Time Series Data for Equipment Reliability Analysis With Deep Learning

BAOTONG CHEN[1], **YAN LIU**[2], **CHUNHUA ZHANG**[1],
AND ZHONGREN WANG[3], **(Member, IEEE)**

[1]School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China
[2]North Automatic Control Technology Institute, Taiyuan 030006, China
[3]School of Mechanical and Automotive Engineering, Hubei University of Arts and Science, Xiangyang 441053, China

Corresponding author: Chunhua Zhang (chhzhang@scut.edu.cn)

ABSTRACT With the deep integration of cyber physical production systems in the era of Industry 4.0, smart workshop dramatically increases the amount of data collected by smart device. A key factor in achieving smart manufacturing is to use data analysis methods for evaluating the equipment reliability and for supporting the predictive maintenance of equipment. Based on these insights, this paper proposes a deep learning-based approach that uses time series data for equipment reliability analysis. First, a framework of the TensorFlow-enabled deep neural networks (DNN) model for equipment reliability analysis is presented. Secondly, using time series equipment data, an evaluation strategy of equipment reliability based on deep learning is proposed. Finally, the reliability of a cylinder, an important part of the small trolley in automobile assembly line, is evaluated in a case study. Compared with the traditional reliability analysis method such as PCA and HMM, the prediction results show a significant improvement in prediction accuracy. This work contributes to promoting artificial intelligence algorithms for realizing highly efficient manufacturing.

INDEX TERMS Reliability analysis, time series data, deep learning, smart manufacturing.

I. INTRODUCTION

With the evolution of Industry 4.0, workshop tends to become cyber-physical production system with the characteristics of real-time perception, dynamic control and information services [1], [2]. The reliability of manufacturing equipment is the basic guarantee for a stable and continuous production in the workshop. The health status of an equipment affects its utilization and production efficiency. Although modern manufacturing systems have high redundancy, unpredictable failure can still break through this redundancy protection. Discrete manufacturing industries such as automobiles, aviation, and household appliances require efficient and continuous operations. Accidental failures will delay the order completion time and cause serious economic losses. It is crucial to model the equipment degradation process by using time series data, analyze the reliability of the equipment, and evaluate the risk of its failure.

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

In an intelligent manufacturing environment, industrial sensor networks acquire data at a high frequency [3], [4]. With the rising popularity of multi-functional low-power sensor nodes and the increase of data interfaces of the intelligent equipment, manufacturing enterprises can easily get the data sets of the whole life cycle of an equipment. However, there are a few shortcomings in sensor data acquired from the equipment such as being time limited, fragmented and inconsistent, which have poor correlation with other data types, and existing random interference noise [5]. Intelligent fault prediction or reliability analysis is an open issue for predictive maintenance of the industrial equipment [6], [7]. Advanced machine learning algorithms and statistical theories are required to mine knowledge from the equipment data. The health status of an equipment should be identified, and scientific reference should be provided for reliable operation and predictive maintenance.

Model-driven condition monitoring plays an important role in understanding and tracking the degradation process during the life cycle of the manufacturing equipment [8], [9], such

as operation, standby, maintenance, partial failure, and complete failure. This paper introduces a TensorFlow-enabled deep neural networks (DNN) model for equipment reliability analysis based on time series data. The target of this model is to achieve active maintenance of the equipment rather than the conventional maintenance. The contributions of this paper are as follows:

- Considering that a shallow learning model cannot effectively characterize the complex mapping relationship between signals and health status of the equipment in the context of big data, we propose a DL-based multi-classification prediction method by using time series sensor data for short-term status monitoring.
- TensorFlow, an open source machine learning framework, is introduced in the system modeling. With the advantage of its analysis efficiency and convergence speed, the equipment status can be predicted in real-time based on the time series data.
- Based on the equipment degradation model, the reliability of the equipment in the next stage change can be evaluated by the time series prediction results. An equipment reliability strategy for early warning is established to provide scientific guidance for the predictive maintenance of the equipment.

This paper is structured as follows: Section II presents the TensorFlow-enabled deep neural network (DNN) system architecture for equipment reliability analysis. Equipment reliability evaluation strategy based on time series data is proposed in section III. Section IV conducts a case study of the cylinder of a small trolley in the automobile assembly line. Section V concludes the paper.

II. LITERATURE REVIEW

Equipment reliability analysis is mainly conducted to quantify the probability of equipment failure. Poor reliability of equipment will lead to a high probability of equipment failure. Yang *et al.* [10] proposed a simple yet effective supervised deep hash approach, which constructed binary hash codes from labeled data for large-scale image search. Makantasis *et al.* [11] proposed a deep supervised learning-based classification method that hierarchically constructs high-level features in an automated way. These references are the main motivation behind the research work presented in this paper.

Deep learning is a method for representing data and for learning data in machine learning. TensorFlow was used to integrate one-dimensional or two-dimensional convolutional neural networks (CNN) in [12]. Considering the complexity of the reliability analysis model and the objectivity of the equipment data set, we propose a TensorFlow-enabled DNN model to simulate the degradation process of the equipment. According to Zio [1], the knowledge, information, and data available for the modeling, computations, and analyses done in reliability engineering are rapidly increasing. The degradation model for health management of equipment is increasingly made of heterogeneous and highly interconnected elements. Lei *et al.* [13] proposed an intelligent fault diagnosis method using unsupervised feature learning for mechanical big data. The proposed unsupervised two-layer neural network achieved high diagnosis accuracies for the motor bearing dataset, compared to existing methods. Gal and Ghahramani [14] used dropout as a Bayesian approximation to estimate uncertainty with a DNN model. Compared with traditional reliability analysis methods, machine learning methods (e.g., DNN) have been applied widely with the features of parallel processing, fault tolerance, self-learning and self-monitoring.

Time series data are measurement sequences that describe the behavior of time-varying systems or equipment. The application of time series-based prediction methods in the fields of medicine, aerospace, finance, commerce, meteorology and entertainment were introduced in [15], [16]. Khodayar *et al.* [17] developed a DNN structure based on stacked autoencoder and stacked denoising autoencoder for ultra-short-term and short-term wind speed predictions. The experiment results showed that the DL model was feasible for short-term predictions. Deb *et al.* [18] summarized state of the art machine learning methods for predicting time series-based energy consumption. The authors concluded that a hybrid model comprised of two or more prediction techniques was more effective for time series prediction. Considering the randomness of equipment deterioration, in this paper, we evaluate the risk of equipment failure through short-term and medium-term predictions. The motivation behind the work presents in this paper is to discover the critical time node and support active maintenance when the running status of equipment changes.

III. SYSTEM ARCHITECTURE

In this section, the system architecture of equipment reliability analysis based on time series data is presented in Fig. 1, which shows the system architecture from three aspects, namely, data collection, model training, and model serving. The data collection is the base layer to support the model training. Model training is the foundation of the model serving layer.

A. DATA COLLECTION

Data collection is an important step in equipment reliability analysis, which includes bottom data collection and data preprocessing. Industrial IoT is characterized by heterogeneity, diversity and dynamic. Cognitive technology-enabled IIoT helps to realize semantic representation and sensor data correlation [19]. MQTT, A TCP-based protocol characterized by its concise and lightweight features, is suitable for low broadband, high latency, and unstable network environments [20]. InfluxDB is an open source distributed database of time series events and metrics. It is popular for storing real-time IoT data. Kafka Streams, a Java library relying on Apache Kafka, is used to build distributed flow processing programs. In this scenario, Kafka Streams is deployed as a distributed flow platform to connect MQTT and InfluxDB. In the database,
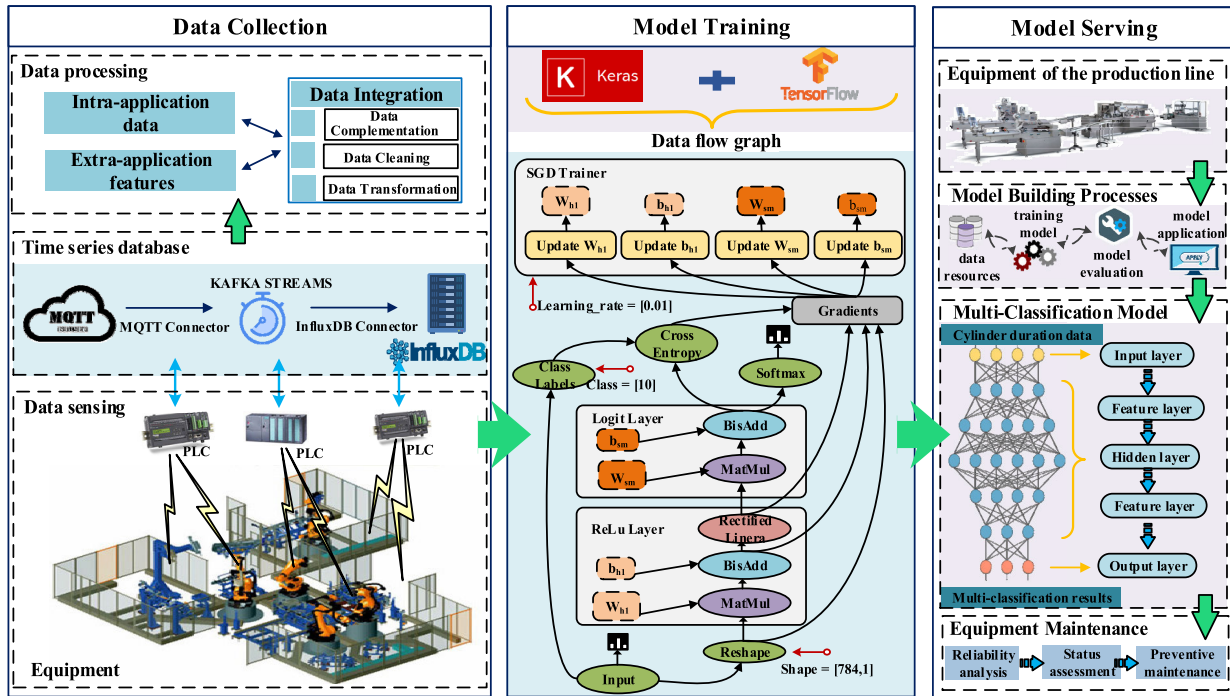
FIGURE 1. System architecture of reliability analysis for equipment time series data.

unexpected data loss or the noise in the sensor-read data from equipment should be pre-processed with combining internal data with external features.

And the data preprocessing includes data cleaning, data completion, and data transformation. Data cleaning is used to delete duplicate information and correct existing errors. Data completion completes missing data for ensuring consistency. Data transformation enables the collected data to meet statistics calculation or analysis need.

B. MODEL TRAINING
TensorFlow is an open source software library for numerical computation using data flow graphs. In TensorFlow, computations are carried out using tensors. A tensor is defined as a vector or n-dimensional matrix representing the data. Human-computer interactive mode of computation is realized by building computational structural graph, which reduces the difficulty in the development of DL models. The data flow graphs describe mathematical computations with directed graphs of nodes and edges. When the input tensors are ready, the nodes will be assigned to processing units to make asynchronous parallel computing. Computation graphs can also be compiled and be optimized to separate the definition of graphs from the actual computing process. We use Keras to build the related models, which is an integrated tool for building the neural network model based on backends including TensorFlow, Theano and CNTK. Keras includes highly modular and easily expandable neural networks API, and supports most kinds of ANN models.

C. MODEL SERVING
TensorFlow provides SavedModel mechanism to export. The trained models are saved as external files for external services

or subsequent applications. Using real-time state data, the equipment status is evaluated with a multi-classification prediction model online. With the prediction performance status of the equipment acquired from time series data, we can track the degradation process of the equipment and provide condition-based maintenance according to its performance degradation.

Intelligent manufacturing puts high emphasis on equipment collaboration and efficiency. Unexpected downtime caused by equipment failure will cause serious economic loss to manufacturing enterprises. Therefore, it is crucial to adapt a reasonable condition-based maintenance strategy. The interval between abnormal events of the equipment is predicted to provide scientific guidance for on-demanded maintenance, and the prediction result is an important reference for equipment health management.

In this paper, we conduct research on the cylinder of a small trolley in automobile production line. The duration of the cylinder operation such as lifting or falling is closely related to its performance. Considering that the operation duration of an aging cylinder will gradually increase, the operation duration can be used to monitor the process of performance decay of the cylinder. Based on the operation duration of the cylinder, we can predict the operation status using the DNN model with time series data. It is helpful in the maintenance of the cylinder before a failure occurs.

IV. EQUIPMENT RELIABILITY ANALYSIS MECHANISM
This section establishes the equipment reliability analysis mechanism based on equipment time series data. First, we build a TensorFlow-enabled DNN model for multi-classification in equipment status monitoring. Then,
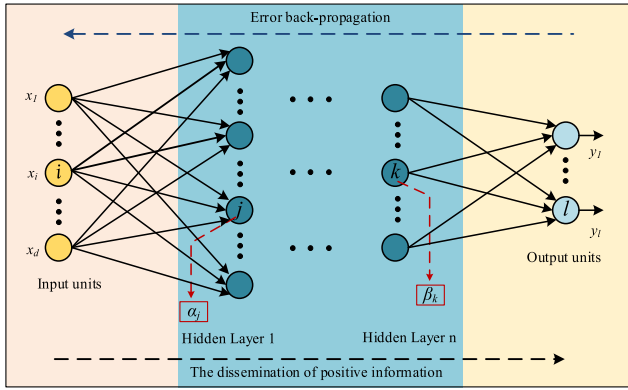
FIGURE 2. Multilayer feedforward neural networks.

real-time series data are used to monitor the equipment status by using database interface technique. Finally, based on the predicted equipment status distributed in time series, we propose the reliability evaluation strategy for equipment.

A. MULTI-CLASSIFICATION MODEL OF EQUIPMENT STATE
1) DEEP NEURAL NETWORK
Deep neural network is a representation learning method in machine learning, which is based on a multi-layer network. The learning ability of multi-layer networks is stronger than a single layer or several layers perceptron. Multi-layer feedforward neural network is a typical example [21] as shown in Fig. 2. The feedforward network is a static non-linear mapping. By compound mapping of simple non-linear processing units, the multi-layer feedforward network can approximate continuous functions with a high accuracy, and has strong ability for classification and pattern recognition.

A training set used for multi-layer feedforward neural network is shown in (1).

$$D = \{(x_1, x_2, x_3, \ldots, x_i, \ldots)(y_1, y_2, y_3, \ldots, y_i, \ldots)\},$$
$$x_i \in R^d, y_i \in R^l \quad (1)$$

where d denotes the number of input neurons and l denotes the number of output neurons.

Let the connection weight be υ_{ij} between the ith neuron in the input layer and the jth neuron in the hidden layer. Then the input received by the jth neuron in the hidden layer is given as follows:

$$\alpha_j = \sum_{i=1}^{d} \upsilon_{ih} x_i. \quad (2)$$

By denoting the connection weight between the kth neuron in the hidden layer and the lth neuron in the output layer as ω_{kl}, the input received at the output layer is (3).

$$\beta_l = \sum_{h=1}^{q} \omega_{kl} b_h. \quad (3)$$

The total number of neurons in the hidden layers is given by q. The learning process of the neural networks adjusts the weights such as υ_{ij} and ω_{kl}, and the threshold θ_i by using the training data. For the q-layer feedforward neural network,

the transform relation between the input and output can be denoted by (4).

$$s_i^{(q)} = \sum_{j=0}^{n_q-1} \omega_{ij}^{(q)} x_j^{(q-1)}, (x_0^{(q-1)} = \theta_i^{(q)}, \omega_{i0}^{(q-1)} = -1)$$
$$i = (1, 2, \ldots, n_q); \quad j = (1, 2, \ldots, n_{q-1});$$
$$q = (1, 2, \ldots, Q) \quad (4)$$

2) KEY TECHNOLOGIES
Based on the time series data in InfluxDB, we conduct the pre-processing process including regularization and feature scaling for the incomplete observed data. The DNN model is built to track the state change of the equipment with Keras. We present the key techniques used in the model as follows.

a: ONE-HOT ENCODING
In the process of model training, character features represented by numerical values can influence the prediction model. To solve this problem, one-hot encoding n-bit state register to encode n states. This encoding process carries out the binarization of the features. Thus converts them into a form that could be provided to the machine learning algorithms. When it comes to deep supervised model, this method solves the problem that the classifier cannot handle the attribute information, and it plays the role in expanding the data labels or the features of the model. As an example, consider a natural state code *[000, 001, 010, 011, 100, 101]*. After one-hot encoding, the code becomes *[000001, 000010, 000100, 001000, 010000, 100000]*. Each indexed data reflects three running states, namely, *[good, fair, poor]*. The one-hot coding for these states is given by *[001, 010, 100]*.

b: ACTIVATION FUNCTION ReLU
Activation function introduces nonlinearity into the neurons and helps the neural network to approximate a non-linear function. The choice of activation function has an important influence on the quality of the neural network. Rectified Linear Unit (ReLU) is a commonly used activation function in ANNs, which is expressed by $f(x) = \max(0, x)$.

Other common activation functions are sigmoid and tanh. The gradients of sigmoid and tanh are very close to zero in the saturation zone. This can easily lead to the problem of vanishing gradient, and slow down the convergence rate. On the other hand, the gradient of ReLU is constant in most cases, which helps to solve the convergence problem in a deep network.

c: SOFTMAX FUNCTION
Softmax is used to map the output of multiple neurons into probability in a network classifier. It calculates the probability of all the possible classes or features considered in a machine learning model. SoftMax is used to map the output value of multiple neurons in the sum of (0,1) as 1. The probability value for the ith class is shown in (5). When there is a discrete

variable with n possible values, the softmax function in the output of a classifier represents n different probabilities in deep supervised learning. When the network gives the output, the maximum value (probability maximum) is selected as the prediction value. Therefore, SoftMax function may be more suitable for a multi-classification problem.

$$S_i = \frac{e^i}{\sum_j e^j}. \tag{5}$$

d: CROSS-ENTROPY LOSS FUNCTION

Cross entropy loss is a measure of the precision of the model in the training process. It is one of the most widely used loss functions in DL, and is mostly applied to multi-classification problems. The functional form of cross-entropy loss is as shown in (6).

$$Loss = H_{y'}(y) = -\sum_i y'_i \log(y_i). \tag{6}$$

The predicted results are given by y_i, and y'_i refers to the ground truth. The cross-entropy loss function reflects the similarity of y_i and y'_i.

3) REAL-TIME MONITORING OF EQUIPMENT STATUS

Equipment degradation is often a slow but gradual process. It is impossible to monitor the equipment at fixed intervals or only when a certain threshold is out of bounds. The real-time monitoring model based on time series data provides important support for equipment health management. The real-time monitoring of equipment status is divided into two stages: (1) First, the TensorFlow-enabled DNN model is trained offline using the equipment history data stored in InfluxDB; and (2) based on the generalization ability of the DL model, online prediction is conducted using time series data obtained by the InfluxDB client.

The equipment status is labeled into three categories. The predicted results of the three types are given by $Y = \langle good, fair, fault \rangle$. The historical data set used for training is referred to by $D_{His} = \{x_1, x_2, x_3, \ldots, x_n, y\}$. For the introduction of equipment reliability evaluation mechanism, the prediction results are quantified as $Q = \{-1, 0, 1\}$. Real-time data $D_{Live} = \{x_1, x_2, x_3, \ldots, x_n\}$ is imported into the trained deep supervised learning model, which is used to monitor the equipment status in real-time. The training process is shown as algorithm 1.

B. EVALUATION STRATEGY FOR EQUIPMENT RELIABILITY

As it is well known, aging or wearing out of components in an equipment leads to an increase in the equipment's action time. This degradation is gradual when the equipment is in use. We propose an evaluation strategy for equipment reliability based on time series data. A DL model is used to evaluate the running state in real time with live data. The equipment reliability is analyzed using the statistic of the sequences running status during continuous functioning. The output data of the equipment is random and obeys a certain distribution.

For instance, the action time of the cylinder including rising, falling and resetting is normally distributed. Therefore, a single moment of abnormal state cannot be used to judge operation reliability of the equipment. The uniformity of equipment working hours is introduced in (7).

Algorithm 1 Real-Time Condition Monitoring Model for Equipment

Input_1: D_{Live} denotes live data set of equipment from InfluxDB client
Input_2: D_{His} denotes historical data set of equipment from InfluxDB
Output_1: Neural network parameters for condition monitoring model
Output_2: Multi-classification results for equipment operation states

---
1 Begin
2   Online data prediction for equipment operation status
3   Initialization
4   Input D_{Live}
5   Offline data fusion model for equipment operation status
6   Initialization
7   Input $D_{His} \leftarrow \{D_{Train} + D_{Test}\}$
8   Step 1: Data preprocessing
9   One-hot encoding for $Q = \{-1, 0, 1\}$
10   Standard Scaler
       $\bar{D}_{His} = \{x_1, x_2, x_3, \ldots, x_n\}$
11   Step 2: Define the DNN model
12   Build fully connected layer
13   Choose *ReLU* for activation function, *cross entropy* for Loss, and *simple_adam* for optimizer
14   Step 3: Training the model
15   Set the batch size N_{size} and the iterations
16   Fitting the training data set D_{Train}
17   Step 4: Evaluation the model
18   Evaluate the loss and accuracy of D_{Test}
19   If (obtain the satisfied evaluation results):
20   Return the data fusion model
21   Else:
22   return to the Step 1
23   End if
24   end
25   Return Multi-classification results $q_i \in Q$
26 end

In (7), q_i denotes the predictive value of DNN model. N_o denotes the number of forecasting instants of DNN model in a given time interval T. φ_T denotes the uniformity of the equipment in a given time interval T.

$$\varphi_T = \frac{1}{N_o} \sum_{i=1}^{N_o} q_i. \tag{7}$$

The evaluation period of the equipment working time is divided into different time intervals, and the time window is 1 hour. Let Φ_T be defined as the collection of long working intervals, $\Phi_T^{(n)} = \{D_1, D_2, D_3, \ldots, D_n\}$. To ensure the equipment reliability, the minimum reliability measure is selected in the same prediction sequence as $\min(\varphi_T) \rightarrow D_\tau$. As shown in (8). τ denotes the input value of the equipment operation duration. D_τ denotes the maximum value of event prediction in time series. M denotes the number of input characteristics for collected operation duration.

$$D_\tau = \frac{1}{N_o M} \sum_{i=1}^{N_o} \sum_{j=1}^{M} \tau_{ij}. \tag{8}$$

The data acquired from the equipment are discrete, however, the decay of the running state of the equipment is continuous. Therefore, a single time interval cannot accurately reflect the degree of the decay. Liu *et al.* [22] proposed weight coefficients $c_{i,t}$ of the probability estimation in the past time, which satisfy (9) and can be calculated by (10). However, it is difficult to use the equipment status from past time windows for equipment evaluation, especially when those past intervals are far from the current time.

$$c_{i,t} \geq c_{i,t-1} \geq 0, \quad t = 1, 2, \ldots, n_i - 1 \tag{9}$$

$$c_{i,t} = c_{i,1} + (t-1)\frac{2 - 2c_{i,1}n_i}{(n_i - 1)n_i}, \quad t = 1, \ldots, n_i \tag{10}$$

The weight coefficient is adjusted according to the difference between the previous time and the current time interval. We evaluate the current status of equipment based on the previous work time in (11). As a result, the influence of long-time interval data is reduced on the running state transformation of equipment. In (11), T_N denotes the current forecast time of working time. T_s denotes the start time of the working time. T_i denotes the i th time interval from the start time. D_T denotes the operation duration for the condition assessment in time T.

$$D_T = \sum_{i=1}^{H_o} \left(\exp\left(-\frac{T_N - T_i}{T_N - T_s}\right) \cdot D_\tau^{(i)}\right). \tag{11}$$

We have retrieved the maintenance records of the equipment and obtained the working hours when the equipment should be maintained by using systematic sampling method. The threshold D_Δ is used to decide whether the equipment is reliable or not. We calculate the real-time maintenance value $RTM(t)$ in (12).

$$RTM(t) = 1 - \exp(D_T - D_\Delta). \tag{12}$$

The calculated RTM value is then compared with the maintenance boundary of the time interval, as shown in (13). PTM^{in} denotes the lower limit of the preventive maintenance value, and $PTM^{\lim}$ denotes the upper limit of the preventive maintenance value. $R(t)$ denotes the best position of all swarm particles. The prediction results of real-time interval are divided into three levels, namely, normal, alarm, and fault. If the current maintenance value is between predefined lower

---

Algorithm 2 DNN-Based Cylinder Reliability Analysis and Evaluation

Input: $X_t^T = [\tau_{\sigma 1}, \tau_{\sigma 2}, \tau_{\sigma 3}, \tau_{\sigma 4}]$ in a given time interval from InfluxDB

Output: $Y_t^T = R(t)$ evaluation results for cylinder operation states

1 Begin
2   Initialization   *// on a continuous working day*
3   for $i \leftarrow 1$ to H_o
4        $Q_{sum} \leftarrow Q_{sum} = 0$
5   for $j \leftarrow 1$ to N_o
6           $q_j \leftarrow DNN(X_t^T = [\tau_{\sigma 1}, \tau_{\sigma 2}, \tau_{\sigma 3}, \tau_{\sigma 4}], q_\sigma)$
7           $Q_{sum} \leftarrow Q_{sum} + q_j$
8   end for
9        $\varphi_i \leftarrow Q_{sum} / N_o$ *// calculate the uniformity*
10   end for
11   for $i \leftarrow 1$ to H_o *// get the minimum uniformity*
12       $\varphi_T, \eta \leftarrow (\varphi_T = \varphi_0, \eta = 0)$
13   if $\varphi_T < \varphi_i$:
14          $\varphi_T = \varphi_i$
15          $\eta = i$
16   end if
17   end for
18   while $(i == \eta)$
19       $D_{sum} = 0$
20   for $j \leftarrow 1$ to N_o
21   for $l \leftarrow 1$ to 4
22             $D_\tau += \tau_{j,l}$
23   end for
24   end for
25       $D_\tau = D_{sum}/N_o$
26   end while
27   for $i \leftarrow 1$ to H_o
28       $D_T += \exp(-(T_N - T_i)/(T_N - T_S)) \cdot D_\tau^{(i)}$
29   end for
30    $RTM(t) = 1 - \exp(D_T - D_\Delta)$
31    $R(t) \propto P(PTM^{in}, PTM^{\lim})$ *// get the range of R(t)*
32 end

---

and upper maintenance limits, the system will be at alarm level. If it is below the lower maintenance boundary, it is at normal level. Failure occurs when the maintenance value is above the upper limit of maintenance threshold.

$$R(t) = P(PTM^{in} \leq RTM(t) \leq PTM^{\lim}). \tag{13}$$

The core algorithm is shown in algorithm 2.

V. CASE STUDY

In this section, we evaluate the reliability of a cylinder based on the proposed technique. The cylinder is an important part of a small trolley in the automobile assembly line, as shown in Fig. 3. The cylinder is marked by the red arrow. Reliable operation of the cylinder guarantees that the trolley functions

FIGURE 3. The data collecting site with a cylinder on the trolley.

TABLE 1. Details of data samples in verification.

	Sample number of the resetting process of the cylinder	Sample number of the falling process of the cylinder	Sample number of the rising process of the cylinder	Proportion of training set in the whole data set
Good	2976	3107	3270	0.68
Fair	1563	1543	1225	0.72
Poor	213	102	257	0.65

in a stable state, which directly affects the stability of long-term operation of the whole automobile production line.

During the long-term operations, insufficient lubrication can lead to a wear off in the cylinder, and the ageing rubber seals lead to leaks in the cylinder. The action time of the cylinder, including rising, falling and resetting is an important reference to indicate an aged cylinder.

A. TENSORFLOW-ENABLED DNN MODEL

As shown in Algorithm 1, the proposed DNN model consists of three layers: input, hidden and output layers, and a SoftMax function. The input layer consists of four neurons, which correspond to four features in the InfluxDB data set. We present the part of the labeled sample data through pre-defined threshold from the InfluxDB as shown in Fig. 4, which includes 1000 training data. We choose ReLU as the activation function, cross entropy as the loss function, and Adam as the iterative optimizer. The connection weights and biases of each layer are initialized randomly. In table 1.

We set the maximum number of epochs as 20 and recorded the accuracy and loss of training set, and the accuracy and loss of test set in each epoch.

As shown in Figs. 5(a) and 5(b), the accuracy of the training set improved from 0.607 to 0.9929, and the loss reduced significantly from 0.9406 to 0.0666. The results with test set given in Figs. 5(c) and 5(d) show similar trends as well. The increase of accuracy shows the effectiveness of the training process, and the decreases of loss underlines the DL model nearing convergence.
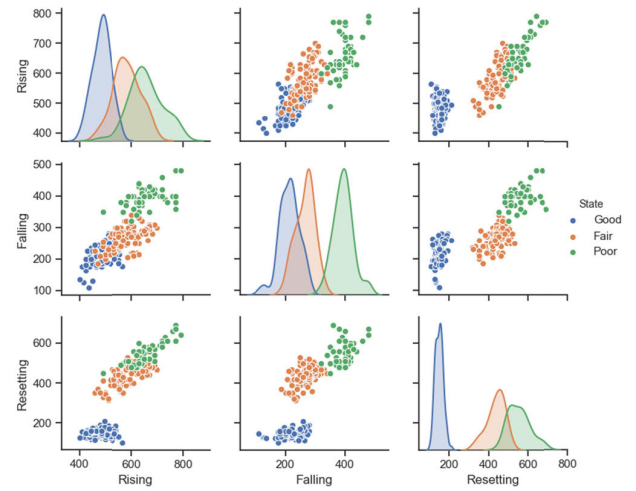

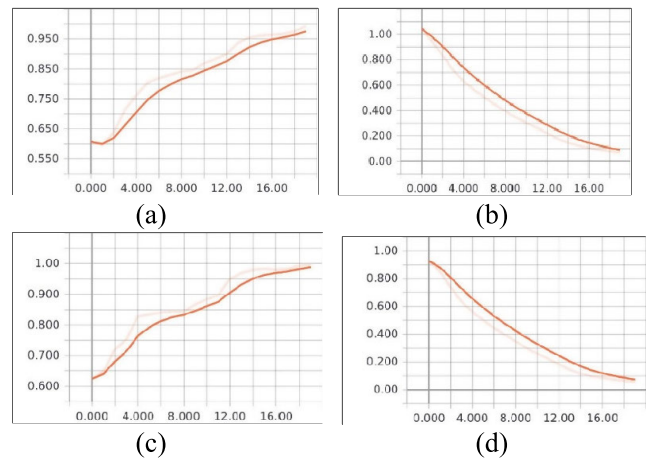**FIGURE 4.** Part of the labeled sample data from the InfluxDB.

FIGURE 5. Training results: (a) training set accuracy; (b) training set loss; (c) test set accuracy; and (d) test set loss.

B. MODEL RESULTS

The degradation process of the equipment follows a monotonous attenuation trend. We conducted the equipment reliability analysis based on the action time of the cylinder. In Fig. 6, the blue curve shows that the reliability value of cylinder operation decreases with the running time of the cylinder. For instance, the more total time the cylinder costs in service, the lower its reliability value becomes. The cylinder has a long total service life, and its performance decay is a slow process. We monitored the condition of the cylinder over two months to verify the proposed model. At the end of two months, the cylinder was almost non-functional. The yellow curve shows that the operation time of the cylinder increased with the degradation of the cylinder.

The maintenance boundary is an important parameter for characterizing the state transition the functional status to the non-functional status. We select the intersection of the curves representing the cylinder operation time and its reliability as the maintenance boundary, as shown in Fig. 6. The relevant formulas are given in (8)-(12).
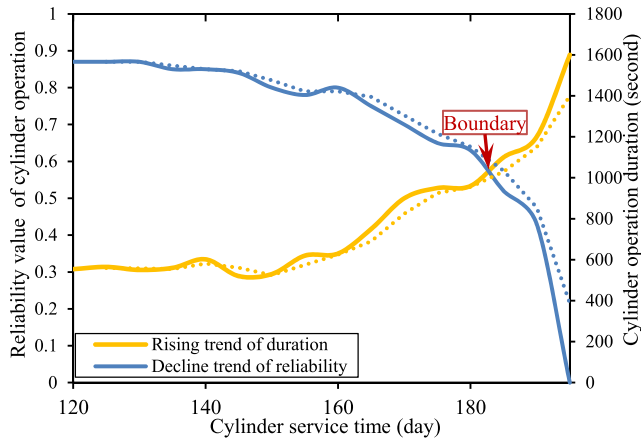
FIGURE 6. The reliability threshold of cylinder related to its operation duration.

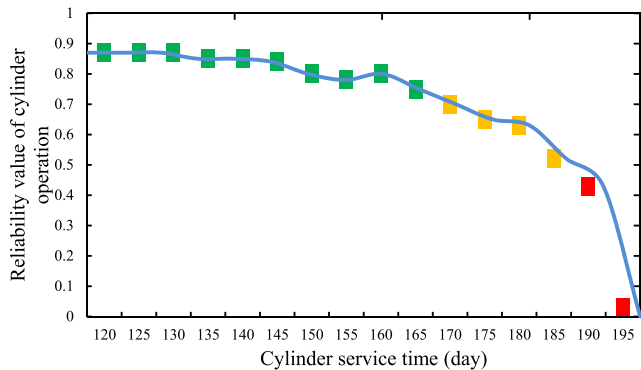**FIGURE 8.** Prediction results of the five trials using the PCA, HMM and the proposed method.

FIGURE 7. The discrete predicted results for the decay process of the cylinder.

In Fig. 7, the colored boxes represent the three conditions of cylinder operation, namely, normal, warning, and fault. Since the action time of the cylinder including rising, falling and resetting is normally distributed. Statistics over a certain time can eliminate the effect of abnormal fluctuations. Based on the running time of the cylinder, which is at least 10 hours a day, the working state of the cylinder is predicted using the time series data. The predicted results approximate the decay process of the cylinder. As shown in Fig. 7, cylinder failure occurred nearly 20 days after the decay started. Once a scientific active maintenance plan is made based on the predicted cylinder states, sudden shutdowns of the automobile production line caused by cylinder failure can be avoided. This mitigates the production capacity decline caused by failure of the components in the production line.

C. MODELS EVALUATION
We identify the hidden state changes of the equipment by monitoring the multi-temporal observation sequence. Based on these identifications, we distinguish the reliability level of the equipment operation, such as failure, maintenance, replacement of parts or switching operation from the state transition process. In an intelligent manufacturing environment, the state transition of equipment failure is usually considered as a memoryless random process. Under specific
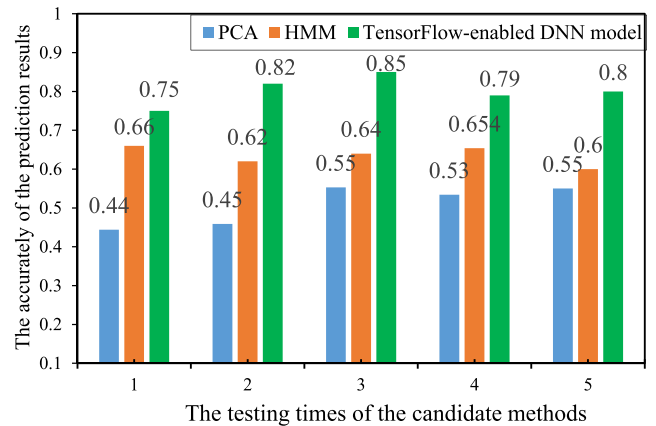
conditions, the stochastic process describing system evolution can be described by a Markov process. The most well-known equipment reliability analysis method is based on the hidden Markov model [1], [23]. However, traditional reliability analysis methods are often limited by poor data sets, and there is less knowledge available to evaluate the equipment reliability.

The shallow learning models cannot effectively reflect the complex mapping relationships between signals and the equipment health status in big data. Similarly, simple linear fitting method of expert systems cannot accurately predict the dynamic performance degradation of equipment.

We used principal component analysis (PCA), hidden Markov model (HMM) and deep learning to map the relationship between the action time and reliability of the cylinder. PCA and HMM have been widely applied to failure prediction. They are also the current mainstream classification algorithm behind the fault prediction application. In the verification, we took them as the standard approach as contrasts for the proposed methods. The three algorithms are tested in order to evaluated the accuracy of the status of the cylinder. Five sets of 10-hour statistical data are given to predict the result as shown in Fig. 8. Traditional preventive maintenance methods focus on the dynamic working state and the change of parameters in complex systems. There is a critical need for a better solution to multi-source heterogeneous data processing. The development of intelligent equipment had led to high speed, high precision and high efficiency manufacturing. This leads to the factors influencing the reliability of equipment operation very complex.

Currently there are sufficient data sets available for equipment reliability evaluation. Although the multi-source sensing data collected by large-scale industrial processes are increasing rapidly, it is still difficult to model the complex performance degradation process. In the context of industrial big data, data-driven algorithms of equipment reliability prediction are of great significance in both academic and engineering environments. The deep prediction model for equipment reliability analysis can realize a composite

evaluation method with multi-type data. It can get rid of the dependence on signal processing technology and diagnostic experience such as data standardization and data normalization. For instance, DL can used to build a multi-sensor data model based on current, temperature, vibration, etc.

VI. CONCLUSION

With the acceleration of industrial 4.0, equipment data collected in the industrial field have grown at unprecedented rates. Data bring opportunities for innovative industrial applications. However, there are also challenges toward new theories and the related optimization algorithms for data processing. As a new research achievement in the field of pattern recognition and machine learning, deep learning theory is seldom applied to the reliability analysis of equipment.

In this paper, we used deep learning theory to achieve the reliability analysis with a deep neural network model, which explored the deployment of advanced machine learning methods for the preventive maintenance of industrial equipment. The main work can be concluded in bullet points: (1) we proposed a TensorFlow-enabled deep neural networks model using time series data for equipment reliability analysis. (2) we used Keras to rapidly deploy the DNN framework and conducted the reliability evaluation of a cylinder with the industrial database InfluxDB. and (3) We validated the performance of our proposed method based on measured data and further showed its superior performance compared to the related algorithms. Considering that the present data type only contains the cylinder action time, in future, recurrent neural networks will be more suitable for constructing the life cycle equipment model by using multi-type time series data. The new deep prediction model will be further applied to enhance the data analytics in industrial 4.0.

REFERENCES

[1] E. Zio, "Some challenges and opportunities in reliability engineering," *IEEE Trans. Rel.*, vol. 65, no. 4, pp. 1769–1782, Dec. 2016.

[2] B. Chen, J. Wan, L. Shu, P. Li, M. Mukherjee, and B. Yin, "Smart factory of industry 4.0: Key technologies, application case, and challenges," *IEEE Access*, vol. 6, pp. 6505–6519, 2018.

[3] H.-N. Dai, H. Wang, G. Xu, J. Wan, and M. Imran, "Big data analytics for manufacturing Internet of Things: Opportunities, challenges and enabling technologies," *Enterprise Inf. Syst.*, pp. 1–25, Jun. 2019. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/17517575.2019.1633689?src=recsys, doi: 10.1080/17517575.2019.1633689.

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

[5] Y. Lei, N. Li, S. Gontarz, J. Lin, S. Radkowski, and J. Dybala, "A model-based method for remaining useful life prediction of machinery," *IEEE Trans. Rel.*, vol. 65, no. 3, pp. 1314–1326, Sep. 2016.

[6] Y. Xu, Y. Sun, J. Wan, X. Liu, and Z. Song, "Industrial big data for fault diagnosis: Taxonomy, review, and applications," *IEEE Access*, vol. 5, pp. 17368–17380, 2017.

[7] J. Liu, J. Wan, D. Jia, B. Zeng, D. Li, C.-H. Hsu, and H. Chen, "High-efficiency urban traffic management in context-aware computing and 5G communication," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 34–40, Jan. 2017.

[8] M. Xia, T. Li, T. Shu, J. Wan, C. W. de Silva, and Z. Wang, "A two-stage approach for the remaining useful life prediction of bearings using deep neural networks," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3703–3711, Jun. 2019.

[9] J. Xiong, Q. Zhang, J. Wan, L. Liang, P. Cheng, and Q. Liang, "Data fusion method based on mutual dimensionless," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 2, pp. 506–517, Apr. 2018.

[10] H.-F. Yang, K. Lin, and C.-S. Chen, "Supervised learning of semantics-preserving hash via deep convolutional neural networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 2, pp. 437–451, Feb. 2018.

[11] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 4959–4962.

[12] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 778–782, May 2017.

[13] 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.

[14] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1050–1059.

[15] S. Aminikhanghahi and D. J. Cook, "A survey of methods for time series change point detection," *Knowl. Inf. Syst.*, vol. 51, no. 2, pp. 339–367, May 2017.

[16] J. Liu, J. Wan, B. Zeng, Q. Wang, H. Song, and M. Qiu, "A scalable and quick-response software defined vehicular network assisted by mobile edge computing," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 94–100, Jul. 2017.

[17] M. Khodayar, O. Kaynak, and M. E. Khodayar, "Rough deep neural architecture for short-term wind speed forecasting," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2770–2779, Dec. 2017.

[18] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 74, pp. 902–924, Jul. 2017.

[19] B. Chen, J. Wan, Y. Lan, M. Imran, D. Li, and N. Guizani, "Improving cognitive ability of edge intelligent IIoT through machine learning," *IEEE Netw.*, vol. 33, no. 5, pp. 61–67, Sep. 2019.

[20] A. Stanford-Clark and H. L. Truong, "MQTT for sensor networks (MQTT-SN) protocol specification," *Int. Bus. Mach. (IBM) Corp. Version*, vol. 1, pp. 1–2, Nov. 2013.

[21] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[22] K. Liu and S. Huang, "Integration of data fusion methodology and degradation modeling process to improve prognostics," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 1, pp. 344–354, Jan. 2016.

[23] X.-Y. Li, J.-P. Wu, L. Liu, M.-L. Wen, and R. Kang, "Modeling accelerated degradation data based on the uncertain process," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 8, pp. 1532–1542, Aug. 2019.
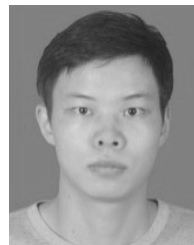
BAOTONG CHEN is currently pursuing the Ph.D. degree from the School of Mechanical and Automotive Engineering, South China University of Technology (SCUT). His research interests include data-driven preventive maintenance for smart equipment, ML algorithms in reliability analysis, edge computing in the IIoT, and artificial intelligence applications in industry.

YAN LIU received the master's degree from the China Research and Development Academy of Machinery Equipment of System Engineering. She is currently working with the North Automatic Control Technology Institute. Her research interests include networked control systems and deep learning.
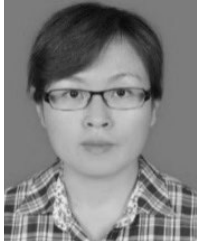
CHUNHUA ZHANG is currently an Associate Professor with the School of Mechanical and Automotive Engineering, South China University of Technology, China. She directed three research projects, including the Applied Science and Technology Research and Development Program of Guangdong Province. She has authored or coauthored more over 30 scientific papers. Her research interests include networked control systems, industrial big data, and cyber-physical systems. She was a recipient of the First Prize for Science and Technology Development of Guangdong Province, in 2009.

ZHONGREN WANG (Member, IEEE) was born in 1974. He received the B.E. degree from the Beijing University of Chemical Technology, China, in 1997, the M.E. degree from the South China University of Technology, Guangzhou, in 2003, and the Ph.D. degree in mechanical engineering from the South China University of Technology, in 2009. Since 2003, he has been with the School of Mechanical and Automotive Engineering, Hubei University of Arts and Science. He is currently a Professor and the Leader of the Laboratory of Intelligent Manufacturing and Machine Vision. Over the last ten years, he has published over 50 journal articles and conference papers. His current research interests include smart factory, machine vision, and image processing. He is a Senior Member of the Chinese Mechanical Engineering Society.

● ● ●