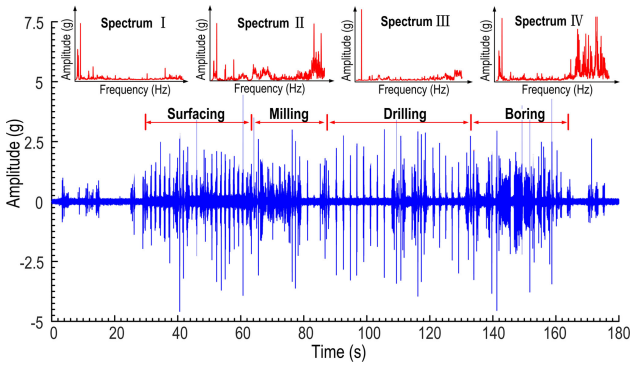


Fig. 1. Vibration signal and its spectrum of the machine tool under time-varying working condition.

stationary based methods ineffective. When the signal features change, it is difficult to determine whether they are caused by an abnormal mechanical failure or a normal change under working condition. To date, the reported literature on the topic of fault detection under nonstationary condition is still very limited. In general, two major challenges exist for early fault detection in real industrial environments:

1) Early fault detection means mining sensitive fault features from large-scale vibration data in long-term running. Most existing analysis methods of these complex and diverse data are manual methods, which are inefficient and time-consuming. It is necessary to introduce an intelligent data preprocessing method for automatic analysis of the large-scale vibration data.
2) In real industrial application, fault features are often weakened and disturbed by the time-varying harmonics and noise during machining. These nonstationary conditions will result in different feature spaces and different distributions. It is difficult to select sensitive fault features that are unaffected by the time-varying working conditions.

To fill this research gap, this paper proposes a novel method for early fault detection under time-varying working conditions. This method is performed in three steps. First, a deep learning model is used to automatically select the impulse responses from the long-term vibrations. Then, an algorithm is introduced to identify the dynamic properties of the machine tool from the impulse responses. Finally, an indicator based on the dynamic properties is *introduced to predict* the health status of the machine tool. The contributions of this paper are summarized as follows.

1) The dynamic properties were used for health condition monitoring in machine tools for the first time. The dynamic properties are inherent characteristics of a machine tool; hence, they are not affected by external time-varying working conditions. Compared to traditional fault features, they are more suitable for health monitoring in industrial applications.
2) A data mining method based on deep learning model was used to identify the dynamic properties, rather than simply pattern recognition and classification of existing faults, like most current deep learning models do.
3) The health status of the machine tool under time-varying operation condition was diagnosed and monitored for the first time.
4) The effectiveness of the proposed method was demonstrated by massive data in long-term running of an actual machine tool in a real industrial plant.

The rest of this paper is organized as follows. Section II reviews the related research works. In Section III, the framework and details of the proposed method for early fault detection under time-varying condition are briefly introduced. In Section IV, details of the validation experiment and data acquisition are illustrated. Then, the effectiveness of the proposed method is discussed in Section V. Finally, Section VI concludes the paper with a discussion of the future work.

II. Review Related Work

Signal-based fault detection and diagnosis methods are widely used in real-time monitoring and diagnosis for motors, buildings, power converters, and other mechanical systems [13]. These methods include three parts: signal acquisition, feature extraction, and fault recognition. The most challenging part is the extraction of sensitive fault features from the collected noisy signals [14]. Currently, two types of approaches exist for selecting sensitive features from the collected noisy signal. The first approach is the signal processing approach, such as statistical analysis [15], Fourier transform [16], wavelet transform (WT) [17], empirical mode decomposition [18], and sparse representation [14]. Besides the signal processing approach, dimensionality reduction is another type of approach to select features, which include PCA [19], local discriminant analysis [19], etc. However, all these approaches require manual selection of features, which has two issues in real industry applications. First, measurement noise, operator errors, nonlinear dynamics, etc., lead to significant uncertainty in the signal [20]. Second, the number of fault features in collected signal is very large but most of them are distorted or indifferent to health status [1], which are hard to be identified through manual selection. Due to these two issues, manually extracting a set of sensitive features is extremely difficult and time-consuming [21], thus it is difficult to apply in real industrial applications.

Different from traditional methods, deep learning is a new proposed machine learning, which has the great capacity to automatically learn the valuable features from the raw data. The deep learning models for fault diagnosis can be divided into three main types: CNN, DBN, and DAE [22]. Ince *et al.* [23] used a one-dimensional CNN for motor fault diagnosis and fused the feature extraction and classification phases into a single learning body. Liu *et al.* [24] constructed a Gaussian–Bernoulli DBN to learn sensitive features for rotor system diagnosis based on vibration imaging. Jia *et al.* [25] constructed a deep DAE to reduce data dimension to extract useful feature from the measured vibration for fault diagnosis. However, most of the deep learning models work well only under a general assumption: the training and testing data must have the same distribution and feature space [26]. It implies that the performances of these models would drop dramatically when the distribution and

feature space change [27]. Unfortunately, this situation happens frequently under time-varying working conditions, which severely challenges the validity of machine learning methods used in fault diagnosis outside of controlled time-invariant laboratory conditions. In order to solve this problem, Wen *et al.* [28] proposed a DAE model for bearing fault diagnosis based on deep transfer features. Liu *et al.* [29] designed a dislocated time series convolutional neural architecture to learn electric machine fault features from the vibration signal under nonstationary condition. Liu *et al.* [26] proposed a novel deep neural network model with domain adaptation for fault diagnosis. All these methods were proposed for simple pattern recognition and classification of existing faults. However, PdM requires mining and tracing the slight fault features in the early stage from large-scale vibration data, and monitoring the evolution of the degradation process. These existing deep learning models used for classification are not suitable in this case. In order to solve this problem, we proposed a data mining method to detect early fault features from vibration in long-term running. Different from other deep learning methods for fault diagnosis, our method used deep learning for dynamics identification rather than classifying the already existing faults.

III. Proposed Method

A. General Procedure of the Proposed Method

In this paper, a novel method for early fault detection under time-varying working conditions is proposed. The framework of this method is shown in Fig. 2 and the general procedures are summarized as follows.

Step 1: Collect the vibration signal of the machine tool in industry environment in long-term running.

Step 2: Take samples of a vibrational signal using a sliding frame. The aim of the sliding frame is to divide the overall vibration signal into smaller samples with a fixed length.

Step 3: Randomly select samples to build training dataset and testing dataset for a deep learning model. The samples are labeled into two categories: impulse response and nonimpulse response.

Step 4: The training dataset and testing dataset are used to train and test the deep learning model that consists of the SAE layer and the BPNN layer.

Step 5: The trained deep learning model is used to recognize and select the impulse responses from the measured vibration signals automatically in long-term running.

Step 6: The dynamic properties of the machine tool are identified from the impulse responses over time.

Step 7: Finally, a health indicator based on the similarity of dynamic properties is used to indicate and monitor the health status of the machine tool.

B. Impulse Response and Deep Learning

All vibration signals could be classified into two categories: impulse response and nonimpulse response. The impulse response is the vibration signal caused by the reaction of the dynamic system in response to the impulse. In the machining process, the impulse is a fast-acting force or impact, such as inertial force due to a sudden change of the movement of drives [30], pulse-like cutting force [31] caused by a sudden change of cutting conditions, etc. Mathematically, the impulse can be modeled as a Dirac delta function and has a wide excitation spectrum in the FD. Therefore, the response of the impulse contains important information for dynamics identification and has been widely used in engineering. On the other hand, the nonimpulse responses are the vibrations excited by ordinary force in machining. Normally, the harmonics in nonimpulse responses is significant, which makes it impossible to identify the dynamic properties.

In order to automatically recognize and select impulse responses from a complex vibration signal in long-term running in high accuracy, a deep learning model consisting of the SAE layer and BPNN layer was adopted. The SAE layer is an unsupervised learning neural network that consists of multiple layers of stacked SAEs. Details about the SAE are given in the literature [25] and [32]. Pretrained by the unlabeled dataset, it is used to reduce data dimension of the input. The BPNN layer is the last layer of the deep learning model. By supervised learning of the BPNN, the error back-propagation is adopted to fine-tune the SAE layer. In this way, the time-varying harmonics in nonimpulse responses could be eliminated and the dynamic properties could be identified easily from the impulse response.

C. Dynamic Identification Algorithm

The dynamic properties can be estimated from the impulse response using the state-space model. The identification algorithm is first proposed to estimate the dynamics of a civil building [33]. Similar to the civil building, the dynamics of the machine tool can also be represented by a state-space model:

$$\begin{aligned} \dot{x}(t) &= A_C x(t) + B_C u(t) \\ y(t) &= C_C x(t) + D_C u(t) \end{aligned} \tag{1}$$

where A_C is a $2n$-by-$2n$ designated state matrix, $x(t)$ is the state vector, which contains the displacements and the velocities vectors of the dynamic system, and $y(t)$ is the output vector that consists of the accelerations of the dynamic system. C_C is the output matrix and D_C is designated direct transmission matrix. In [33], it is demonstrated that state matrixes could be further expressed as follows:

$$\begin{aligned} A_C &= \Psi \Lambda_C \Psi^* \\ \Lambda_C &= \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda^* \end{bmatrix}, \quad \Psi = \begin{bmatrix} \Theta & \Theta^* \\ \Theta\Lambda & \Theta^*\Lambda^* \end{bmatrix} \\ \Theta &= \begin{bmatrix} \cdots & \phi_k & \cdots \end{bmatrix}, \quad k = 1, \ldots, n \end{aligned} \tag{2}$$

where $\bullet^*$ means a complex conjugate, λ_k is related to the structure frequency ω_k and damping ratio ξ_k according to the following expression:

$$\lambda_k = -\xi_k \omega_k + i\sqrt{1-\xi_k^2}\,\omega_k. \tag{3}$$

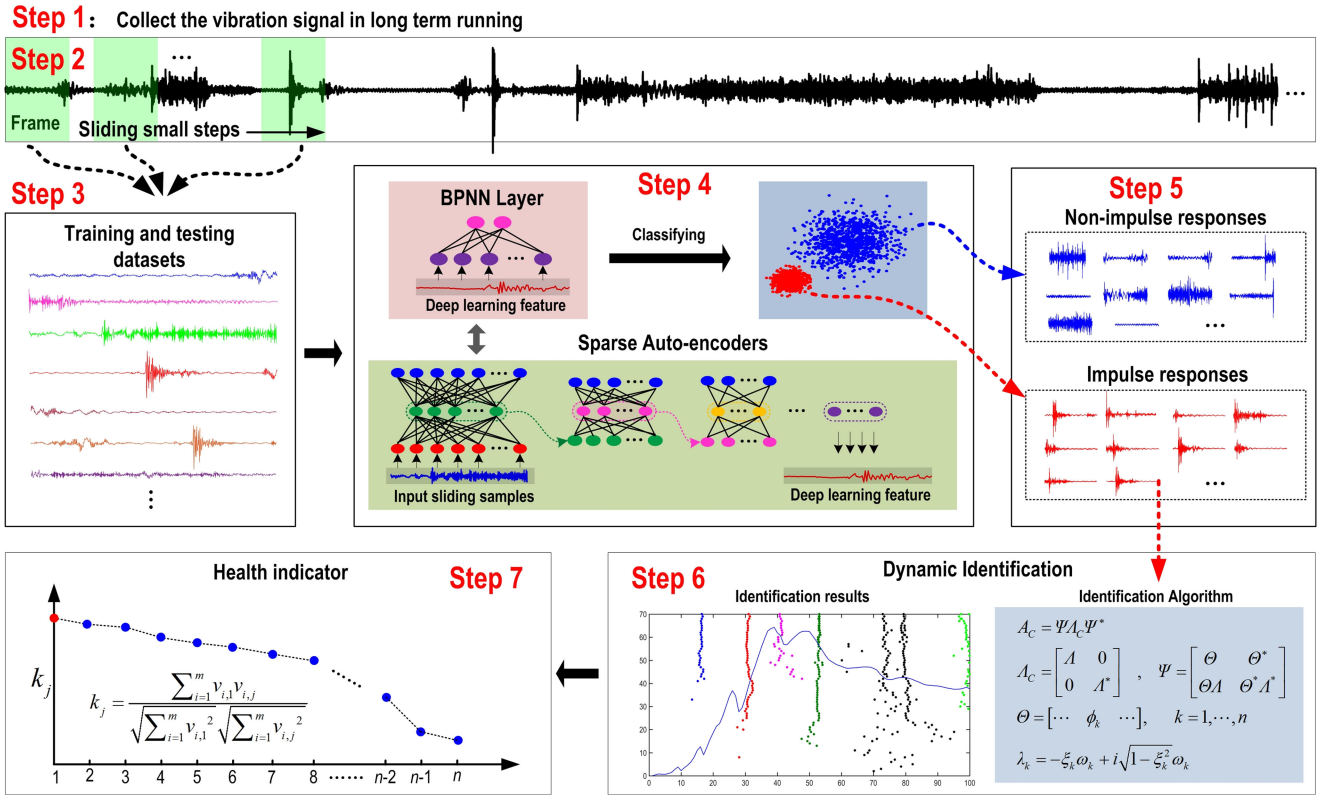


Fig. 2. Framework of the early fault detection method based on deep learning and dynamics identification.

Thus, once state matrix A_C has been estimated, it is easy to extract the natural frequencies and the damping ratios from the state matrix. In order to acquire the state matrix A_C, Reynders has given the relation between the state matrix and the polynomial matrices [33]:

$$A_c = \begin{bmatrix} -A_p^{-1}A_{p-1} & -A_p^{-1}A_{p-2} & \cdots & -A_p^{-1}A_1 & -A_p^{-1}A_0 \\ I & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}$$

$$B_C = \begin{bmatrix} A_p^{-1} \\ 0 \end{bmatrix}$$

$$C_C = \begin{bmatrix} B_{p-1} - B_p A_p^{-1} A_{p-1} & \cdots & B_0 - B_p A_p^{-1} A_0 \end{bmatrix}$$

$$D_C = B_p A_p^{-1}. \tag{4}$$

The polynomial matrices can be calculated by a right matrix fraction model acquired from the impulse response:

$$S_{yy}^{+}(\omega_j) = B(A)^{-1} = \left[\sum_{r=0}^{p} B_r e^{i\omega_j \Delta tr}\right]\left[\sum_{r=0}^{p} A_r e^{i\omega_j \Delta tr}\right]^{-1} \tag{5}$$

where B_r and A_r are matrices for dynamic identification, p is the order of the polynomials, and Δt is the interval time determined by the sampling rate. The goal of our identification algorithm is to find matrices B_r and A_r that minimize the differences between the half-spectrum matrix of the impulse response and the theoretical half-spectrum matrix given by

$$E^{NLS}(\omega_j) = \left[\sum_{r=0}^{p} B_r e^{i\omega_j \Delta tr}\right]\left[\sum_{r=0}^{p} A_r e^{i\omega_j \Delta tr}\right]^{-1} - \hat{S}_{yy}^{+}(\omega_j) \tag{6}$$

where $\hat{S}_{yy}^{+}(\omega_j)$ is a half-spectrum matrix of the impulse response. Then, the polynomial matrices B_r and A_r can be determined using the least squares cost function, which is obtained by adding all the squared elements of the error matrix evaluated at all the discrete frequency values from ω_1 to ω_{nf}:

$$\varepsilon = \sum_{o=1}^{n_o}\sum_{r=1}^{n_r}\sum_{j=1}^{nf} E_{o,r}(\omega_j) E_{o,r}(\omega_j)^* \tag{7}$$

where $E_{o,r}(\omega_j)$ is a general element in o line and r column of the error matrix. If the polynomial basis functions evaluated at ω_j are organized in one row vector with $(p+1)$ components

$$\begin{aligned} \Omega(\omega_j) &= \begin{bmatrix} \Omega_0(\omega_j) & \Omega_1(\omega_j) & \cdots & \Omega_p(\omega_j) \end{bmatrix} \\ &= \begin{bmatrix} e^{i\omega_j \Delta t0} & e^{i\omega_j \Delta t1} & \cdots & e^{i\omega_j \Delta tp} \end{bmatrix}. \end{aligned} \tag{8}$$

Then, one general line o of the error matrix can be calculated with the following equation:

$$E_o(\omega_j) = \Omega(\omega_j) \bullet \begin{bmatrix} B_{0o} \\ B_{1o} \\ \vdots \\ B_{po} \end{bmatrix} + \left[\Omega_0(\omega_j)\ \Omega_1(\omega_j) \cdots \Omega_p(\omega_j)\right] \hat{S}^+_{yyo} \begin{bmatrix} A_0 \\ A_1 \\ \vdots \\ A_p \end{bmatrix} \tag{9}$$

where B_{ro} represents the o line of matrix B_r and $\hat{S}^+_{yyo}$ represents the o line of matrix $\hat{S}^+_{yy}$. Introducing the following definitions:

$$\beta_o = \begin{bmatrix} B_{0o} \\ B_{1o} \\ \vdots \\ B_{po} \end{bmatrix} \text{ with } o = 1,\ 2, \ldots, n_o \text{ and } \alpha_o = \begin{bmatrix} A_0 \\ A_1 \\ \vdots \\ A_p \end{bmatrix}. \tag{10}$$

Equation (10) can be further generalized to all the discrete frequencies values

$$E_o(\beta_o, \alpha) = \left[X_o\ Y_o\right] \begin{bmatrix} \beta_o \\ \alpha_o \end{bmatrix}. \tag{11}$$

Finally, the scalar cost function can be calculated with the following matrix equation:

$$\begin{aligned} \varepsilon(\beta_\mathbf{o}, \alpha) &= \sum_{\mathbf{o}=1}^{\mathbf{n_o}} tr\left\{\mathbf{E_o}(\beta_\mathbf{o}, \alpha)^\mathbf{H} \mathbf{E_o}(\beta_\mathbf{o}, \alpha)\right\} \\ &= \sum_{\mathbf{o}=1}^{\mathbf{n_o}} tr\left\{\left[\beta_\mathbf{o}^\mathbf{T}\ \alpha^\mathbf{T}\right] \begin{bmatrix} \mathbf{R_o} & \mathbf{S_o} \\ \mathbf{S_o^T} & \mathbf{T_o} \end{bmatrix} \begin{bmatrix} \beta_\mathbf{o} \\ \alpha \end{bmatrix}\right\} \end{aligned} \tag{12}$$

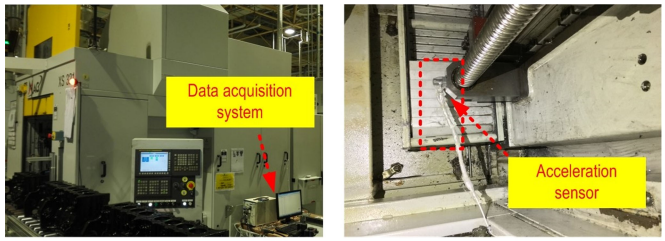
where $R_o = \mathrm{Re}(X_o^H X_o)$, $R_o = Re(X_o^H X_o)$, $T_o = Re(Y_o^H Y_o)$, and $tr\{\bullet\}$means the sum of the elements in the main diagonal of a matrix, and $Re(\bullet)$ represents selection the real part of a complex number. The minimum of the cost function is then determined by forcing its derivatives with respect to elements B_r and A_r matrices to be zero

$$\begin{aligned} \frac{\partial \varepsilon(\beta_o, \alpha)}{\partial \beta_o} &= 2(R_o \beta_o + S_o \alpha) = 0 \\ \frac{\partial \varepsilon(\beta_o, \alpha)}{\partial \alpha} &= 2\sum_{o=1}^{n_o} \left(S_o^T \beta_o + T_o \alpha\right) = 0. \end{aligned} \tag{13}$$

By solving the above-mentioned partial differential equation, α and β can be calculated to construct the state matrix. With this step, the identification problem is solved.

IV. Data Acquisition

In this study, the vibrational data are collected continuously from a CNC machining center in an automobile factory. The machining center is used to manufacture the engine blocks. It handles the following different machining operations: surfacing, boring, lifter bore machining, drilling, spot facing, cylinder chamfering, and O-ring grooving. Thus, the working conditions of the machine tool vary significantly over time. Due to high-speed movement of the feed drive, ball screw wear is the most common failure in daily manufacturing. In order to detect the early fault, a three-phase acceleration sensor is fixed on the feed drive base. Fig. 3 shows the sensor location and the data acquisition system. To digitize the continuous vibration signals, a Dewesoft data acquisition system is used. Detailed specifications of the data acquisition system are presented in Table I. In this experiment, the vibration data were collected continuously for 288 days with the data volume of 1.086 Tb.



(a) (b)

Fig. 3. Data acquisition system on CNC machining center in a real manufacturing factory.

TABLE I
SPECIFICATIONS OF THE DATA ACQUISITION SYSTEM

Acceleration sensor (Dytran:3097A1)	1) Sensitivity: 10 mV/g sensitivity 2) Range F.S. For ±5 Volts Output: ±500 g 3) Frequency response: 0.3–10 000 Hz 4) Resonant frequency: >40 kHz 5) Temperature range: –60 to +250 F
Data acquisition system (Dewesoft: HD-ACC)	1) 24 bit 100 kHz/5 kHz conversion 2) Gain accuracy: ±0.05% of reading 3) 2000 samples/s rate
Acquisition time	1) From 2016.5.14 17:36:41 2) To 2017.2.27 09:56:27
Data volume	1) 1.086 Tb

V. Experimental Results

In this section, the impulse responses were selected by the DAE model first. Then, the impulse responses were used to identify the dynamic properties for early fault detection.

A. Effectiveness of the Deep Learning Model

A DAE model was trained to recognize and select the impulse responses from the long-term vibration data. In order to evaluate the performance of the model, 10 000 samples were collected from the measured signals by a sliding frame and each frame contains 1024 data points. The samples were classified into two categories: the impulse response and nonimpulse response. Among all the response samples, 9000 samples are randomly selected for training, and the remaining 1000 for testing.

TABLE II
PARAMETERS OF THE DEEP LEARNING MODEL

Description	Value
The length of the sliding frame	1024
The number of the auto-encoders	4
The size of the first auto-encoders	1024*256*1024
The size of the second auto-encoders	256*128*256
The size of the third auto-encoders	128*64*128
The size of the fourth auto-encoders	64*32*64
Initial learning rate/Increasing factor/Decreasing factor	0.08/1.05/0.9

TABLE III
CLASSIFICATION RESULTS BASED ON DIFFERENT FEATURES OF THE IMPULSE RESPONSE

Method	Average accuracy	Standard deviation
DAE features	97.3%	0.058
WT features	86.5%	0.354
FD features	79.7%	0.412
TD features	64.2%	1.277

TABLE IV
PERFORMANCE OF DAE MODELS WITH DIFFERENT NUMBERS OF AUTO-ENCODERS

Auto-encoder number	2	3	4	5
The final accuracy	94.5%	97.3%	97.5%	98.1%
The training time (Minutes)	152.4	245.7	372.1	513.8

In the experiments, the DAE model with four auto-encoders was constructed. The parameters of the DAE model are listed in Table II. Meanwhile, the traditional signal-based features were selected manually as comparison, including TD features, FD features, and WT. Detailed descriptions of the WT features, TD features, and FD features can be found in [34], [35], and [36], respectively. In order to validate the results, 15 trials were performed for training and testing the methods.

The average accuracy of the 15 trials is listed in Table III. It can be seen that the DAE features have higher accuracy and better stability than traditional signal-based features. Besides, the traditional methods are labor-intensive and time-consuming when dealing with massive vibrational data. It can be expected that any possible potential improvements of traditional methods would cost large amount of labor and time.

To futhur study the performance of DAE models, deep learning models with different numbers of auto-encoders were built and compared. The results are presented in Table IV. In general, the final accuracy is improved gradually as the number of auto-encoders increases. However, it costs longer time to train the deep learning model with the increased number of auto-encoders. Thus, the balance between model accuracy and its training time needs to be considered in a specific case.

B. Effectiveness of a Dynamics Identification Algorithm

In order to verify the dynamics identification algorithm, the natural frequencies and damping ratios of the machine tool are identified from two different impulse responses A and B.

TABLE V
RESULTS OF IDENTIFICATION

	Natural frequency (Hz)				Damping			
Mode	First	Second	Third	Fourth	First	Second	Third	Fourth
Response A	13.02	24.12	39.47	62.01	0.139	0.064	0.036	0.017
Response B	12.89	23.94	39.55	63.29	0.108	0.115	0.043	0.028
Change rate	1.0%	0.8%	0.2%	2.0%	28%	44%	16%	39%

The impulse responses A and B are selected under two different working conditions, such as drilling and milling. Fig. 4 presents the stability diagram of the results that identified from impulse responses A and B. The stability diagram is an intuitive approach that presents the model order of the algorithm and the solutions found for the order. In Fig. 4(b) and (e), the clusters with different colors represent the calculated dynamic property results, in which the natural frequencies are shown on the abscissa axis, while the model orders of the algorithm are displayed on the ordinate axis. Usually if a calculated frequency were to appear at any order, there would be a high probability that it was a natural system frequency. Therefore, the clusters in blue, red, pink, and green colors indicate the estimated natural frequencies of the first, second, third, and fourth modes of the machine tool, respectively. Similarly, the clusters with the corresponding colors in Fig. 4(c) and (f) represent the estimated damping ratios of those four modes of the machine tool. As shown in the four stability diagrams in Fig. 4, the identifications gradually converge into linear clusters as the order number increases. Once the model order is high enough, most clusters in the stability diagrams are fairly clear, which indicates that the results of the dynamics identification algorithm are believable and the method is valid.

Based on the stability diagrams, the mean values of natural frequencies and damping ratios of the four modes that identified from impulse responses A and B are calculated and presented in Table V. As seen, under different working conditions, the change rates of the damping ratios are much larger than those of the natural frequency change rates. The change rate of natural frequency of each mode is very small, which implies that the natural frequencies are almost unaffected by external time-varying working conditions. Therefore, in this paper, only natural frequencies are used for indicating the health condition of the machine tool.

C. Heath Indicator Based on Dynamics Identification

The similarity between feature vectors is used to represent the health status. Existing methods for computing the similarity between the two vectors include Hamming distance, Euclidean distance, Hausdorff distance, Cosine distance, Bray–Curtis similarity, Pairwise-adaptive similarity, Jaccard coefficient, etc. [37]. Among them, the cosine distance is one of the most widely used methods for similarity measurement in fault diagnosis and its effectiveness has been proved in the literature [38].

In this paper, the Cosine distance is adopted as the health indicator. As shown in Fig. 5, the first feature vector with red color is defined as a standard vector that represents the initial

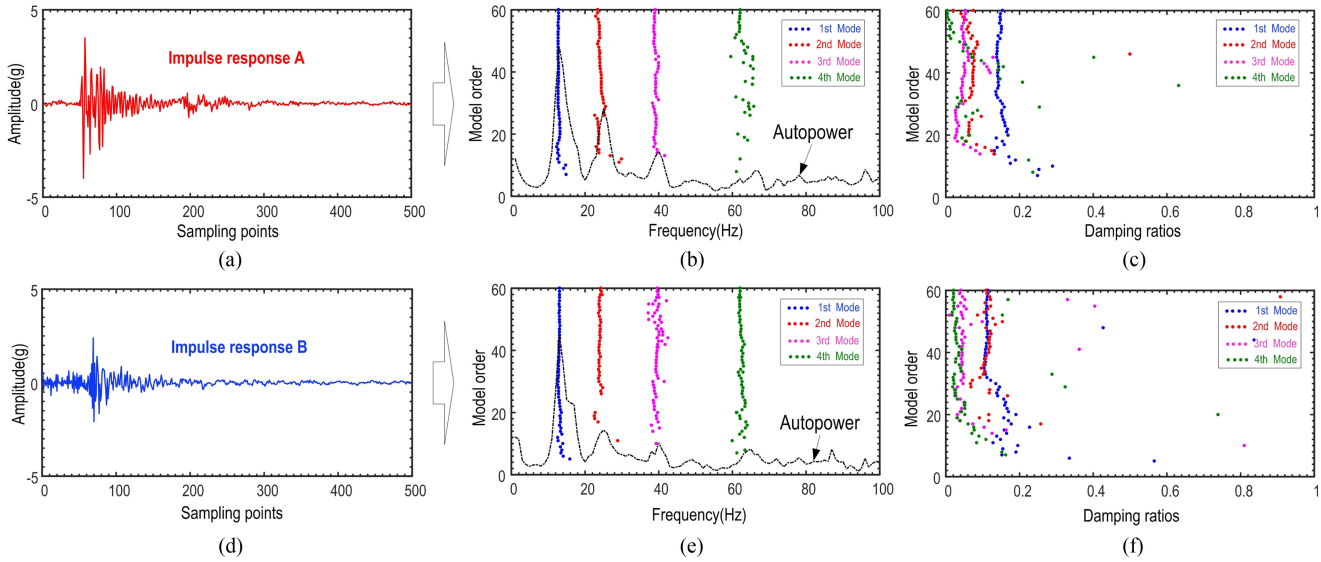


Fig. 4. Stability diagrams of the identification results.

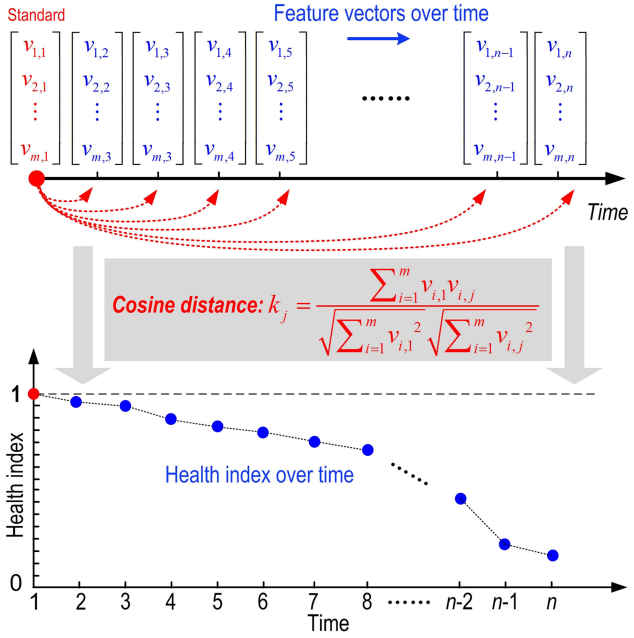


Fig. 5. Health index based on the similarities of the feature vectors over time.

health status. The blue feature vector with subscript n indicates the current health status. The cosine of the angle between two vectors could be used to measure how "similar" they are: two vectors with the same orientation have a cosine similarity of 1, which indicates current health status does not change; two vectors in orthogonal directions have a similarity of 0, which indicate the health status changes a lot from its original status; and in-between values indicate intermediate similarity or dissimilarity. If the cosine similarity value between the standard vector and current vector decreases, it implies that the current health status is becoming worse than the initial health status.

In this experiment, the vibrational signal has been collected continuously for 228 days. The vibration data volume is huge; thus, the computational burden is considerable. It is impractical and unnecessary to use all impulse responses for dynamics identification. In order to solve this problem, only a partial of impulse responses are selected for dynamics identification and health condition monitoring. The overall process is as follows:

Step 1: To reduce data volume, a portion of them are selected for health condition monitoring. Although more impulse responses are selected means higher the accuracy and resolution of the health index, more impulse response also implies more computation. Thus, the balance between accuracy and computing resource should be considered in a specific case. In our experiment, four impulse samples were randomly selected every day. Thus, there were totally $228 \times 4 = 1152$ impulse responses.

Step 2: The natural frequencies of the first four modes of the machine tool are identified in this study. Besides, in order to validate their efficiency, TD features, FD features, and DAE features are also extracted from the impulse response as comparison.

Step 3: Four kinds of feature vectors are constructed based on natural frequencies, TD features, FD features, and DAE features.

Step 4: Cosine distances are calculated as the health indicators based on the four different feature vectors.

Fig. 6 presents the four heath indicators based on different feature vectors. The dots represent the calculated health indexes, while the curve indicates their mean values. From the four figures, it can be seen that only the natural frequencies vector can indicate health status of the machine tool, since the other three methods did not show the health status change during the degradation process. The reason for this is that the natural frequencies are inherent dynamic characteristics of the machine tool. They are unaffected by the variation of the external time-varying working conditions. But the damage in a mechanical component will introduce local flexibilities and then changes in dynamic characteristics [39]. Fig. 7(a) presents

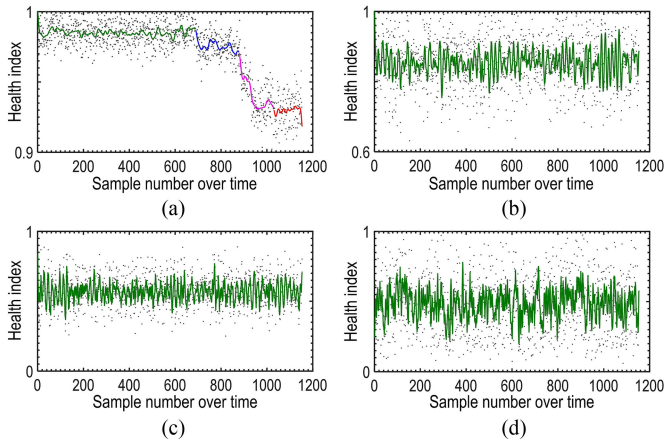


Fig. 6. Health indicator based on different feature vectors. (a) Natural frequency features, (b) DAE features, (c) FD features, and (d) TD features.

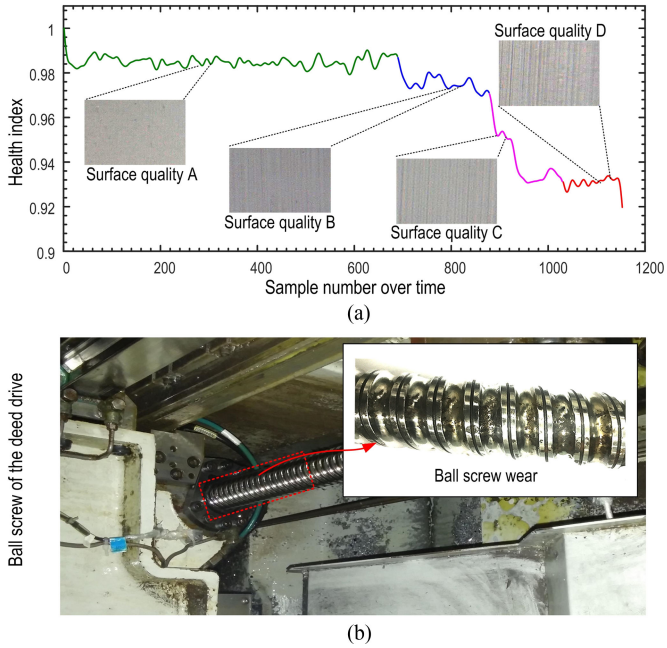


Fig. 7. Performance of the machine tool in the overall degradation process. (a) The surface quality in different health stages, and (b) the final failure of the feed drive.

the performance of the machine tool in the overall degradation process. As seen, there are four stages in the health index curve: health, slight deterioration, rapid deterioration, and severe deterioration. The performance of the machine tool is verified by the surface qualities of the workpieces at these four stages. As shown in Fig. 7(a), the roughness of the machined surface deteriorates gradually. And Fig. 7(b) shows the picture of the faulty feed drive of the machine tool at the final stage. As seen, serious abrasions occurred on the surface of the ball screw, shown by the areas highlighted in the red box.

VI. Conclusion

In this paper, a novel data mining method was proposed for early fault detection of the CNC machine tool under time-varying conditions. A deep learning model was built and trained to automatically select the impulse responses from massive vibration data in long-term running. And an algorithm was introduced to identify the dynamic properties from the impulse samples. Finally, based on the similarity between the dynamic properties, a health index was used to indicate the slow and gradual deterioration process of the machine tool. To validate the proposed method, long-term experiment was carried out on a CNC machine tool in an automobile factory. Massive data were collected and analyzed, and the results showed that the method can effectively indicate the health status of the machine tool.

The uniqueness and contribution of this study are fourfold. First, the dynamic properties were used for early fault detection in machine tools for the first time. Second, a deep learning model was used to identify the dynamic properties, rather than simply pattern recognition and classification of existing faults, like most current deep learning models do. Third, the health status of the machine tool under time-varying operation condition was diagnosed and monitored for the first time. Fourth, this method was validated by actual industrial machining data collected continuously in a time-span of 288 days.

For the future work, three possible directions are suggested here. First, the prediction of the product quality using deep learning methods can be further investigated. In this study, the deep learning models successfully identified the health status of the machine tool. However, the product quality of the machine tool is also highly interesting, thus needs further research. Second, the optimal number of selected samples of impulse response needs further investigation to reach the balance between efficiency and accuracy. Third, different deep learning models can be investigated in early fault detection.

References

[1] R. Teti, K. Jemielniak, G. O'Donnell, and D. Dornfeld, "Advanced monitoring of machining operations," *CIRP Ann., Manuf. Technol.*, vol. 59, no. 2, pp. 717–739, Jul. 2010.

[2] M. Canizo, E. Onieva, A. Conde, S. Charramendieta, and S. Trujillo, "Real-time predictive maintenance for wind turbines using big data frameworks," in *Proc. 2017 IEEE Int. Conf. Prognostics Health Manage.*, 2017, pp. 70–77.

[3] R. Duan and F. Wang, "Fault diagnosis of on-load tap-changer in converter transformer based on time–frequency vibration analysis," *IEEE Trans. Ind. Electron.*, vol. 63, no. 6, pp. 3815–3823, Jun. 2016.

[4] A. Naha, A. K. Samanta, A. Routray, and A. K. Deb, "Low complexity motor current signature analysis using sub-Nyquist strategy with reduced data length," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 12, pp. 3249–3259, Dec. 2017.

[5] Y. Wang, C. Xue, X. Jia, and X. Peng, "Fault diagnosis of reciprocating compressor valve with the method integrating acoustic emission signal and simulated valve motion," *Mech. Syst. Signal Process.*, vol. 56, pp. 197–212, 2015.

[6] L. Frosini, C. Harlişca, and L. Szabó, "Induction machine bearing fault detection by means of statistical processing of the stray flux measurement," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1846–1854, Mar. 2015.

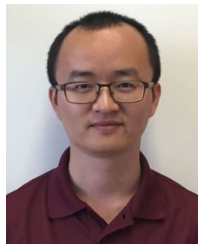
[7] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3757–3767, Jun. 2015.

[8] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, "Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3398–3407, Aug. 2013.

[9] X. Jin, Y. Sun, Z. Que, Y. Wang, and T. W. Chow, "Anomaly detection and fault prognosis for bearings," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 9, pp. 2046–2054, Sep. 2016.
[10] Z. Feng, M. Liang, and F. Chu, "Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples," *Mech. Syst. Signal Process.*, vol. 38, no. 1, pp. 165–205, 2013.
[11] M. S. Kan, A. C. Tan, and J. Mathew, "A review on prognostic techniques for non-stationary and non-linear rotating systems," *Mech. Syst. Signal Process.*, vol. 62, pp. 1–20, 2015.
[12] G. Quintana and J. Ciurana, "Chatter in machining processes: A review," *Int. J. Mach. Tools Manuf.*, vol. 51, no. 5, pp. 363–376, 2011.
[13] G. Hackmann, W. Guo, G. Yan, Z. Sun, C. Lu, and S. Dyke, "Cyber-physical codesign of distributed structural health monitoring with wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 1, pp. 63–72, Jan. 2014.
[14] M. Riera-Guasp, J. A. Antonino-Daviu, and G.-A. Capolino, "Advances in electrical machine, power electronic, and drive condition monitoring and fault detection: State of the art," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1746–1759, Mar. 2015.
[15] B. R. Nayana and P. Geethanjali, "Analysis of statistical time-domain features effectiveness in identification of bearing faults from vibration signal," *IEEE Sens. J.*, vol. 17, no. 17, pp. 5618–5625, Sep. 2017.
[16] M. A. Moussa, M. Boucherma, and A. Khezzar, "A detection method for induction motor bar fault using sidelobes leakage phenomenon of the sliding discrete Fourier transform," *IEEE Trans. Power Electron.*, vol. 32, no. 7, pp. 5560–5572, Jul. 2017.
[17] J. Seshadrinath, B. Singh, and B. K. Panigrahi, "Investigation of vibration signatures for multiple fault diagnosis in variable frequency drives using complex wavelets," *IEEE Trans. Power Electron.*, vol. 29, no. 2, pp. 936–945, Feb. 2014.
[18] J. B. Ali *et al.*, "Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals," *Appl. Acoust.*, vol. 89, pp. 16–27, 2015.
[19] X. Jin, M. Zhao, T. W. S. Chow, and M. Pecht, "Motor bearing fault diagnosis using trace ratio linear discriminant analysis," *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2441–2451, May 2014.
[20] S. Simani, F. Saverio, and P. Castaldi, "Fault diagnosis of a wind turbine benchmark via identified fuzzy models," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3775–3782, Jun. 2015.
[21] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3137–3147, May 2016.
[22] G. Jiang, H. He, P. Xie, and Y. Tang, "Stacked multilevel-denoising autoencoders: A new representation learning approach for wind turbine gearbox fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2391–2402, Sep. 2017.
[23] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016.
[24] Z. Liu, Z. Jia, C. M. Vong, S. Bu, J. Han, and X. Tang, "Capturing high-discriminative fault features for electronics-rich analog system via deep learning," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1213–1226, Jun. 2017.
[25] F. Jia *et al.*, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mech. Syst. Signal Process.*, vol. 72, pp. 303–315, 2016.
[26] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang, "Deep model based domain adaptation for fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2296–2305, Mar. 2017.
[27] R. Zhang, H. Tao, L. Wu, and Y. Guan, "Transfer learning with neural networks for bearing fault diagnosis in changing working conditions," *IEEE Access*, vol. 5, pp. 14347–14357, 2017.
[28] L. Wen, L. Gao, and X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.
[29] R. Liu, G. Meng, B. Yang, C. Sun, and X. Chen, "Dislocated time series convolutional neural architecture: An intelligent fault diagnosis approach for electric machine," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1310–1320, Jun. 2017.
[30] B. Li, B. Luo, X. Mao, H. Cai, F. Peng, and H. Liu, "A new approach to identifying the dynamic behavior of CNC machine tools with respect to different worktable feed speeds," *Int. J. Mach. Tools Manuf.*, vol. 72, pp. 73–84, Sep. 2013.
[31] N. Tounsi and A. Otho, "Dynamic cutting force measuring," *Int. J. Mach. Tools Manuf.*, vol. 40, no. 8, pp. 1157–1170, 2000.
[32] H. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2727–2736, Mar. 2018.
[33] B. Peeters and H. Van derAuweraer, "PolyMAX: A revolution in operational modal analysis," in *Proc. 1st Int. Oper. Modal Anal. Conf.*, Copenhagen, Denmark, Apr. 2005, pp. 1–12.
[34] J. Lin and L. Qu, "Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis," *J. Sound Vib.*, vol. 234, no. 1, pp. 135–148, 2000.
[35] Y. Lei, Z. He, and Y. Zi, "EEMD method and WNN for fault diagnosis of locomotive roller bearings," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 7334–7341, 2011.
[36] Y. Lei, Z. He, and Y. Zi, "A new approach to intelligent fault diagnosis of rotating machinery," *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1593–1600, 2008.
[37] S.-S. Choi, S.-H. Cha, and C. C. Tappert, "A survey of binary similarity and distance measures," *J. Syst., Cybern. Informat.*, vol. 8, no. 1, pp. 43–48, 2010.
[38] Y. Dong, Z. Sun, and H. Jia, "A cosine similarity-based negative selection algorithm for time series novelty detection," *Mech. Syst. Signal Process.*, vol. 20, no. 6, pp. 1461–1472, 2006.
[39] F. Magalhães, A. Cunha, and E. Caetano, "Vibration based structural health monitoring of an arch bridge: From automated OMA to damage detection," *Mech. Syst. Signal Process.*, vol. 28, pp. 212–228, 2012.

Bo Luo received the Bachelor's degree in machinery manufacturing and automation and the Ph.D. degree in mechatronic engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2008 and 2014, respectively.

He is currently a Postdoctoral Researcher with the School of Mechanical Science and Engineering, Huazhong University of Science and Technology. His current research interests include intelligent manufacturing, sensor data mining, signal processing, and machine learning.



Haoting Wang received the B.S. degree in mechanical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2008, the M.S. degree in mechanical engineering from New York University, New York, NY, USA, in 2011, and the Ph.D. degree in aerospace engineering from Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, in 2017.

His current research interests include thermal management and heat transfer, vehicle aerodynamics, and machine learning in industrial applications.

Hongqi Liu received the Ph.D. degree in mechatronic engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2008.

He is currently an Associate Research Professor with the School of Mechanical Science and Engineering, Huazhong University of Science and Technology. His current research interests include machinery health monitoring, diagnosis and prognosis, complex systems failure analysis, quality and reliability engineering, and manufacturing systems design, modeling, scheduling, and planning.

Bin Li received the B.S., M.S., and Ph.D. degrees in mechanical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1982, 1989, and 2006, respectively.

He is currently a Professor with the School of Mechanical Science and Engineering, Huazhong University of Science and Technology. His current research interests include intelligent manufacturing and computer numerical control machine tools.

Fangyu Peng received the Bachelor's and Ph.D. degrees in mechanical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1994 and 2000, respectively.

He is currently a Professor with the School of Mechanical Science and Engineering, Huazhong University of Science and Technology. His current research interests include machining process optimization, robot machining, and ultraprecision machining.