Dislocated Time Series Convolutional Neural Architecture: An Intelligent Fault Diagnosis Approach for Electric Machine

Ruonan Liu, *Student Member, IEEE*, Guotao Meng, Boyuan Yang, *Student Member, IEEE*, Chuang Sun, and Xuefeng Chen, *Member, IEEE*

***Abstract*—In most current intelligent diagnosis methods, fault classifiers of electric machine are built based on complex handcrafted features extractor from raw signals, which depend on prior knowledge and is difficult to implement intelligentization authentically. In addition, the increasingly complicated industrial structures and data make handcrafted features extractors less suited. Convolutional neural network (CNN) provides an efficient method to act on raw signals directly by weight sharing and local connections without feature extractors. However, effective as CNN works on image recognition, it does not work well in industrial applications due to the differences between image and industrial signals. Inspired by the idea of CNN, we develop a novel diagnosis framework based on the characteristics of industrial vibration signals, which is called dislocated time series CNN (DTS-CNN). The DTS-CNN architecture is composed of dislocate layer, convolutional layer, sub-sampling layer and fully connected layer. By adding a dislocate layer, this model can extract the relationship between signals with different intervals in periodic mechanical signals, thereby overcome the weaknesses of traditional CNNs and is more applicable for modern electric machines, especially under nonstationary conditions. Experiments under constant and nonstationary conditions are performed on a machine fault simulator to validate the proposed framework. The results and comparison with respect to the state of the art in the field is illustrated in detail, which highlights the superiority of the proposed method in industrial applications.**

***Index Terms*—Convolutional neural architecture, dislocated time series, electric machine, intelligence fault diagnosis, nonstationary condition.**

I. Introduction

FAULT diagnosis has always been an attractive and critical task in electric machines due to its ability to reduce maintenance costs and prevent the accidents [1], [2]. As electric machines are widely used in modern society [3], numerous algorithms and techniques have been proposed for fault diagnosis of them, which vary from system to system [3]–[7].

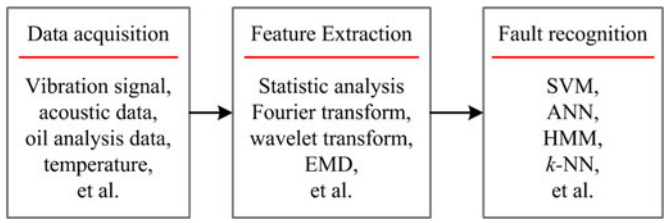

Fig. 1. Three key steps in traditional fault diagnosis.

According to the current literature, the framework of traditional fault diagnosis includes three steps: signal acquisition; feature extraction and dimensionality reduction (DR); and fault recognition, as shown in Fig. 1 [8], [9]. Data acquisition is a process of collecting and storing useful data (information) from targeted physical equipments [10]. The data can be very versatile and vibration signals are utilized dominantly because they provide the most intrinsic information of the equipments [11]. Since the variability and richness of natural data make it almost impossible to build an accurate recognition system entirely via early pattern recognition algorithms, most fault-diagnosis system are built with a combination of feature extraction and machine learning algorithms. To extract representative features from the complex and nonstationary noisy signal, numerous vibration signal processing approaches have been developed such as statistic analysis, Fourier transform [12], wavelet transform [13], empirical mode decomposition (EMD) [14], and sparse representation [15], which is developed recently. To increase the performance of the diagnosis system, DR methods should be employed to deal with high-dimensional data such as principle component analysis and local discriminant analysis [16]. In the last step, the extracted features are used to train the machine learning models such as support vector machine (SVM) [4], artificial neural networks (ANN) [17], hidden Markov model [18], and k-nearest neighbor [19].

Well developed and diverse as traditional intelligent diagnosis systems are, they still suffer from three main weaknesses. First, the original signal cannot be used to train the traditional artificial intelligent model directly. Therefore, it is important to design

Manuscript received June 25, 2016; revised September 15, 2016 and November 27, 2016; accepted December 21, 2016. Date of publication December 26, 2016; date of current version June 1, 2017. Paper no. TII-16-0560. *(Corresponding author: X. Chen.)*

R. Liu, B. Yang, C. Sun, and X. Chen are with the State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: liuruonan0914@stu.xjtu.edu.cn; yangboyuanxjtu@163.com; ch.sun@xjtu.edu.cn; chenxf@mail.xjtu.edu.cn).

G. Meng is with the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: kakaguotao@hotmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2016.2645238

a feature extractor, which can represent original signal by low-dimensional vectors that can be easily matched or compared. However, during feature extraction process, much information will be lost. And the feature extractor contains most of the prior knowledge and is often designed entirely hand crafted [20]. As a result, the diagnosis accuracy is largely depended on prior knowledge and the ability of designer. Second, to come up with appropriate set of features, the precise physical model and a comprehensive understanding of the fault are needed. In addition, feature extractor is specific to the task, therefore it must be redesigned for every new problem. However, with the rapid development of mechanical manufacturing industry, mechanical models have become more and more precise, complex and automatic, which makes it much more difficult, even impossible to design a specific and suitable feature extractor for every diagnosis system. Last but not least, in practical applications, electric machines are always working under nonstationary conditions. However, because the signals used for classification are variable as the operating condition changing, the fault diagnosis and recognition of electric machines under nonstationary condition is always a hard nut to crack.

Therefore, to overcome the problems above, a novel framework of intelligent fault diagnosis which is fed with raw information is proposed. In this framework, inspired form the work of LeCun *et al.* [20], we proposed an improved convolutional neural network (CNN) architecture based on the characteristics of mechanical signals, which is called the dislocated time-series CNN (DTS-CNN). The basic idea behind the proposed DTS-CNN method is that the periodic fault information between nonadjacent signals can be extracted by continuously dislocating the input raw signals. In this framework, the DTS-CNN is considered as a black box model, and different kinds of faults can be predicted automatically from a training set of dislocated raw signals. The reduction of the demand for prior knowledge, physical model, and human labor makes this diagnosis framework more robust and adaptive in modern industrial system and sophisticated electric machines. And this ability to automatically learn powerful and proper features will make the mechanical fault diagnosis system real artificial intelligence and become increasingly important as the amount and kind of industrial system continue to grow [21].

The rest of this paper is organized as follows. In Section II, the CNN architecture is described in detail. In Section III, the proposed DTS-CNN method and the intelligence fault diagnosis framework is illustrated. The vibration data of gearbox are studied by the proposed method in Section IV. Conclusion and future work of the research are presented in Section V.

II. CNN Architecture

Begins with neuroscientific experiments, CNN have been mainly influenced by Hubel and Wiesel's early work on the vision cortex working mechanisms of mammalian brain [22], [23]. A CNN layer is designed to capture the following three properties of the primary visual cortex (V1).

1) V1 has a two-dimensional (2-D) structure mirroring the structure of the image in the retina.

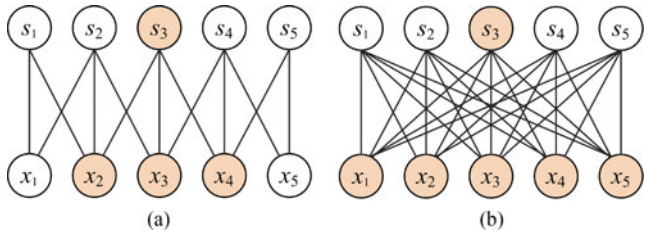

Fig. 2. Sparse connectivity: (a) three inputs affect s_3 when s is formed with a kernel of width 3; (b) all inputs affect s_3 when it is not a sparse connection.

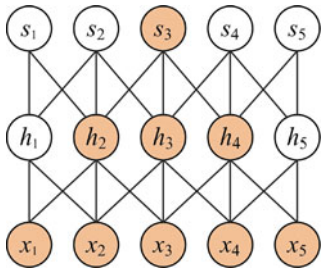

Fig. 3. Local receptive field of the units in the deeper layer.

2) A simple cell's activity can be characterized by a linear function of the image in a small region.
3) V1 also contains many complex cell that are invariant to small shifts in the position [24].

Inspired by these properties, CNNs are introduced and successfully used by LeCun *et al.* [20], [25], [26], which are powerful bioinspired hierarchical multilayered neural networks that combine three architectural ideas: sparse connectivity (or local receptive fields), shared weights, and pooling (spatial subsampling).

A. Sparse Connectivity

Traditional ANNs use a matrix multiplication to describe the interaction among each input unit and each output unit. However, CNN exploits sparse connectivity by making the kernel smaller than the input and enforcing a local connectivity pattern among neurons of adjacent layers. In other words, the inputs of hidden units in layer m are from a subset of units in layer $m-1$, which can be graphically illustrated in Fig. 2 [24].

Even though direct connections in a CNN are sparse, units in the deeper layers may indirectly interact with a larger portion of the input, as shown in Fig. 3. By sparse connectivity, the complicated interactions between units can be described more efficiently, and the overfitting risk can also be reduced due to the less parameters.

B. Shared Weights

Shared weights refers to using the same parameters for more than one function in a model. In traditional ANN, each weight matrix is multiplied by one element of the input and used only once. In a CNN, each kernel (or weight matrix) is used across the entire visual field. The shared weights are learnt only once instead of learning a separate set of weights for every position [24], [27].

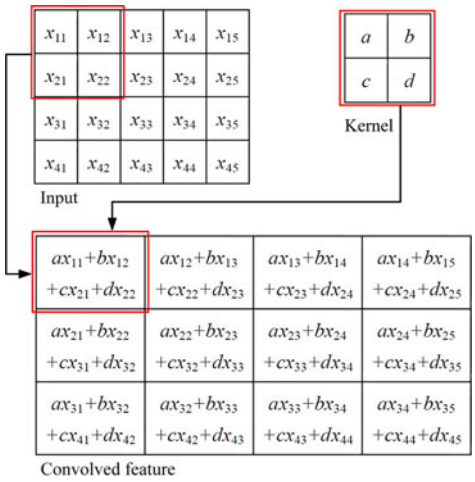

Fig. 4. Operating mechanism of shared weights and local receptive field.

The graphical depiction of how shared weights work is shown in Fig. 4. For example, suppose the features have been learnt on $a \times b$ patches, and there is an $m \times n$ image. To get the convolved features, we need to convolve the learned $a \times b$ feature detector with every $a \times b$ region in the image. And this would result in a $(m-a+1) \times (n-b+1)$ array of convolved features.

For the multilayered input and multilayered kernel, for example a $d \times m \times n$ input and a $c \times a \times b$ kernel matrix, the convolution operation will be more complicated, which can be summed up in three steps:

1) use a kernel to convolute with every layer of the input and result in a $d \times (m-a+1) \times (n-b+1)$ matrix;
2) sum up the results of the matrix and a $1 \times (m-a+1) \times (n-b+1)$ matrix will be obtained;
3) repeat the first two steps for each kernel. As a result, the output will be a $c \times (m-a+1) \times (n-b+1)$ array of feature maps.

This step can be described by a convolution kernel and requires $m \times n \times (a+1) \times b$ floating point operations (one addition per output pixel), which would take $m \times n \times m \times n$ floating point operations in a dense ANN. Therefore, convolution is an extremely efficient way that apply the same linear transformation of a local region across the entire input to describe transformations.

C. Pooling

There are two main stages in CNN. In the first stage, several convolutions are performed in parallel. Another important stage of CNN is pooling, which is a form of subsampling. A pooling operation is replacing the output of the convolution layer with a summary statistic. The most commonly used in CNN is max-pooling, which partitions the input image into a set of nonoverlapping rectangles and for each such subregion outputs the maximum value [27]. Therefore, pooling is useful in CNN for the following two reasons.

1) It makes the representation become invariant to small translation of the input. That is, if we translate the input by a small translation or rotation, the values of pooling outputs will not change.
2) By reporting summary statistics for pooling regions spaced k pixels apart rather than 1 pixel apart, the final pool has a smaller size. Therefore, it reduces computation for upper layers and improves the computational efficiency of the network.

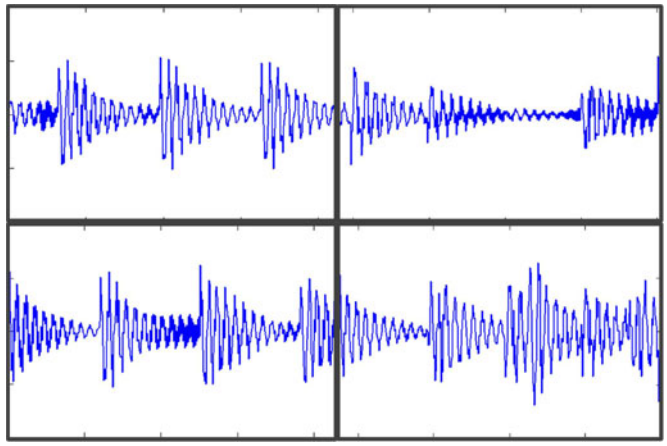

Fig. 5. Example of pooling invariance.

Not only the high computing efficiency, but also the invariance to local translation can be a very useful property in industrial signal processing because we care more about whether some feature is present than exactly where it is. For example, the Laplace impulse signal is one of the typically vibration response of a bearing fault [4]. In fault diagnosis systems, what we need is to detect the impulse component no matter how the initial phase or operating environment change. As shown in Fig. 5, if each of these filters drive units that appear in the same max-pooling region, then the pooling unit will detect the same value in any initial phase or fault types.

Sparse convolutional layer and max-pooling layer are at the heart of the CNN. In lower layers, CNN composed by alternating the convolutional layer and pooling layer. After extracting the convolutional features, the upper layers are consists of full-connected layer to classify different inputs.

III. Proposed Diagnosis Framework

A. Dislocated Time Series Convolutional Neural Architecture

CNN has been used successfully in many different applications, such as handwriting recognition [25], [28], machine-printed character recognition [29], and face recognition [30], [31]. For the purpose of those recognitions, the correlation and information in a local neighbor region are requiring more thought. Therefore, state-of-the-art CNNs mainly focus on the feature extraction and pattern recognition in a local region.

However, mechanical signal is different from image signal. Because rotating machineries abound in electric machines, the collected signals are often time series signals with strong periodicity and deep correlations among different time points. There is also a lot of valuable information hiding behind these periodicity

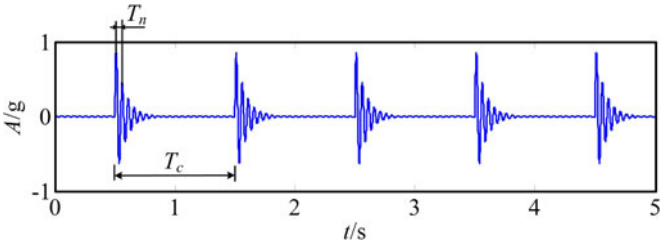

Fig. 6. Typical vibration response of a surface fault.

and correlations. For example, it is known that the components of impact vibration are divided into two parts, as shown in Fig. 6 [4]. The first one is the vibration of low frequency with period of T_c, which will generate repeated impacts due to the passing of the rolling elements over the defect. The other one is the vibration of high frequency with period of T_n, which is caused by the impact. If we just consider the local region in a period, the result will be easily influenced by random factor. It is more considerable to compare the relationship between the amplitude of an impulse point and another impulse point. Of course, amplitude correlation is only a drop in the ocean of the information behind the relationship between mechanical signals. Therefore, it will make full use of all information in the signal if we also consider the relationship between signals in different locations.

To overcome the limitations of CNN and make it more suitable for industrial applications, a DTS-CNN architecture is proposed according to the characteristics of mechanical signal in this paper. In the DTS-CNN architecture, instead of using the measured signal as the input of CNN to convolve with kernel directly, we add another layer to dislocate the input. Suppose the input signal is $x(t)$, which is a time series signal, then the output of dislocated layer can be represented as

$$D(m,:) = x(t_{m1} : t_{m2}) \tag{1}$$

where

$$\begin{aligned} t_{m1} &= 0.5m(m+1) \times k \\ t_{m2} &= t_{m1} + n \\ m &= 0, 1, 2, \ldots \end{aligned} \tag{2}$$

where D is the output matrix of dislocated layer, which is an $m \times n$ matrix, k is the dislocate step, and n is the length of the intercepted signal from original signal. That is, we cut out n points from row signal as a sample in each dislocate operation. m is the sample size of D. It means that the raw signal was dislocated for m times, and we will obtain an $m \times n$ matrix as the output of dislocated layer. The graphical demonstration is shown in Fig. 7.

From Fig. 7, it can be seen that we extract a signal with a length of n each time. Then, translate k distance and extract another signal with the same length. Next translating $2 \times k$ distance, next $3 \times k$. In the ath operation, we translate $a \times k$ distance and extract another signal. Then, combine the extracted signals into the matrix D. The main parameters in DTS-CNN are m, n, and k. n is comparatively easy to determine, just make sure it is bigger than the length of one period signal. m influences the effect of the proposed method. The larger the m is, the

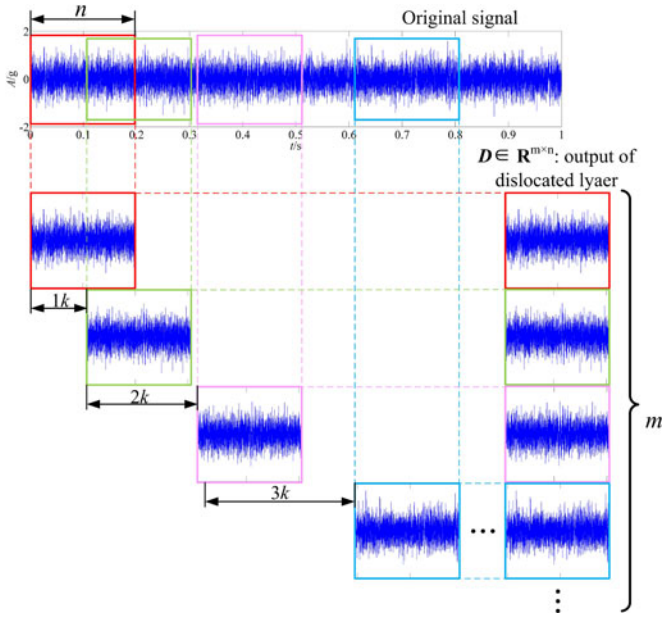

Fig. 7. Detail operation in dislocated layer.

more the layers D1 layer generates, therefore, the more precise the results are. However, on the other hand the computational capacity increases. Fortunately, m does not have to be very large in most cases if the value of k is proper. There is no accurate way to determine the value of k. Considering mechanical signal is highly cyclical, it would be better if k or its multiple (smaller than $m \times k$) is the multiple of period. Actually, according to the experiments, the result of DTS-CNN is robust for the choice of k if it is multiple of period.

The specific architecture of DTS-CNN applied in this paper is graphical illustrated in Fig. 8. There are three axis in the cubes, as shown in Fig. 8. x-axis and y-axis represent the width and height of the feature map, respectively; z-axis represents the number of kernels. Each kernel, like a filter, slides on the input and produces one feature map. Thus, each feature map can be regarded as the response of a certain pattern. Often there are several filters in one layer, so each layer produces a 3-D output contains several feature maps as much as the filters. In the architecture shown in Fig. 8, two convolution layers are added first. The description of convolution kernel $16@3 \times 4$ represents that there are 16 kernels with the size of 3×4. Then, a pooling operation is applied in the following S4 subsampling layer. This leads to the same number of C3 feature maps with a half-spatial resolution. Then, convolution kernels with the size of $32@3 \times 4$ are applied in C5 and C6, respectively. Next, a maxpooling layer S7 is applied with a 2×2 pooling region. Finally, a fully connected neural network is added. We first flatten the output of S7 to a 1-D feature maps in F8. Then, the output of F8 is used as the input of F9 to dense the feature maps and reduce the dimension. Suppose the number of hidden units in F9 is a, therefore, an $a \times 1$ vector is used as the input of softmax classifier and the final output is the number of categories. In the proposed architecture, dislocated layer and CNN are used for feature extraction. Some strategies such as sparse connectivity, shared weights, and drop out operation endow DTS-CNN

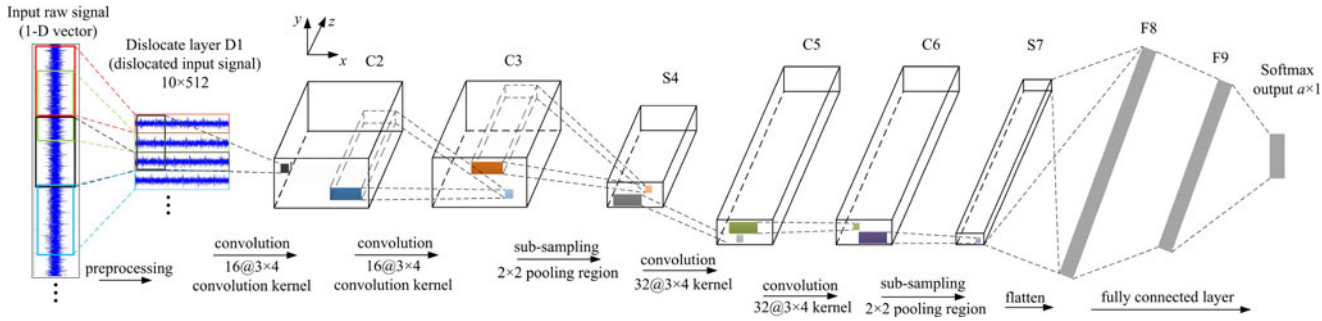

Fig. 8. Architecture of DTS-CNN. For the cubes: x-axis and y-axis represent the width and the height of the feature map, respectively; z-axis represents the number of kernels.

architecture with the ability to reduce the risk of overfitting [32], [33]. Because the number of units is declined layer by layer in the last two layers, F8 and F9 are also two-feature reduction layers, whose dimensions are reduced during the mapping process. Since both the dislocated layer, CNN, and the full-connected network are belong to the overall network, the optimization objective of the objective function and iterative algorithm are the connection weights. That is, weights of both CNN and softmax classifier are trained together in order to minimize the error between the output of the network and label of the sample for better classification accuracy. In order to achieve better classification effect, the feature extraction layers of the deep learning framework are trend to eliminate the similarities and keep the differences between the signals of different classes. Therefore, most of the redundant or useless information of the features has already been reduced when the features are used as the input of softmax classifier. However, the number of input feature for softmax classifier will influence the performance of the proposed method and increase the risk of overfitting, because there is no DR step between CNN and softmax classifier. Although the influence can be reduced by the strategies of sparse connectivity, shared weights, and drop out operation, the architecture of the DTS-CNN and the input feature dimensionality of softmax still should be designed carefully to ensure the best effect. Moreover, since the DTS-CNN method is designed based on deep learning theory, it will be more advantageous for industrial big data.

B. Parameters Optimization Algorithm

In the DTS-CNN method, the rectified linear units (ReLU) function is used as an activation function instead of traditional activation function, because in terms of training time with gradient descent, the saturating nonlinearities sigmoid or tanh function are much slower than the nonsaturating nonlinearity ReLU function [34]. The ReLU function is defined as

$$f(z) = \begin{cases} 0, & \text{if} \quad z < 0 \\ z, & \text{otherwise.} \end{cases} \tag{3}$$

Then, back-propagation (BP) algorithm is used to learn the connective weights. According to the ReLU activation function, we define the cost function of a single sample $(x^{(i)}, y^{(i)})$ to be

$$J(W, b; x^{(i)}, y^{(i)}) = \frac{1}{2}\|h_{W,b}(x^{(i)}) - y^{(i)}\|^2 \tag{4}$$

where $(x^{(1)}, y^{(1)}), \ldots, (x^{(m)}, y^{(m)})$ is training sets with m samples. The overall cost function is given as

$$J(W, b) = \frac{1}{m}\sum_{i=1}^{m} J(W, b; x^{(i)}, y^{(i)}) + \frac{\lambda}{2}\sum_{l=1}^{n_l - 1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(W_{ij}^{(l)})^2 \tag{5}$$

where $W_{ij}^{(l)}$ is the weight associated with the connection between unit j in layer l and unit i in layer $l+1$. s_l is the number of units in layer l. The first term in $J(W, b)$ is an average sum-of-squares error term. The second term is a regularization term. λ is weight decay parameter. Based on (5), the gradient term of output layer (the n_lth layer) can be calculated as

$$\delta^{(n_l)} = -(y - a^{n_l}) \times f'(z^{n_l}) \tag{6}$$

where a_i^l is the output of the ith unit in lth layer and $l = 1, 2, \ldots, n_l$. Therefore, the gradient term of the $l = n_l - 1$, $n_l - 2, \ldots, 2$ hidden layer can also be calculated

$$\delta^{(l)} = [(W^{(l)})^T \delta^{l+1}] \times f'(z^l). \tag{7}$$

Then, we can iteratively update the parameters by the stochastic gradient descent algorithm. First, set $\triangle W_{(1)}^{(l)} = 0, \triangle b_{(1)}^{(l)} = 0$ for all layers. Then, repeat for $i = 1$ to m

$$\triangle W_{(i+1)}^{(l)} = \triangle W_{(i)}^{(l)} + \delta^{(l+1)}(a^{(l)})^T \tag{8}$$

$$\triangle b_{(i+1)}^{(l)} = \triangle b_{(i)}^{(l)} + \delta^{(l+1)} \tag{9}$$

where $\triangle W_{(i+1)}^{(l)}$ and $\triangle b_{(i+1)}^{(l)}$ are the value of $\triangle W^{(l)}$ and $\triangle W^{(l)}$ at ith iteration, respectively. Update the parameters until the termination condition is satisfied

$$W_{(i+1)}^{(l)} = W_{(i)}^{(l)} - \alpha\left[\left(\frac{1}{m}\triangle W^{(l)}\right) + \lambda W^{(l)}\right] \tag{10}$$

$$b_{(i+1)}^{(l)} = b_{(i)}^{(l)} - \alpha\left(\frac{1}{m}\triangle b^{(l)}\right) \tag{11}$$

where $W_{(i+1)}^{(l)}$ and $b_{(i+1)}^{(l)}$ are the value of $W^{(l)}$ and $b^{(l)}$, respectively, at the ith iteration. α is the learning step. And the convolution kernel (namely weights) can be calculated by the BP algorithm.

To sum up, in the proposed framework, because of the local receptive fields and weights sharing strategy, the overfitting problem in traditional fully connected ANN is mitigated. And the periodic fault information between nonadjacent signals can be extracted due to the time series dislocation. The dislocate

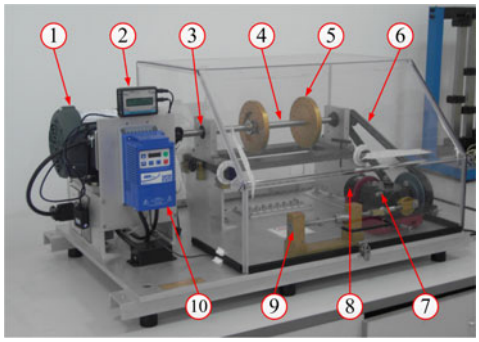

Fig. 9. Experimental setup: (1) induction motor; (2) tachometer; (3) bearing; (4) shaft; (5) load disc; (6) belt; (7) bevel gearbox; (8) magnetic load; (9) reciprocating mechanism; and (10) variable speed controller.

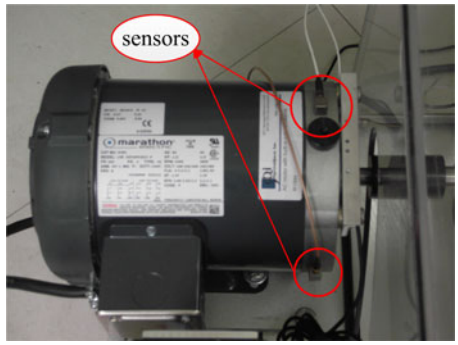

Fig. 10. Location of sensors.

layer can be added in anywhere of a CNN network as the convolution layer and pooling layer. If it is added in the beginning, the input raw signal will be dislocated. If it is added in the middle of the CNN network, the feature map will be dislocated.

In addition, DTS-CNN can extract the relationship between signals in different time series. When a is small, the extracted signals are close to each other, and the time range is small, therefore, the features in high rotation speed can be extracted. Similarly, when the translate distance is large, the extracted signals are far away from each other, and the time range will be large, as a result the features in low rotation speed can be extracted. Due to this structural characteristics of DTS-CNN, combining with the strong capability to extract local features of CNN and the invariance to local translation of pooling operation, the proposed method is also suitable for nonstationary conditions. Therefore, the DTS-CNN is a powerful method that can overcome the shortcomings of traditional CNNs and extract high-level features in industrial applications.

IV. Experimental Demonstrations

To verify the effectiveness of the proposed method, two experiments on an electric machine fault simulator with different induction motor conditions were conducted. The first one is operating under nine different conditions with constant speed, and the second one is operating under six motor conditions with variable speed.

A. Fault Diagnosis of Induction Motor Under Constant Rotation Speed

The experiment with constant speed under nine different induction motor conditions were conducted, as shown in Fig. 9. The vibration signals are acquired by two accelerometers in vertical and horizontal directions. Fig. 10 shows the location of sensors in the induction motor and the vertical vibration signals are used. The vibration data of the simulator under the power supply frequency of 50 Hz are collected with the sampling frequency of 6400 Hz. In this experiment, nine different conditions have been simulated. The descriptions of different induction motor conditions are listed in Table I. There are 35 417 samples in total; out of which 25 000 samples are selected randomly for training and 10 417 samples for testing. The numbers of samples for each class are shown in Table II.

TABLE I
Description of the Nine Motor Experiment Conditions

Label	Condition	Description
1	Broken rotor bar	An intentionally broken rotor bar
2	Bowed rotor	Rotor bent in center 0.01
3	Faulted bearing	One bearing with an inner race fault
4	Unbalanced rotor	Intentionally altered rotor caused unbalance
5	Normal motor	Healthy, no defect
6	Angular misalignment	Displacing one end more than the other
7	Parallel misalignment	Displacing bearing same amount each end
8	Low impedance	Caused by short circuit of the coil
9	High impedance	Caused by open circuit of the coil

TABLE II
Number of Instances for Each Class

	1	2	3	4	5	6	7	8	9	sum
Training	2550	2436	2444	2462	2892	3007	3059	2610	3540	25000
Testing	1081	1020	962	957	1239	1249	1222	1171	1516	10417

After acquiring the vibration signals, the DTS-CNN method is then applied to predict the nine different faults. The length of one period signal can be calculated as 256 points. Therefore, n is set to be 512 to make sure it is longer than a period signal. m is set to be 10. The selection of k is researched due to its significance and uncertainty. In this experiment, k is set to be 3, 8, 10, 15, and 20, respectively. The recall ratio r and precision ratio p are used for performance analysis, which are defined as

$$r = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \qquad p = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{12}$$

where true positive (TP) is correctly classified as positives samples, false positive (FP) is misclassified as positives samples, true negative (TN) is correctly classified as negatives samples, and false negative (FN) is misclassified as negatives samples. To permit an individual evaluation of each class and have a consideration of negative class, the accuracy of each class is also used for comparison, which can be calculated as

$$\text{accuracy} = 1 - \frac{\mathrm{FP} + \mathrm{FN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}}. \tag{13}$$

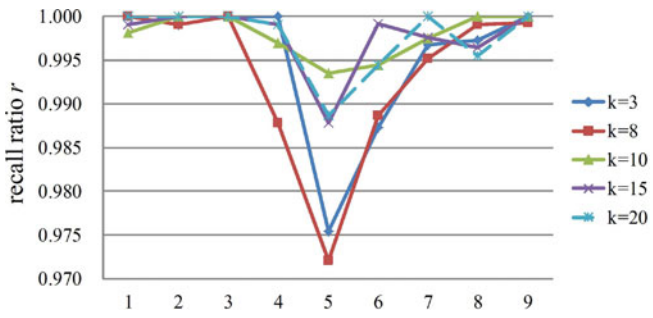

Fig. 11. Recall ratio r with different values of k.

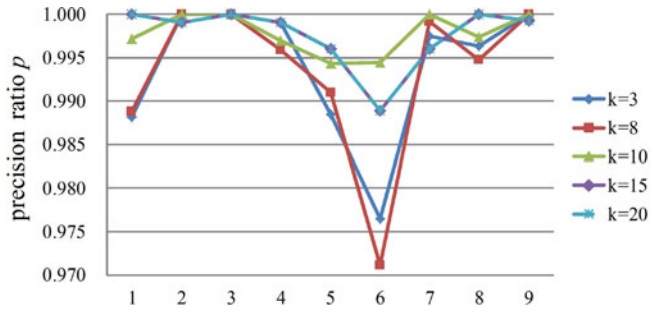

Fig. 12. Precision ratio p with different values of k.

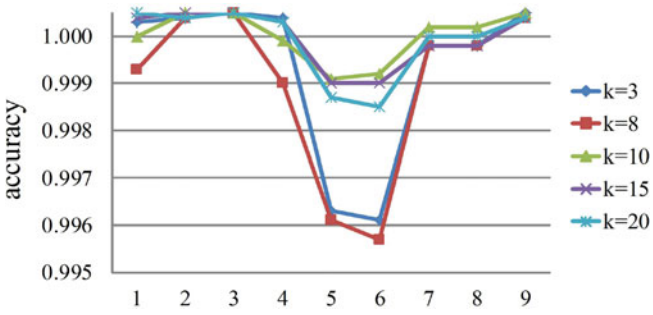

Fig. 13. Accuracy of each class with different values of k.

The results of r and p with different k are shown in Figs. 11 and 12. The accuracy of each class with different values of k is illustrated in Fig. 13.

First of all, it can be seen from Figs. 11–13 that the DTS-CNN is a data-driven method that has excellent performance in fault diagnosis of industrial system. This framework can also learn the complex, high dimensional, nonlinear feature maps from the raw signal directly and automatically from the nine different conditions. Although the accuracy rate is lowest when $k = 3$ and 8, it is still better than 97%.

In addition, with different values of k, DTS-CNN shows similar results and outstanding accuracy rate. When $k = 3$ and 8, the results are worse, because when k is too small, the dislocated distance will also be short. Therefore, signals of each row in D matrix overlap too much, the information in D is not sufficient. As shown in Figs. 11– 13, when $k = 10$, 15, and 20, both the recall ratio r and precision ratio p are higher than 98.5%, the accuracies are better than 99.8%, which are close to the perfect classification results if the random factor is ignored. We will only discuss the results of these three conditions. Fig. 11 shows that except class 5 (normal motor), r of other classes are all better than 99.5%, which means that there are less than 5 samples in 1000 test samples which are misclassified. These misclassified samples can be caused by noise, environment mutation, or other random factors and will not expected to have impact on the final results. The recall ratio of class 5 is 99.35%, 98.78%, and 98.87%, which are relatively lower than other classes. It can be found that this discrepancy is caused by some confusion of class 5 and class 6. As a result, in Fig. 12, k of class 6 is the smallest because some samples belong to class 5 are error classified to class 6. As from Fig. 11, p of other classes are better than 99.44%. After investigation, we found that the confusion of class 5 and class 6 is mainly caused by the installation accuracy and assembly error of the test rig. There is a flexible coupling between the output shaft of the motor and shaft 4 shown in Fig. 9. After several assemble and dismantle processes during the experiment, a slight misalignment will exist inherently in the test rig even with a health motor.

In this experiment, k is determined to be 10. The confusion matrix is shown in Table III. The number from 1 to 9 in the first column represents the test data labels in nine different conditions. The number from 1 to 9 in the first row represents the classification result of test data. For example, the number 1020 in the second row represents that 1020 test samples which labeled 2 (the number in first column is 2) are classified to label 2 (the number in first row is 2), which means that the 1020 test samples are classified correctly. On the other hand, the number 2 in the first row represents that two test samples which labeled 1 is misclassified to label 4. To sum up, the diagonal elements in Table III represent the number of test sets which are classified correctly, and the wrongly classified numbers in other areas are underlined. In addition, the last three columns show recall ratio r, precision ratio p, and the accuracy of each class, respectively.

To illustrate the advantage of the DTS-CNN, the same vibration signal is processed using two state-of-the-art methods: traditional CNN, wavelet packet and SVM (WPT-SVM). Because WPT-SVM is widely used for fault diagnose, and traditional CNN is the DTS-CNN approach without dislocate layer. Therefore, the advantages of dislocate layer can be analyzed by the comparison of DTS-CNN and CNN methods. In traditional CNN, sections of original signal were used as the inputs with the length of 512 points. In WPT-SVM, the vibration signals of each class were divided into signal sections with 2048 points so that wavelet packet transform can achieve better performance. Each signal sample was decomposed by Symmlet 8 wavelet packet and the degree of frequency partition is set to be 4. Therefore, 16 reconstructed signals can be obtained in different frequency bands. The energy proportions of these reconstructed signals were used as features, and the features of all samples were used as the inputs of SVM. Radial basis function is used as the kernel function for the SVM implementation, which can be represented as

$$\kappa(x_i, x_j) = \exp(3.7321\|x_i - x_j\|^2). \tag{14}$$

To deal with the multiclass classification problem, the one-versus-one strategy is applied in SVM. The comparison results are shown in Figs. 14–16.

TABLE III
DTS-CNN RESULTS WHEN $k = 10$

Conditions	1	2	3	4	5	6	7	8	9	r	p	Accuracy
1	1079	0	0	2	0	0	0	0	0	99.81%	99.72%	99.95%
2	0	1020	0	0	0	0	0	0	0	100%	100%	100%
3	0	0	962	0	0	0	0	0	0	100%	100%	100%
4	3	0	0	954	0	0	0	0	0	99.69%	99.69%	99.94%
5	0	0	0	1	1231	7	0	0	0	99.35%	99.43%	99.86%
6	0	0	0	0	7	1242	0	0	0	99.44%	99.44%	99.87%
7	0	0	0	0	0	0	1219	3	0	99.75%	100%	99.97%
8	0	0	0	0	0	0	0	1171	0	100%	99.74%	99.97%
9	0	0	0	0	0	0	0	0	1516	100%	100%	100%

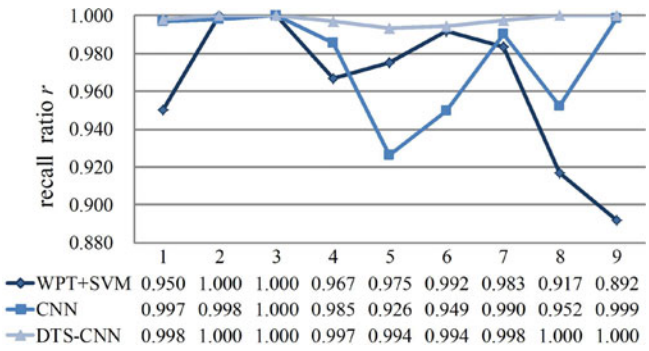

Fig. 14. Recall ratio r of the three classification methods.

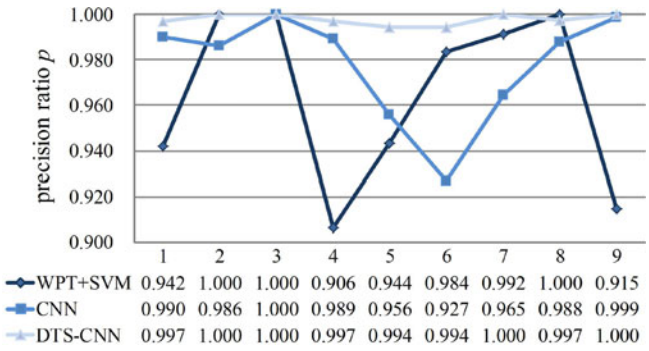

Fig. 15. Precision ratio p of the three classification methods.

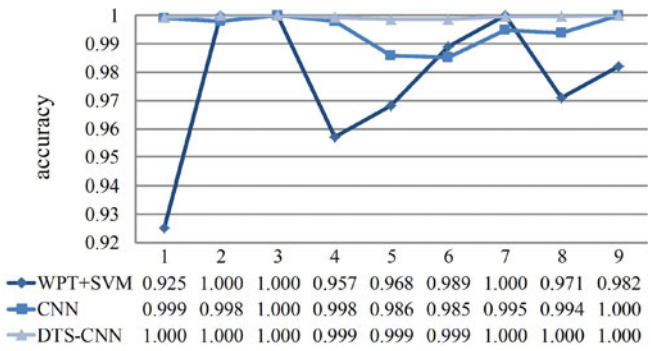

Fig. 16. Accuracy of each class via the three classification methods.

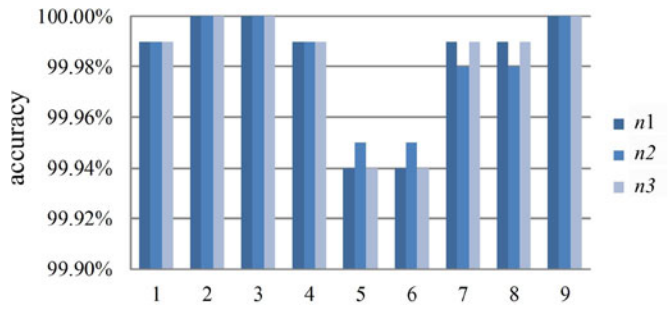

Fig. 17. Accuracy of each class when the signals are added with noise.

It can be seen from Figs. 14–16 that performance of DTS-CNN is the best and the classification accuracy of every category is better than 99%. And the confusion of normal motor and angular misalignment also appears in CNN but is more serious. It is because by adding a dislocate layer, DTS-CNN can extract more profound characteristics of signals and find the essential difference between the two classes. Traditional CNN can only extract the features in an adjacent region and ignore the relationship between nonadjacent signals, therefore some valuable information is missed. In addition, it can also be found that the results of WPT-SVM mainly suffer from misclassification of classes 1, 4, and 9. Because WPT-SVM is a classifier at feature level and the basic idea is the difference between frequency band energy. Therefore, a lot of information will be lost in the process of feature extraction. The results illustrate that the frequency band energy distribution of these three classes is similar and therefore WPT-SVM cannot distinguish them. By contrast, DTS-CNN performs better than WPT-SVM and CNN, because it is a classifier at signal-level that uses signals as the input directly, instead of using extracted features as the inputs, which can keep all information of the signals. And features can be learnt by the proposed framework at multiple levels from raw data automatically and directly, without depending completely on human-crafted features. Consequently, release researchers from the complex feature extraction algorithms and reduce the demand of prior knowledge.

To further illustrate the reliability and robustness against noise of DTS-CNN, the original experiment signals are added with three different degree of noises and analyzed by the proposed method. n 1 is the noise with 0.1 mean and 1 variance, n 2 with 0.2 mean and 1 variance, n 3 with 0.3 mean and 1 variance, with energy corresponding to about 0.1, 0.2, and 0.3 times, respectively, of the original signal energy. The accuracy of each class is shown in Fig. 17. It can be seen that the accuracy of each class is still better than 99.94%, even though considerable

TABLE IV
NUMBER OF INSTANCES FOR EACH CLASS

	1	2	3	4	5	6	Sum
Training	993	1020	972	1082	955	978	6000
Testing	301	274	297	312	289	291	1764

noises are added. And for one test sample, the running time of the classification procedure by DTS-CNN is 6.04 ms. Therefore, it can be concluded that either the traditional CNN or WPT-SVM has its limitations and cannot be suited under every condition, whereas DTS-CNN has strong robustness, real time capability, and excellent performance.

B. Fault Diagnosis of Induction Motor Under Nonstationary Condition

In practical applications, the working condition of an electric machine is always nonstationary. However, the fault diagnosis and recognition of the electric machine under nonstationary condition is always a difficult issue and the research is scanty in this area. Because the signals used for classification are variable as the operating condition change, consequently makes it harder for fault identification. Studies about the variable working conditions mostly focus on feature extraction of a section signal, which is specific to the task and time consuming. And these features often cannot used for classification because they are changing over time. In this paper, the proposed DTS-CNN method shows its advantages in dealing with nonstationary signals intelligently.

The experiment with the nonstationary condition was conducted on the same fault simulator. The experiment setup is shown in Fig. 9. Six induction motor were conducted, namely, bowed rotor, broken rotor bar, faulted bearing, normal motor, high impedance, and unbalanced rotor, as listed in Table I. The sample frequency is 12 800 Hz. The rotation speed is controlled manually, which ranges from 0 to 3600 r/min. The vertical vibration signals are processed using the proposed framework. There are 6000 samples used for training and 1764 samples for testing. The numbers of samples for each class are shown in Table IV. Fig. 18 shows the waveforms of the whole vibration signals and two local samples which begin at 20 and 27 s under six conditions.

In this experiment, n is set to be 512, and m is set to be 10 as with the previous case, and k is set to be 1, 3, 5, and 8. The results with different k are shown in Table V. It can be found that the four overall accuracies are all larger than 92% even if the rotation speed varied greatly. In this experiment, the best result is obtained when $k = 3$ because the rotation speed is changing from 0 to 3600 r/min, therefore the vibration signal is also changing dramatically. As mentioned before, k should be small when the signal changes greatly. When $k = 3$, considering both n and the dislocate distance, the length of a sample is 647 points.

For comparison purpose, a traditional CNN is also used to process the same signal. The results of the two methods are shown in Table VI. The overall accuracies of the two methods are 83.39% (CNN) and 96.32% (DTS-CNN). In this experiment, we did not show the comparison results of other methods because most fault recognition methods are not suitable for variable working condition.

It can be observed from Fig. 18 that the vibration signals are constantly changing, sample 2 is only 7 seconds later than sample 1, but the signals are much different from the signals of sample 1. From Table VI, it can be found that the results of

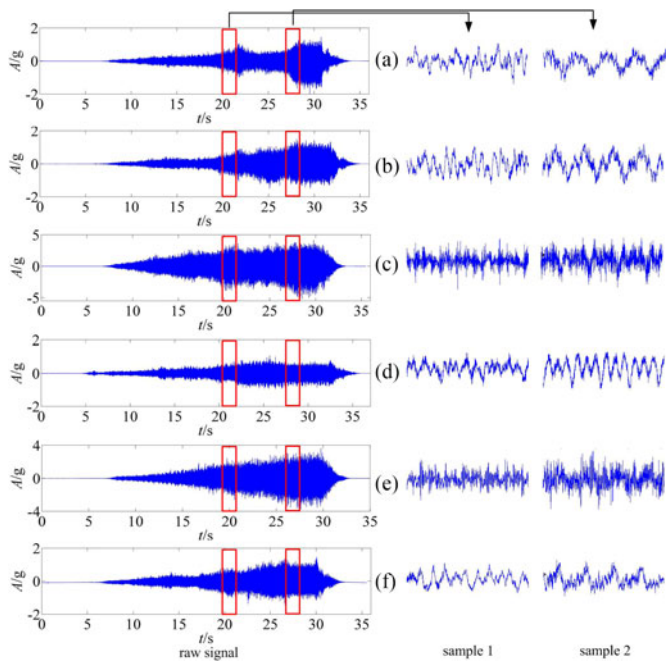

Fig. 18. Vibration signals of the motor under six conditions. (a) bowed rotor, (b) broken rotor bar, (c) faulted bearing, (d) normal motor, (e) high impedance, and (f) unbalanced rotor. The left is the whole signal and the right is the vibration signal of a sample.

TABLE V
DTS-CNN RESULTS WITH DIFFERENT k UNDER NONSTATIONARY CONDITION

Condition	1	2	3	4	5	6
$k = 1$ (accuracy = 94.10%):						
r	95.85%	89.03%	95.45%	95.68%	94.36%	94.60%
p	93.17%	95.95%	97.67%	87.08%	96.17%	97.05%
accuracy	98.02%	97.34%	98.98%	96.60%	98.58%	98.70%
$k = 3$ (accuracy = 96.32%):						
r	99.34%	98.91%	96.97%	93.59%	95.85%	93.47%
p	93.15%	92.81%	98.29%	96.37%	98.58%	99.27%
accuracy	98.64%	98.64%	99.21%	98.24%	99.09%	98.81%
$k = 5$ (accuracy = 93.20%):						
r	91.80%	91.67%	94.86%	93.81%	93.05%	94.04%
p	85.37%	95.16%	97.19%	92.38%	95.63%	95.04%
accuracy	95.86%	97.79%	98.70%	97.45%	98.36%	98.24%
$k = 8$ (accuracy = 92.86%):						
r	90.16%	88.00%	98.15%	94.48%	93.24%	93.04%
p	89.58%	94.90%	95.68%	90.03%	95.83%	92.36%
accuracy	96.49%	97.39%	99.04%	96.88%	98.19%	97.73%

TABLE VI
COMPARISON RESULTS OF CNN AND DTS-CNN

Condition	1	2	3	4	5	6
CNN (accuracy = 83.39%):						
1	219	75	6	11	3	12
2	54	207	0	13	0	8
3	3	7	259	4	2	1
4	9	16	6	280	0	1
5	11	5	1	4	263	2
6	23	10	1	5	0	243
r	67.18%	73.40%	93.84%	89.74%	91.96%	86.17%
p	68.65%	64.69%	94.87%	88.33%	98.13%	91.01%
accuracy	88.27%	89.34%	98.24%	96.09%	98.41%	96.43%
DTS-CNN (k = 3, accuracy = 96.32%)						
1	299	0	0	1	1	0
2	2	271	0	1	0	0
3	2	2	288	4	0	1
4	6	12	1	292	1	0
5	3	1	4	3	277	1
6	9	6	0	2	2	272
r	99.34%	98.91%	96.97%	93.59%	95.85%	93.47%
p	93.15%	92.81%	98.29%	96.37%	98.58%	99.27%
accuracy	98.64%	98.64%	99.21%	98.24%	99.09%	98.81%

CNN are mainly influenced by misclassification of classes 1, 2, and 6. Because classes 1, 2, and 6 are all mechanical failure, and they all have strong periodicity. In addition, both the bowed rotor and the broken bar can cause unbalance, therefore these three conditions are influenced by small unbalance. And from Fig. 18, it can be observed that these three signals have similar local features. Therefore, these three conditions are similar and it is reasonable for CNN to confuse them. On the contrary, the DTS-CNN method still can get excellent effect under such nonstationary conditions and performs much better than CNN in the classification of every category, as shown in Table VI. And the overall accuracy of DTS-CNN is also increased by more than 12 percentage points. Because in the practical industrial applications, raw information of electromechanical systems has strong local structure. Compared with traditional fault diagnosis methods, the DTS-CNN method can enforce on the extraction of local features by restricting the receptive fields of hidden units to be local. And the performance of the fault in an electric machine is mostly cyclical due to the rotating components. Therefore, the relationship of signals between different intervals can be as helpful as the adjacent signals. By the DTS-CNN method, information and correlations between different time points can be extracted by means of dislocating, no matter whether they are border or not. Furthermore, signal in a period is always similar with another period in a short time due to its strong periodicity. By changing the length of dislocation (different times of k), richer information can be extracted between different locations in their period from the raw signal. And due to the deep architecture, the framework can learn complex, high dimensional, non-linear feature maps from large collections of train datasets. As a result, benefit from the specific structure, DTS-CNN shows superior performance in fault recognition under nonstationary conditions, which is always be a tough nut. From Tables II and IV, it can also be found that the testing samples are large enough in the two experiments to demonstrate the generalization of the proposed method. Therefore, the DTS-CNN is a powerful fault diagnosis method for large testing samples in nonstationary conditions.

V. CONCLUSION

A new framework (DTS-CNN) has been presented for fault diagnosis of electric machines. Benefit from the dislocated time series operation in dislocate layer (D1), the DTS-CNN not only keeps the advantages of CNN, but also solves the inapplicability of CNN for mechanical periodic signal. The critical features of DTS-CNN are presented as follows.

1) The proposed method is fed up with raw signals directly and with less dependence on prior knowledge. Compared with traditional diagnosis methods, the proposed framework is operated as a black box and liberates human labor from designing different feature extractors. Moreover, the proposed framework is built based on deep network for industrial big data, which means that its performance is increasing with the increase of samples.
2) The overfitting problem can also be mitigated because of the sharing weights and sparse connectivity. Due to the dislocated time series operation in the dislocate layer, both the high-level complicated local and holistic features can be extracted intelligently. According to the experiments verification, it also shows excellent performance in nonstationary conditions in mechanical systems and strong generalization ability for large testing samples.
3) There are three main parameters in the application of DTS-CNN: m, n, and k and the selection of k will decide the scope of D1 and the final results. If signals changes quickly, k should be small; otherwise k should be large. There is no parameter optimization methods for k yet and it needs further research. In this experiment, k is determined by comparing the results with different k.

REFERENCES

[1] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—Part 1: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3757–3767, Jun. 2015.
[2] J. Antonino-Daviu, S. Aviyente, E. G. Strangas, and M. Riera-Guasp, "Scale invariant feature extraction algorithm for the automatic diagnosis of rotor asymmetries in induction motors," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 100–108, Feb. 2013.
[3] X. Jin, M. Zhao, T. W. Chow, and M. Pecht, "Motor bearing fault diagnosis using trace ratio linear discriminant analysis," *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2441–2451, May 2014.
[4] R. Liu, B. Yang, X. Zhang, S. Wang, and X. Chen, "Time–frequency atoms-driven support vector machine method for bearings incipient fault diagnosis," *Mech. Syst. Signal Process.*, vol. 75, pp. 345–370, 2016.
[5] H. Zhou, J. Chen, G. Dong, and R. Wang, "Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model," *Mech. Syst. Signal Process.*, vol. 72, pp. 65–79, 2016.
[6] J. Tian, C. Morillo, M. H. Azarian, and M. Pecht, "Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with k-nearest neighbor distance analysis," *IEEE Trans. Ind. Electron.*, vol. 63, no. 3, pp. 1793–1803, Mar. 2016.
[7] O. O. Ogidi, P. S. Barendse, and M. A. Khan, "Fault diagnosis and condition monitoring of axial-flux permanent magnet wind generators," *Elect. Power Syst. Res.*, vol. 136, pp. 1–7, 2016.
[8] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, "Bearing fault detection by a novel condition-monitoring scheme based on

statistical-time features and neural networks," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3398–3407, Aug. 2013.
[9] Y. Shatnawi and M. Al-Khassaweneh, "Fault diagnosis in internal combustion engines using extension neural network," *IEEE Trans. Ind. Electron.*, vol. 61, no. 3, pp. 1434–1443, Mar. 2014.
[10] A. K. 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, 2006.
[11] W. Li, S. Zhang, and S. Rakheja, "Feature denoising and nearest-farthest distance preserving projection for machine fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 393–404, Feb. 2016.
[12] N. E. Huang *et al.*, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proc. Roy. Soc. London A, Math. Phys. Eng. Sci.*, vol. 454, no. 1971, 1998, pp. 903–995.
[13] 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, 2014.
[14] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mech. Syst. Signal Process.*, vol. 35, no. 1, pp. 108–126, 2013.
[15] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.
[16] M. Van and H.-J. Kang, "Bearing defect classification based on individual wavelet local fisher discriminant analysis with particle swarm optimization," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 124–135, Feb. 2016.
[17] P. D. Wasserman, "Neural computing, theory and practice," New York, NY, USA: Van Nostrand Reinhold, 1989.
[18] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, "A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models," *IEEE Trans. Rel.*, vol. 61, no. 2, pp. 491–503, Jun. 2012.
[19] J. Yu, "Local and nonlocal preserving projection for bearing defect classification and performance assessment," *IEEE Trans. Ind. Electron.*, vol. 59, no. 5, pp. 2363–2376, May 2012.
[20] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[21] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
[22] D. H. Hubel and T. N. Wiesel, "Receptive fields of single neurones in the cat's striate cortex," *J. Physiol.*, vol. 148, no. 3, pp. 574–591, 1959.
[23] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiol.*, vol. 195, no. 1, pp. 215–243, 1968.
[24] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
[25] Y. Le Cun *et al.*, "Handwritten digit recognition with a back-propagation network," in *Proc. Adv. Neural Inf. Process. Syst.*, 1990, pp. 396–404.
[26] Y. LeCun *et al.*, "Generalization and network design strategies," in *Connections in Perspective.*North-Holland, Amsterdam: Elsevier, 1989, pp. 143–55.
[27] Y. LeCun and M. Ranzato, "Deep learning tutorial," in *Tutorials in International Conference on Machine Learning (ICML'13)*, Citeseer, 2013.
[28] G. L. Martin, "Centered-object integrated segmentation and recognition of overlapping handprinted characters," *Neural Comput.*, vol. 5, no. 3, pp. 419–429, 1993.
[29] J. Wang and J. Jean, "Multiresolution neural networks for omnifont character recognition," in *Proc. IEEE Int. Conf. Neural Netw.*, 1993, pp. 1588–1593.
[30] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, Jan. 1997.
[31] C. Garcia and M. Delakis, "Convolutional face finder: A neural architecture for fast and robust face detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 11, pp. 1408–1423, Nov. 2004.
[32] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
[33] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," 2012. [Online]. Available: http://arxiv.org/abs/1207.0580
[34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inform. Process. Syst.*, 2012, pp. 1097–1105.

Ruonan Liu (S'15) received the B.S. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2013, where she is currently working toward the Ph.D. degree in mechanical engineering.

Her current research focuses on machine learning, deep learning and sparse signal representation for mechanical signal processing, diagnosis and prognosis for industrial systems.

Guotao Meng is currently working toward the Graduate degree in the Department of Computer Science and Technology at Xi'an Jiaotong University (XJTU), Xi'an, China.

He is a Research Intern in the Institute of Artificial Intelligence and Robotics, XJTU. His research interest include machine learning, computer vision, and robotics.

Boyuan Yang (S'16) received the B.S. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2013, where he is currently working toward the Ph.D. degree in mechanical engineering.

His current interests include wind turbine fault diagnostics, signal sparse representation and time–frequency representation, data-driven prognostics, and health management.

Chuang Sun received the Ph.D. degree in 2014 in mechanical engineering from Xi'an Jiaotong University, Xi'an, China.

From March 2015 to March 2016, he was a Postdoctoral Researcher at Case Western Reserve University, Cleveland, OH, USA. He is currently an Assistant Research Fellow in the Collaborative Innovation Center for High-End Manufacturing Equipment, Xian Jiaotong University. His research interests include operational reliability assessment, fault pattern recognition and prognosis.

Xuefeng Chen (M'12) received the Ph.D. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2004.

He is currently a Professor of mechanical engineering at Xi'an Jiaotong University. His current research interests include finite-element method, machinery condition monitoring, and prognostics.

Prof. Chen received the National Excellent Doctoral Dissertation of China in 2007, the Second Awards of Technology Invention of China in 2009, and National Science Fund for Distinguished Young Scholars in 2012.