Fault Detection of Harmonic Drive Using Multiscale Convolutional Neural Network

Guo Yang, Yong Zhong, *member*, *IEEE*, Lie Yang, Ruxu Du

***Abstract*— Harmonic drive is a key component in industry robot. Its complex design is highly sensitive to the manufacturing and assembly errors. Even a small error could cause excessive vibration jeopardizing the performance of the robot. Moreover, the vibration is dependent on the operation conditions. As a result, it is necessary to carry out a comprehensive test to detect possible manufacturing and assembling faults. This paper presents an intelligent approach that can automatically identify the different health conditions of harmonic drives. First, the acquired data from multiple acceleration sensors are cascaded in sequence after several preprocessing steps. Then, the multiscale convolutional neural network (MSCNN) architecture is used for disposing the inherent multiscale characteristics of the harmonic drive vibration signal, which can simultaneously perform multiscale feature extraction and classification. Finally, a large number of experiments were conducted on a real industrial robot vibration test platform to evaluate our approach. Based on the experiment results, the fault detection classification accuracy is 96.79%, which is higher than the other compared methods. This demonstrates that the presented method is effective for shop floor applications.**

***Index Terms*—Harmonic drive, intelligent fault detection, Multiscale Convolutional Neural Network (MSCNN), Multiple working conditions.**

I. INTRODUCTION

HARMONIC drives are widely used in industrial robots to increase the amount of torque transferred to or from the end of one shaft due to the advantages of high transmission ratio, no backlash, high compactness and light weight, good resolution and excellent repeatability [1]. Because of its complex design, the harmonic drive is sensitive to the manufacturing and assembly errors. Even a small error could result in excessive vibration jeopardizing the performance of the robot. Moreover, its behavior is dependent on the operation conditions. Therefore, it is necessary to carry out comprehensive tests to detect possible manufacturing and assembly faults.

Currently, the fault detection of harmonic drive is mainly done by shop floor technicians using vibration analysis instrument. The decision is objective and often inaccurate, and it is usually not clear how to carry out repair and/or improvement. Besides, there are only a few studies about the condition monitoring and fault detection for harmonic drives, and few satisfactory results come out due to the complex structures which include high-precision rigid and flexible components [2]. Limited research has studied the disassembled parts of the harmonic drive [3-5], such as reduction gears and bearings. These investigations have a number of limitations. First, in these papers, the harmonic drives are tested as an isolate component. Mounting a harmonic drive on an industry robot and running the robot with different operation conditions could result in rather different outcomes. Second, most of the studies focus on the bearings or gears of harmonic drives, not the entire drive. The manufacturing error and assembly error are not considered. Third, we noticed that the harmonic drive is highly nonlinear. Its behavior is dependent on many factors, such as its operating conditions. Therefore, more comprehensive investigation is needed.

There are many fault detection methods for rotating machinery equipment (except for harmonic drives), such as the conventional time domain and time-frequency domain methods as well as the machine learning methods [6]. For example, Yan *et al*. [7] showed an example of bearing defect detection through unified time-scale frequency analysis. Liu *et al*. [8] used a time-domain analysis of vibration signals to diagnose gearbox faults. Though, these time domain and time-frequency domain methods request some prior knowledge and considerable amount of online computational resources, which limits their effectiveness and flexibility [9]. The machine learning methods have been also widely used in the field of machinery fault detection. In general, machine learning methods include four parts: data acquisition, feature extraction, learning and classification. Based on literatures, a large number of machine learning methods are available, such as Support Vector Machine (SVM), Backpropagation (BP), Decision Tree (DT), K-Nearest Neighbor (KNN) and etc. For example, the raw signal was processed and feature extracted by wavelet denoising, and then adopted SVM to fault detection [10,11]. However, these machine learning techniques need to manually extract representative features from the vibration signals, which requires a lot of time and some prior knowledge [12–14]. Unlike machine learning methods mentioned above, deep learning methods such as convolutional neural networks (CNN) and recursive neural networks (RNN) only rely on a large amount of original data and can directly obtain the classification

Research was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (class A) (Grant No. XDA22040203), the Fundamental Research Funds for the Central Universities (Grant No. 2019XX01), Guangdong Marine Economic Development Project (2020), and the Natural Science Foundation of Guangdong Province (Grant No. 2020A1515010621) (*Corresponding author: Yong Zhong*)

Guo Yang, Yong Zhong, Lie Yang, and Ruxu Du are with Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou, China (email: 1318458639@qq.com, zhongyong@scut.edu.cn, 201810100415@mail.scut.edu.cn, duruxu@scut.edu.cn)

results from the raw signals. These methods include feature extraction and classification modules, which can avoid the loss of representative features without manually designing features. Recently, CNN is popularized in the field of fault detection, especially for abnormal detection of rotating machinery and equipment [15,16]. Nevertheless, most of the current studies are conducted using public data sets or ideal experimental environments, which obtain poor result in actual engineering scenarios.

In this paper, the harmonic drive is well known as a very delicate and complex mechanical component. It consists of flexspline, circular spline, and wave generator, and the collected vibration signals are usually coupled with other subsystems, such as motor, pulley, and cam shaft. Thus, measured vibration signals will contain multiple intrinsic oscillatory modes due to the interaction and coupling effects. In the experimental design stage, vibration signals from multiple sensors should be fused to create more representative information. For example, in [17], a multi-sensor feature-level fusion method was presented for bearing fault classification. In [18], a feature-level fusion technique based on Stacked Autoencoder (SAE) and Deep Neural Network (DNN) were presented for rotating machinery. Additionally, harmonic drive usually works in variable operation conditions, especially under different speeds or loads. Thus, the vibration signals are characterized with nonlinearity and nonstationarity, along with strong environment noise. In summary, vibration signals from harmonic drives on practical working scenarios usually exhibit multiscale characteristics and contain complex patterns at multiple time scales [19–22]. Therefore, we use multi-scale extraction which is better than only using the original signal and can improve the accuracy of fault classification. In terms of network structure, researchers prefer to use the 2-D CNN model [15,16] to diagnose bearing faults, which is good at processing high-dimensional data. Thus, we adopt 2-D CNN in our approach.

In summary, in this study, an intelligent method was presented to detect the fault of the harmonic drive and tested in the real industrial robot test platform. First, the collected vibration signals from multiple acceleration sensors were preprocessing and cascaded in sequence. Then, the multiscale decomposition is employed to capture complementary and rich feature information of the raw vibration signals at different scales. Finally, the reconstructed multiscale signals are fused to many specified 2-D signal matrices to match the structure of 2-D CNN which is utilized to automatically recognize features. From the actual experimental results and method comparison, we can conclude that our method can enhance the feature learning capability and therefore improve the fault detection performance. The main contributions of this paper are summarized as follows.

1) The proposed multiscale convolutional network (MSCNN) method provides an end-to-end learning method without additional signal processing and diagnostic expertise for harmonic drive fault detection.
2) The multiscale learning scheme in MSCNN can capture complementary and rich diagnostic characteristics of the vibration signal under multiple working conditions. The test shows that MSCNN is superior to traditional methods in terms of robustness and classification performance.
3) This is the very first intelligent fault diagnosis method for an intact harmonic drive, not just for disassembled parts of it, which shows considerable potential for practical applications.

The rest of this paper is organized as follows. Section II details our proposed approach for harmonic drive fault detection. Section III presents industrial robot vibration test platform. Section IV analyzes the performance of multiscale feature extraction and demonstrates the experimental results to evaluate the effectiveness of the proposed model and compared it with other methods. Finally, the conclusion and outlook of the paper are given in Section V.

II. Materials and Methods

This section presents the Spearman rank correlation, data preprocessing, the proposed MSCNN architecture, and the online and offline diagnostic systems. This method is used for fault detection of harmonic drive under multiple working conditions. The flowchart of the proposed method is shown in Fig. 1. First, the goal of correlation analysis is to choose which data set to use among the three channels. Second, data preprocessing is applied to maximize the information acquired from the sensor signal and normalize the data. Third, MSCNN is used to detect the faulty harmonic drive. Finally, through online and offline diagnostic systems, a large number of experiments were conducted on the shop floor.

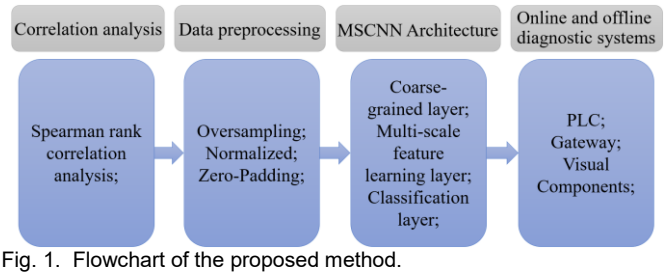

Fig. 1. Flowchart of the proposed method.

A. Spearman Rank Correlation Analysis

There are three common processing methods for correlation coefficients on data. Pearson method is used for linear correlation and normal distribution data. Kendall method is used for non-normally distributed data of an unordered sequence. Spearman correlation coefficient method is used for non-linear and non-normally distributed data. The experimental platform is often equipped with many sensors, and some of the collected feature data may have little effect, which not only increases the amount of data storage and the complexity of the model but also reduces the classification accuracy [23]. The vibration data of different working conditions are non-linearly correlated, thus the Spearman correlation coefficient ranking method is adopted to select the useful channel.

The computation procedure of Spearman method is as follows: First, sort the two data set $\{x_i\}$ and $\{y_i\}$ to get the ordered data set $\{x'_i\}$ and $\{y'_i\}$. Next, compute $d_i = x'_i - y'_i$, these are called the ranks. If there is no repetitive rank, the Spearman correlation value ρ_s can be calculated by [24]

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3024355, IEEE Transactions on Instrumentation and Measurement

$$\rho_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{1}$$

where n is the number of variables in the data sets. If the repetitive rank exists, then following formula shall be used:

$$\rho_s = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{2}$$

The calculation of the importance of each channel through the Spearman coefficient method will help us to choose the appropriate channel for neural network calculation and optimization.

B. Data Preprocessing

The larger the number of data samples, the more conducive to the training of the neural network. To ensure that the data quantity is sufficient, one-dimensional continuous vibration acceleration data is overlapped and sampled [15]. The principle of this method is achieving data augmentation by moving a fixed sampling window. Fig. 2 shows the collected one-dimensional acceleration vibration signal. Assuming that the total number of signals is N_1, and set N_2 consecutive signals as a sample unit. If the data does not overlap, only N_1 / N_2 samples can be constructed. The purpose of overlapping sampling is to increase the number of sample units by offsetting some data and constructing sample units with the same data size.

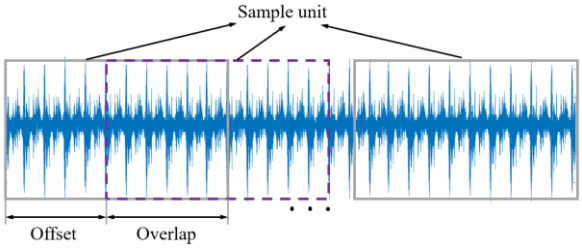

Fig. 2. Oversampling illustration.

The raw vibration acceleration data collected from the sensor needs to be normalized before being used as input to the neural network. The reconstructed input signals are stacked into many 2-D signal matrices to form the images. The value range of an image pixel is 0 ~ 255, and the value range of the vibration acceleration data in this paper is 0 ~ 50. Therefore, the vibration acceleration data needs to be normalized to the range of 0 ~ 255 according to Equation (3). Each input data normalized to 0 ~ 1 can help speed up the convergence of the training network [25]. Therefore, Equation (4) was used to further process the input data to the neural network.

$$Result = \frac{255\times(Value + Range)}{2\times Range} \tag{3}$$

$$scaled\ value = \frac{value\times 2}{255} - 1 \tag{4}$$

The zero-padding method is used to make full use of the information and avoid the loss of data dimension, which can improve the accuracy of the neural network. During the training and testing process of the neural network, the data are processed by zero-padding method to ensure that the analysis results are not affected by the increase of edge data. Let M be the input size, N be the output size, F be the filter width, and S be the stride. The number of padding on left PL and right PR can be calculated by the following equations [15].

$$N = ceil(\frac{M}{S}) \tag{5}$$

$$PT = (N-1)\times S + F - M \tag{6}$$

$$PL = floor(\frac{PT}{2}) \tag{7}$$

$$PR = PT - PL \tag{8}$$

where ceil(.) and floor(.) are the ceil and floor functions, respectively. To vividly explain the principle of this method, a simple example of the one-dimensional filling process is given below, as shown in Fig. 3. The parameters are M = 3, S = 1, F =5, then the padding result will be N = 3, PL = 2, PR = 2.

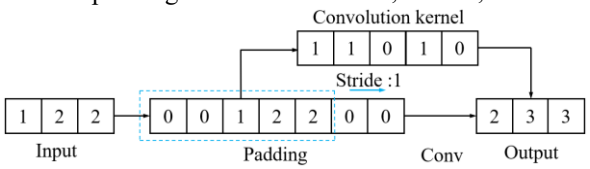

Fig. 3. Zero-Padding.

C. MSCNN Architecture

The MSCNN architecture consists of four parts: a multisensor data fusion layer, coarse-grained layer, a multi-scale feature learning layer, and a classification layer. The multisensor data fusion layer is a reliable measure for vibration test analysis. The core of the MSCNN method is to add the multiscale feature extraction step to the traditional CNN method, which means adding the multiscale coarse-grained layer. The system framework of MSCNN is shown in Fig. 4.

1) Multisensor data fusion layer:

In the vibration test platform, the mechanical components need to be tested in three orthogonal directions. Firstly, the vibration signals collected from each sensor are preprocessed according to Equations (3) and (4). Then, the vibration acceleration data are successively cascaded to construct each sample unit. The fusion data is used as input to the next module, which is beneficial to provide comprehensive information for subsequent detection and analysis.

2) Multiscale Coarse-grained layer:

The multiscale coarse-grained layer is used to extract complementary and rich feature information of the raw vibration signals at different scales. For a given measured signal $x = \{x_1, x_2, ..., x_N\}$, x_i is the value corresponding to the timestamp i, and each signal has a total of N timestamp values. By using a multiscale factor s, a continuous coarse-grained signal $\{y^{(s)}\}$ can be obtained from the average data from the original signal. This paper uses a 4-layer scale decomposition, and each element of the coarse-grained signal can be formulated as

$$y_j^{(s)} = \frac{1}{s}\sum_{i=(j-1)s+1}^{js} x_i, 1 \le j \le \frac{N}{s}. \tag{9}$$

After obtaining vibration information at 4 scales, the core issue is to fuse these signals effectively. In this module, we construct a 2D-CNN network structure. The feature extraction parameters are given in Fig. 4. The data fusion conversion of multiscale 1-D signal to 2-D bitmap input is shown in Fig. 5.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3024355, IEEE Transactions on Instrumentation and Measurement

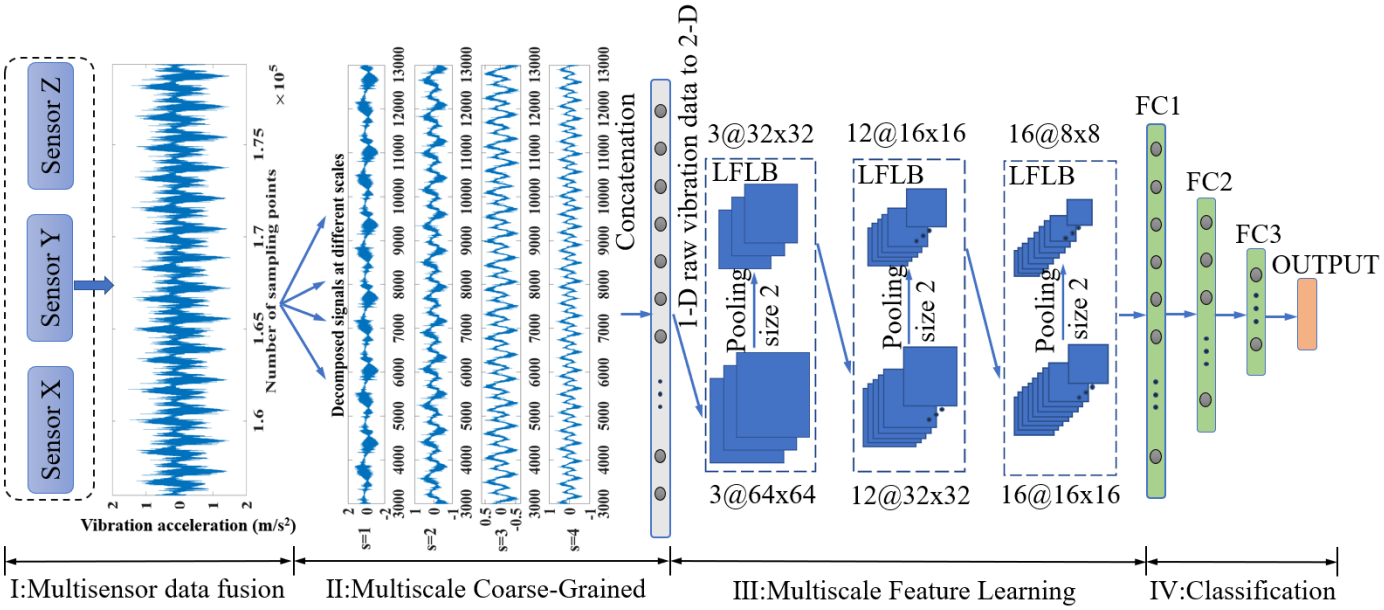

Fig. 4. The architecture of MSCNN.

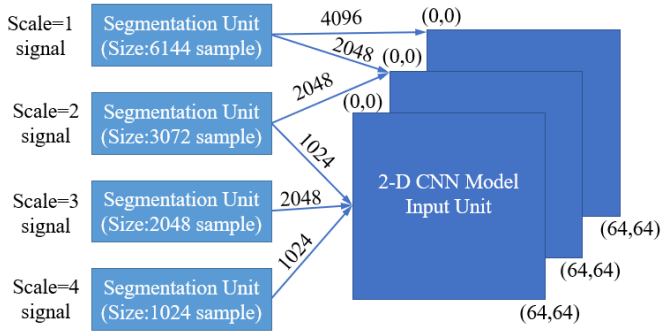

Fig. 5. The data fusion of multiscale 1-D signal to 2-D bitmap input.

3) Multi-scale Feature Learning layer:

The goal of this layer is to extract high-level effective fault features from multiscale coarse-grained signals. The multiscale feature learning layer consists of many pairs of convolution and pooling modules. Concretely, the convolutional module convolves the input local regions with filter kernels and each kernel is convoluted across the input vector and producing a feature vector. To reduce computational complexity and effectively, the function of pooling operation is used to reduce the spatial dimension [26].

The local feature learning blocks (LFLBS) module consists of a convolutional layer and a pooling layer. In this architecture, there are 3 LFLBS in the multi-scale feature learning layer.

4) Classification layer:

The classification task for the harmonic drive fault detection in this paper is a binary classification problem. The multiscale representation obtained in the feature learning layer is directly fed to the classification layer. The fully connected layer with Rectified Linear Unit (RELU) units and a softmax function are used to output the class [27]. Class 0 means that the harmonic drive is in a normal condition, and class 1 means abnormal.

The neural network structure has 3 pairs of convolution and pooling layers and 3 fully connected layers. The details of the structural parameters of the MSCNN model are shown in Table I. The kernel size, stride, and padding parameters are obtained through debugging experiences and a lot of validated experiments. These parameters were used to train and get the fault classification results. In a data sample unit, the 4 dimensions of the coarse-grained layer occupy 6144, 3072, 2048, and 1024 data, respectively，which related to the input size of neural nets.

TABLE I
STRUCTURAL PARAMETERS OF THE MSCNN MODEL

Layer name	MSCNN Models
MS Layer	[6144 3072 2048 1024]
Conv1	Conv2d(3,3,kernel size=(5,5),stride=(1,1),padding=(2,2))
Conv2	Conv2d(3,12,kernel size=(5,5),stride=(1,1),padding=(2,2))
Conv3	Conv2d(12,16,kernel size=(5,5),stride=(1,1),padding=(2,2))
Pooling1/2/3	Kernel size=(2,2)
Fc1	Linear(in features=1024,out features=512,bias=True)
Fc2	Linear(in features=512,out features=80,bias=True)
Fc3	Linear(in features=80,out features=2,bias=True)

It shall be concluded that the MSCNN architecture includes the following steps: multisensor data fusion, multiscale coarse-grained feature layer to extract the inherent characteristics of the vibration signal in different time scale, using multiple pairs of convolutions and pooling layers, as well as the fully connected layers to detect the faulty harmonic drive. The key difference between traditional multi-time CNN (MCNN) and our proposed MSCNN lies in the multiscale operation. Concretely, in MSCNN, each branch corresponds to a coarse-grained subsignal differed only by the scale factors, and performs moving average smoothing and downsampling simultaneously. Besides, our method can improve the signal-to-noise ratio by averaging the collected periodic data, thereby improving the quality of the data. In MCNN, there are multiple subscale signals corresponding to various degrees of smoothness and downsampling rates.

D. Online and offline diagnostic systems

The intelligent fault detection system can operate both offline and online. The system supports importing off-line collected vibration data text format files and then using multiscale

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3024355, IEEE Transactions on Instrumentation and Measurement

convolutional neural network models for training and testing. At the same time, we can also use Visual Components software to build a digital twin system to achieve online monitoring, detection, and control parameter optimization to achieve data networking and closed-loop diagnostic early warning of neural network systems. The framework of an online data collection, monitoring and early warning system built in an enterprise is shown in Fig. 6. It consists of three sequential stages: field layer, middle layer, and application layer. The field layer includes industrial robots, non-standard equipment, etc. The signals of these field devices are first connected to the PLC (OMRON CJ2M). Then, the gateway (ADVANTECH ECU-1251) acts as a middle layer and plays the role of the signal relay. Its input end interacts with the PLC through TCP/IP protocol, and its output end uses the standard communication protocol OPC-UA to interact with the Visual Components software of the application layer. This software is used to display the operating data of field devices in real-time and use the embedded MSCNN module to implement intelligent fault detection and control of field devices.

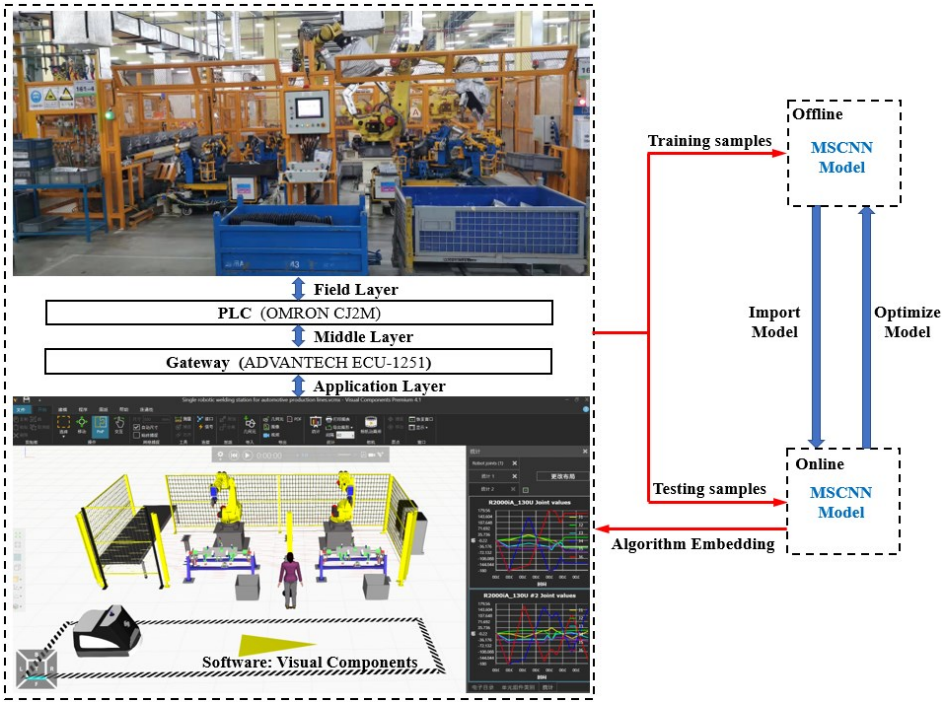

Fig. 6. Flowchart of the developed MSCNN-based fault detection systems for robot drive.

III. EXPERIMENTS

In this section, the MSCNN-based harmonic drive fault detection method is verified on a robot vibration test platform. It contains the setup and data collection, the details of harmonic drive, and fault dataset construction.

A. Setup and Data Collection

The test platform includes an industrial robot with a rated load of 8 kg, 8 harmonic drives, an SKF vibration analyzer, 3 vibration acceleration sensors (SKF MODEL CMSS2200), and a laptop. The robot vibration experiment platform is shown in Fig. 7. The SKF vibration analyzer is set to Hanning-window sampling. The sampling frequency is 512 Hz and the lower bound cut off frequency is 0.1Hz. According to the statistics of defective products provided by the manufacturer, the harmonic drive on the fifth axis of the robot has the highest probability of malfunction. Therefore, the experimental studies only focus on the operating state of the harmonic drive installed on the fifth axis of the robot.

The fifth axis is used to perform a reciprocating rotary motion. A schematic diagram of the transmission chain of the fifth axis of the industrial robot is given in Fig. 8. The transmission route of Joint 5 of the industrial robot is that the motor drives the pulley, and transmits the motion to the harmonic drive through the belt drive, and finally transmits the motion to the tool end connected to the harmonic drive. The three acceleration sensors are respectively fixed on the end-load of the robot according to the mutually orthogonal directions of X, Y, and Z, and are used to record the vibration acceleration data of the end of the industrial robot during the movement.

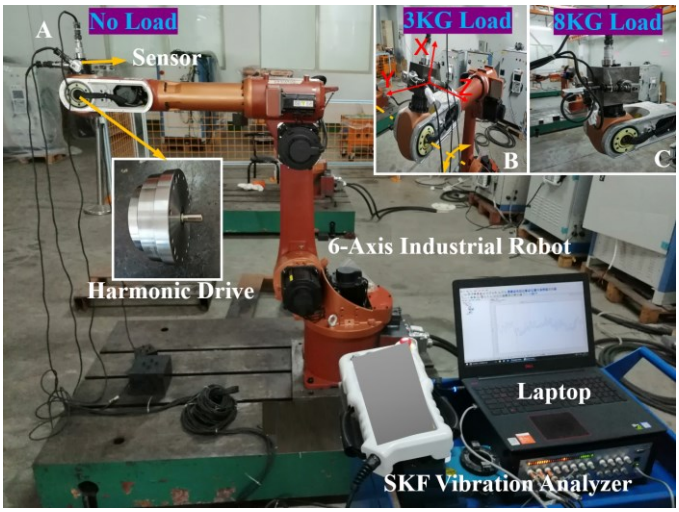

Fig. 7. Robot vibration experiment platform.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3024355, IEEE Transactions on Instrumentation and Measurement

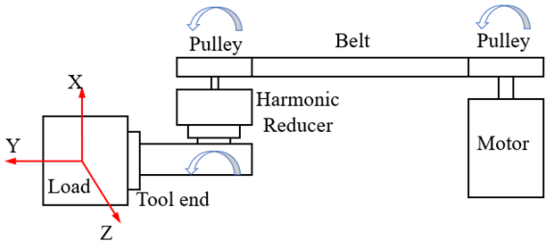

Fig. 8. Schematic diagram of the fifth axis of the robot.

The loads to be tested are no-load, 3 kg, and 8 kg. The 5th-axis motor is tested under four levels of speeds, which are super high (S), high (H), medium (M), and low (L) corresponding to 4985, 3739, 2492, and 1247 r/min respectively. The 5-axis reduction ratio provided by the manufacturer is 84.1666.

B. The Details of Harmonic Drive

The harmonic drive in this paper is an integrated harmonic drive manufactured by a Chinese company. The transmission is mainly composed of three basic components, namely a wave generator, a flexspline, and a circular spline. The structural parameter diagram of the harmonic drive is shown in Fig. 9.

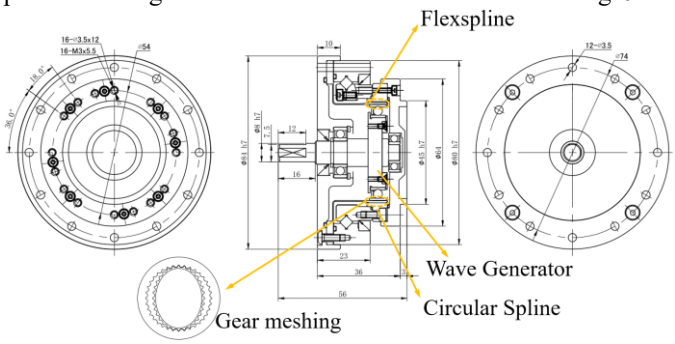

Fig. 9. Structural parameter diagram of harmonic drive.

The wave generator is usually installed at the input end of the drive. The number of teeth of the drive's flexspline is 200, and the number of teeth of the circular spline is 202. Fig. 9 describes the gear meshing of the circular spline and the flexspline. The flexspline and the drive body are fixed by bolts, and the circular spline serves as the output end of the drive. The key performance parameters of the harmonic drive are shown in Table II.

TABLE II
KEY PERFORMANCE PARAMETERS OF HARMONIC DRIVE

Reduction ratio	Start-stop max torque (Nm)	Average payload max torque (Nm)	Allowable instantaneous maximum torque (Nm)	Allowable maximum input speed (r/min)	Design life (Hour)
100	66	49	134	7000	15000

Generally, the fault characteristics of a harmonic drive are high-frequency vibrations. However, failures also occur at low frequencies. The characteristic frequency mainly includes the rotation frequency and the meshing frequency. Given the input speed V_{motor} of the motor, the reduction ratio i_1 of the belt drive and the reduction ratio i_2 of the harmonic drive, the output speed n_i of the circular spline on the harmonic drive can be calculated. Then, the base frequency f_i of the circular spline can be further calculated by

$$f_i = n_i / 60 = V_{motor} / i_1 / i_2 / 60 \tag{10}$$

The gear meshing frequency f_{zi} is equal to the base frequency f_i multiply the number of teeth Z_i and can be depicted as

$$f_{zi} = f_i \times Z_i \tag{11}$$

During operation, the normal sound of the harmonic drive is about 60db, and the maximum normal value is 78db. If the sound produced by the harmonic drive exceeds 78db, the product is considered abnormal. Generally, the harsh noise of abnormal products is often accompanied by obvious vibration.

C. Fault Dataset Construction

It should be noted that the robot controller only drives the motor of the fifth axis to move under 12 operating conditions, and the other joints of the robot remain stationary. In the experiment, four normal harmonic drives and four abnormal drives were tested. The description of the health status of the harmonic drive is shown in Table III. The rotation range is between -105˚ and +105˚.

TABLE III
DESCRIPTION OF HARMONIC DRIVE HEALTH CONDITIONS

Label	Fault description	Motor speed (r/min)	Load (kg)	Train sample	Test sample
0	Normal	4985/3739/2492 / 1247	0/3/8	2945	820
1	Abnormal	4985/3739/2492 / 1247	0/3/8	755	310

In the data set, all kinds of normal data under different speeds and load conditions are classified into class 0, and abnormal data under various working conditions are classified into class 1. A total of 4380 data sample units were collected, including 12 different operating conditions. Each sample unit contains 3072 data. The minimum speed of the fifth axis motor is 1247r / min, and the reduction ratio is 84.1666. Therefore, the output angular speed of the tool end is 1247*360˚/84.1666/60≈88.895 ˚/s. The rotation movement angle of Joint 5 for one cycle is 105 ˚ * 4 = 420˚. In each motion cycle, the time of the robot's 5th axis acceleration and deceleration is very small, thus the approximate motion time of the entire process is 420 / 88.895 ≈4.725 seconds. The sampling frequency of the sensor is 512Hz, and the sampling time of one sample unit is 3072/512 = 6 s. The time of one sample unit is longer than a motion cycle time of the fifth axis of the robot. All collected data is divided into many consecutive sample units, each of which is used to feed the neural network. To ensure the completeness and comprehensiveness of the sample unit information, it is necessary to ensure that the data amount of each sample unit is greater than the data amount of a motion cycle, regardless of the starting point of the sample unit.

IV. RESULTS AND DISCUSSION

This section shows the result of channel correlation and spectrum analysis and the classification for each channel. The optimized MSCNN algorithm was applied to the test platform and compared the results using different methods. The models are written in Python 3.7 with Pytorch and run on the same computer with a GTX 1660 TI GPU.

A. The Results of Channel Correlation Analysis

The sensors on channel 1 and channel 2 and channel 3 are

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3024355, IEEE Transactions on Instrumentation and Measurement

named sensor 1 and sensor 2 and sensor 3 respectively, which correspond to the X, Y, and Z directions of the workpiece. Spearman rank correlation analysis was used to obtain the results of the three rankings. ρ_i represents the Spearman correlation value between the *i-th* channel data and the labeled result. The calculation result of the three channels is $\rho_2 \approx 1.12 * \rho_3 \approx 22.86 * \rho_1$. Therefore, the order of importance of three channels is Channel 2> Channel 3> Channel 1. The results show that Channel 2 has the highest correlation with health status. At the same time, Channel 2 is the easiest channel to collect on-site data. Because this channel collects vibration data is installed in the rotation direction of the harmonic drive, and the installation of the sensor is not easily interfered.

B. The spectrum results of harmonic drive

The motor has 4-speed inputs. Taking V_{motor} equal to 2492 r / min as an example, the base frequency and meshing frequency of the circular spline at the output of the harmonic drive are f_i =2497/84.1666/60 ≈ 0.4935 (Hz) and f_{zi} =0.4935×202=99.687 (Hz) respectively. The collected vibration acceleration signal contains the frequency information of the motor, belt pulley, camshaft, flexspline and circular spline. The different input velocity of the motor runs under 3 payloads (0/3/8Kg). Therefore, the frequency component in Table IV can calculated by Equations (10) and (11). These frequency components are conducive to our subsequent multi-scale analysis.

TABLE IV
FREQUENCY COMPONENT DIAGRAM OF VIBRATION ACCELERATION SIGNAL

Motor speed (r/min)	Motor base frequency (Hz)	Pulley/Cam shaft speed (r/min)	Pulley/Cam shaft base frequency (Hz)	Flexspline meshing frequency (Hz)	Circular spline base frequency (Hz)
1247	20.78	1049.56	17.49	50.5	0.25
2492	41.53	2097.43	34.96	98.98	0.49
3739	62.32	3146.99	52.45	149.48	0.74
4985	83.08	4195.71	69.93	199.98	0.99

To compare the difference between normal and fault signals, we selected 4 groups of vibration acceleration signals under different working conditions to perform FFT (Fast Fourier Transformation) processing, which is used to extract useful features of the raw vibration signal. The frequency of the normal state and the abnormal condition using spectrum analysis is shown in Fig. 10.

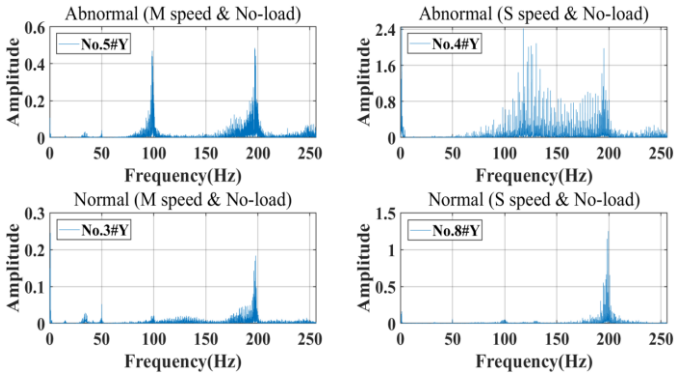

Fig. 10. Frequency domain of harmonic drive under normal and abnormal conditions.

According to the experimental report provided by the manufacturer, the vibration frequency of the harmonic drive is difficult to distinguish normal or abnormal and it generally relies on expert experience. The vibration signal of the harmonic drive contains complex characteristic information of different components. That means the fault features are distributed in multiple signals irregularly, and this kind of complicated nonlinear relationship is hard to be analyzed by the traditional time-frequency method. Therefore, we need to overcome these disadvantages and solve the black box problem by using the end-to-end technique. Thus, it is necessary to use multiscale decomposition and 2-D CNN techniques to process the raw signal of multiple working conditions.

C. The classification results for each channel

The optimal hyperparameters selection combination of a neural network is shown in Table V, which were obtained through debugging experiences and a lot of validated experiments. These hyperparameters were used to train and test the samples.

TABLE V
THE OPTIMAL HYPERPARAMETERS OF THE NEURAL NETWORK

Optimizer	Criterion	Learning rate	Epochs
Adam	Cross Entropy	0.001	50

To further analyze the impact of each channel on the health of the harmonic drive. The convolutional neural network (CNN) method was used to analyze the data of the three-channel sensors and the multi-sensors data, and the training and testing results obtained are shown in Fig. 11 and Fig. 12.

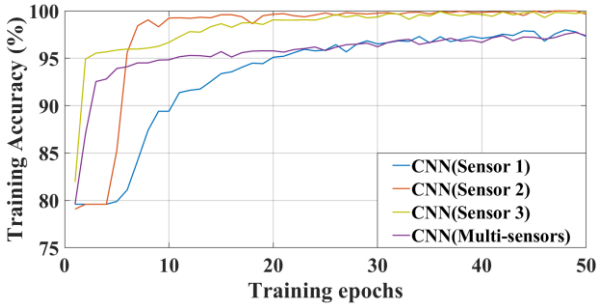

Fig. 11. Training accuracy of CNN(Sensor1), CNN(Sensor2), CNN(Sensor3) and CNN(Multi-sensors).

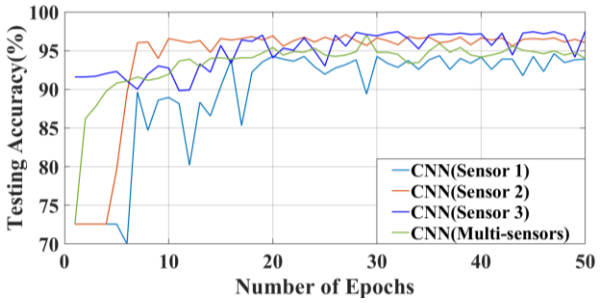

Fig. 12. Testing accuracy of CNN(Sensor1), CNN(Sensor2), CNN(Sensor3) and CNN(Multi-sensors).

It can be observed from Fig. 11 and 12 that when training and testing the convolutional neural network, Channel 2 first reaches convergence and has the smallest fluctuation range and the best effect after convergence. Moreover, the analysis results from data collected by multi-channel sensors (multi-sensors data) are worse than some single channels (Channel 2 and Channel 3). This indicates that the unreasonable sensor arrangement and the interference of noise sources will decrease

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3024355, IEEE Transactions on Instrumentation and Measurement

the classification accuracy. Therefore, the data of Channel 2 are selected to perform multiscale convolutional neural network optimization for better fault detection results.

D. The optimization results of MSCNN

The frequency spectra of the abnormal and normal vibration acceleration signal under the same working conditions are shown in Fig. 13 and Fig. 14. The multiscale layer was used as a filter to highlight useful information on different time scales, which can capture the abundant characteristics of the raw signal and avoid noise interference. s = 1 represents the original vibration signal, s = 2, 3, 4 represents the signals of multi-scale decomposition processing.

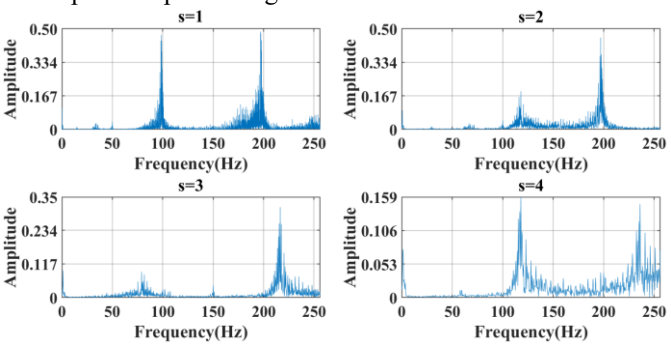

Fig. 13. Abnormal frequency domain of the MS layer.

The goal is that the extracted feature information of different scales is capable to distinguish normal and abnormal samples to the greatest extent, which will help to improve the final fault detection classification accuracy.

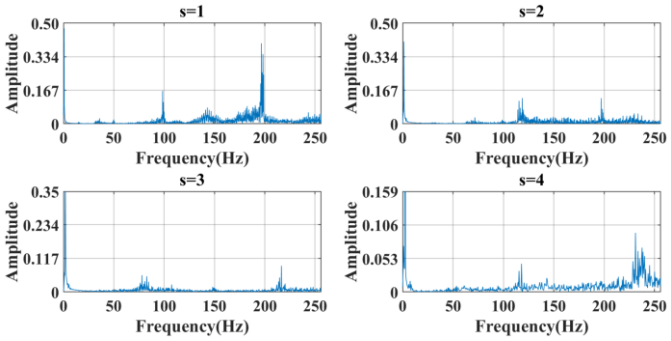

Fig. 14. Normal frequency domain of the MS layer.

It can be seen from Fig. 13 and Fig. 14 that the spectrum information extracted by normal and abnormal samples at various scales is different. Through scale decomposition, the difference between normal and abnormal sample features will become larger, and high-frequency signals will be filtered. Besides, multi-scale decomposition can obtain more spectrum components in Table IV than directly processing the original signals. MSCNN method firstly uses multiscale decomposition on the raw data, and then combines the 2-D CNN, while traditional CNN is directly training using raw data.

The training and testing accuracy of MSCNN and CNN are shown in Fig. 15. The results show that the fault detection accuracy of MSCNN (indicated by the blue dotted line) for the harmonic drive is higher than that of the traditional CNN method (indicated by the purple dotted line). The results obtained by the MSCNN method are superior to the CNN method, which also shows that the multiscale decomposition module plays a role in improving fault classification performance. It also proves the multiscale learning module can capture more useful feature information than a single layer.

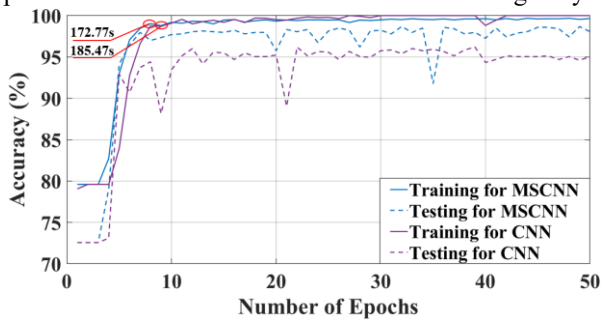

Fig. 15. Performance curve in terms of overall accuracy.

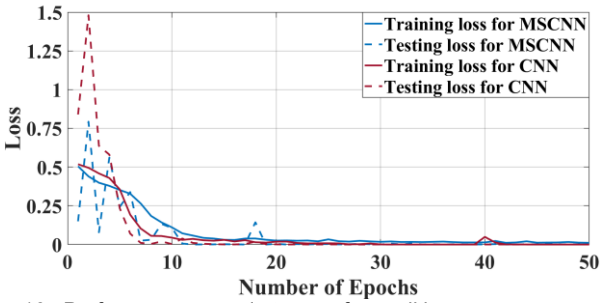

Fig. 16. Performance curve in terms of overall loss.

The loss function can effectively reflect the difference between the predicted value and the true value, it is an index to evaluate the pros and cons of a neural network model. The loss functions for training and testing of the MSCNN and CNN methods are shown in Fig. 16. Both of them converge faster, but MSCNN is better. As shown in Fig. 15, the training process of MSCNN (blue solid line) and CNN (purple solid curve) only requires nearly 7 epochs and 8 epochs respectively to reach convergence, and the corresponding time is nearly 172.77 seconds and 185.47 seconds. After training the samples, the test results are used to evaluate the performance of the algorithm. At the beginning of testing, the testing loss values of MSCNN are obviously smaller than those of traditional CNN. Therefore, this verified the MSCNN is better than the traditional 2-D CNN in the fault detection.

After the MSCNN model is trained, the separate test samples are used to verify the detection performance. The number of test set accounts for about 30% of the total samples. The value of the confusion matrix can visually display the predicted values, include the number of correct predictions and misjudgments for each label of the test samples. Label 0 represents the normal sample, and label 1 represents the abnormal sample. As shown in Fig. 17, among the 820 normal samples, MSCNN can accurately identify 810 samples, while CNN can only accurately identify 802 samples. Among the 310 abnormal samples, MSCNN can accurately identify 298 samples, while CNN can only identify 272 samples. Thus, it can be concluded that the prediction result of the MSCNN method is more accurate than that of CNN. The results show that the application of our method can effectively improve the fault detection performance of the harmonic drive in practical industrial scenarios without technicians' experiences.

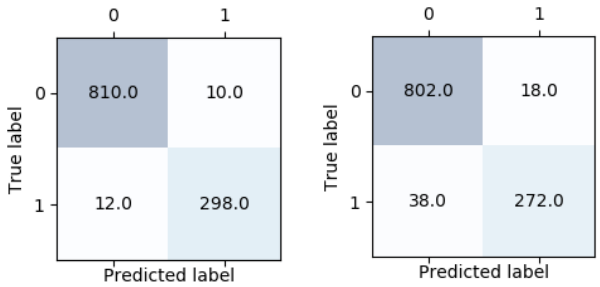

Fig. 17. Confusion matrix results with (a) MSCNN and (b) CNN.

E. Contrast results using different methods

Finally, this paper also compares and analyzes the effect of other learning methods on the fault detection of harmonic drive. We also use the scikit learn library [28] to import the machine learning algorithm module we need, such as support vector machine (SVM), Decision tree (DT) and K-nearest neighbor (KNN). It should be noted that the training set and test set of the various comparison methods are the same. Some key model parameters of these neural networks are given below.

1) In recurrent neural network (RNN), the hidden dimension is 600. The number of hidden layers is 3 and the criterion is Cross Entropy. The optimizer is SGD and the learning rate is 0.001.
2) In SVM, the gamma parameter is selected as auto.
3) In DT, we chose the default parameter Decision Tree classifier, and predict the accuracy score.
4) In KNN, we select the neighbors' numbers parameter as 3 to fit the model, which is determined by experience and test results.

The calculation results of various learning methods on the test set are shown in Table VI.

TABLE VI
CONTRAST THE RESULTS OF DIFFERENT METHODS

Methods	Mean accuracy (%)
MSCNN (Proposed method)	**96.79**
CNN	95.97
RNN	72
SVM	78.79
DT	78.39
KNN	71.54

Through the comparison of the results of the neural network algorithms in Table VI, it can be concluded that the MSCNN in this paper is the most suitable for solving the fault detection of harmonic drives. To ensure the stability and reliability of the experimental results, MSCNN and CNN methods were applied to the test set for 10 repeated experiments.

The highest accuracy obtained by the MSCNN method used in this paper is 98.761%, the lowest accuracy is 95.664%, and the average accuracy of 10 times is 96.79%. However, the highest accuracy rate obtained by the traditional CNN method is 96.991%, the minimum accuracy rate is 94.779%, and the average accuracy rate for the 10 times is 95.97%. Therefore, the MSCNN is not only better than traditional CNN and RNN, but also better than other machine learning methods such as SVM, DT, and KNN.

V. CONCLUSION

To tackle the challenge of fault detection for harmonic drives, we presented an intelligent fault detection method for intact harmonic drives under multiple working conditions, i.e. a new MSCNN architecture. The MSCNN approach based on multisensor data fusion is capable to extract complex vibration signals and detect malfunction of the harmonic drive in the real industrial robot test platform. It can simultaneously perform automatic feature extraction and classification without relying on manual signal processing and empirical knowledge. Moreover, unlike using the public standard data set, we collect data from manufactures and our experimental platform for the harmonic drives. Based on data, we test our approach, and the results indicate that in terms of robustness and classification performance, MSCNN is significantly superior to traditional CNN, RNN, SVM, DT, and KNN methods, especially under receiving multiscale vibration signals. It can be concluded that MSCNN is a useful framework in the field of intelligent fault detection, which can be applied and extended to different industrial systems, such as the gearbox of wind turbine, the slewing bearings of tunnel-boring machine, the electromechanical actuators of aircraft engine, and etc.

In our future work, we will use Visual Components software to build an intelligent digital twin system on an industrial robot production line to achieve online monitoring of the MSCNN algorithm. Besides, we will further study the MSCNN combined with transfer learning to reduce the dependence on scarce label data so that the performance of the learning algorithm can be further improved.

REFERENCES

[1] Tuttle, Timothy D. Understanding and modeling the behavior of a harmonic drive gear transmission. No. AI-TR-1365. MASSACHUSETTS INST OF TECH CAMBRIDGE ARTIFICIAL INTELLIGENCE LAB, 1992.

[2] Park, In-Gyu, Il-Kyeom Kim, and Min-Gyu Kim. "Vibrational characteristics of developed harmonic reduction gear and fault diagnosis by campbell diagram." In *2015 15th International Conference on Control, Automation and Systems (ICCAS)*, pp. 2062-2065. IEEE, 2015.

[3] Adams, Christian, Adam Skowronek, Joachim Bös, and Tobias Melz. "Vibrations of Elliptically Shaped Bearings in Strain Wave Gearings." *Journal of Vibration and Acoustics* 138, no. 2 (2016).

[4] Guo, Yingying, Xuezhi Zhao, Wen-Bin Shangguan, Weiguang Li, Hui Lü, and Chunliang Zhang. "Fault characteristic frequency analysis of elliptically shaped bearing." *Measurement* 155 (2020): 107544.

[5] Zheng, Jianlin, and Wei Yang. "Failure Analysis of a Flexspline of Harmonic Gear Drive in STC Industrial Robot: Microstructure and Stress Distribution." In *IOP Conference Series: Materials Science and Engineering*, vol. 452, no. 4, p. 042148. IOP Publishing, 2018.

[6] M. Kordestani, M. Saif, M. E. Orchard, R. Razavi-Far, and K. Khorasani, "Failure Prognosis and Applications—A Survey of Recent Literature," *IEEE Trans. Rel.,* vol. 68, no. 3, pp. 1-21, Sept. 2019.

[7] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: A review with applications," *Signal Process*., vol. 96, pp. 1–15, Mar. 2014.

[8] L. Hong and J. S. Dhupia, "A time domain approach to diagnose gearbox fault based on measured vibration signals," *J. Sound Vib*., vol. 333, no. 7,pp. 2164–2180, Mar. 2014.

[9] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process*. 115 (2019) 213-237.

[10] Konar, P., & Chattopadhyay, P. (2011). Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs). *Applied Soft Computing*, 11(6), 4203–4211.

[11] Abbasion, S., Rafsanjani, A., Farshidianfar, A., et al. (2007).Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. *Mechanical Systems and Signal Processing*, 21(7), 2933–2945.

[12] Mukhopadhyay R, Panigrahy P S, Misra G, et al. Quasi 1D CNN-based Fault Diagnosis of Induction Motor Drives. *2018 5th International Conference on Electric Power and Energy Conversion Systems (EPECS). IEEE*, 2018: 1–5.
[13] J. Fu, J. Chu, P. Guo, and Z. Chen, "Condition monitoring of wind turbine gearbox bearing based on deep learning model," *IEEE Access*, vol. 7, pp. 57078-57087, 2019.
[14] J. Zhu, N. Chen, and W. Peng, ''Estimation of bearing remaining useful life based on multiscale convolutional neural network,'' *IEEE Trans. Ind. Electron*., vol. 66, no. 4, pp. 3208–3216, Apr. 2019.
[15] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron*., vol. 65, no. 7, pp. 5990-5998, Jul. 2018.
[16] O. Janssens et al., "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib*., vol. 377, pp. 331–345, Sep. 2016.
[17] Z. Chen, W. Li, Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network, IEEE Trans. Instrum. Meas. 66(2017) 1693–1702, https://doi.org/10.1109/TIM.2017.2669947.
[18] J. Liu, Y. Hu, Y. Wang, B. Wu, J. Fan, Z. Hu, An integrated multi-sensor fusion-based deep feature learning approach for rotating machinery diagnosis, Meas. Sci. Technol. 29 (2018), https://doi.org/10.1088/1361-6501/aaaca6.
[19] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, "Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference," *Expert Syst. with Appl*., vol. 37, no. 8, pp. 6077–6085, Aug. 2010.
[20] H. Liu and M. Han, "A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings," *Mech. Mach. Theory*, vol. 75, pp. 67–78, May 2014.
[21] J. Zheng, H. Pan, and J. Cheng, "Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines," *Mech. Syst. Signal Process*., vol. 85, pp. 746–759, Feb. 2017.
[22] G. Jiang, H. He, J. Yan, and P. Xie, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Trans. Ind. Electron*., vol. 66, no. 4, pp.3196-3207, Apr. 2019.
[23] M. Canizo, I. Triguero, A. Conde, and E. Onieva, "Multi-head CNN-RNN for multi-time series anomaly detection: An industrial case study," *Neurocomputing* 363 (2019) 246-260.
[24] D. Yang, Y. Pang, B. Zhou, and K. Li, "Fault diagnosis for energy internet using correlation processing-based convolutional neural networks," *IEEE Trans. Syst., Man, Cybern., Syst*., vol. 49, no. 8, pp. 1739-1748, Aug. 2019.
[25] J. He, Z. Zhang, X. Wang, and S. Yang, "A low power fall sensing technology based on FD-CNN," *IEEE Sensors J*., vol. 19, no. 13, pp. 5110-5118, Jul. 2019.
[26] W. Huang, J. Cheng, Y. Yang, and G. Guo, "An improved deep convolutional neural network with multi-scale information for bearing fault diagnosis," *Neurocomputing* 359 (2019) 77-92.
[27] R. Liu, G. Meng, B. Yang, C. Sun, and X. Chen, "Dislocated time series convolutional neural architecture: An intelligent fault diagnosis approach for electric machine," *IEEE Trans. Ind. Informat*., vol. 13, no. 3, pp.1310-1320, Jun. 2017.
[28] *Scikit-learn library.* Accessed: Jan. 14, 2020. [Online]. Available: https://scikit-learn.org/stable/.

Guo Yang received the M.S degree in mechanical engineering from Guangdong University of Technology, Guangzhou, China, in 2018.

He is currently a Ph.D. candidate with the Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou, China. His main research focuses on intelligent monitoring, diagnosis and control of advanced manufacturing systems.

Yong Zhong (S'12–M'17) received the B.Eng. degree in mechanical design, manufacturing and automation from the Huazhong University of Science and Technology, Hubei, China, in 2011, and the M.Eng. degree in control engineering from the University of Chinese Academy of Sciences, Beijing, China, in 2014, and the Ph.D. degree in Mechanical and Automation Engineering from the Chinese University of Hong Kong, Shatin, Hong Kong, in 2017. He was a research fellow at National University of Singapore from 2017 to 2019.

He is currently an Assistant Professor with Shien-Ming Wu School of Intelligent Engineering, South China University of Technology. His research interests bioinspired robots, soft robots, intelligent diagnosis and control.

Lie Yang received the B.Eng. degree in mechanical design, manufacturing and automation from School of Mechanical Engineering, Wuhan Polytechnic University, Wuhan, China, in 2015.

He is currently a Ph.D. candidate with the Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou, China. His main focuses on artificial intelligence, computer vision, and electroencephalogram decoding.

Ruxu Du received the B.S. degree in electrical engineering, the M.S. degree in automatic control from the South China University of Technology, Guangzhou, China, in 1983, and the Ph.D. degree in mechanical engineering from the University of Michigan, Ann Arbor, MI, USA, in 1989. From 1991 to 2001, he taught at the University of Windsor, Windsor, ON, Canada; at the University of Miami, Coral Gables, FL, USA; and at Chinese University of Hong Kong (CUHK), Shatin, Hong Kong. Since 2002, he has been a Professor with the Department of Mechanical and Automation Engineering, and the Director of the Institute of Precision Engineering, CUHK.

He is currently a Professor with Shien-Ming Wu School of Intelligent Engineering, South China University of Technology. His research interests include precision engineering, condition monitoring, fault diagnosis, manufacturing processes (metal forming, machining, plastic injection molding, etc.), and robotics.

Dr. Du became a Fellow of the America Society of Mechanical Engineers (ASME) in 2009, a Fellow of the Society of Manufacturing Engineers (SME), and the Hong Kong Institute of Engineers (HKIE) in 2012, and an Academician of Canadian Academy of Engineering in 2017.