IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 26, 2020, accepted June 14, 2020, date of publication June 19, 2020, date of current version June 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3003723

A Product Quality Monitor Model With the Digital Twin Model and the Stacked Auto Encoder

SHIZHENG ZHANG, CUNFENG KANG, ZHIFENG LIU, JUAN WU, AND CHUNMIN MA
College of Mechanical Engineering and Applied Electronics Technology, Beijing University of Technology, Beijing 100124, China

Corresponding author: Zhifeng Liu (lzf@bjut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51975019.

ABSTRACT With the development of intelligent manufacturing and computer science, the system of equipment in the workshop has become more and more complex. In the intricate environment, the state of device changes constantly, which could affect the accuracy of methods since they cannot adapt the changing context. Recently, Digital Twin (DT) has received great focus among academic world and industrial world, which provides a new normal form for solving problems. In this paper, the structure of DT is discussed and a DT and Stacked Auto Encoder (SAE) Based Model is proposed to monitor the product quality. Based on the classical structure of DT, the digital model of DT is further divided into two parts, a task-achieved part and a self-update part. The former that comprises an encoder network that is a part of SAE and an Artificial Neural Network (ANN)-based classifier could check whether products are qualified. And a decoder network, another part of SAE, and a parameters-update rule make up the self-update part that could detect the accuracy of the task-achieved part and retrain the neural networks as the accuracy is poor. Furthermore, a new loss function is put forward as a training criterion in order to magnify the tiny difference between input data and result. In order to emulate the changing environment, the experimental data are collected at two different points in time. The data are then input to the proposed model and two other traditional methods to test the ability of accuracy and the adaptability for changing context. The comparisons show that the proposed method has got improvements, especially in where the effect of working environment is significant.

INDEX TERMS Digital twin, stacked auto encoder, parameters-update rule, product quality monitor.

I. INTRODUCTION

As the process of globalization continues to accelerate and the concept of intelligent manufacturing increases, the business environment pushes the manufacturing industry to improve its product quality [1]. The product quality is an important aspect for ensuring the productivity and economy of production [2] such that a quality monitor is crucial in the machining process to guarantee the yield. Conventionally, off-line quality detection methods, like nondestructive testing method, manual inspiration and so on, have been used to control the distribution of product quality. With the increasing degree of automation in industry, online testing has been studied for achieving further productivity.

Moreover, the development of data-acquisition systems, information technology (IT), and network technologies has brought out a new normal form to industrial [4], [5] into the era of big data. A large number of historical data provide the foundation for data-driven approaches, also known as knowledge-based methods. Bayesian networks, Principal Component Analysis (PCA) [7]–[9], Artificial Neural Network (ANN) [10], Extreme Learning Machine (ELM) [11], Support Vector Machine (SVM) [12] and so on, have improved the accuracy and efficiency of product inspection.

Whereas, an assumption that the probability distribution of training data and that of actual data is similar is convention when data-driven methods are used for tasks. But factors including the processing environment, the state of the workpiece, and etc. easily lead to differences between the training data and the actual data collected. In other words, the accuracy of the data-driven model would not be maintained for a long term, because of the factors.

The Digital Twin (DT) methods, recently emerged and developed rapidly, has provided a new idea for solving the challenges. The DT is an integrated multi-physics, multi-scale, multi-disciplinary attribute with real-time synchronization, faithful mapping, high-fidelity, and the ability to

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

implement the technical means of interaction and integration between the world and the information world [13]. The DT promotes the digitalization in industry and provides a new path to solve problems in data-driven methods [14].

With the help of features of Digital Twin (DT), this paper combines DT and Stacked Auto Encoder (SAE) to propose a DT and SAE Based Model (DSBM) for products quality detection. Comparing to using only a data-driven method, the method this paper proposed has an ability to update parameters through calculating the trend of the stored historical data, which could be suitable well with the data disturbed by the work environment. To summarize, the contributions of this paper are shown as follows:

- The architecture of the DSBM is introduced and the digital model of DT is divided into a task-achieved part and a self-update part.
- A new loss function used for SAE and a parameters-update rule are studied to improve the precision of the model and maintain the fidelity on a long term.
- A case study is presented to validate the proposed DSBM method has better accuracy than traditional data-driven methods without DT model under the changing work environment.

The rest of this paper is organized as follows. Section II introduces the related works. Section III presents the methodology. Section IV presents the testing result of the proposed method on a CNC bending machine instance. Section V presents the conclusion and future works of the proposed method.

II. RELATED WORKS

The related study about DT method, product quality monitor and Auto Encoder are discussed in this section.

A. THE DIGITAL TWIN METHOD

Firstly, NASA put forward the conception of "Digital Twin" from the Apollo Project, which helped staff to predict the remaining useful life of spacecraft [4]. So that the ability to emergency management in aerospace mission was increasingly enhanced [14]. With the improvement of computer performance, Computer-Aided Design (CAD), Computer-Aided Manufacturing (CAM), and Computer Simulation (CS), have been the most focus on the DT. The 3-Dimension geometry model and process simulation have been used in the DT methods [15]. With the further development of Computer Science, Sensor Technology, and Internet Technology, the informatization in the shop floor has been vastly developed, which makes the industrial data blowout [16]. There are a large number of data collected during machining processing nowadays, which contains a mass of information about machining and states of equipment. To extract useful information from a bulk of data, data-driven approaches are receiving the attention of researchers, which has become an important research direction of the DT. Many virtual models of DT have been constructed by data-driven method.

In terms of architecture establishment, Tao *et al.* [17] put forward a Five-Dimension Digital Twin (5-DDT) architecture. There are five parts – physical entity model, virtual equipment model, services model, DT data model, and connection model, embedded in the architecture. It makes the structure of the DT clearer comparing with the initially DT architecture that contained only three parts – the physical model, the virtual model, and the connection between models [18]. Besides, the DT was categorized into three categories [19], which makes the definition of DT more clearly.

The application of DT is wildly distributed in Product Lifecycle Management (PLM) [20], Product Life Management (PHM) [17], [21], [22], structure design [23], [24], and real-time monitor [25]. In [26], a DT-based method using deep transfer learning was offered. The deep transfer learning in the method aimed to ensure the accuracy of the model in the absence of historical data. The method implemented by the experiment reached a better result than the common stacked sparse auto-encoder model in fault diagnosis task. Moreover, Debroy *et al.* [27] built a series of simulation models to verify the mechanical properties through the DT instead of the experimental methods, which reduced the cost of time and money. Alam and El Saddik [28] investigated a DT architecture model for cloud-based cyber-physical systems, which was applied to a vehicle driving assistance system.

The DT comprises four parts generally - the physical model, the digital model, the bidirectional data connection, and the DT data. The physical model is an equipment in the Physical Space; the DT data is the Twin of the physical model in the Data Space; the digital model is a map which has the ability to map the data to the Task Space from the Data Space; and the bidirectional connection is the connection between them. More details about these are discussed at the Section III.

B. PRODUCT QUALITY MONITOR

Factory automation needs to ensure the quality of machining workpieces. Initially, nondestructive testing was the major method to detect the product defect [2]. However, the method is difficult to be deployed on some automated production lines (like large mass production situations). So, the machine vision method has attracted researchers' focus. Reference [29] proposed a method to monitor the friction stir welding surface quality by surface image, which used Maximally Stable Extremal Regions method to detect the blob and express the flaw, at the detected weld joint. Liu *et al.* [30] proposed an image analysis method to detect the defect at addictive manufacturing. The method consists of a textural analysis-based image classification algorithm and a Proportion Integral Differential (PID) equation-based feedback quality control system. Wang *et al.* [31] used the deep Convolution Neural Network (CNN) to process images, which could detect the unqualified product. And defective products could be detected through three major stages: image preprocessing, a region of interest extraction, and image identification.

Whereas, the machine vision method couldn't work effectively where the light condition and space are limited. With the trend of industrial big data has become irresistible, the data-driven method has become a great solution. In [32], a data-driven based method for detecting the surface roughness at addictive manufacturing was introduced. The data about extruder temperature and build plate vibration of a 3D printer were collected for predicting the surface roughness through an ensemble learning algorithm. Reference [33] introduced a method that could predict the machine tool's health condition with the help of multi-sensor fusion technology to ensure machining quality.

Whereas, researchers [26] found that the data used for training models and the actual data show different distribution in some cases, and the gap between them will be enlarged over time. In other words, the accuracy of the model is degraded for a long term.

C. AUTO ENCODER

Auto Encoder (AE), a method for nonlinearity dimensionality reduction, was known as auto-association before, which is a 3-layers neural network and it annoyed many researchers for the training method in 1990s [34]. Hinton and Salakhutdinov [35] summed up the method for layer-wise training of AE. From then on, the AE method was studied diffusely [34]. In 2007, Bengio *et al.* [36] introduced a deep AE called Stacked Auto Encoder (SAE) that is powerful in dimensionality reduction. Reference [37] investigated a teacher and supervise dual stacked auto-encoder (TSSAE) whose feature extraction and model construction are implemented by two neural networks such that the quality indicator is guaranteed. And the AE in [37] is composed of a nonlinear encoder network and a linear encoder network, which is different from the usual case. Reference [38] proposed a label and sparse regularization AE (LSRAE) by integrating label and sparse constraints to update the structure of the AE. The method enhanced the performance of the classifier in depth.

The AE, composed of an encoder network and a decoder network, has showed a robust performance. The encoder network is used for dimensionality reduction, and the decoder network for reconstructing the data compressed by the encoder network. Whereas, the decoder network which makes it impossible to ensure reconstruction error in the application only works during the training phase. In other words, there is no measure to supervise the accuracy of the AE in the application. So, in this paper, the decoder network is used to measure the reconstruction errors at the training phase as well as at the application phase to ensure the model is suitable for changing data.

III. METHODOLOGY

This section presents the proposed DSBM method. First, the architecture of DT and the DSBM method were introduced. Then, the structure of SAE and a new loss function was presented. Moreover, the method was also proposed to maintain the accuracy of the model.

A. THE ARCHITECTURE OF THE DSBM METHOD

DT achieves a particular function or goal, in the virtual environment, through describing the actual system. The key properties of the DT are the high-fidelity model and bidirectional data connection. These features greatly improve the performance of applications such as DT-based fault prediction, DT-based planning system, and DT-based product quality monitor.

The DT comprises physical model, digital model, bidirectional data connection, and DT data, in a broad sense.

- The physical model normally can be divided into two parts: entity and nonentity. The entity means a physical system such as processing equipment, workpieces, workers and so on, whose information is normally inputted by workers or collected form sensors. And the nonentity denotes the non-physical environments such as processing environment, market conditions and so on. Furthermore, the physical model is the data source of DT data and the calculating goal of the digital model.
- A map or equation that can display a particular state of the physical model is called a digital model. The digital model of DT does not take artificial signal as input like traditional simulation model but take true states from a physical model to complete a task. In other word, before a physical model is finished, the DT model couldn't be built whereas a simulation model could be. It could not operate without a machine or a processing equipment. Furthermore, on the time dimensionality, it can be classified into two groups, the synchronous model and the asynchronous model, according to time difference by the input and output. If the input to the model is the data occurred at this time, but the result describes the state of the device at the next moment, then the model is regarded as an asynchronous model. Conversely, if the input of the digital model and result describe the states of the same time period, then the model is a synchronous model. For example, at fault prediction, the input is historical data and the result is a value that describes the future state of the device, so an asynchronous model is used here. In the quality monitor this paper proposed, the input is motor current data and the result is the product quality state occurred at the same time with the input, so the synchronous model is used.
- DT data is a database which in accordance with practical applications could contain device states (such as geometry, state of critical parts of the device, workpiece topology information, etc.), processing conditions (such as schedule, bill of materials, yield, etc.), model data (such as test results, model results, process strategies, etc.), and others. It is the data source of the digital model and the twin of the physical model in the Data Space.

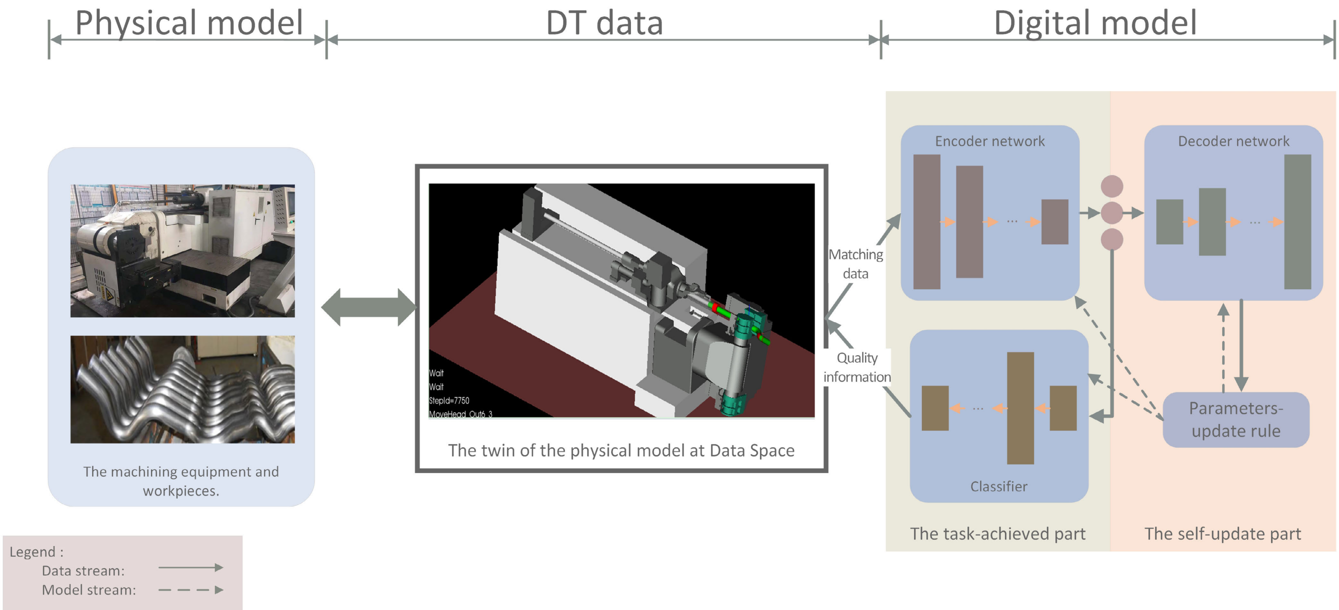


FIGURE 1. The architecture of the DSBM.

- The bidirectional data connection is the bridge between physical model, digital model and DT data. It serves to connect the various parts of the DT. And the bidirectional data connection decides the performance of real-time which in industrial application is completely important.

Comparing with the data-driven method that contains some analogous parts similar to the digital model in DT model, there are some differences between them. Firstly, the data-driven method is an isolated part without physical parts and the connection between the physical model and the digital model. Secondly, the data-driven model normally can be as a part of the digital model of the DT model or be directly the whole digital model. In other word, the digital model of the DT model, which is application-oriented, may have more than one method to complete a task, whereas the data-driven method is just one method. That is, the digital model can be a set of methods rather than only containing one method. And a DT model is tended to be a system rather than a method.

Based on the conception of the DT, this paper proposes a DSBM method as Fig. 1 shown. The frame is similar with the DT standard model. In detail, the physical model is a machining equipment and workpieces; the DT data is a database stored the quality information of workpieces, the machining data collected from sensors in the processing equipment, and the result of the digital model. In addition, the digital model is composed by a task-achieved part and a self-update part. The former is made up by a encoder network which is a part of SAE and used to reduce the data dimension and an Artificial Neural Network (ANN) used to classify the production quality; the latter comprises a decoder network and a parameters-update rule that can update the parameters of the model, including the encoder network, the decoder network, and the classifier, once the reconstruction error of SAE is too large over a certain period.

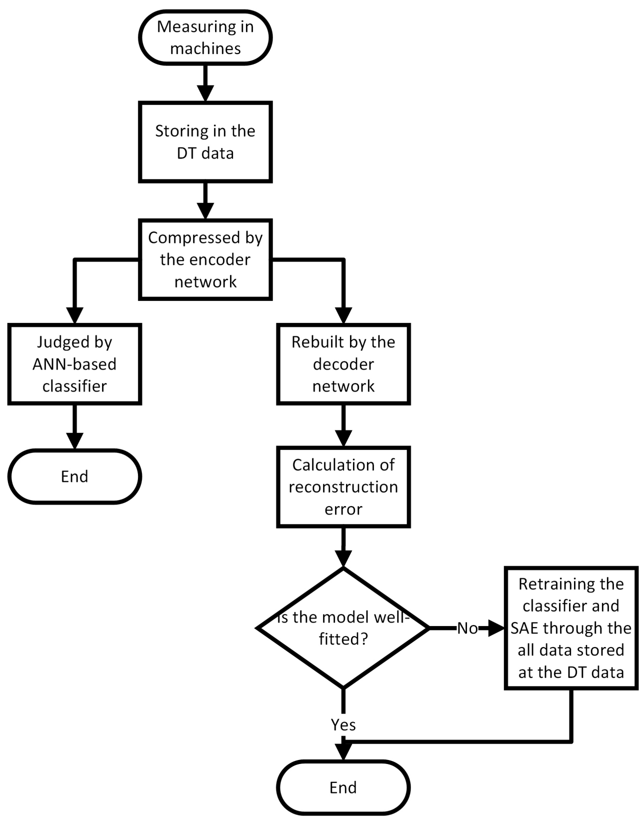


FIGURE 2. The flowchart of the DSBM.

As Fig.2 shown, after a working step is completed, the time-sequenced data generated during the working step will be stored at the DT data and then they will be copied to the digital model as the input data of the encoder network of SAE. The encoder network will reduce the dimension of the data. And the dimensionality-reduced data will be copied

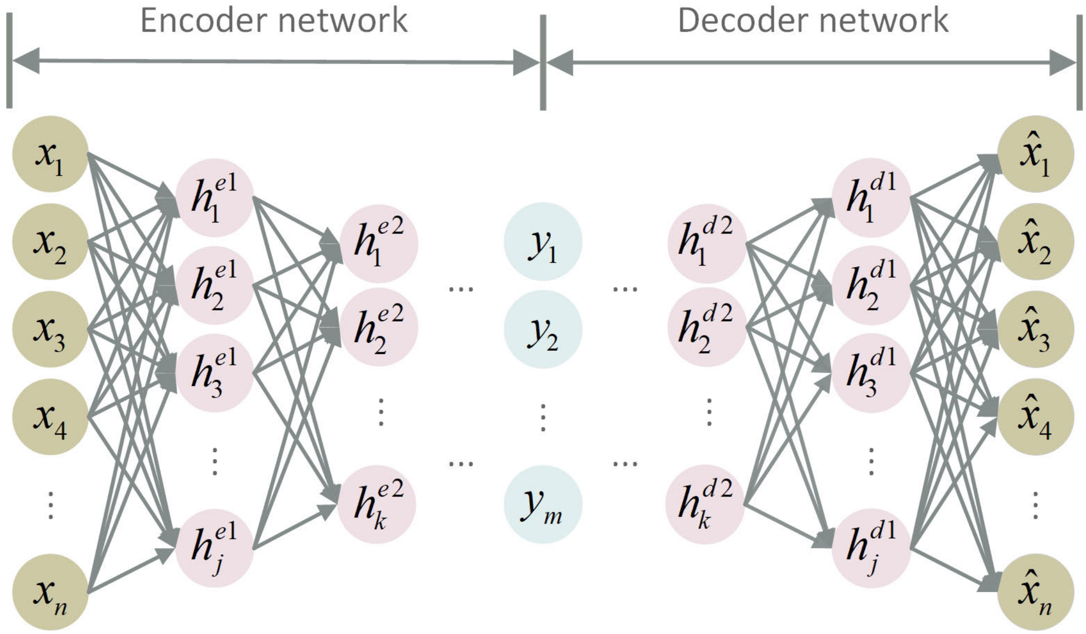


FIGURE 3. The structure of SAE.

separately to the classifier and the decoder network. The classifier will determine whether the product is qualified and then the quality information will be stored in the DT data. The dimension of the dimensionality-reduced data will be restored by the decoder network and the parameters-update rule will check whether the model is still well-fitted. If the rule determines the model cannot fit the data, the models will be retrained with all of data stored at the DT data to update the weights and bias.

B. THE TASK-ACHIEVED PART OF DT

The task-achieved part of DT is composed by the encoder network of SAE and a ANN-based classifier. The former is used to reduce the dimension of the input data, and the latter is used to determine whether the product is qualified.

SAE, composed by an encoder network and a decoder network, is a type of deep AE whose hidden layers are more than one, as shown in Fig. 3. The structure of the decoder network and that of the encoder network are symmetrical. And the number of neurons of the encoder network is decreasing layer by layer. In contrast, the number of neurons in decoder network is increasing layer by layer. The function of the encoder network is reduction of the dimension of input data. And the decoder network is used to reconstruct the compressed data.

The output of SAE is $\hat{\boldsymbol{x}} = \left(\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \cdots, \hat{\boldsymbol{x}}_n\right)$ and the input of SAE is $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, where n is the dimension of the input and the output. The output of the first layer of the encoder network is shown in (1).

$$\boldsymbol{h}^{e1} = f\left(\boldsymbol{W}^{e1}\boldsymbol{x} + \boldsymbol{b}^{e1}\right). \tag{1}$$

The $\boldsymbol{h^{e1}} = \left(h_1^{e1}, h_2^{e1}, \cdots, h_j^{e1}\right)$ is the output of the first layer of the encoder network where j is the dimension of the $\boldsymbol{h^{e1}}$ and $j < n$. The $\boldsymbol{W^{e1}}$, a matrix of j rows and n columns, is the weight matrix of the first layer and the $\boldsymbol{b^{e1}} = \left(b_1^{e1}, b_2^{e1}, \cdots, b_j^{e1}\right)$, a j dimensional sequence vector, is the bias vector of the first layer of the encoder network. And $f(\cdot)$, a activation function, denotes a ReLU function in this paper. $\boldsymbol{y} = (y_1, y_2, \cdots, y_m)$, output of the encoder network, is compressed from the $\boldsymbol{x}$.

The SAE is a self-supervised method. The input data are mapped to a low dimensional space by the encoder network and then mapped to the original space by the decoder network. Minimizing reconstruction error between $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ is the training criterion. Mean-Square Error (MSE) is usually used to assess the reconstruction error of SAE. Whereas, in many instances, the number of fault data is much less than the amount of good data, which means the data are unbalanced. The MSE, measuring the abstract error, will weaken the reconstruction error of machining data. That is, if the output of the model is all similar to good data, the result of MSE is small either. So, a proportional MSE function is proposed, which can magnify the tiny difference between input data and result. To prevent zero from being in the denominator position, a regular term, which is small near the zero point and close to zero in others position, is added on the denominator.

$$Loss\left(\boldsymbol{x},\hat{\boldsymbol{x}}\right) = \sum_{i=0}^{n}\left(\frac{\left(x_i - \hat{x}_i\right)^2}{x_i^2 + \frac{b}{a\cdot e^{x_i^2}}}\right). \tag{2}$$

where a and b are all hyper parameter that determine the degree of influence of the regular item. $Loss(\boldsymbol{x}, \hat{\boldsymbol{x}})$ is the result of the proportional MSE that shows the reconstruction error of SAE.

After the training of SAE completed, the weights and bias in the SAE are all fixed. For the data only need to be compressed at the task-achieved part, the output of the encoder network will be directly copied to the ANN-based classifier rather than to the decoder network, as Fig. 4 shown.

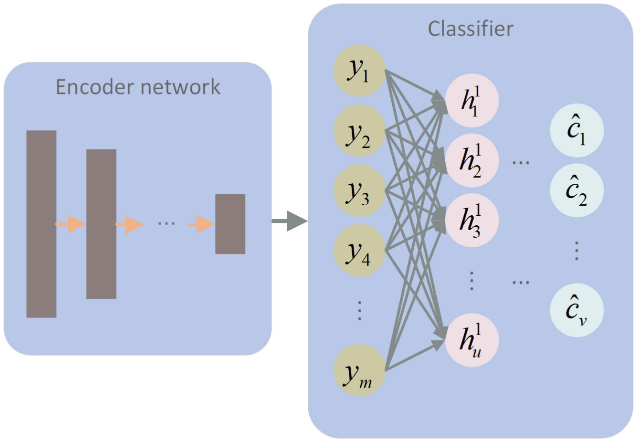


FIGURE 4. The ANN-based classifier.

The output of the classifier is $\hat{\boldsymbol{c}} = \left(\hat{c}_1, \hat{c}_2, \cdots, \hat{c}_v\right)$ that denotes the classifier appraises which group, such as qualified or unqualified, the data belong to, where v is the dimension of $\hat{\boldsymbol{c}}$. And the output of the first layer of the classifier, $\boldsymbol{h}^{\mathbf{1}}$, can be shown as (3).

$$\boldsymbol{h}^1 = g\left(\boldsymbol{W}^1\boldsymbol{y} + \boldsymbol{b}^1\right). \tag{3}$$

where $\boldsymbol{W}^{\mathbf{1}}$ is the weights matrix and $\boldsymbol{b}^{\mathbf{1}}$ is a bias vector. And $g(\cdot)$ denotes the tanh activation function in this paper. The activation function of the last layer of the classifier is the Sigmoid function that is widely used in classification tasks.

The parameters in the classifier are learned by minimizing the cross entropy, as (4) shown.

$$CrossEncropy = \frac{1}{v}\sum_{i=1}^{v} c_i \cdot log\left(\hat{c}_i\right). \tag{4}$$

where $\boldsymbol{c} = (c_1, c_2, \cdots, c_v)$ is the label vector of data. The elements c_i in the $\boldsymbol{c}$ are bigger than any other elements if the data belong to the group i.

Finally, all models are fine turned. The data from DT data should be compressed by the encoder network and then the classifier will determine which group they are belong to.

C. THE PARAMETERS-UPDATE RULE

Over time, data changes due to the material quality, states of critical components of the equipment, and other factors. As shown in Fig. 5, with time going, the reconstruction error between results calculated by decoder network and input data is increasing. That means the accuracy of the model drops, which can lead to the accuracy down.

The decoder network is a crucial part to indicate whether the encoder network is well-fitted. Normally there is no need to use the decoder network that is only used in the training phase, which means that there are no measures to indicate

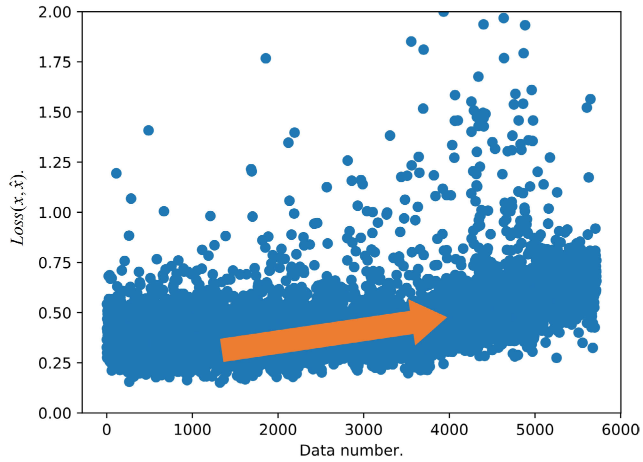


FIGURE 5. The tendency of reconstruction error.

the accuracy of the model during the execution of SAE. In this paper, the decoder network, well-trained before at the training phase in the task-achieved part, and (2) are used to detect the accuracy of the model at any stage, not just in the training phase. The dimensionality-reduced data will be copied from the encoder network to the decoder network to be reconstructed and then (2) will calculate reconstruction error.

As can be seen from Fig. 5, updating the model once a certain value is exceeded is a waste of computing resources. Because irregular drift occurs among reconstruction errors. Based on the phenomenon, the parameters-update rule used to calculate the fidelity of the SAE are proposed:

$$\boldsymbol{tt} = \gamma \cdot \left(\frac{\boldsymbol{t}}{\left\|\boldsymbol{t}\right\|_\infty} - \frac{\boldsymbol{ones}}{2}\right). \tag{5}$$

$$TH = \alpha \cdot \sum_{i=1}^{p} \frac{\beta \cdot e^{tt_i} - e^{-tt_i}}{\beta \cdot e^{tt_i} + e^{-tt_i}} \cdot l_i \tag{6}$$

where $\boldsymbol{t} = (1, 2, \cdots, p)$ whose length is p that is the number of results of (2) stored at DT data is a time vector, and $\boldsymbol{ones} = (1, 1, \cdots, 1)$ is a vector whose dimension is p and elements are all one. tt_i $(i = 1, 2, \ldots, p)$ denotes an element of normalized time vector $\boldsymbol{tt}$. l_i $(i = 1, 2, \ldots, p)$ is the t_i-th result of $Loss(\boldsymbol{x}, \hat{\boldsymbol{x}})$. The TH is the result of (6). It is based on hyperbolic tangent function (tanh), because the dependent variable of tanh changes drastically nearby zero point and the result of points far from the zero point is almost constant. So, those properties can reduce the impact of l_i according to time and prevent the influence of irregular drift. To make (6) obtain an ability to adjust the sensibility to suit more situation, three variables are added. α is a variable that controls the amplitude of the function; β is a sensitive regulation factor determining the time-sensitivity; γ determines the sensitive degree to the data. The reasonable α, β and γ can avoid updating the model caused by randomness errors increases. Namely, the impact of short-term fluctuations of data on the influence of the functions is reduced.

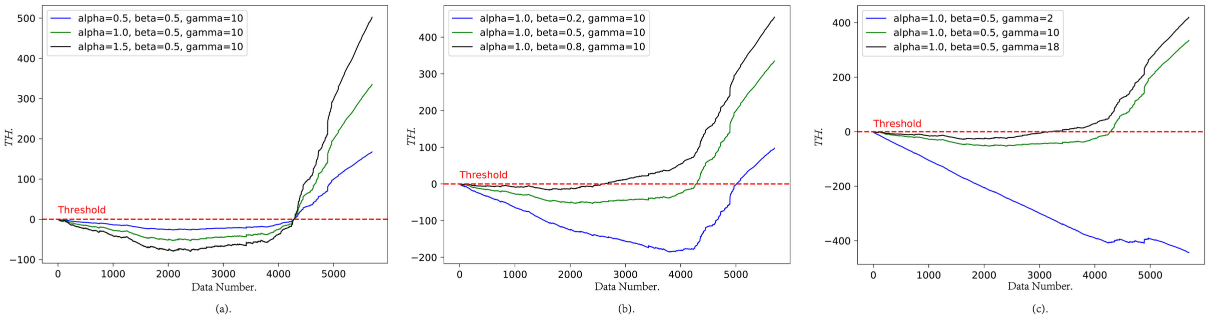


FIGURE 6. The effect from (a) α, (b) β, and (c) γ.

The effect of different variable values is shown in Fig. 6. The threshold in the figures is set to zero. And the $\boldsymbol{t}$ in (5) is the data from Fig. 5.

If the *TH* is greater than or equal to the threshold, the SAE model will be no longer applicable to the present data, which means the parameters of the SAE, i.e. the weights and bias in the neural network, need to be updated. So, the SAE model and the ANN-based classifier will be retrained by data that contain those used to train models before and the new collected data stored in the DT data.

IV. EXPERIMENT

This section presents an instance study about the proposed method in a CNC bending machine. The SAE model and the ANN model are written in Python 3.7 with TensorFlow 1.14 and run on Windows 10 (x64) 1903 with a GTX 1070 GPU.

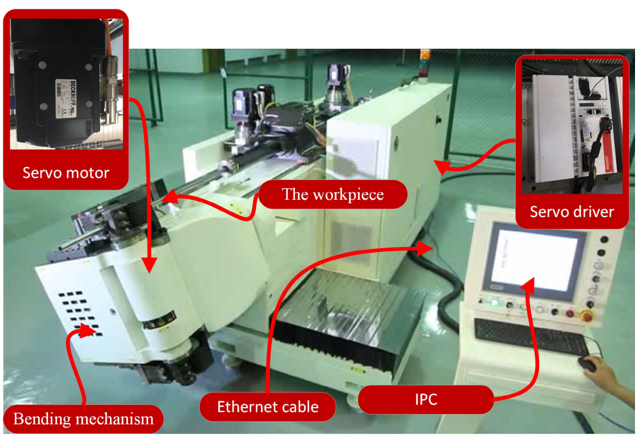


FIGURE 7. The CNC bending machine.

The CNC bending machine is shown in Fig. 7. The bending mechanism is drived by a BECKHOFF AM3072 servo motor whose max output torque is 117 N·m and its driver is a BECKHOFF AX5118 that has a 8-bit current-measuring model used to measure the current of the motor in real time. And the collected current data will be stored at an industrial PC (IPC). The data connection between the driver and the IPC is EtherCAT. The control system is TWINCAT2, a kind of soft PLC, that is installed at the IPC. Further, the material of workpieces is QSTE380TM. The material was chosen because it is considered to be one of the most wildly used automobile steel.

A. THE DESCRIPTION OF DATASETS

Cracking of pipe fittings and wrinkling of pipe fittings are the most likely forms of workpiece failure during machining. While these conditions occur, the servo motor current, proportional to the load, will change suddenly, as shown in Fig. 8.

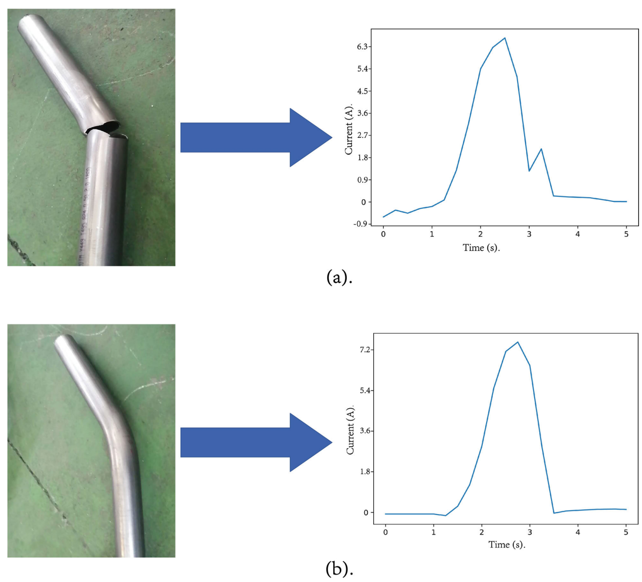


FIGURE 8. Motor's current under (a) unqualified product and (b) qualified product.

The datasets used in this paper are current data of the servo motor jointed with the bending mechanism, collected by the built-in measuring model of the servo driver in the real production environment, when the machine is processing. And the sampling frequency is 40 Hz. A dataset contains a data matrix and a label matrix. Each column vector, the input of SAE, in the data matrix is several groups of data that contains 5 seconds of current data, which means that there are 200 current data points in a group of data. And each column of the label matrix has two elements. If the column vector denotes that the processing is qualified, the first element of the column is one and the second is zero, otherwise the first is zero and the second is one. It is used to check whether the result of the classifier is correct. There are some samples received from 5 different CNC bending machines and divided into three groups. The details are shown in the Table 1. And the dimension of data matrix of each dataset, in total, is 200*2500, 200*1200, 200*900 respectively. The dimension of label matrix is 2*2500, 2*1200, 2*900 respectively.

TABLE 1. The details of the datasets.

	Dataset I.	Dataset II.	Dataset III.
Num of good data.	2000	1000	750
Num of fault data.	500	200	150

In the Table 1, good data refer to the processing data of qualified pipe fittings. Similarly, fault data refer to the data for substandard products. The datasets I and II were collected from five machines used less than 200 hours. The dataset III is the data after 2000 hours of use of the five machines. It could be regarded as the situation influenced by the changing environment, because the states of the machine change for wear after 2000 hours using. All datasets include a certain number of good data and fault data. The dataset I is used to train the model at the beginning, and the datasets II and III are used to verify the accuracy.

B. THE DESCRIPTION OF THE PROPOSED METHOD

At first, an SAE model and an ANN-based classifier are built. The six SAE networks with different numbers of layers and neurons were trained 10 times respectively, and the super parameters in (2), i.e. a and b, are 1.5 and 0.5 respectively by the engineering tuning method. SAE networks were trained by the dataset I and tested by the dataset II. The details are shown in the Table 2 where max, min, mean, and var denote the max loss, min loss, mean loss, and the variance of loss respectively. The name represents the structure of the model. For example, "3L-100-50-25" means a model where the encoder network has three hidden layers with 100, 50, and 25 neurons per layer. And the number of neurons is 200 in each model, which is determined by input data. The model "3L-100-50-25" manifests the best performance at the dimensional reduction task.

The Table 3 shows the performance of the ANN-based classifier with different structures. Besides, the results of the encoder network with the datasets I and II will be used to train and test the ANN-based classifiers, respectively. And the dimension of the output layer is 2, which is determined by the label matrix of datasets. The model "2L-40-40" shows higher accuracy in the classifier task.

The super parameters, i.e. α, β, γ, in parameters update rule are set to 1, 0.5, and 10, respectively, by the engineering tuning method. And the threshold is set to 0. After training, the models could be used in detection tasks.

TABLE 2. The results of the SAE with different structures.

Name.	Max.	Min.	Mean.	Var.
3L-100-50-25	0.561	0.417	**0.475**	0.002
3L-150-100-50	0.861	0.458	**0.675**	0.23
3L-170-120-50	1.671	0.467	**0.814**	0.102
4L-100-50-25-12	6.155	0.406	**1.465**	3.378
4L-120-90-60-30	5.497	0.526	**2.154**	3.057
4L-180-150-120-80	63.721	1.074	**17.435**	410.785

TABLE 3. The results of the classifier with different structures.

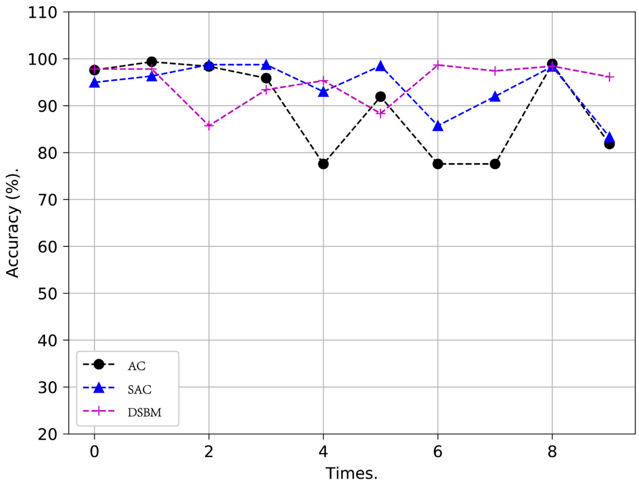
Name.	Max.	Min.	Mean.	Var.
1L-20	99.07%	77.58%	**90.96%**	8E-3
1L-30	99.07%	91.95%	**97.10%**	3E-4
1L-40	99.07%	77.58%	**95.03%**	5E-3
2L-20-20	99.69%	94.14%	**98.44%**	2E-4
2L-30-30	99.45%	94.18%	**97.81%**	3E-4
2L-40-40	99.30%	97.91%	**98.79%**	2E-5

C. COMPARE AND DISCUSSION

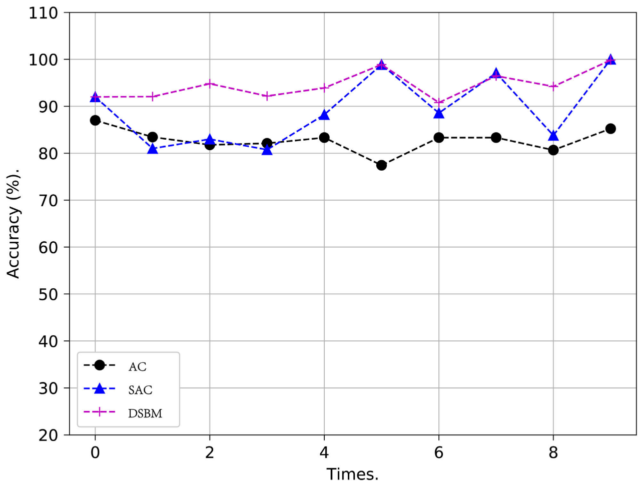
To show the advanced performance of the proposed method in accuracy and generalization, two other methods were selected to compare the prognosis accuracy in this case. They are a traditional ANN-based Classifier (AC) and a combination method of SAE and an ANN-based classifier, shorten as SAC, that has the same neural network structure of the encoder network and the classifier in DSBM. And, the AC that the input data is the current data has three hidden layers with 150, 100, and 50 neurons per layer. The two methods and proposed method were trained by the dataset I and tested by the datasets II and III.

Each method was tested ten times for avoiding randomness. The testing results under datasets II and III are shown in Fig. 9 (a) and Fig. 9 (b) respectively. Since the datasets I and II are collected from the same work environment, namely the parameters-update rule is not established under the dataset II because reconstruction error of model is not so large and SAC has the same structure of the neural network as DSBM. Both show similar performance under the dataset II. The accuracy of AC is poorer under the dataset II than the others. Under the dataset III, the accuracy of both AC and SAC shows varying degrees of degradation. Benefiting from the parameters-update rule, the recession of accuracy of the model is prevented while the rule is established.

As the Table 4 shown, under the dataset II, SAC has a similar performance to DSBM, with an average of 94.80% and 94.74%, and variances of 1.4E-3 and 1.9E-3, respectively. The performance of AC under the dataset II is weaker than other methods, the average is 89.67%, and the variance is



(a).



(b).

FIGURE 9. Three results of methods under (a) dataset II and (b) dataset III.

TABLE 4. The comparison between three methods.

Method.	Max.	Min.	Mean.	Var.
AC - dataset II.	99.38%	77.58%	**89.67%**	7.8E-3
SAC - dataset II.	98.76%	83.33%	**94.80%**	1.4E-3
DSBM - dataset II.	98.68%	85.73%	**94.74%**	1.9E-3
AC - dataset III.	92.11%	77.44%	**85.77%**	1.8E-3
SAC - dataset III.	98.89%	80.74%	**88.00%**	3.6E-3
DSBM - dataset III.	99.83%	90.78%	**94.51%**	7.6E-4

7.8E-3. Moreover, under the dataset III, only DSBM maintains a good performance; other methods have experienced different levels of degradation. These results validate the performance of the proposed DSBM method.

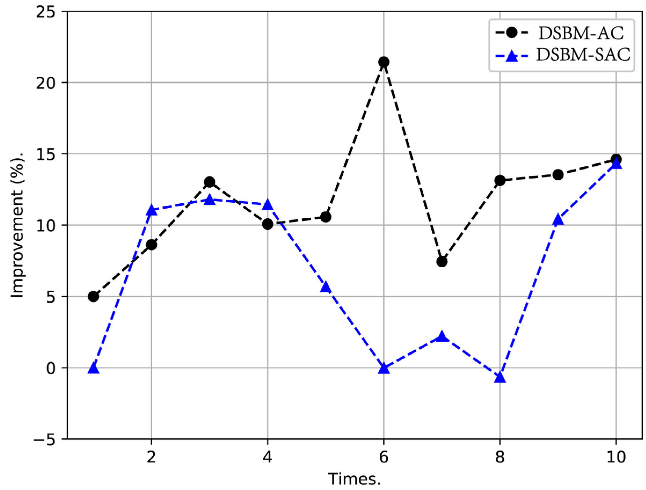


FIGURE 10. The comparison of DSBM with AC and SAC.

The differences in accuracy between DSBM and AC and between DSBM and SAC are shown in Fig. 10. The horizontal axis is the different testing time and the vertical axis is the difference between different methods. DSBM is 5% or more accurate than AC each time. For SAC, the improvement in the first, sixth, eighth and tenth experiments is not obvious. And Fig. 9 (b) shows that the corresponding accuracy of the four SACs under the dataset III is more than 90%, which means that the accuracy of the four experiments does not decrease much. So, the impact of the parameters-update rule is small, which is in line with the intuitive understanding of the proposed method. Therefore, it is concluded that the proposed method can achieve a better result, especially in the case where the accuracy is lowered more than the case in which the accuracy is decreased diminutively.

V. CONCLUSION AND FUTURE WORKS

In this paper, the structure of DT which contains the physical model, the digital model, the bidirectional data connection, and DT data was discussed and the different between DT methods and conventional data-driven methods was indicated. The changing of workshop circumstance, such as the material quality, states of critical components of the equipment, and so on, can cause the actual data to be deviated from the ideal data, which reduces the accuracy of many kinds of methods. Based on the structure of DT, a DT and SAE based quality monitor model was proposed. To reduce the impact of the changing environment in the workshop, the digital model of DT was further divided into two parts, a task-achieved part and a self-update part which are made by an SAE, an ANN-based classifier, and a parameters-update rule. After a working step is completed, the task-achieved part will judge whether it is qualified and the self-update part will check the accuracy of the task-achieved part. If the reconstruction error is too large to meet the parameters-update rule, all neural networks will be retrained and the parameters will be updated next.

And the proposed method was applied to a case of CNC bending machine. To emulate the influence of the changing circumstances of the workshop, we collected three groups data at different working time from five machines. The data were used to test the proposed approach and other two data-driven methods. The results demonstrated that the proposed method achieved a more accuracy than two other conventional methods and had the ability to maintain stability especially in the situation where the accuracy was affected hugely by the working context.

The limitation of the proposed method includes the following aspects. First, there is a strong real-time requirement in the industrial application generally. Whereas, calculating new arguments during the update process would require much time. The time cost depends on the performance of the computer. Besides, like other data-driven methods, the method cannot distinguish the unknown faults, therefore data corresponding to a certain problem must be collected before. Based on these limitations, future research can address these problems. First, there is a need to investigate an improved model update method to increase speed, such as an iterative based approach. Second, the DT based approach can be modified to handle unknown situations, for example using other types of classifiers.

REFERENCES

[1] J. Liu, X. Du, H. Zhou, X. Liu, L. Li, and F. Feng, "A digital twin-based approach for dynamic clamping and positioning of the flexible tooling system," *Procedia CIRP*, vol. 80, Jan. 2019, pp. 746–749.

[2] R. X. Gao, X. Tang, G. Gordon, and D. O. Kazmer, "Online product quality monitoring through in-process measurement," *CIRP Ann.*, vol. 63, no. 1, pp. 493–496, Mar. 2014.

[3] A. Wirtz, C. Gächter, and D. Wipf, "From unambiguously defined geometry to the perfect quality control loop," *CIRP Ann.*, vol. 42, no. 1, pp. 615–618, 1993.

[4] E. Negri, L. Fumagalli, and M. Macchi, "A review of the roles of digital twin in CPS-based production systems," *Procedia Manuf.*, vol. 11, pp. 939–948, Jan. 2017.

[5] S. Yin, S. X. Ding, X. Xie, and H. Luo, "A review on basic data-driven approaches for industrial process monitoring," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6418–6428, Nov. 2014.

[6] L. Bennacer, Y. Amirat, A. Chibani, A. Mellouk, and L. Ciavaglia, "Self-diagnosis technique for virtual private networks combining Bayesian networks and case-based reasoning," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 1, pp. 354–366, Jan. 2015.

[7] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometr. Intell. Lab. Syst.*, vol. 12, pp. 37–52, Aug. 1987.

[8] V. Choqueuse, M. E. H. Benbouzid, Y. Amirat, and S. Turri, "Diagnosis of three-phase electrical machines using multidimensional demodulation techniques," *IEEE Trans. Ind. Electron.*, vol. 59, no. 4, pp. 2014–2023, Apr. 2012.

[9] C. F. Alcala and S. J. Qin, "Reconstruction-based contribution for process monitoring with kernel principal component analysis," *Ind. Eng. Chem. Res.*, vol. 49, no. 17, pp. 7849–7857, Sep. 2010.

[10] Z. Zhang and J. Zhao, "A deep belief network based fault diagnosis model for complex chemical processes," *Comput. Chem. Eng.*, vol. 107, pp. 395–407, Dec. 2017.

[11] K. Javed, R. Gouriveau, X. Li, and N. Zerhouni, "Tool wear monitoring and prognostics challenges: A comparison of connectionist methods toward an adaptive ensemble model," *J. Int. Manuf.*, vol. 29, pp. 1890–1970, Dec. 2018.

[12] C. Gao, L. Jian, and S. Luo, "Modeling of the thermal state change of blast furnace hearth with support vector machines," *IEEE Trans. Ind. Electron.*, vol. 59, no. 2, pp. 1134–1145, Feb. 2012.

[13] P. Wu, M. Qi, L. Gao, W. Zou, Q. Miao, and L.-L. Liu, "Research on the virtual reality synchronization of workshop digital twin," in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, Chongqing China, May 2019, pp. 24–26.

[14] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen, "About the importance of autonomy and digital twins for the future of manufacturing," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 567–572, Jun. 2015.

[15] J. Wang, L. Ye, R. X. Gao, C. Li, and L. Zhang, "Digital twin for rotating machinery fault diagnosis in smart manufacturing," *Int. J. Prod. Res.*, vol. 57, no. 12, pp. 3920–3934, Jun. 2019.

[16] F. Tao, Q. Qi, L. Wang, and A. Y. C. Nee, "Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison," *Engineering*, vol. 5, no. 4, pp. 653–661, Aug. 2019.

[17] F. Tao, M. Zhang, Y. Liu, and A. Y. C. Nee, "Digital twin driven prognostics and health management for complex equipment," *CIRP Ann.*, vol. 67, no. 1, pp. 169–172, 2018.

[18] M. Grieves. *Digital Twin: Manufacturing Excellence through Virtual Factory Replication, White Paper*. [Online]. Available: https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication

[19] A. K. Ghosh, A. S. Ullah, and A. Kubo, "Hidden Markov model-based digital twin construction for futuristic manufacturing systems," *Artif. Intell. Eng. Design, Anal. Manuf.*, vol. 33, no. 03, pp. 317–331, May 2019.

[20] B. Schleich, N. Anwer, L. Mathieu, and S. Wartzack, "Shaping the digital twin for design and production engineering," *CIRP Ann.*, vol. 66, no. 1, pp. 141–144, Apr. 2017.

[21] C. Zhuang, J. Liu, and H. Xiong, "Digital twin-based smart production management and control framework for the complex product assembly shop-floor," *Int. J. Adv. Manuf. Technol.*, vol. 96, nos. 1–4, pp. 1149–1163, Apr. 2018.

[22] E. J. Tuegel, A. R. Ingraffea, T. G. Eason, and S. M. Spottswood, "Reengineering aircraft structural life prediction using a digital twin," *Int. J. Aero. Eng.*, vol. 2010, pp. 1687–5966, Oct. 2011.

[23] R. Söderberg, K. Wärmefjord, J. S. Carlson, and L. Lindkvist, "Toward a digital twin for real-time geometry assurance in individualized production," *CIRP Ann.*, vol. 66, no. 1, pp. 137–140, Apr. 2017.

[24] D. Wu, S. Liu, L. Zhang, J. Terpenny, R. X. Gao, T. Kurfess, and J. A. Guzzo, "A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing," *J. Manuf. Syst.*, vol. 43, pp. 25–34, Apr. 2017.

[25] R. M. Soares, M. M. Camara, and T. Feital Pinto JC, "Digital twin for monitoring of industrial multi-effect evaporation," *Processes*, vol. 7, pp. 537–550, Aug. 2019.

[26] Y. Xu, Y. Sun, X. Liu, and Y. Zheng, "A digital-twin-assisted fault diagnosis using deep transfer learning," *IEEE Access*, vol. 7, pp. 19990–19999, 2019.

[27] T. DebRoy, W. Zhang, J. Turner, and S. S. Babu, "Building digital twins of 3D printing machines," *Scripta Mater.*, vol. 135, pp. 119–124, Jul. 2017.

[28] K. M. Alam and A. El Saddik, "C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems," *IEEE Access*, vol. 5, pp. 2050–2062, 2017.

[29] S. Sudhagar, M. Sakthivel, and P. Ganeshkumar, "Monitoring of friction stir welding based on vision system coupled with machine learning algorithm," *Measurement*, vol. 144, pp. 135–143, Oct. 2019.

[30] C. Liu, A. C. C. Law, D. Roberson, and Z. Kong, "Image analysis-based closed loop quality control for additive manufacturing with fused filament fabrication," *J. Manuf. Syst.*, vol. 51, pp. 75–86, Apr. 2019.

[31] J. Wang, P. Fu, and R. X. Gao, "Machine vision intelligence for product defect inspection based on deep learning and Hough transform," *J. Manuf. Syst.*, vol. 51, pp. 52–60, Apr. 2019.

[32] Z. Li, Z. Zhang, J. Shi, and D. Wu, "Prediction of surface roughness in extrusion-based additive manufacturing with machine learning," *Robot. Comput.-Integr. Manuf.*, vol. 57, pp. 488–495, Jun. 2019.

[33] A. P. Kene and S. K. Choudhury, "Analytical modeling of tool health monitoring system using multiple sensor data fusion approach in hard machining," *Measurement*, vol. 145, pp. 119–129, May 2019.

[34] Y. Wang, H. Yao, and S. Zhao, "Auto-encoder based dimensionality reduction," *Neurocomputing*, vol. 184, pp. 232–242, Apr. 2016.

[35] G. E. Hinton, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[36] Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle, and U. Montreal, "Greedy layer-wise training of deep networks," in *Proc. NIPS*, vol. 19, Jan. 2007, pp. 4–7.

[37] S. Yan and X. Yan, "Design teacher and supervised dual stacked auto-encoders for quality-relevant fault detection in industrial process," *Appl. Soft Comput.*, vol. 81, Aug. 2019, Art. no. 105526.
[38] Z. Chai, W. Song, H. Wang, and F. Liu, "A semi-supervised auto-encoder using label and sparse regularizations for classification," *Appl. Soft Comput.*, vol. 77, pp. 205–217, Jan. 2019.

SHIZHENG ZHANG received the B.E. degree in mechanical engineering and automation from Ludong University, China, in 2018. He is currently pursuing the master's degree with the Beijing University of Technology, China. His research interests include data-driven approaches, machine learning, and electromechanical control.

CUNFENG KANG received the bachelor's degree in mechanical manufacturing and process equipment, the master's degree in mechanical manufacturing and automation, and the Ph.D. degree in mechanical design and theory from the Beijing University of Technology, in 1993, 2001, and 2008, respectively.

From August 1998 to January 2004, he served as a Lecturer with the College of Mechanical and Electrical Engineering, Beijing University of Technology, where he has been serving as an Associate Professor with the School of Mechanical and Electrical Engineering, since January 2004. He has written one textbook and more than 20 articles, and holds more than 20 patents and software copyrights. He is mainly engaged in high-end numerical control equipment, robot control technology, metal forming technology, and other research works. He had many experiences in numerical control technology, computer testing technology, numerical analysis method, signal analysis, and processing.

ZHIFENG LIU received the Ph.D. degree in mechanical design from the College of Mechanical Engineering and Automation, Northeastern University, in 2001.

He is currently an Associate Dean of the College of Mechanical Engineering and Applied Electronics Technology, Beijing University of Technology, and the Director of the Beijing Key Laboratory of Advanced Manufacturing Technology. His research interests include robotic motion control, machine vision technology, networked manufacturing, and manufacturing system information.

JUAN WU received the B.E. degree in mechanical engineering and automation from Jinan University, China, in 2016. She is currently pursuing the master's degree with the Beijing University of Technology, China. Her research interests include electromechanical control and machine learning.

CHUNMIN MA received the Ph.D. degree from the College of Mechanical and Electrical Engineering, Beijing University of Technology, in 2013. He remained in the university as a teacher. He engaged in the research of NC control technology and mainly engaged in NC equipment control software programming and high-level language programming. His projects are Electric Injection Molding Machine, NC Cross-Cutting Machine, NC Shaftless Printing Machine, NC Flying Saw Machine, NC Pipe Bend Machine, and Internal High Pressure Equipment.