

Received August 26, 2019, accepted September 7, 2019, date of publication September 12, 2019, date of current version October 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2940769

# Motor Fault Detection and Feature Extraction Using RNN-Based Variational Autoencoder

YANG HUANG<sup>ID1</sup>, CHIUN-HSUN CHEN<sup>ID2</sup>, AND CHI-JUI HUANG<sup>2</sup>

<sup>1</sup>Institute of Information Management, National Chiao Tung University, Hsinchu 30010, Taiwan

<sup>2</sup>Department of Mechanical Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan

Corresponding author: Chiun-Hsun Chen (chchen@mail.nctu.edu.tw)

This work was supported by the Ministry of Science and Technology of Taiwan under Grants MOST 107-2622-8-009-020 and MOST 108-2221-E-009-059-MY2.

**ABSTRACT** In most of the fault detection methods, the time domain signals collected from the mechanical equipment usually need to be transformed into frequency domain or other high-level data, highly relying on professional knowledge such as signal processing and fault pattern recognition. Contrary to those existing approaches, we proposed a two-stage machine learning analysis architecture which can accurately predict the motor fault modes only by using motor vibration time-domain signals without any complicated preprocessing. In the first stage, the method RNN-based VAE was proposed which is highly suitable for dimension reduction of time series data. In addition to reducing the dimension of sequential data from 150\*3 to 25 dimensions, our method furthermore improves the prediction accuracy evaluated by several classification algorithms. While other dimension reduction methods such as Autoencoder and Variational Autoencoder cannot improve the classification accuracy effectively or even decreased. It indicates that the sequential data after dimension reduction via the RNN-based VAE still can maintain the high-dimensional data information. Furthermore, the experimental results demonstrate that it can be well applied to time series data dimension reduction and shows a significant improvement of the prediction performance, even with a simple double-layer Neural Network can reach over 99% of accuracy. In the second stage, Principal Components Analysis (PCA) and Linear Discriminant Analysis (LDA) are used to further perform the second dimension reduction, such that the different or unknown fault modes can be clearly visualized and detected.

**INDEX TERMS** Motor fault detection, feature extraction, recurrent neural network, variational autoencoder.

## I. INTRODUCTION

In recent years, the global manufacturing has faced with the environmental tests such as rapidly changed market and increasingly personalized demands. In addition, under the general lack of labor force, it is difficult for manufacturers to maintain their industrial status if they cannot flexibly adapt to the production conditions. The “smart manufacturing” rising with the stack of the above factors is rapidly reshaping the appearance of the global manufacturing and factories. For example, in semiconductor process, due to continuous pursuit of lightweight, efficiency and endurance of electronic products, the competition on semiconductor wafer process technology among all major factories is becoming increasingly fierce. In the well-developed area of materials and products, it is more likely to encounter bottlenecks and technical challenges are getting tougher. If various machine parameters

of the process control can be successfully used for analysis, there will be an opportunity to find breakthroughs.

In the traditional way of machine fault diagnosis, simple measuring instruments are used to collect signals, so that maintenance personnel can conduct diagnosis based on their knowledge and experience. The principle of machine management is to conduct regular and routine maintenance, or to shut down for maintenance after faults occur, or even direct replacement. However, in such a way, the capacity losses arising from shutdown shall be borne, and the fault factors may have irreparably damaged machines. Hence, in smart manufacturing, the use of fault detection and classification system (FDC) is specially emphasized [4].

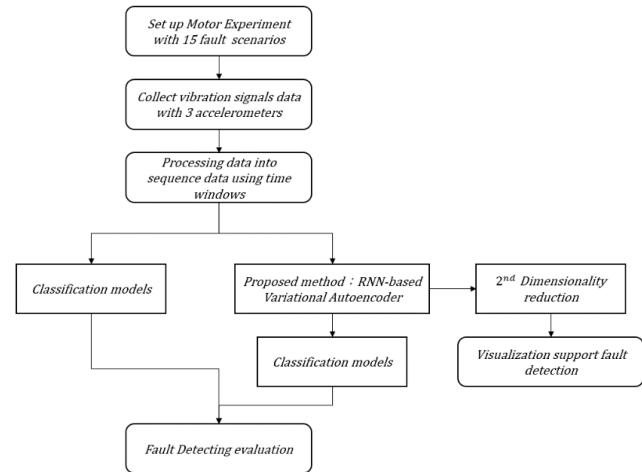
According to the McKinsey Report [26], big data analysis has great commercial application potentials in the maintenance industry. Different from the previous mode of routine inspection and break-down repair, in FDC, big data is used to analyze health conditions of machine, to predict possible faults of machine and to prevent further loss

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

through preventive maintenance. In E-maintenance proposed by Arnaiz *et al.* [2], the main function of big data analysis is assisting in making maintenance decisions. Generally speaking, maintenance modes can be divided into two types: corrective type and preventive type, and the latter can be divided into look-ahead type and condition-based type (predictive type). Look-ahead maintenance can prevent damages to machines and equipment through fixed intervals, but is more likely to cause waste resources due to excessive maintenance. Predictive maintenance predicts the probability of failure based on the health conditions of machine, which depends on the analysis of machine data collected by a large number of sensors. Not only internal parameters of machine but external factors, such as temperature, humidity and vibration, will become the important basis for determination and modeling, to achieve the diagnosis effect in cooperation with appropriate models. As for the production line machine, the most vital component is the motor providing power. In practice, motor faults are most likely to have serious consequences, so that most of the current studies focus on motor fault detection [1], [10], [11], [33]. However, most of the current studies rely on professional knowledge such as high-level signal processing [9] and fault identification, and input data are entered into the classification model after feature extraction. Therefore, this study attempted to propose a model that can effectively reduce the dimension of time series by simply inputting the time domain data of original signals, which can improve the accuracy of fault detection in addition to feature extraction and dimension reduction.

In this study, a motor experiment platform was designed to simulate 15 fault scenarios and to collect the time domain vibration signals at 3 different locations. The experimental results show that the prediction accuracy is not ideal when the original time domain data are input into all the classification models. However, if the RNN-based VAE model proposed in this study is used to extract features firstly and then the dimension of all sequential data with the dimension of (150, 3) are reduced to the non-sequential data with 25 dimensions. In addition to great reduction in hardware burden of computation, a high fault detection accuracy can be obtained through simple classification models. The major contributions of this study can be divided into 4 aspects:

- 1) A feature extraction and dimension reduction model suitable for time series data was proposed
- 2) In most of the fault detection methods, the time domain signals are transformed into frequency domain or other high-level data, highly relying on professional knowledge such as signal processing and fault mode identification. Whereas in this model, the fault mode can be successfully identified by simply inputting the time domain data of original vibration signals.
- 3) The RNN-based VAE model was proposed. The experimental results show that it can effectively conduct dimension reduction and feature extraction on sequential vibration signals, then the accuracy of fault



**FIGURE 1.** Study flowchart.

detection using simple classification models can be significantly improve.

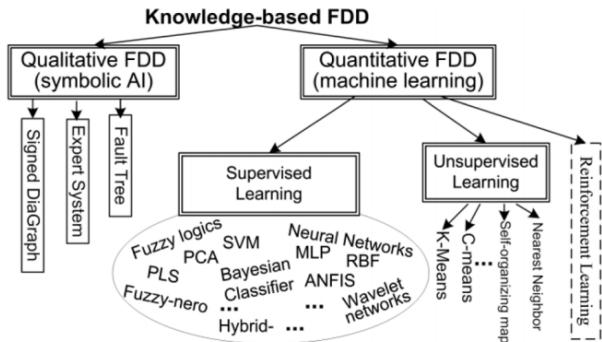
- 4) At the last part of the experiment, the data after dimension reduction were reduced to 2 dimensions using PCA and LDA, and the visualization results show that the identification effects are good and can be used to assist in determining unknown fault modes.

The research structure is shown in Fig 1. The remainder of this paper is organized as follows: Section 2 reviews the existing fault detection and predictive maintenance, and presents studies related to machine learning used for feature extraction and fault mode identification; Section 3 introduces the method to propose RNN-based VAE; Section 4 describes the experiment platform establishment, data collection and preprocessing; Section 5 compares the differences in effects before and after using our method, and discusses the visualization results of dimension reduction of the two stages.

## II. RELATED WORK

Fault detection technology plays an important role in Industry 4.0, among which, the most important part is data processing and algorithm design. In this section, the studies related to fault detection and the performance of various algorithms on fault detection are organized.

Comstock *et al.* [7] proposed 7 steps for the operation of predictive maintenance: data capture, data preprocessing, condition stability, health assessment, prognostics, decision support and presentation. It is shown that the first 4 steps are the basic elements constituting fault detection diagnosis (FDD). Dai and Gao [8] divided FDD into: physical mode-based method, signal-based method, and knowledge-based and historical data-navigated method, as shown in Fig 2. In physical mode-based method, by cooperating with the empirical data, the governing equations deduced for the physical characteristics of equipment systems are used to obtain the empirical laws. In signal-based method, specific signals are used to define different FDD scenarios.



**FIGURE 2.** Knowledge-based and historical data-navigated method [8].

In recent years, the machine learning algorithm is popular due to the rapid development of software and hardware, so that the knowledge-based FDD has become a hot study topic.

According to the statistical results of the study conducted by Thomson and Fenger [31], when motors are abnormal, the abnormalities in core components, such as stators, rotors and bearings, account for 88% in all motor abnormalities, and the variables commonly used in motor abnormality detection are current, electromagnetic field, vibration and other signals. Based on Li and Mechefske [23], the signals of motor stators, such as current, vibration and sound, are collected, and various motor speeds and load test scenarios of multiple groups are made, after cross test, it is found that stator current is highly sensitive to rotor damages and vibration signals are more suitable for detecting bearing abnormalities. Thomson *et al.* [30] used power spectrum analysis to diagnose a high-power and high-voltage three-phase induction motor, the results show that current spectrum data can effectively control the occurrence of air-gap eccentricity.

Fault detection can be roughly divided into two steps: feature extraction and fault mode identification. Glowacz and Glowacz [14] collected the sound signals of unidirectional induction motors and obtained the new feature vectors through extraction of the signal features, and on this basis, established a k-nearest neighbor algorithm model. The experiment results show that the goal of fault detection can be successfully achieved by inputting the new feature vectors into the model. Zhong *et al.* [38] applied empirical mode decomposition (EMD) [17] and teager energy operator (TEO) [37] to analyze the periodic pulse signals of vibration inside and outside races of fault bearings, and the results show that the faults inside and outside races of fault bearings can be detected more effectively by using the periodic pulse signals of vibration than by using the complete spectrum signals. Xiao *et al.* [36] used support vector machine (SVM) in cooperation with principle component analysis (PCA) to compare with the three-layer neural network, which mentioned that the dimension reduction technology of PCA can dramatically decrease the model training time and provide better accuracy. Gauch extracted the data features in PCA-based method to eliminate noises and improve the performance of dimension reduction results [12].

In recent years, artificial neuron network (ANN) has become the most dominant algorithm. Due to its advantages such as high dimensional data processing, strong ability in reasoning and elastic structure design, more and more people have applied neural network to mechanical equipment fault detection and signal feature extraction [35]. Bailey [3] used artificial neuron network to conduct fault detection analysis of screw chillers, and proposed that the misclassification rate of the optimal model is 20%, through direct classification of normal modes and fault modes in the literature. Liyanagedera *et al.* [24] made 4 different fault scenarios of internal combustion engines, used fast Fourier transform to transform the collected vibration signals into frequency domain data and established a neural network model to train data. 250 input nodes, 150 hidden layers and elastic back-propagation are used to optimize the training course to verify that this model can successfully predict the fault types, after testing multiple sets of mixed fault scenarios.

Many people used Autoencoder which is a method based on neural network to extract features. Thirukovalluru *et al.* [29] input many handcrafted features into autoencoder to extract features and then into the two models including SVM and Random Forest, to obtain good prediction accuracy. Chen and Li [5] input the statistics of bearing signals into Sparse-autoencoder and used Deep belief network for classification prediction. Shao *et al.* [28] input spectrum data into SAE for feature extraction and into restricted Boltzmann machine (RBM) to conduct fault detection of induction motors. In addition, some scholars used RNN [25] and Autoencoder [32] to remove noises while extracting features.

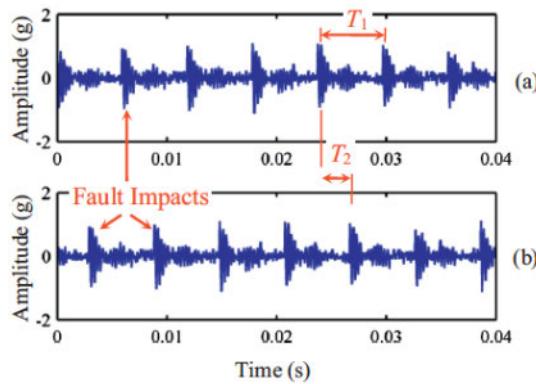
To sum up, it can be found that most of the existing studies are about the pre-processed data such as input spectrum and statistics but not the original time domain data [19]. Additionally, if the time windows of the original time domain data are directly slide and segmented, it will be found that the fault types appear at different time points, which will affect the interpretation effect (see Fig 3) [18]. Therefore, the above study can achieve a good accuracy, but one major problem still remains: the signal feature extraction relies on professional knowledge such as high-level signal processing and fault identification, which consumes considerable costs before actual input into the classification model.

### III. METHODOLOGY

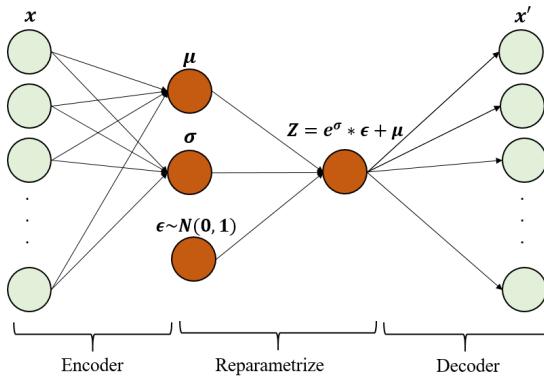
An implementation scheme of our proposed method is shown in Fig 8. We use RNN as the encoding and decoding models in VAE, in order to extract the latent features of the sequential data. These methods will be explained in this section.

#### A. VARIATIONAL AUTOENCODER

Autoencoder is an unsupervised learning algorithm based on neural network, with the main purpose of extracting low-dimensional features (also known as code) which can represent the original data, so as to achieve the effects of dimension reduction. Autoencoder is comprised of two neural networks, respectively encoder and decoder. The input data



**FIGURE 3.** Fault signal displacement caused by time window [18].



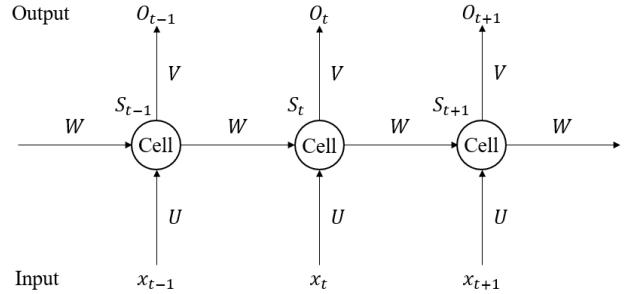
**FIGURE 4.** Structure of variational autoencoder.

will be compressed into lower-dimensional codes after passing through the encoder, and the resulting codes will pass through the decoder again and restored to the original data by using the decoder.

The weights of Autoencoder are updated according to the output and input reconstruct losses, the original data and the difference between decoded output data can be minimized through iterative update, so that the compressed codes can retain data information in lower dimensions, to achieve the purpose of dimension reduction and feature extraction. The code size, number of layers of encoder and decoder, and number of neurons can be self-designed.

Kingma and Welling proposed Variational Encoder (VAE) [22] in 2013, as shown in Fig 4. Compared with the general Autoencoder, in VAE, variational inference and parameterization trick are added to strengthen the model's ability in feature extraction and maintain the diversity of generated data, instead of simple reconstruction of input data. VAE is one of the generative models because of its purpose to maximize the log likelihood of the data  $x$  generated by a given latent variable  $z$ .

In VAE, assume the data distribution can be expressed in latent variable  $z$  obtained from encode and  $z$  is subject to standard Gaussian distribution  $N \sim (0, 1)$ . Let the distribution function of  $z$  be  $p(z)$ , the process of the data  $x$  generated by decoder can be expressed by (1). In other words, to generate



**FIGURE 5.** Diagram of RNN structure.

the data  $x$ , firstly randomly sample  $z$ , and then generate  $x$  from  $P_\theta(x|z)$ .

$$P(X) = \int P(X|z)P(z)dz \quad (1)$$

The objective function of VAE is shown in (2).  $KL(q||p)$  denotes the Kullback-Liebler divergence of the two probability distributions  $p$  and  $q$ . Please refer to section 3.3 for the detailed description of loss function.

$$\text{maximize } E_{q_\phi(z|x)} - KL(q_\phi(z|x) || P_\theta(z)) \quad (2)$$

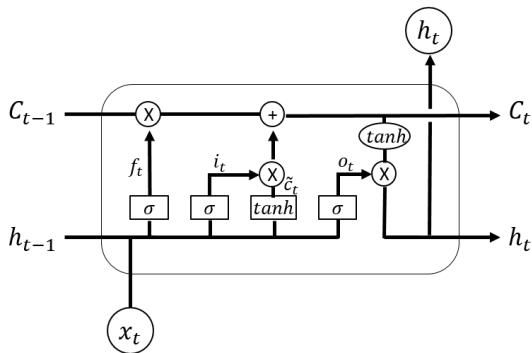
### B. RECURRENT NEURAL NETWORK(RNN)

The vibration signals used in this study are time series data, so RNN is more suitable for processing. Before input into RNN, data are transformed into two-dimensional according to the size of time windows to maintain their sequentiality. In addition, in computing the state  $S_t$  of current time  $t$ , RNN not only multiplies the current input value  $x_t$  by the weight  $U$  of input layer, but also sums up with the result obtained by multiplying by the common weight  $W$ , so as to retain the previously memorized information (see Fig 5).  $W$  is the common weight,  $U$  and  $V$  are the weight matrix of input layer and output layer respectively, and  $o_t$  is the output value at this point of time.

$$S_t = f(U * x_t + W * S_{t-1}) \quad (3)$$

$$o_t = f(s_t * V) \quad (4)$$

The recurrent neural network can keep data to be relevant, but due to weight reuse, gradient vanish or gradient explode is more likely to happen in the face of a longer time step. Gradient explode is easy to be solved, gradient clipping [27] can be added to regulate the modified gradient. The way to solve gradient vanish comes from LSTM (Long Short-Term Memory) [16] proposed by Hochreiter and Schmidhuber in 1997, and LSTM was optimized by Felix Gers' team in 2000 (see Fig 6.) [13]. The solution is to set three thresholds, and the values are controlled between 0 and 1 by sigmoid function to control the proportion of input, memory and output. 0 denotes blocking all information and 1 denotes accepting all information. If the value of the input gate  $i_t$  is approximately 0, the input value at this time is blocked. If the value of the forget gate  $f_t$  is approximately 0, and the value of the memory state at this time is forgotten. If the value



**FIGURE 6.** Diagram of LSTM structure.

of the output gate  $o_t$  is approximately 0, the memory won't be output.  $x_t$  is the input at the time t and  $C_t$  is the memory unit at the time t, and  $h_t$  is the state of the hidden layer.

It is worth noting that relevant literature [15], [20] proposed forget gate is normally the most important of all gates, and the larger the initial value is, the better. Please refer to (5) to (10) for the formula of LSTM:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f + \text{forgetbias}) \quad (5)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (6)$$

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (7)$$

$$C_t = i_t * \tilde{C}_t + f_t * C_{t-1} \quad (8)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (9)$$

$$h_t = o_t * \tanh(C_t) \quad (10)$$

Cho et al. proposed a modified version of LSTM model in 2014: Gated Recurrent Unit (GRU) [6], and its structure is shown in

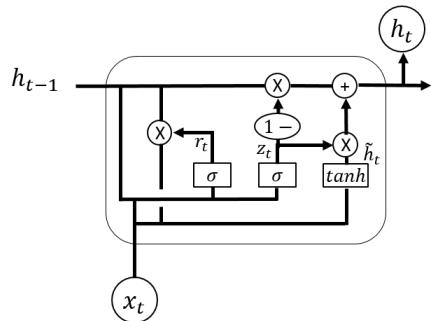
Cell state and hidden state are combined and the output gate  $o_t$  is cancelled, and the three control gates are reduced to two: the update gate z and the reset gate r. The reset gate r determines the proportion of  $h_{t-1}$  in the previous step to become the candidate  $\tilde{h}_t$  in the current state, and all of them are to be reset if r is 0. The update gate z determines the proportion of updated data, and the output  $h_t$  at the current time t will equal to the output  $h_{t-1}$  at the previous point of time if z is 1, which indicates that the input data at the current time are not adopted at all.

Compared with LSTM, GRU decreases the number of parameters to be trained by reducing gates, which can effectively reduce the occurrence of overfitting in addition to improving the speed of model training and convergence.

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (11)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (12)$$

$$h_t = z_t * h_{t-1} + (1 - z_t) * (\tanh(W_h x_t + U_h (r_t * h_{t-1}) + b_h)) \quad (13)$$



**FIGURE 7.** Diagram of GRU structure.

### C. PROPOSED METHOD: RNN-BASED VAE

In RNN-based Autoencoder, two RNN networks are used to respectively replace the general fully connected layers as the encoder and the decoder, so that the model can incorporate the temporal characteristics in computation and reduce the dimension of the time series data more effectively [34]. In this study, Autoencoder in the structure was replaced by Variational Autoencoder (VAE), and variance and noise were added to make the generation of model more real, rather than just to reconstruct the input data.

The structure of RNN-based VAE proposed in this study is shown in Fig 8. The time domain vibration signals of motor are input into the Encoder, and then mean and variance of the latent variable are obtained after passing through RNN and the dense layer. After that, the latent variable is obtained by adding up the results gained based on the equations in Fig 8. Finally, the input data are decoded by using the dense layer and RNN. During the training, batch normalization is used to improve the model performance, and the latent space computed in the process is used as the input data of the subsequent classification model.

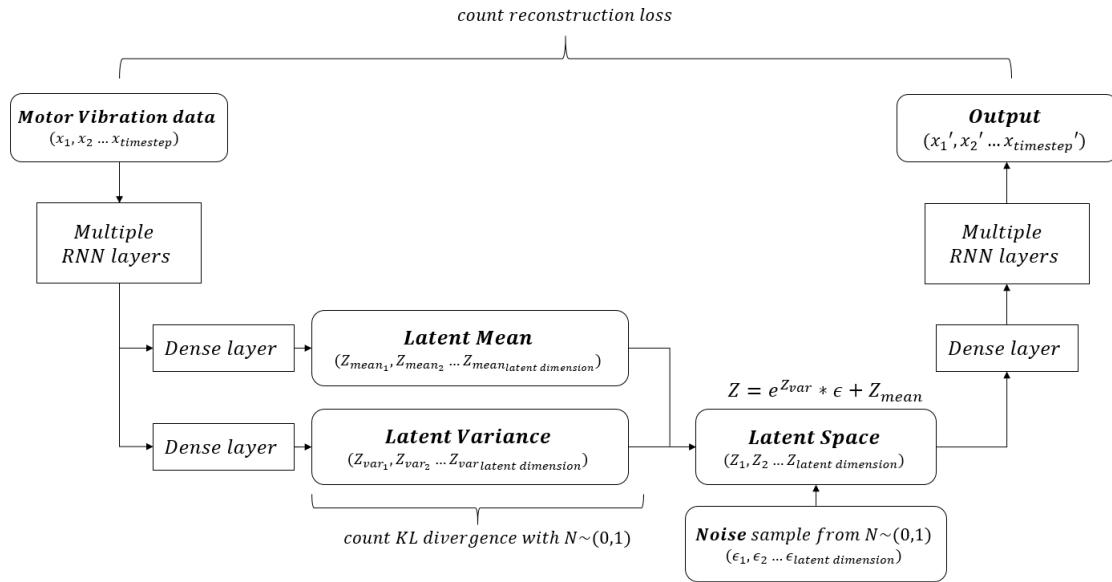
The loss function used in RNN-based VAE is designed as the sum of reconstruction loss and KL divergence. Reconstruction loss is calculated as the cross entropy of the probability distribution of the model's output value and true value, as shown in Eq. (14). Symbol descriptions:  $y$  denotes true probability,  $a$  denotes predict probability, and  $N$  denotes the number of data.

$$\text{Cross entropy} = -\frac{\sum_{n=1}^N [y * \ln a + (1 - y) * \ln(1 - a)]}{N} \quad (14)$$

KL divergence is to calculate the differences between latent variable distribution and standard normal distribution, to achieve the function of regularization. The equation is shown in (15),  $P$  denotes real distribution, and  $Q$  denotes predicted distribution:

$$KL(P \parallel Q) = -\sum_i P(i) * \ln \frac{Q(i)}{P(i)} \quad (15)$$

For optimizer, after the experiment in this study, Kingma et al. [21] has the best effect. Integrating the gradient

**FIGURE 8.** Proposed RNN-based VAE structure.

descent with momentum and the advantages of Adagrad, Adam is the most commonly used optimizer at present, and the mathematical equations are shown in (16) to (20). In the equation  $\theta$  denotes the parameter,  $\eta$  denotes the learning rate,  $J$  denotes the loss function,  $i$  denotes the index of  $\theta$ ,  $t$  denotes the number of iteration, and  $\varepsilon$  denotes the smooth term.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)g_t \quad (16)$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)g_t^2 \quad (17)$$

$$\hat{m}_t \leftarrow \frac{m_t}{1 - \beta_1^t} \quad (18)$$

$$\hat{v}_t \leftarrow \frac{v_t}{1 - \beta_2^t} \quad (19)$$

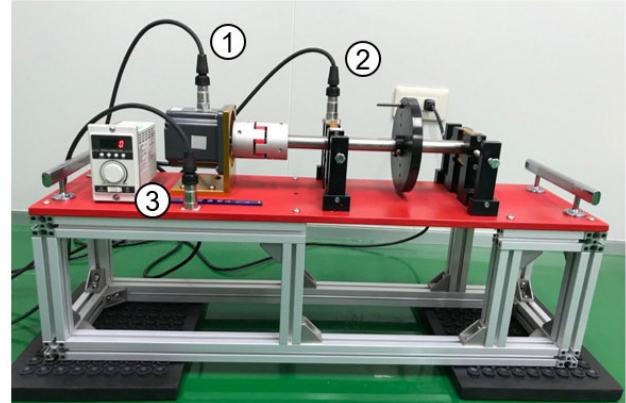
$$\theta_{t+1} \leftarrow \theta_t - \left( \frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon} \right) \hat{m}_t \quad (20)$$

## IV. EXPERIMENT

### A. DATASET

In this study, a motor experiment platform was designed to collect vibration signals. Featured as low temperature rise, low noise, small size and low maintenance cost, a brushless gear reduction motor was used, and the detailed specification is shown in Table 1. As shown in Fig 9, two black cushions are placed under the corner base of the motor, aiming at reducing the vibration arising from the slightly uneven floor. In the choice of coupling, a flexible coupling is used to connect the motor with the rotating shaft, and the reason why rigid coupling is not used is to avoid damages to the shaft due to excessive vibration, but the disadvantages are absorbing vibration and suppressing signal intensity and performance.

The way to collect signals is to fix the accelerometer at the position where the vibration is to be measured, to receive and measure the g value of the accelerated speed generated by the vibration, and the sampling rate is set to 4500 Hz. Totally

**FIGURE 9.** Positions of accelerometers in the motor experiment platform.

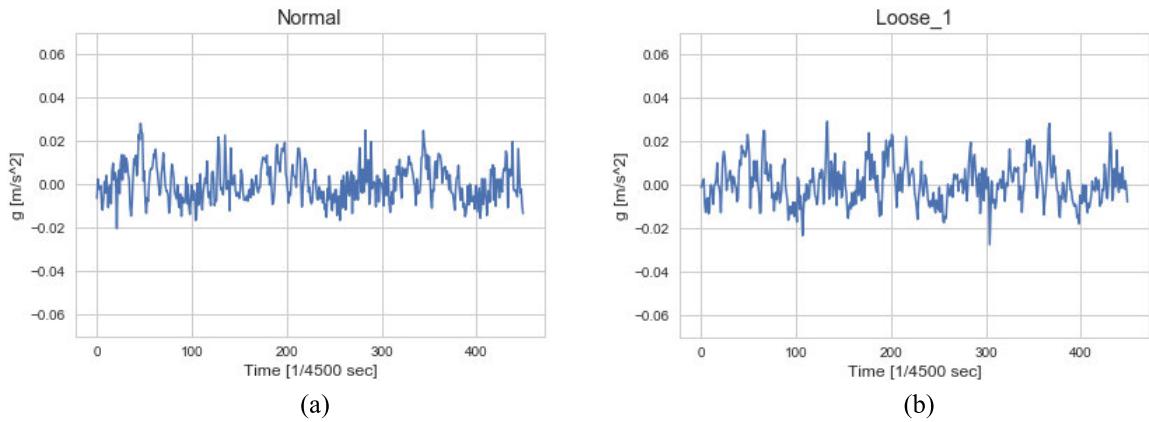
three accelerometers were used to measure the vibration of the motor, the bearing and the experiment platform respectively, as shown in Fig 9.

The rotation speed was set at 1800rpm, which is 30Hz. There are 4 designed experiment scenarios, respectively: normal, dynamic unbalance, loosening and mixing, and a total of 15 categories were created after different degrees of severity were adjusted, the details are described below.

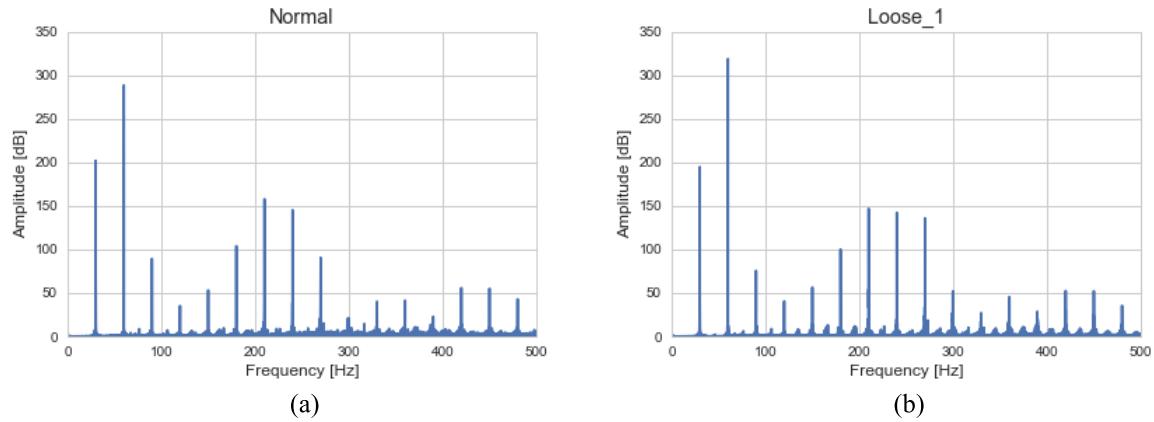
- 1) Normal: no external load was added on the motor experiment platform, and all the operation conditions were normal without any other interference.
- 2) Dynamic unbalance: if the central axis of mass of a rotating component is not on the same line with the central axis of rotation, then vibration will be caused when the axis rotates, that is dynamic unbalance. In this study, screws and two nuts were used as a set of load, with a total weight of 10g. There were respectively

**TABLE 1.** Motor specification.

Rated speed	Maximum speed	Machine length	Rated power	Input voltage	Phase	Rated torque	Maximum torque
3000 rpm	3500 rpm	80 mm	200 W	AC 220	3	0.66 N·M	1.32 N·M



**FIGURE 10.** Time domain vibration signals of motor collected in (a) normal scenario (b) loosening scenario.



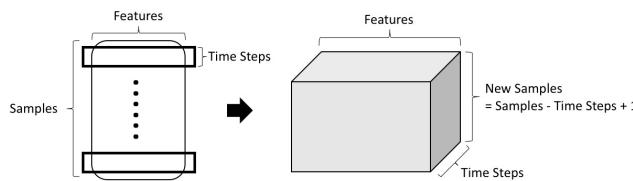
**FIGURE 11.** Transformed the data in Fig 10 into spectrum signals (a) normal scenario (b) loosening scenario.

3 screw holes in all directions from the inner ring to the outer ring on the load disc, therefore, various rotations with dynamic unbalance could be created when different number of screws were locked on different positions. Firstly, 1 to 3 screws were locked in Y direction, and then 1 to 3 screws were also locked in X direction. A total of 6 scenarios with dynamic unbalance were created.

- 3) Loosening: the screws on the bearing blocks were loosened to simulate the situation of loosening. A total of 2 loosening scenarios were created based on different positions.
  - 4) Mixing: the screws on the bearing blocks were loosened from one side and the load was added, to create 6 abnormal scenarios mixing dynamic unbalance with loosening.

The sampling rate is set to 4500 Hz and the speed of 1800 rpm represent that 150 data points can be collected per revolution (that is why the time step was set at 150). The “normal” scenario and the “loosening” scenario are taken as the examples, time domain data are drawn as shown in Fig 10, and FFT frequency domain data are drawn as shown in Fig 11.

Normally, the dominant frequency should appear at the position of double frequency. However, as can be seen from Fig 11, with too many noises in the spectrum, the spectrum features of all experiment scenarios could not be clearly observed, so that it was difficult for the staff to use traditional knowledge to determine the fault scenarios. The algorithm model proposed in this study can avoid this problem, and the fault scenarios can be distinguished immediately after direct input of time domain signals.

**FIGURE 12.** Time window slide and segmentation.**TABLE 2.** Loss of rnn-based VAE using different encoder/decoder.

	RNN	LSTM	GRU
Loss	1.384	0.325	0.215

**TABLE 3.** Loss of RNN-based VAE using different hidden unit numbers.

	16	32	34	128	256
Loss	0.51	0.49	0.24	0.038	0.0044

A total of 12,150,000 data were collected, with 3 features. Time step was set to 150, because 150 data points could be collected every revolution of the motor, which took 1/30 second. The data were converted into time windows which were slide and segmented by 150 time steps, and then transformed from 2-dimensional data into 3-dimensional data, so as to conform to the input structure of the recurrent neural network, as shown in Fig 12.

Before input into the model, the data were randomly split into training set, validation set and testing set, with the proportion of 60:20:20. The training set was used to train the model and update the weights. Validation set was used to evaluate training curve of RNN-based VAE. After RNN-based VAE converged or met the conditions of early stopping, the weight of this model was fixed. After that, the feature vectors of dimension reduction obtained after the training data passing through this model were used for training of the subsequent classification model.

### B. HYPERPARAMETER TUNING

This part is to determine the parameters of RNN-based VAE. RNN, LSTM and GRU are the three models that most widely used in processing sequential data. Firstly, the latent space was set at 75 dimensions and time step was set at 150 (Experiments show that time step below 150 will greatly reduce the model performance), in order to find out the most suitable model for encoder/decoder. The final results of the three models are shown in Table 2, and during the course, 5-fold cross-validation was used for verification. It can be found that GRU has the best effect and rapid convergence, so GRU is used as the encoder and the decoder for RNN-based VAE in the subsequent experiments.

The comparison on the number of hidden units in GRU is shown in Table 3, and it can be found that loss continuously decreases as the number of hidden units increases. However, in consideration of the fact that overfitting is easy to occur with long computation time and large number of parameters,

**TABLE 4.** Loss of RNN-based VAE using different batch sizes.

	64	128	256	512
Loss	0.144	0.159	0.038	0.187

**TABLE 5.** Loss and accuracy of RNN-based VAE using different latent dimensions.

	3	5	10	25	50	75	100
Loss	0.482	0.121	0.067	0.0453	0.083	0.038	0.072
Accuracy	0.495	0.862	0.956	0.984	0.994	0.997	0.992

and it is found in subsequent experiments that hidden units have slight effects on classification accuracy after they are increased to a certain amount, 128 is finally determined to be the number of neurons in encoder/decoder of GRU.

Table 4 shows the results obtained by changing the batch sizes of all trainings, and it can be seen from the table that the minimum training loss can be obtained if the batch size is set at 256.

In the study structure, the most important part was the latent space obtained after encoding. Table 5 shows the loss calculated by the latent space which different dimensions are set, then input into the neural network for comparison of classification accuracy. It can be found that the information of original data are hard to be saved at a low latent space dimension, resulting in unsatisfactory subsequent classification accuracy. After weighing, the dimension of the latent space was set at 25, because increasing the dimension further will not be obviously beneficial for loss and accuracy, but contrary to the purpose of “dimension reduction”.

After the structure and the parameters were established according to the above experiment results, the RNN-based VAE model was designed based on this parameter combination to reduce the dimension of the motor vibration data, and then: (1) they were input into the subsequent classification model to assess accuracy and (2) visualization analysis was conducted after another dimension reduction. Encoder/Decoder: GRU with 128 hidden units, Dimension of latent space: 25, Time step of input data: 150, Batch size: 256.

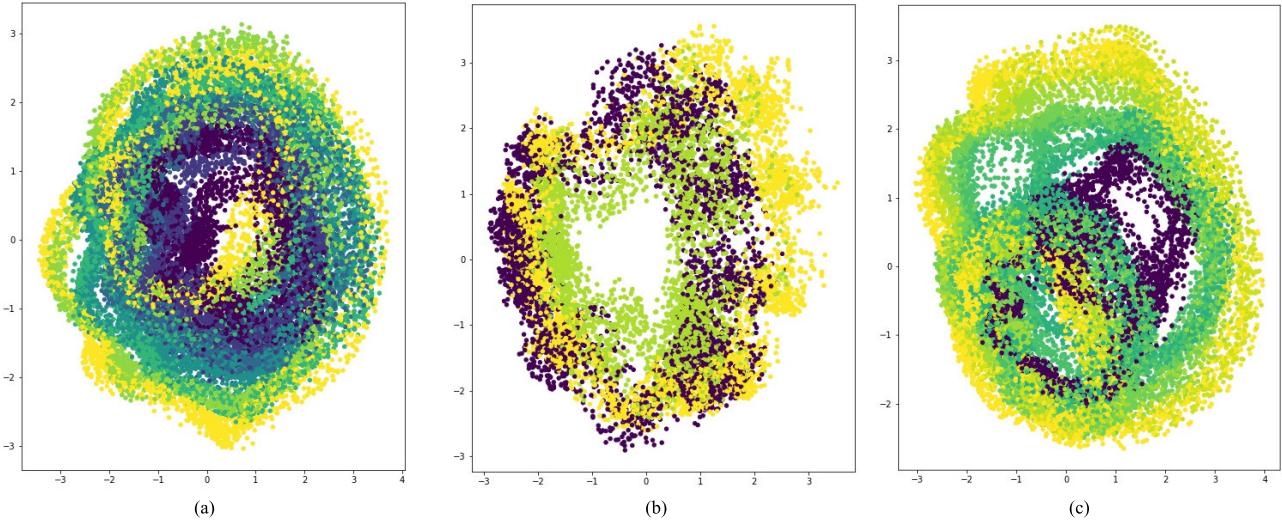
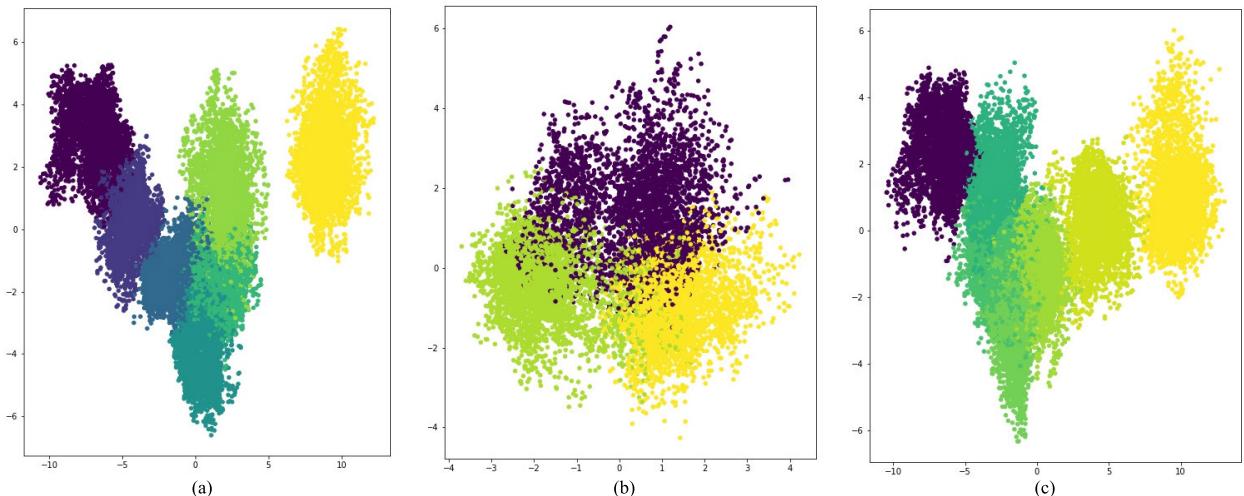
### C. MODEL EVALUATION

In this section, the prediction accuracy calculated by several state-of-the-art algorithms before and after different dimension reduction methods are compared, as shown in Table 6. In row 1, the data are slide by the time window with 150 time steps, which has a total of 450 features. In row 2, 3 and 4, the sequential original data are respectively input into Autoencoder, Variational Autoencoder and RNN-based VAE, being reduced to 25 dimensions. Because the results of dimension reduction obtained from Autoencoder, Variational Autoencoder, RNN-based VAE are not sequential data, so RNN, LSTM and GRU are not used for prediction.

According to the results in Table 6, it can be found that our method achieves the best fault prediction accuracy among other methods. RNN is the model most widely used in time series data forecasting, but based on our experimental results,

**TABLE 6.** Results of accuracy evaluated using 7 classification models, while data are conducted different dimension reduction methods.

	SVM	Random Forest	XGBoost	Neural Network	RNN	LSTM	GRU
None (original data)	0.565	0.555	0.940	0.850	0.601	0.833	0.948
Autoencoder	0.461	0.554	0.914	0.821			
Variational Autoencoder	0.504	0.597	0.932	0.848			
RNN-based VAE (ours)	<b>0.792</b>	<b>0.889</b>	<b>0.966</b>	<b>0.998</b>			

**FIGURE 13.** Visual results obtained from reducing from 25 dimensions to 2 dimensions by PCA, and purple denotes the normal scenario. (a) 6 dynamic imbalance scenarios and normal scenario (b) 2 loosening scenarios and normal scenario (c) 6 mixing scenarios and normal scenario.**FIGURE 14.** Visual results of reducing 25 dimensions to 15 dimensions by PCA, then reducing 15 dimensions to 2 dimensions by LDA. The purple points denote the normal scenario. (a) 6 dynamic imbalance scenarios and normal scenario (b) 2 loosening scenarios and normal scenario (c) 6 mixing scenarios and normal scenario.

the performance of RNN is less than ideal, and even LSTM and GRU can only achieve the accuracy of 83.3% and 94.8%. This result may be attributed to the sequence length being too long, causing the information to be diluted during the propagation.

However, after dimension reduction by RNN-based VAE, a simple double-layer neural network can achieve the high accuracy of 99.8%, increased by 14.8%. Other dimension

reduction methods, such as Autoencoder and Variational Autoencoder, cannot effectively extract the features of the sequential data, only the accuracy of random forest is slightly improved, while the others all get lower accuracy. As for sequential original data, by using our method, the prediction accuracy of SVM, random forest and XGBoost can be improved by 22.7%, 33.4% and 2.6% respectively.

## D. LATENT SPACE VISUALIZATION

In order to visually represent the dimension reduction results, the data set of new features with the dimensions reduced to 25 by RNN-based VAE are further reduced to 2 dimensions for visualization by using the dimension reduction algorithm, such as PCA and LDA. Different categories (load, loosening and mixing) are compared with normal categories respectively through drawings. Fig 13 shows the visual results of reducing the data from 25 dimensions to 2 dimensions by PCA, and it can be found that the data points of normal scenarios (purple) are mixed with other fault scenarios and are not well separated. In Fig 14, PCA is firstly used to reduce 25 dimensions to 15 dimensions, and then LDA is used to reduce 15 dimensions to 2 dimensions to get the visual results. The effect is significantly better than that achieved by simply using PCA for dimension reduction. In addition that normal scenario and other fault scenarios can be clearly distinguished, the fault scenarios in the same category with different order of severity also can be distinguished. In the future, such visual results of dimension reduction can determine which fault scenario is close to, by observing the positions of new data points in the 2-dimensional space, so as to assist in determining unknown fault modes.

## V. CONCLUSION

In most of the fault detection methods at present, time domain signals are required to be transformed into frequency domain or statistics and other high-level data, highly relying on the professional knowledge such as signal processing and fault mode pattern recognition. While, in this paper, we can achieve the purpose of fault detection by input the original time domain vibration data of motor and extract the feature based on the proposed RNN-based VAE model, so that the computing cost will be reduced, and on top of that, the classification accuracy of fault detection can be well improved. In this study, we designed a motor experiment platform, 15 fault scenarios were simulated. Accelerometers were placed in 3 different positions to collect vibration signals, and there are totally 12,150,000 data. The experiment results show that the traditional classification models, such as SVM and Random Forest, and even XGBoost, RNN, LSTM and GRU cannot achieve ideal accuracy. If Autoencoder and Variational Autoencoder are used to reduce the dimension of sequential data, the classification accuracy cannot be improved effectively or even decreased, indicating that the two commonly used dimension reduction methods cannot effectively extract the features of time series data. However, after the dimensions of signals are reduced by RNN-based VAE, the prediction effects of all classification models can be significantly improved. Moreover, by cooperating with the dimension reduction in the subsequent stages, such as PCA and LDA, the visual results can be used to assist in determining the scenarios of unknown faults.

For future work, we plan to extend our experiment using the motor with larger power and implement more fault

scenarios to strengthen the stability and robustness of our model. At last, we will work with the company to obtain real data from the production line for further evaluation.

## ACKNOWLEDGMENT

The authors would like to thank the editor and anonymous reviewers for constructive comments. In memory of Dr. Wu-Chi Ho, who has always inspired us with his research passion and creative ideas.

## REFERENCES

- [1] J. Antonino-Daviu, S. Aviyente, E. G. Strangas, and M. Riera-Guasp, “Scale invariant feature extraction algorithm for the automatic diagnosis of rotor asymmetries in induction motors,” *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 100–108, Feb. 2013.
- [2] A. Arnaliz, B. Jung, A. Adgar, T. Naks, A. Tohver, T. Tommingas, and E. Levrat, “Information and communication technologies within E-maintenance,” in *E-maintenance*. London, U.K.: Springer, 2010, pp. 39–60.
- [3] M. B. Bailey and J. F. Kreider, “The design and viability of a probabilistic fault detection and diagnosis method for vapor compression cycle equipment,” Ph.D. dissertation, Dept. Civil Eng., Univ. Colorado Boulder, Boulder, CO, USA, 1999.
- [4] G. G. Barna, S. W. Butler, D. A. Sofge, and D. A. White, “Virtual sensor based monitoring and fault detection/classification system and method for semiconductor processing equipment,” U.S. Patent 5 864 773 A, Jan. 26, 1999.
- [5] Z. Chen and W. Li, “Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network,” *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1693–1702, Jul. 2017.
- [6] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” Jun. 2014, *arXiv:1406.1078*. [Online]. Available: <https://arxiv.org/abs/1406.1078>
- [7] M. C. Comstock, J. E. Braun, and R. Bernhard, *Experimental Data From Fault Detection and Diagnostic Studies on a Centrifugal Chiller*. West Lafayette, IN, USA: Purdue Univ., 1999.
- [8] X. Dai and Z. Gao, “From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis,” *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2226–2238, Nov. 2013.
- [9] W. Deng, S. Zhang, H. Zhao, and X. Yang, “A novel fault diagnosis method based on integrating empirical wavelet transform and fuzzy entropy for motor bearing,” *IEEE Access*, vol. 6, pp. 35042–35056, 2018.
- [10] J. Faiz, V. Ghorbanian, and B. M. Ebrahimi, “EMD-based analysis of industrial induction motors with broken rotor bars for identification of operating point at different supply modes,” *IEEE Trans. Ind. Informat.*, vol. 10, no. 2, pp. 957–966, May 2014.
- [11] V. Fernandez-Cavero, D. Morinigo-Sotelo, O. Duque-Perez, and J. Pons-Llinares, “A comparison of techniques for fault detection in inverter-fed induction motors in transient regime,” *IEEE Access*, vol. 5, pp. 8048–8063, 2017.
- [12] H. G. Gauch, Jr., “Noise reduction by eigenvector ordinations,” *Ecology*, vol. 63, no. 6, pp. 1643–1649, 1982.
- [13] F. A. Gers, J. Schmidhuber, and F. Cummins, “Learning to forget: Continual prediction with LSTM,” in *Proc. 9th Int. Conf. Artif. Neural Netw. (ICANN)*, Sep. 1999, pp. 850–855.
- [14] A. Glowacz and Z. Glowacz, “Diagnosis of stator faults of the single-phase induction motor using acoustic signals,” *Appl. Acoust.*, vol. 117, pp. 20–27, Feb. 2017.
- [15] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, “LSTM: A search space odyssey,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017.
- [16] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [17] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

- [18] F. Jia, Y. Lei, L. Guo, J. Lin, and S. Xing, "A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines," *Neurocomputing*, vol. 272, pp. 619–628, Jan. 2018.
- [19] S. B. Jiang, P. K. Wong, R. Guan, Y. Liang, and J. Li, "An efficient fault diagnostic method for three-phase induction motors based on incremental broad learning and non-negative matrix factorization," *IEEE Access*, vol. 7, pp. 17780–17790, 2019.
- [20] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *Proc. Int. Conf. Mach. Learn.*, Jul. 2015, pp. 2342–2350.
- [21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Dec. 2014, *arXiv:1412.6980*. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [22] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," Dec. 2013, *arXiv:1312.6114*. [Online]. Available: <https://arxiv.org/abs/1312.6114>
- [23] W. Li and C. K. Mechefske, "Detection of induction motor faults: A comparison of stator current, vibration and acoustic methods," *J. Vib. Control*, vol. 12, no. 2, pp. 165–188, 2006.
- [24] N. D. Liyagedera, A. Ratnaweera, and D. I. B. Randeniya, "Vibration signal analysis for fault detection of combustion engine using neural network," in *Proc. IEEE 8th Int. Conf. Ind. Inf. Syst.*, Dec. 2013, pp. 427–432.
- [25] A. Maas, Q. V. Le, T. M. O'Neil, O. Vinyals, P. Nguyen, and A. Y. Ng, "Recurrent neural networks for noise reduction in robust ASR," in *Proc. 13th Annu. Conf. Int. Speech Commun. Assoc.*, Jan. 2012, pp. 22–25.
- [26] J. Manyika. (2011). *Big Data: The Next Frontier for Innovation*. [Online]. Available: [http://www.mckinsey.com/Insights/MGI/Research/Technology\\_and\\_Innovation/Big\\_data\\_The\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation)
- [27] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proc. Int. Conf. Mach. Learn.*, Jun. 2013, pp. 1310–1318.
- [28] S.-Y. Shao, W.-J. Sun, R.-Q. Yan, P. Wang, and R. X. Gao, "A deep learning approach for fault diagnosis of induction motors in manufacturing," *Chin. J. Mech. Eng.*, vol. 30, no. 6, pp. 1347–1356, 2017.
- [29] R. Thirukovalluru, S. Dixit, R. K. Sevakula, N. K. Verma, and A. Salour, "Generating feature sets for fault diagnosis using denoising stacked auto-encoder," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2016, pp. 1–7.
- [30] W. T. Thomson, D. Rankin, and D. G. Dorrell, "On-line current monitoring to diagnose airgap eccentricity in large three-phase induction motors—industrial case histories verify the predictions," *IEEE Trans. Energy Convers.*, vol. 14, no. 4, pp. 1372–1378, Dec. 1999.
- [31] W. T. Thomson and M. Fenger, "Current signature analysis to detect induction motor faults," *IEEE Ind. Appl. Mag.*, vol. 7, no. 4, pp. 26–34, Jul. 2001.
- [32] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn.*, Jul. 2008, pp. 1096–1103.
- [33] Y. Wang, F. Zhang, T. Cui, and J. Zhou, "Fault diagnosis for manifold absolute pressure sensor(MAP) of diesel engine based on Elman neural network observer," *Chin. J. Mech. Eng.*, vol. 29, no. 2, pp. 386–395, 2016.
- [34] T. Wong and Z. Luo, "Recurrent auto-encoder model for large-scale industrial sensor signal analysis," in *Proc. Int. Conf. Eng. Appl. Neural Netw.*, Jul. 2018, pp. 203–216.
- [35] D. Xiao, Y. Huang, L. Zhao, C. Qin, H. Shi, and C. Liu, "Domain adaptive motor fault diagnosis using deep transfer learning," *IEEE Access*, vol. 7, pp. 80937–80949, 2019.
- [36] F. Xiao, S. Wang, X. Xu, and G. Ge, "An isolation enhanced PCA method with expert-based multivariate decoupling for sensor FDD in air-conditioning systems," *Appl. Therm. Eng.*, vol. 29, no. 4, pp. 712–722, 2009.
- [37] S. Yin, N. Hu, L. Chen, and L. Hu, "Teager energy operator and its application in the study of induction motor rotor broken bars fault," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM)*, Oct. 2015, pp. 1–5.
- [38] C. López, W. Zhong, F. Cong, and V. Hidalgo, "Bearing fault diagnosis based on ensemble empirical mode decomposition and teager energy operator," in *Proc. IEEE 13th Int. Colloq. Signal Process. Appl. (CSPA)*, Mar. 2017, pp. 55–60.



**YANG HUANG** was born in Taichung, Taiwan, in 1994. He received the B.S. degree in power mechanical engineering and the master's degree from National Tsing Hua University, in 2016 and 2018, respectively. He is currently pursuing the Ph.D. degree with the Institute of Information Management, National Chiao Tung University, Hsinchu, Taiwan. His research interests include machine learning, smart manufacturing, and recommender systems.



**CHIUN-HSUN CHEN** received the B.S. degree in mechanical engineering from National Taiwan University, in 1978, and the M.S. and Ph.D. degrees in mechanical and aerospace engineering from Case Western Reserve University at Cleveland, OH, USA, in 1983 and 1985, respectively. He was a Research Associate with the National Research Council, NASA Lewis (now Glenn) Research Center, Cleveland, OH, USA, from 1985 to 1987. After that, he was recruited as an Associate Professor by the Department of Mechanical Engineering, National Chiao Tung University, Taiwan, and promoted to a Professor, in 1990. From 1998 to 2002, he served as the department chair. From 2009 to 2015, he served as the Dean of the Engineering College. He is currently the senior vice president. In addition, he is leading a task force team with Intelligent the Autonomous Vehicles System Research Center. He is also in coordination of setting up The Robotic Institute in NCTU together with Colleges of Engineering, Electrical and Computer Engineering, and Computer Science. He has published around 80 journals and 90 conference papers. He holds three patents and a computer code (authorized by Taiwan Construction and Planning Agency). He is currently the in charge of the green campus subproject of 3DNet technologies and applications for smart campus project, whose PI is President F. Chang. Such project fund is US\$ 4 million/year awarded by the Ministry of Science and Technology and private sector for three years. His research interests include energy technologies, such as energy saving and management and biomass combustion, and fire safety engineering and combustion fundamentals. He received the National Youth Award, in 1991, and the Outstanding Engineering Professor Award from the Chinese Institute of Engineers, in 2003.



**CHI-JUI HUANG** was born in Taipei, Taiwan, in 1994. He received the B.S. degree in mechanical engineering. He is currently pursuing the M.S. degree in mechanical engineering with National Chiao Tung University, Hsinchu, Taiwan. His research interests include machine learning and thermal-fluid sciences.