IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 3, 2019, accepted February 22, 2020, date of publication March 3, 2020, date of current version March 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2977892

Unsupervised Anomaly Detection of Industrial Robots Using Sliding-Window Convolutional Variational Autoencoder

TINGTING CHEN[1], XUEPING LIU[1], BIZHONG XIA[1], WEI WANG[2], AND YONGZHI LAI[2]
[1]Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China
[2]Sunwoda Electronic Company, Ltd., Shenzhen 518108, China

Corresponding author: Bizhong Xia (xiabz@sz.tsinghua.edu.cn)

This research was funded by the Shenzhen Economic, Trade and Information Commission of Shenzhen. Municipality Strategic Emerging Industries and Future Industrial Development ''Innovation Chain + Industrial. Chain'' Project (2017).

ABSTRACT With growing dependence of industrial robots, a failure of an industrial robot may interrupt current operation or even overall manufacturing workflows in the entire production line, which can cause significant economic losses. Hence, it is very essential to maintain industrial robots to ensure high-level performance. It is widely desired to have a real-time technique to constantly monitor robots by collecting time series data from robots, which can automatically detect incipient failures before robots totally shut down. Model-based methods are typically used in anomaly detection for robots, yet explicit domain knowledge and accurate mathematical models are required. Data-driven techniques can overcome these limitations. However, a major difficulty for them is the lack of sufficient fault data of industrial robots. Besides, the used technique for anomaly detection of robots should be required to not only capture the temporal dependency in collected time series data, but also the inter-correlations between different metrics. In this paper, we introduce an unsupervised anomaly detection for industrial robots, sliding-window convolutional variational autoencoder (SWCVAE), which can realize real-time anomaly detection spatially and temporally by coping with multivariate time series data. This method has been verified by a KUKA KR6R 900SIXX industrial robot, and the results prove that the proposed model can successfully detect anomaly in the robot. Thus, this work presents a promising tool for condition-based maintenance of industrial robots.

INDEX TERMS Anomaly detection, industrial robots, sliding window, variational autoencoder, convolutional neural network.

I. INTRODUCTION

Nowadays, industrial robots are playing an increasingly important role in manufacturing as they greatly improve productivity and quality. According to the reports by the International Federation of Robotics, the number of robot installations is beyond 40,0000 units per year in 2018 [1]. Industrial robots are widely used in repetitive and continuous works, such as pick-and-place, welding, painting and so on. However, with growing dependence of industrial robots, a failure of a robot may cause a significant interruption in the entire production line and can quickly deteriorate into a catastrophe while some faults that have potential to lead to a failure are not easy to be recognized. Especially some soft and hardware-related faults in the early stage will not instantly result in a failure since the control algorithms can adapt the subtle changes caused by those faults [2]. Hence, it is very essential to maintain industrial robots to ensure high-level performance.

Regular preventative maintenance for industrial robots is typically performed in many industries which can prevent costly unplanned downtime due to mechanic failures. But in most cases, scheduling maintenance can be inefficient and unnecessary and it may require to shut down the whole production. Therefore, it is widely desired to constantly monitor robots and automatically detect incipient failures by gathering time series data from robots, which is called condition-based maintenance (CBM). CBM can greatly minimize unscheduled breakdown and production losses [3], [4]. Especially with the growing of the Internet of Thing (IoT)

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

producing large amounts of data, CBM has been widely studied in recent years in the field of prognostics and health management (PHM).

The main task of CBM is to perform real-time anomaly detection from the gathered time series data, which takes note of indicative fault data that do not conform to some explicit laws or historical patterns. There are two general techniques usually used for anomaly detection: model-based techniques [5]–[8] and data-driven techniques [2], [9]–[20]. Model-based techniques have been developed due to their good performance in predicting specific faults of the robots. However, they require to construct an accurate mathematical model for the robot, which demands domain knowledge and a lot of time to tune parameters. In contrast, data-driven approaches become more appealing as they do not require prior knowledge about the robot and detect faults entirely from collected data. Both supervised and unsupervised methods among data-driven approaches have been explored for anomaly detection. However, a major challenge of data-driven based approaches is the difficulty in obtaining sufficient fault data and labelling data accurately. Besides, the used technique for anomaly detection of robots requires the capabilities of capturing not only the temporal dependency in collected time series data, but also the inter-correlations between different metrics.

In this paper, an unsupervised anomaly detection method based on sliding-window convolutional variational autoencoder (SWCVAE) for industrial robots is proposed, which automatically learns normal patterns from time series data in training. A fully trained model can then be used for detecting anomalies spatially and temporally in input data and taking note of them correspondingly. Finally, a KUKA KR6R 900 SIXX industrial robot has been used to test the effectiveness of the proposed method. The angle configuration data and current data in six joints of the robot during normal operation are collected and trained in our proposed model. Then we tested the method by injecting faults artificially on the robot. The results show that the approach achieves satisfactory performance. Thus, this work presents a promising tool for condition-based maintenance of industrial robots.

II. RELATED WORKS

A. CHALLENGES DEFINITION

Anomaly detection for industrial robots recognize indicative fault by taking note of data that do not conform to some explicit laws or historical patterns [17]. In general, there are three main challenges in detecting anomalies for industrial robots.

- Highly unbalanced dataset. The number of anomalous samples are very limited because industrial robots are under normal condition in most cases. Besides, recording instances of faults in industrial robots is costly and dangerous.
- The transitory and non-stationary nature of robot data [14]. Each joint in the robot will rotate at different angles and require different currents in different robot motions. Besides, the signals emitted from a robot's mechanical or electrical parts are most likely transient with great fluctuation.
- Requirement for real-time responding. As mentioned above, a fault in robots could result in a catastrophe. Therefore, online anomaly detection so as to respond as soon as possible is required.

The first challenge makes supervised methods like classification-based methods unfeasible as they require an adequate amount of normal data and fault data. The second problem requires a reliable anomaly detection algorithm which can extract latent features that are sensitive to faults instead of robot motions. The third issue requests the proposed algorithm should be suitable for real-time anomaly detection instead of analyzing accumulated data.

B. APPROACHES

Model-based approaches are typically used in anomaly detection of industrial robots. These techniques are required to establish a precise model for the robot to predict some estimates. The deviations between the estimates and measured values, which are so-called residuals, are processed to perform fault detection of industrial robots. Since these methods are based on prior knowledge of robot systems, the mentioned problems have little impact on these methods. Nevertheless, obtaining accurate model requires explicit domain knowledge and it is a time-consuming task. Additionally, the performance of model-based techniques is prone to model accuracy.

Therefore, we turn to data-driven approaches. Compared to model-based methods, data-driven approaches involve less limitations and have received increasing interest. Data-driven approaches can be divided into three general methods: Statistical methods, Signal-analysis based methods, and Machine learning based methods.

Statistical methods are frequently used due to their computational efficiency. They detect data points that deviate from the distribution of the historical data. These methods used for robots include Statistical Control Charts (SCCs) [9], Principal Component Analysis (PCA) based method [10], Partial Least Squares (PLS) based approach [11] and so on. However, most of these methods require that all the data have to be accumulated before faults can be detected, which make them unsuitable for real-time anomaly detection. Furthermore, they assume that normal data are generated from a known distribution.

Another common method for anomaly detection of industrial robots is signal-analysis methods. With the use of additional sensors like current sensor [12], acoustic sensor [13], or accelerometers [14], these methods based on integral transforms like Fourier or Wavelet transform can easily extract the features of signal in the transformed domain. However, the biggest limitation of signal-analysis methods is that they can only handle one dimension signal. As for multivariate signals, they have to process each signal individually without considering inter-connection of signals. Besides, using additional sensors to perform anomaly detection is not feasible in many industrial environments because it increases the cost and complexity and requires extra space. Moreover, some additional sensors are sensitive to sensor locations and environmental noise.

Recently, machine leaning based methods for anomaly detection have received widespread attention due to their promise to automatically derive underlying rules from the data itself. Machine learning based methods comprise supervised methods and unsupervised methods. Supervised methods like classification-based methods [18] are only feasible when there is a high volume of labeled data including sufficient normal and abnormal data, because they only learn from examples where the desired outcomes are already known.

Therefore, unsupervised methods are desired. Among unsupervised methods, density-based methods like K-Nearest Neighbors (KNN) [21] and distance-based methods [20] may work well yet face the constraints of time and computational load when handling high dimensional data. One-Class Support Vector Machines (OCSVM) [22] can also be feasible for highly unbalanced dataset as it only looks at the distribution of normal data. Angle-based Outlier Detection (ABOD) [23] method detect outliers by considering variable correlations. Reconstruction-based approaches including Autoencoder [24], Variational autoencoder [25], which learn latent feature of normal data and then reconstruct it, are frequently used for anomaly detection. However, methods mentioned above only focus on the spatial anomalies in input data, without considering temporal dependencies of input data, which reduces the potential to detect the operational anomalies of industrial robots. Some improved models based on these methods can overcome this limit and capture temporal dependency by handling time series data. For instance, Bayer [17] introduced STORN, which combined variational inference and RNN to model time series data, and applied to anomaly detection of robot. Xu *et al.* [26] applied a variational autoencoder to anomaly detection of KPI time series data. Pereira and Silveira [27] proposed an unsupervised anomaly detection method using variational recurrent autoencoders with attention which was applied to energy time series data.

In this work, we are interested in realizing anomaly detection of industrial robots in terms of time series effect. We introduce an online unsupervised anomaly detection method for industrial robots by an unsupervised method based on sliding-window convolutional variational autoencoder (SWCVAE). It is worth remarking that the input data are gathered from the robot controller without installing extra sensor. Besides, compared to model-based method, this method doesn't require any domain knowledge of robot system. Furthermore, unlike some unsupervised methods which only focus on spatial anomalies, this method can detect spatial and temporal anomalies in data by dealing with time series data, and thus may help for recognizing the deviation of workflow under repetitive operation.

III. BACKGROUND

A. VARIATIONAL AUTOENCODER (VAE)

VAE, an important generative model, has similar network frame as Autoencoder (AE), which consists of two parts: an encoder and a decoder. In Autoencoder, the encoder defines a mapping from input data $\mathbf{x} \in \mathbb{R}^{d_x}$ to a latent variable $\mathbf{z} \in \mathbb{R}^{d_z}$, while the decoder defines a mapping back from the latent variable $\mathbf{z}$ to input space, which outputs the reconstructed $\hat{\mathbf{x}}$. The training objective of AE is to make the reconstructed term $\hat{\mathbf{x}}$ as close as the original one $\mathbf{x}$, forcing AE to learn latent features of normal data. In VAE, the latent variable $\mathbf{z}$ is constrained to be distributed according to a prior distribution $p_\theta(\mathbf{z})$, usually multivariate unit Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{I})$, forcing the model to learn the distribution of input data. However, when mapping from input data $\mathbf{x}$ to latent variable $\mathbf{z}$, according to the equation (1), $p_\theta(\mathbf{z}|\mathbf{x})$ is usually intractable since $p_\theta(\mathbf{x})$ is also intractable.

$$p_\theta(\mathbf{z}|\mathbf{x}) = \frac{p_\theta(\mathbf{x}, \mathbf{z})}{p_\theta(\mathbf{x})} \tag{1}$$

Hence, Variational Inference techniques are used to solve this problem in a tractable way by finding an approximation posterior $q_\phi(\mathbf{z}|\mathbf{x})$.

$$q_\phi(\mathbf{z}|\mathbf{x}) = N\left(\boldsymbol{\mu}_\mathbf{z}, \boldsymbol{\sigma}_\mathbf{z}^2\mathbf{I}\right) \tag{2}$$

where the mean $\boldsymbol{\mu}_\mathbf{z}$ and standard deviation $\boldsymbol{\sigma}_\mathbf{z}$ of the approximation posterior $q_\phi(\mathbf{z}|\mathbf{x})$ are derived by the encoder.

Given an inference model $q_\phi(\mathbf{z}|\mathbf{x})$, the evidence lower bound (ELBO) can be derived as follows:

$$\begin{aligned} logp_\theta(\mathbf{x}) &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log p_\theta(\mathbf{x})\right] && (3)\\ &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log \frac{p_\theta(\mathbf{x}|\mathbf{z})\, p_\theta(\mathbf{z})}{p_\theta(\mathbf{z}|\mathbf{x})}\right] && (4)\\ &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log \frac{p_\theta(\mathbf{x}|\mathbf{z})\, p_\theta(\mathbf{z})}{p_\theta(\mathbf{z}|\mathbf{x})} \frac{q_\phi(\mathbf{z}|\mathbf{x})}{q_\phi(\mathbf{z}|\mathbf{x})}\right] && (5)\\ &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})\right] \\ &\quad + D_{KL}(q_\phi(\mathbf{z}|\mathbf{x})\,||p_\theta(\mathbf{z}|\mathbf{x})) && (6) \end{aligned}$$

In Equation (6), the first term is the Evidence Lower Bound (ELBO) and the second term is the Kullback-Leibler divergence of the approximate $q_\phi(\mathbf{z}|\mathbf{x})$ from the true posterior $p_\theta(\mathbf{z}|\mathbf{x})$. To ensure $q_\phi(\mathbf{z}|\mathbf{x})$ gets closer to $p_\theta(\mathbf{z}|\mathbf{x})$, the KL divergence term between them has to be minimized. According to the equation, minimizing KL divergence can be transformed into the task of maximizing ELBO. Therefore, the loss function of VAE can be written as below:

$$\begin{aligned} \mathcal{L}_{VAE}(\theta, \emptyset; \mathbf{x}) = -\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\big[&\log p_\theta(\mathbf{x}|\mathbf{z}) \\ &+ \log p_\theta(\mathbf{z}) - \log q_\phi(\mathbf{z}|\mathbf{x})\big] \end{aligned} \tag{7}$$

VAE has been applied successfully in different domains. With sliding window, VAE can be used to realize real time anomaly detection temporally in time series data [26]. Yet, standard VAE with sliding window is only able to handle univariate time series data. Hence, the standard VAE needs to be modified to consider the dependencies of all variates.

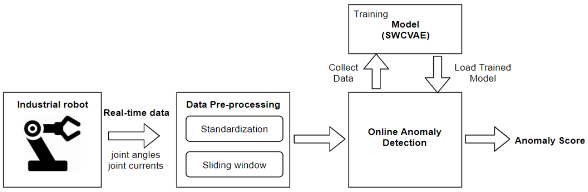

FIGURE 1. Overall architecture of online anomaly detection for industrial robots.

B. CONVOLUTIONAL NEURAL NETWORK (CNN)

Convolutional Neural Network (CNN) is a type of deep neural network, with the capability of extracting useful features by several convolutional operators. It is particularly suitable for 2-dimensional data structures, thus mostly popular for pattern recognition in image processing.

In CNN, as a weighted kernel W slides over every position of input data X, convolution operation of the input data and kernel is triggered, resulting in a feature map:

$$S(i,j) = (X * W)(i,j) \tag{8}$$
$$= \sum_m \sum_n I(i-m, j-n)W(m,n) \tag{9}$$

where S is feature map resulted by input data X and kernel W, * denotes the convolution operation.

Typically, the kernel size is smaller than the input data size, but with more in-depth. That means there are several different kernels applied to the input data at the same time, resulting in the same amount of feature maps. The weights of kernels are adjusted during the training.

Although CNN is mostly applied for analyzing images, it is also successfully explored in multivariate time series data [22]. Since multivariate time series have the same 2-dimensional data structures as image, CNN for analyzing images is suitable for handling multivariate time series as well.

IV. METHODOLOGY

In order to reflect the characteristics of multivariate time series data collected from the industrial robot, sliding window-based convolutional variational autoencoder (SWCVAE) is applied. In this model, we utilize a relatively long sliding window in order to capture temporal dependencies of periodic time series data. Besides, each coming data point instead of batch data in a time window adopted in [28] will be evaluated to be a normal or an anomaly, in order to respond to anomalies as soon as possible. The overall structure is shown in Figure 1.

A. NETWORK STRUCTURE

Since sliding window is applied over the data, the input data for timestep t is a sequence $\mathbf{x}^{(t)}$ with length T (the length of the sliding window), i.e. $\mathbf{x}^{(t)} = (\mathbf{x}_{t-T+1}, \mathbf{x}_{t-T+2}, \ldots, \mathbf{x}_t)$ and each observation in sliding window is a $d_{\mathbf{x}}$-dimensional vector. Therefore, the input data have dimensions $(T, d_{\mathbf{x}})$. Before training, the input data will be standardized to zero mean and unit variance.

The overall network structure is illustrated as Figure 2. The model consists of two main components: an encoder and a decoder, both modeled in CNN structure in order to handle multivariate time series data. In this case, Conv2d are used for extracting features from both time axis and feature axis in observations. Specifically, the encoder comprises three convolutional layers with the rectified linear unit (ReLu) activations and one flatten layer while the decoder correspondingly has three transpose convolutional layers with ReLu activations and one flatten layer. To avoid overfitting, all layers in the encoder and decoder except for flatten layers are applied with L2 regularizers that penalizes large weights in the model.

The prior distribution $p_\theta(\mathbf{z})$ over the latent variables $\mathbf{z}$ is chosen as an isotropic multivariate Normal $N(\mathbf{0}, \mathbf{I})$. The approximation posterior distribution $q_\phi(\mathbf{z}|\mathbf{x})$ and the decoding distribution $p_\theta(\mathbf{x}|\mathbf{z})$ are designed to be multivariate Normal with diagonal co-variance matrix $N\left(\boldsymbol{\mu}_\mathbf{z}, \boldsymbol{\sigma}_\mathbf{z}^2\mathbf{I}\right)$. The Gaussian parameters of these distributions are derived from the hidden features of the network. The means $\boldsymbol{\mu}_\mathbf{z}$, $\boldsymbol{\mu}_\mathbf{x}$ and standard diviations $\boldsymbol{\sigma}_\mathbf{z}$, $\boldsymbol{\sigma}_\mathbf{x}$ are derived from the final hidden state of encoder or decoder using linear layers and softplus layers respectively. The softplus activation is used to ensure the predicted standard variations greater than zero.

$$\boldsymbol{\mu}_\mathbf{z} = f_{linear}(\mathbf{W}_{\boldsymbol{\mu}_\mathbf{z}}\mathbf{h}_\mathbf{z} + \mathbf{b}_{\boldsymbol{\mu}_\mathbf{z}}) \tag{10}$$
$$\boldsymbol{\sigma}_\mathbf{z} = f_{softplus}(\mathbf{W}_{\boldsymbol{\sigma}_\mathbf{z}}\mathbf{h}_z + \mathbf{b}_{\boldsymbol{\sigma}_\mathbf{z}}) \tag{11}$$
$$\boldsymbol{\mu}_\mathbf{x} = f_{linear}(\mathbf{W}_{\boldsymbol{\mu}_\mathbf{x}}\mathbf{h}_\mathbf{x} + \mathbf{b}_{\boldsymbol{\mu}_\mathbf{x}}) \tag{12}$$
$$\boldsymbol{\sigma}_\mathbf{x} = f_{softplus}(\mathbf{W}_{\boldsymbol{\sigma}_\mathbf{x}}\mathbf{h}_\mathbf{x} + \mathbf{b}_{\boldsymbol{\sigma}_\mathbf{x}}) \tag{13}$$

where $\mathbf{h}_\mathbf{z}$ and $\mathbf{h}_\mathbf{x}$ are the vectors in the final hidden layers of the encoder and decoder respectively.

The model output is the reconstruction probabilities of each point in the sliding window. It should be noted that only the reconstruction probabilities for the last observation $\mathbf{x}_t^{(t)}(\mathbf{x}_t^{(t)}$ in the sequence $\mathbf{x}^{(t)} = (\mathbf{x}_{t-T+1}, \mathbf{x}_{t-T+2}, \ldots, \mathbf{x}_t)$ are used for the evaluation of anomaly score at timestep t so as to realize real-time anomaly detection as shown in Figure 3, following the work of Xu *et al.* [26]. In order to let the final result less affected by normal data of some dimensions, the anomaly score for timestep t is the negative of the sum of negative reconstruction probabilities. If the anomaly score is higher than a predefined threshold, then the input data for the current timestep t will be considered as anomaly.

B. TRAINING

During training, the gradients of the loss function are needed for optimization of ELBO. However, it is not easy to differentiate the loss with respect to the variational parameters Ø because the gradients cannot be backpropagated through the latent variable $\mathbf{z}$. Hence, re-parameterization trick following the work in [29] is applied to overcome this problem.

The latent variable z is assumed to be a deterministic function of $\mathbf{x}$ and a random variable $\boldsymbol{\varepsilon}$ sampled from a fixed distribution, $N(\mathbf{0}, \mathbf{I})$. Hence, the undifferentiable random variable

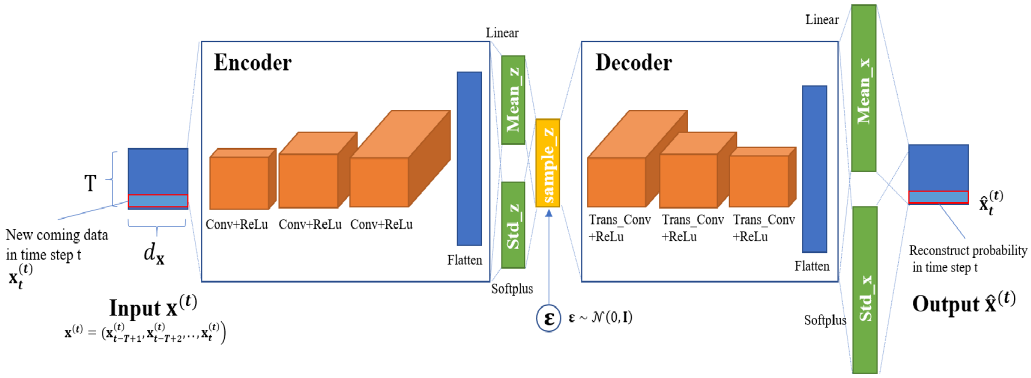

FIGURE 2. Network structure of SWCVAE.

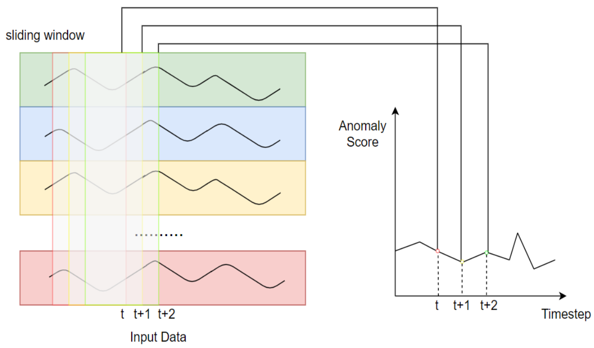

FIGURE 3. Sliding window for real-time anomaly detection.

FIGURE 4. KUKA KR6R 900 SIXX industrial robot.

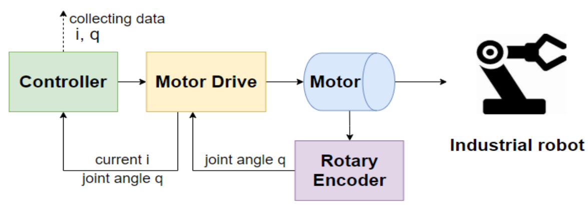

FIGURE 5. Communication mechanism among devices.

$\mathbf{z}$ is converted to a differentiable function of $\mathbf{x}$ and a random $\boldsymbol{\varepsilon}$.

$$\mathbf{z} = \boldsymbol{\mu}_{\mathbf{z}} + \boldsymbol{\sigma}_{\mathbf{z}} \odot \boldsymbol{\varepsilon} \ with\ \boldsymbol{\varepsilon} \sim N\left(\mathbf{0}, \mathbf{I}\right) \tag{14}$$

where $\boldsymbol{\mu}_{\mathbf{z}}$ and $\boldsymbol{\sigma}_{\mathbf{z}}$ are the variational parameters derived from the encoder.

In this case, Stochastic Gradient Variational Bayes (SGVB) algorithm [29] was applied to maximize the ELBO. The ELBO can be written as:

$$\mathcal{L}\left(\theta, \emptyset; \mathbf{x}\right) = \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log p_\theta\left(\mathbf{x}|\mathbf{z}\right) + \log p_\theta\left(\mathbf{z}\right) - \log q_\phi\left(\mathbf{z}|\mathbf{x}\right)\right] \tag{15}$$

$$\approx \frac{1}{L}\sum_{l=1}^{L}\left(\log p_\theta\left(\mathbf{x}\middle|\mathbf{z}^l\right) + \log p_\theta\left(\mathbf{z}^l\right) - \log q_\phi\left(\mathbf{z}^l\middle|\mathbf{x}\right)\right) \tag{16}$$

where $\mathbf{z}^l = \boldsymbol{\mu}_{\mathbf{z}} + \boldsymbol{\sigma}_{\mathbf{z}} \odot \boldsymbol{\varepsilon}^l$ with $\boldsymbol{\varepsilon}^l \sim N\left(\mathbf{0}, \mathbf{I}\right)$

The sampling number L during the training was set to 1 since one sample is already sufficient. With model loss, the negative ELBO, we trained the model using Adam optimizer [30] to update the weights of the model.

C. ONLINE ANOMALY DETECTION

In the online anomaly detection case, we load the model which has been fully trained on the historical normal data as online anomaly detector. The detector maintains a recent history of the input data stream by a time window with the length of T. For each coming data in timestep t, the corresponding window will be fed into the model. Then, the model outputs the reconstruct probability for each point in the window. Only the reconstruction probabilities for the last observation are utilized for the evaluation of anomaly score at timestep t.

V. EXPERIMENTS

A. EXPERIMENTS SETUP

The experimental work has been performed using a KUKA KR6R 900 SIXX, which is an industrial robot with six revolute joints, as shown in Figure 4. The schematic diagram of the robot system is shown in Figure 5. The joints of the industrial robot are driven by the motor, which is connected to the motor drive. The motor drive would send currents

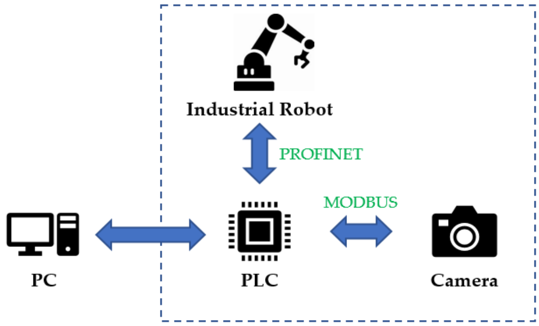

FIGURE 6. Communication mechanism among devices.

feedback to the robot controller. Rotary encoders are mounted on the output shafts of all motors to record the shaft rotating speed. This information together with the reference motion signal are fed into the motion control unit. A feedback control signal is sent back to the adjustable-speed drive to control the rotating speed of the motor. A current sensor is used to measure one-phase current signal between adjustable-speed drive and industrial robot system.

The industrial robot repeatedly performs a pick-and-place task using vision guidance system. Materials are successively placed onto the running conveyor. Then the conveyor brings them to the robot and waits the robot to operate. Next, the robot approaches and takes photo of the closest material to get the accurate target positions for grasping. Then, the robot picks it up with a vacuum gripper and places it to the fixed destination. The robot moves back to the home position and waits for the next material, beginning the next cycle. The robot is connected with Siemens PLC S7-1200 via PROFINET protocol. In the meantime, PLC also communicates with the camera on the robot via MODBUS protocol. PLC reads the accurate target positions for grasping the material from the camera and sent them to robot, then PLC controlls the vacuum grasper to help the robot to pick the item up. In this work, all the robot data are collected from the PLC, which contains robot data acquired from the robot. The overall communication mechanism among devices is shown in Figure 6.

For training and testing, we record the joint angles and joint currents of the six joints of the robot at 33.3 Hz from the robot controller. We choose the joint angles and currents as the source of input data. Joint angles are important to extract dynamics feature for the robot as adopted in [17], but when some faults such as oscillation happen on the robot, there are subtle changes in the angle configuration unless the injected force is big enough. In contrast, compared with the joint angles, joint currents are more sensitive to the external disturbance. Therefore, we choose joint angles and joint currents as input data, which may help for extracting the true dynamics of the robot acutely.

Then, the input data for training at timestep t can be expressed as $\mathbf{x}^{(t)} = (\mathbf{x}_{t-T+1}, \mathbf{x}_{t-T+2}, \ldots, \mathbf{x}_t)$ (T is the length of sliding window) and each observation in input data is a

TABLE 1. The structure of the encoder.

Layer	Output shape	Kernel size	Stride	Padding
$Conv2d$	(None,T,$d_{\mathbf{x}}$,1)	3	1	same
$Conv2d$	(None,T,$d_{\mathbf{x}}$,4)	3	1	same
$Conv2d$	(None,T,$d_{\mathbf{x}}$,16)	3	1	same
$Flatten$	(None,T$*d_{\mathbf{x}}*$16)	-	-	-
$Dense(\boldsymbol{\mu}_{\mathbf{z}})$	(None,d_z)	-	-	-
$Dense(\boldsymbol{\sigma}_{\mathbf{z}})$	(None,d_z)	-	-	-

TABLE 2. The Structure of the decoder.

Layer	Output shape	Kernel size	Stride	Padding
$Conv2d\,Transpose$	(None,T,$d_{\mathbf{x}}$,16)	3	1	same
$Conv2d\,Transpose$	(None,T,$d_{\mathbf{x}}$,4)	3	1	same
$Conv2d\,Transpose$	(None,T,$d_{\mathbf{x}}$,1)	3	1	same
$Flatten$	(None,T$*d_{\mathbf{x}}$)	-	-	-
$Dense(\boldsymbol{\mu}_{\mathbf{x}})$	(None,T$*d_{\mathbf{x}}$)	-	-	-
$Dense(\boldsymbol{\sigma}_{\mathbf{x}})$	(None,T$*d_{\mathbf{x}}$)	-	-	-

12-dimensional vector. For example, the observation at timestep t $\mathbf{x}_t$ comprises of 12 signals including 6 joint angles and 6 joint currents. The input data would be normalized to zero mean and unit variance before training.

We used normal data produced by the robot during normal operation to train the model so as to ensure the model only learns pattern from the normal data. During the experiments, the length of the sliding time window was set to 250. And the size of latent dimension was set to 10. For L2 regularization, the lambda for kernel regularization was set to 10^{-3}. Batch size of 256 and an initial learning rate of 10^{-3} in Adam optimizer were used in the experiments. The detailed network structure is shown in the Table 1 and Table 2.

In order to assess the performance of the proposed model, the model should be tested on both normal data and the realistic fault data, yet which is difficult. In order to simulate faults, collisions have been induced by manually hitting the robot. Finally, 29349 normal samples and 15018 test samples containing anomalies were recorded to develop the anomaly detection model.

B. EXPERIMENTS RESULTS

Figure 7 and Figure 8 show a segment of test data (including anomalies) of the robot. The window presented in red line represents abnormal period. The robot motions are based on visual guidance instead of fixed trajectory, therefore there are slight differences in joint angles and currents in the same operation of different cycles during the pick-and-place task.

From Figure 7 and Figure 8 we could see that as the robot repeatedly performed a pick-and-place task, the joint angles and joint currents show their periodicity. Besides,

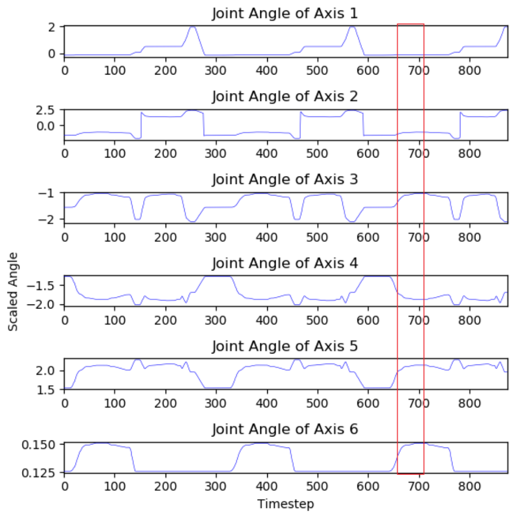

FIGURE 7. Joint angles of the industrial robot under operation.

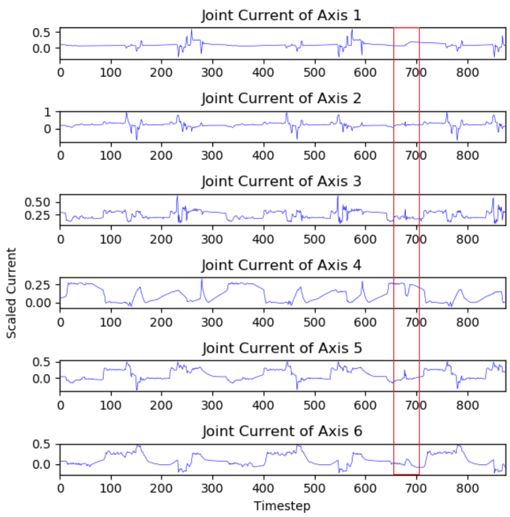

FIGURE 8. Joint currents of the industrial robot under operation.

it is presented that there are always some large transitory fluctuations in both joint currents and joint angles for a short time triggered by different robot motions, which may be a difficulty to distinguish normal pattern from abnormal behaviors. It should be noted that when injecting a fault to the robot (as presented as red background on the charts), joint currents are more sensitive to the external disturbance than joint angles.

For comparison, the experiments were conducted with SWCVAE, Multi-VAEs, PCA-VAE, OCSVM, ABOD, KNN. Multi-VAEs constructed several sliding-window standard VAE models for all variables in input data respectively, and then took the negative of sum of negative parts of reconstruct probabilities of models as the final anomaly score. Also, we exploited PCA to transform multivariate time series data into univariate time series, which would then be handled by standard sliding-window VAE, which is referred to as PCA-VAE. Both Multi-VAEs and PCA-VAE have the same sliding window length of 250 as SWCVAE. We also tested some typical anomaly detection methods like OCSVM, ABOD, KNN by means of a python toolkit PyOD [31].

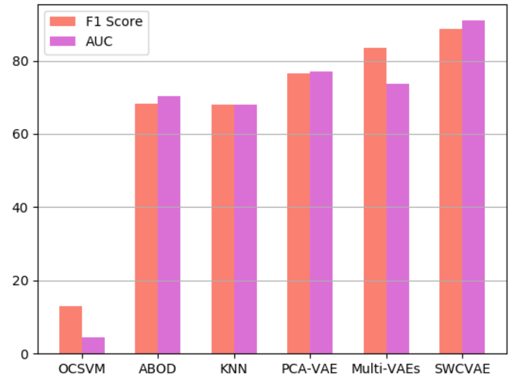

FIGURE 9. The Performance of different methods.

There are several metrics can be used to be the indicator of performance of anomaly detection algorithm. In this paper, we used Precision, Recall, F_1 Score, and the Area Under the Precision Recall Curve (PRAUC) to evaluate the performance of models. High precision relates to low false alarm rate, yet could not guarantee low miss alarm rate. Recall indicates the sensitivity of the model to the anomaly, but high Recall may also leads to high false alarm. F_1 Score is the harmonic mean of Precision and Recall, which can be a reliable indicator of performance of models. PRAUC is a comprehensive metric which can indicate the global performance of the model with considering wide range of threshold. Therefore, we mainly focus on F_1 Score and PRAUC.

$$Precision = \frac{TP}{TP + FP} \tag{17}$$

$$Recall = \frac{TP}{TP + FN} \tag{18}$$

$$F_1 = \frac{2 * Precision * Recall}{Precision + Recall} \tag{19}$$

where TP is the correctly predicted positive values (Predicted: Abnormal, Actual: Abnormal), TN is the correctly predicted negative values (Predicted: Normal, Actual: Normal), FP is falsely predicted positive values (Predicted: Abnormal, Actual: Normal), FN is falsely predicted negative values (Predicted: Normal, Actual: Abnormal).

Finally, the detailed performances of different approaches are reported in Table 3. The Figure 9 shows collision data evaluated with different methods.

As can be seen in Table 3, SWCVAE performs best over all methods in F_1 Score and PRAUC, followed by Multi-VAEs and PCA-VAE. Although Multi-VAEs has higher Recall than SWCVAE, constructing several model for each variate would inevitably reduce efficiency. PCA-VAE performs worse than SWCVAE, which may indicate that some significant information were thrown when using PCA to compress input data into one dimension data. Overall, methods based on

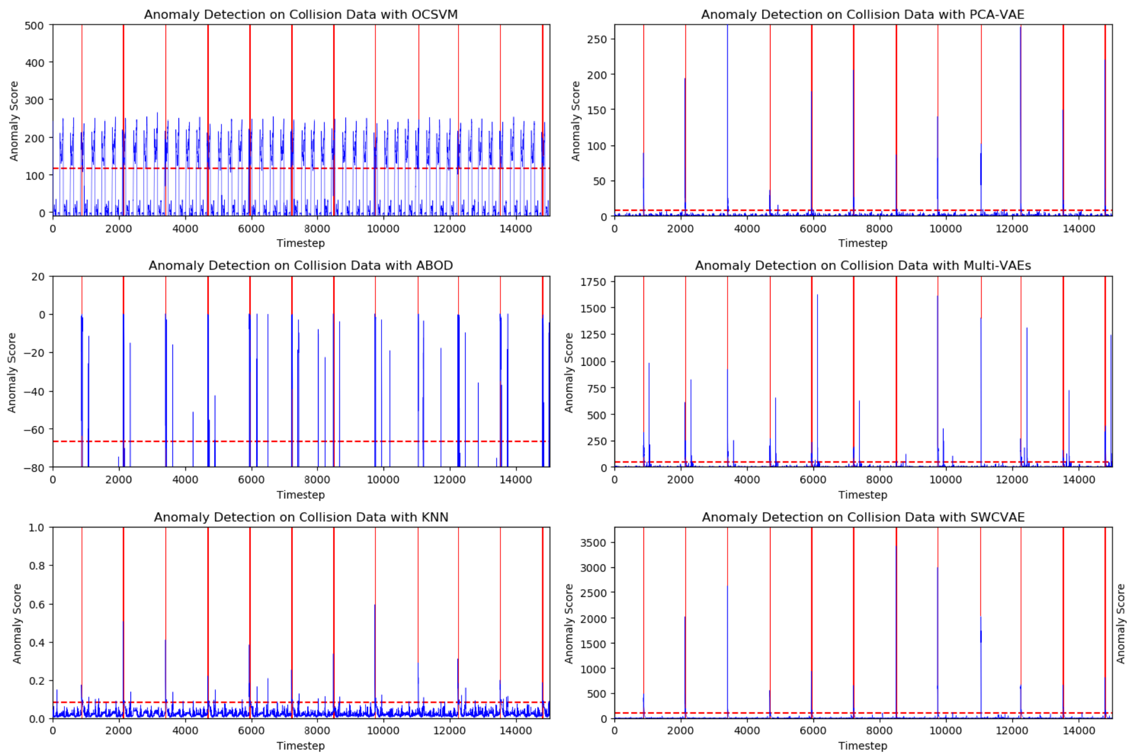

FIGURE 10. Collision data evaluated by different methods. Red background represents anomalies.

TABLE 3. The performances of different methods.

Method	Precision	Recall	F_1 Score	PRAUC
$OCSVM$	6.9	98.53	12.89	4.57
$ABOD$	60.22	78.22	68.28	70.27
KNN	62.78	74.41	68.10	68.02
$PCA-VAE$	87.5	67.94	76.49	77.1
$Multi-VAEs$	76.66	91.76	83.53	73.79
$SWCVAE$	96.54	82.06	**88.71**	**90.93**

sliding-window VAE outperform than traditional methods. Among traditional methods, OCSVM performs worst.

The detailed performances with different models are presented in Figure 10. OCSVM can hardly detect any anomalies in robot data. A possible explanation for this might be that OCSVM could not find internal anomalies within normal boundary. ABOD can detect almost all collisions, but there is also a fairly large amount of false positives.KNN can also identify anomalies well and anomaly scores of abnormal data are much higher than normal data, yet it is still not stable due to a great amount of noises of anomaly scores which lead to many false positives.As for PCA-VAE, its performancce is fairly good and can detect all anomalies with small noises. From the result of Multi-VAEs, we could observe that it provides clearer peaks of anomaly scores in abnormal data. Nonetheless, there are still a lot of false positives in Multi-VAEs. SWCVAE can detect all collisions with clearer peaks without false positives, indicating that SWCVAE successfully extracts latent features sensitive to faults. Overall, SWCVAE comprehensively outperforms other unsupervised algorithms in experiments.

C. DISCUSSION

There are two significant parameters in SWCVAE model, including the length of sliding window and the size of latent dimension. Hence, in this section, the impact of these two parameters would be discussed.

The length of sliding window on one hand would play an important role in the performance of capturing normal pattern from data. On the other hand, the length of sliding window relating with computing cost can also affect the performacne of online detection. Actually, the selection of the length of sliding window may greatly depend on the characteristic of data and their data pattern. Therefore, the discussion would not help to find the best choice for general data. Instead, discussion just provides some references of selection.

Figure 11 presents the performance with different lengths of sliding window (from 50 to 600 with internal of 50) on testing data under **z** dimension of 10 by comparing F_1 Score, PRAUC and detect time (ms). However, it is not easy to recognize patterns from the results. Nevertheless, we could see that SWCVAE can achieve good performances with window length of 200 and 250, with both F_1 Score and PRAUC around 90. When the length is lower than 200, the results are

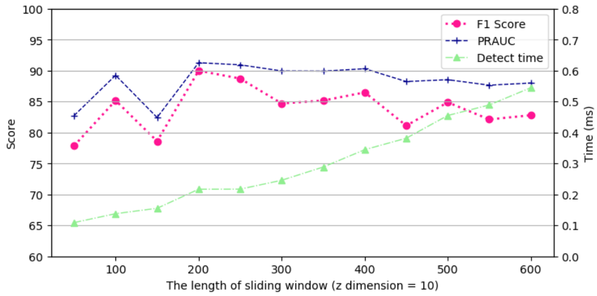

FIGURE 11. Performance of SWCVAE with different lengths of sliding window under z dimension of 10.

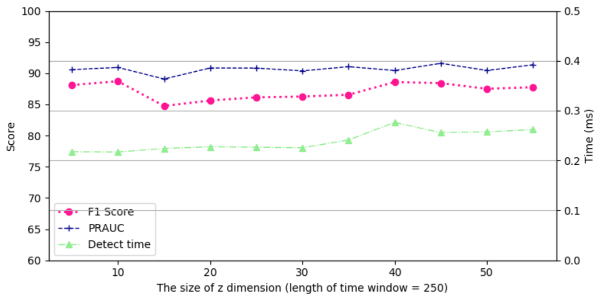

FIGURE 12. Performance of SWCVAE with different sizes of latent dimension under the length of sliding window of 250.

unstable. But when it comes to the length of sliding window between 250 to 400, PRAUC remains relatively stable while F_1 Score has downward trend. When the length of sliding window is bigger than 400, the performance in these two indicators does not perform as well as performances with the length from 200 to 400. Therefore, it is recommended not to choose long sliding window in consideration of its impact and and computing cost on time.

As for the size of latent dimension, the performance of SWCVAE with different sizes of latent dimension (from 5 to 55) under the length of sliding window of 250 is presented in the Figure 12 by comparing F_1 Score, PRAUC and detect time (ms). It can be seen that the performance in terms of PRAUC keeps fairly good among all tested z dimension while high F_1 Score can be achieved with relatively small or big size of latent dimension. Overall, the size of z dimension has relatively small impact on the performance of SWCVAE.

VI. CONCLUSION

In this paper, we presented an online unsupervised anomaly detection method based on VAE with CNN structure for the industrial robot. This method adopts a relatively long sliding window in order to better recognize normal patterns of data. In addition, CNN structure is embedded in the encoder and decoder of VAE model to extract temporal and spatial features. In general, this method has several following contributions. First, this method does not require any additional sensor and any prior domain knowledge for the robot. The robot data are collected from robot controller with relatively low frequency. Besides, the model SWCVAE is an unsupervised method without requiring labeled data. It only learns the normal pattern from normal time series data and detects anomalies by recognizing the unseen pattern of data, which can help to save effort and time for collecting fault data. In addition, this method can be applied online without accumulating data, thus speeding response to anomalies. Furthermore, this method can automatically extract effective features sensitive to the faults instead of robot motions by using CNN structure which helps for capturing the dependencies of the input data. Overall, the experiments show that our model is capable of reliably detecting unknown spatial and temporal anomalies on the industrial robot. Thus, this work presents a promising tool for condition-based maintenance of industrial robots.

However, despite the good performance obtained by our model, there are still some researches to be done in the future. Firstly, the model was tested by only one anomaly scenario and only provided anomaly detection. Therefore, in the future work, more study on testing under different anomaly scenarios and analyzing anomalies would be performed. Secondly, up to now, our model is partially online because the process of training model is offline which means our model is fixed and cannot adapt to new normal condition. Hence, finding a way to improve the capability of the model in online learning is the next research direction in the future.

REFERENCES

[1] International Federation of Robotics(IFR). (2019). *Executive Summary: World Robotics 2019 Industrial Robots*. Accessed: Oct. 1, 2019. [Online]. Available: https://ifr.org/downloads/press2018/Executive%20Summary%20WR%202019%20Industrial%20Robots.pdf

[2] R. Hornung, H. Urbanek, J. Klodmann, C. Osendorfer, and P. van der Smagt, "Model-free robot anomaly detection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 3676–3683.

[3] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, Oct. 2006.

[4] R. C. M. Yam, P. W. Tse, L. Li, and P. Tu, "Intelligent predictive decision support system for condition-based maintenance," *Int. J. Adv. Manuf. Technol.*, vol. 17, no. 5, pp. 383–391, Feb. 2001.

[5] D. Brambilla, L. M. Capisani, A. Ferrara, and P. Pisu, "Fault detection for robot manipulators via second-order sliding modes," *IEEE Trans. Ind. Electron.*, vol. 55, no. 11, pp. 3954–3963, Nov. 2008.

[6] S. Chen, M. Luo, and F. He, "A universal algorithm for sensorless collision detection of robot actuator faults," *Adv. Mech. Eng.*, vol. 10, no. 1, Jan. 2018, Art. no. 168781401774071.

[7] F. Piltan and J. Kim, "FPGA-based ARX-Laguerre PIO fault diagnosis in robot manipulator," *J. Intell. Robotic Syst.*, vol. 2, no. 1, pp. 99–112, 2018.

[8] L. M. Capisani, A. Ferrara, A. Ferreira de Loza, and L. M. Fridman, "Manipulator fault diagnosis via higher order sliding-mode observers," *IEEE Trans. Ind. Electron.*, vol. 59, no. 10, pp. 3979–3986, Oct. 2012.

[9] A. A. Jaber and R. Bicker, "Industrial robot fault detection based on statistical control chart," *Amer. J. Eng. Appl. Sci.*, vol. 9, no. 2, pp. 251–263, Feb. 2016.

[10] V. Sathish, S. Ramaswamy, and S. Butail, "Training data selection criteria for detecting failures in industrial robots," *IFAC-PapersOnLine*, vol. 49, no. 1, pp. 385–390, 2016.

[11] R. Muradore and P. Fiorini, "A PLS-based statistical approach for fault detection and isolation of robotic manipulators," *IEEE Trans. Ind. Electron.*, vol. 59, no. 8, pp. 3167–3175, Aug. 2012.

[12] F. Cheng, A. Raghavan, D. Jung, Y. Sasaki, and Y. Tajika, "High-accuracy unsupervised fault detection of industrial robots using current signal analysis," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2019, pp. 1–8.

[13] E. Olsson, P. Funk, and M. Bengtsson, "Fault diagnosis of industrial robots using acoustic signals and case-based reasoning," in *Advances in Case-Based Reasoning*, P. Funk and P. A. G. Calero, Eds. Berlin, Germany: Springer, 2004, pp. 686–701.
[14] A. A. Jaber and R. Bicker, "The optimum selection of wavelet transform parameters for the purpose of fault detection in an industrial robot," in *Proc. IEEE Int. Conf. Control Syst., Comput. Eng. (ICCSCE)*, Nov. 2014, pp. 304–309.
[15] A. A. Jaber and R. Bicker, "Fault diagnosis of industrial robot gears based on discrete wavelet transform and artificial neural network," *Insight-Non-Destructive Test. Condition Monitor.*, vol. 58, no. 4, pp. 179–186, Apr. 2016.
[16] D. Park, S. Kim, Y. An, and J.-Y. Jung, "LiReD: A light-weight real-time fault detection system for edge computing using LSTM recurrent neural networks," *Sensors*, vol. 18, no. 7, p. 2110, 2018.
[17] M. Soelch, J. Bayer, M. Ludersdorfer, and P. van der Smagt, "Variational inference for on-line anomaly detection in high-dimensional time series," 2016, *arXiv:1602.07109*. [Online]. Available: http://arxiv.org/abs/1602.07109
[18] S. Vallachira, M. Orkisz, M. Norrlof, and S. Butail, "Data-driven gearbox failure detection in industrial robots," *IEEE Trans Ind. Informat.*, vol. 16, no. 1, pp. 193–201, Jan. 2020.
[19] A. Datta, S. Patel, C. Mavroidis, I. Antoniadis, J. Krishnasamy, and M. Hosek, "Fault diagnostics of industrial robots using support vector machines and discrete wavelet transforms," in *Proc. Int. Mech. Eng. Congr. Expo. (ASME)*, 2006, pp. 245–251.
[20] E. Khalastchi, M. Kalech, G. A. Kaminka, and R. Lin, "Online data-driven anomaly detection in autonomous robots," *Knowl. Inf. Syst.*, vol. 43, no. 3, pp. 657–688, Jun. 2015.
[21] F. Angiulli and C. Pizzuti, "Fast outlier detection in high dimensional spaces," in *Principles of Data Mining and Knowledge Discovery*, T. Elomaa, H. Mannila, and H. Toivonen, Eds. Berlin, Germany: Springer, 2002, pp. 15–27.
[22] B. Schölkopf, R. C. Williamson, A. J. Smola, J. Shawe-Taylor, and J. C. Platt, "Support vector method for novelty detection," in *Proc. Adv. Neural Inf. Process. Syst.*, 2000, pp. 582–588.
[23] H.-P. Kriegel, M. S. Hubert, and A. Zimek, "Angle-based outlier detection in high-dimensional data," in *Proc. 14th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2008, pp. 444–452.
[24] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
[25] J. An and S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability," *Special Lect. IE*, vol. 2, no. 1, pp. 1–18, Dec. 2015.
[26] H. Xu, Y. Feng, J. Chen, Z. Wang, H. Qiao, W. Chen, N. Zhao, Z. Li, J. Bu, Z. Li, Y. Liu, Y. Zhao, and D. Pei, "Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in Web applications," in *Proc. World Wide Web Conf. World Wide Web (WWW)*, 2018, pp. 187–196.
[27] J. Pereira and M. Silveira, "Unsupervised anomaly detection in energy time series data using variational recurrent autoencoders with attention," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2018, pp. 1275–1282.
[28] D. Kim, H. Yang, M. Chung, S. Cho, H. Kim, M. Kim, K. Kim, and E. Kim, "Squeezed convolutional variational AutoEncoder for unsupervised anomaly detection in edge device industrial Internet of Things," in *Proc. Int. Conf. Inf. Comput. Technol. (ICICT)*, Mar. 2018, pp. 67–71.
[29] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, *arXiv:1312.6114*. [Online]. Available: http://arxiv.org/abs/1312.6114
[30] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[31] Y. Zhao, Z. Nasrullah, and Z. Li, "PyOD: A Python toolbox for scalable outlier detection," *J. Mach. Learn. Res.*, vol. 20, no. 96, pp. 1–7, Jan. 2019.

TINGTING CHEN received the B.S. degree in mechanical engineering from South China Agricultural University, China. She is currently pursuing the M.S. degree in mechanical engineering with Tsinghua University. Her current research interests include robotics and digital twins.

XUEPING LIU received the Ph.D. degree from Tsinghua University, China, in 2000. He is currently an Associate Professor with the Division of Advanced Manufacturing, Graduate School at Shenzhen, Tsinghua University. His main research interests include green manufacturing process and intelligent control.

BIZHONG XIA received the B.S. degree and the M.S. degree in mechanical engineering from Chongqing University, China, in 1992 and 1995, respectively, and the Ph.D. degree in fluid transmission and control from Zhejiang University. He is currently an Associate Professor with the Division of Advanced Manufacturing, Graduate School at Shenzhen, Tsinghua University. His main research interests include battery management systems (BMS), robotics, and digital twins.

WEI WANG, photograph and biography not available at the time of publication.

YONGZHI LAI, photograph and biography not available at the time of publication.