A Novel Fault Detection Method for Running Gear Systems Based on Dynamic Inner Slow Feature Analysis

YU SONG, SHENG YANG, CHAO CHENG, (Member, IEEE), AND PU XIE

1School of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China
2National Engineering Laboratory, CRRC Changchun Railway Vehicles Co., Ltd., Changchun 130062, China

Corresponding author: Chao Cheng (chengx415@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61903047 and in part by the Jilin Province Development and Reform Commission under Grant 2019C040-3.

ABSTRACT Process and sensor noise often affect the monitoring of slowly changing faults in high-speed trains, which seriously increases the difficulty of slow-change fault detection (FD). By introducing the idea of dynamic inner into the framework of multiple statistics, a novel dynamic inner slowness feature analysis (DISFA) is proposed in this article. This method considers both dynamic and static conditions and improves the detection speed of slow-change faults of running gears, and improves the detection rate of slow-change faults. Compared with other traditional methods, this method is more sensitive to slow changes, can effectively analyze fault-related data, and improve the detection efficiency of slow-changing faults. First, the validity of the method proposed in this article is proved by mathematical deduction, then it is verified by actual operation devices.

INDEX TERMS Slow-change fault, fault detection (FD), dynamic inner slow feature analysis (DISFA), high-speed trains.

I. INTRODUCTION

In recent years, the research on fault detection of running gears of high-speed trains has received extensive attention. The running gear serves as the support of the car body, transmits the load, providing the key equipment for traction and braking. Because of long-time operation, running gears are prone to faults [1]. However, the slow change of faults in running gears, various unknown interference, and noise effects pose difficulties in detecting faults. Successful detection of faults in the running gear earlier can earn the operator more time to take remedial measures, but the detection of slowly changing faults is still a difficult task.

There are many types of faults in the running gear, which are mainly divided into bearing faults and gear faults [2]–[4]. Metal particles and sand particles are squeezed with the gear, the dynamic interaction between the wheelset and the axle box bearing, the oil film ruptures when the meshing tooth surface is relatively sliding, the unit area load on the tooth surface is too large, and the data collection is performed under the condition of interference, these will increase the probability of fault [5]. In addition, long-term high-load use of gears can also cause faults. Due to uneven material or local scratches, contact fatigue will also occur, spalling will occur, and large-scale faults will eventually occur.

The main difficulties of fault detection in running gear include the following three points

1) The vibration amplitude of the running gear is relatively small, and it is difficult to detect the fault when it occurs, but it is likely to cause serious consequences over time, e.g. bearing faults.
2) Running gear produces transients after a period of slowly change. When a subsystem fails, a cascading effect will soon occur, and the fault will spread across the entire system in an instant, and finally cause the entire system to be paralyzed, e.g. System circuit aging.
3) The running gear has a short time lag characteristic. When there is a potential fault, if it is not detected in time, it will cause serious consequences, e.g. broken gear.
In order to solve the slowly changing faults, some of the work has focused on the following aspects: the inspector effectively extracts the slowly changing faults [6] and the inspector fully analyzes and utilizes the historical data through the data-driven method [7]. With the rapid development of science and technology, the complexity and interference of engineering systems have also increased significantly, and it is difficult for researchers to establish accurate analysis models, e.g. the main components of the running gear of a high-speed train include fuel tank, frame, wheel set, basic brake, motor drive and active suspension. Each of these components also includes one or more sub-components [8], [9].

Many researchers have done numerous work on data-driven methods for detecting slow-change faults [10]-[12]. Since most of the measured signals of high-speed trains are non-Gaussian, a non-Gaussian process fault detection method based on generalized gauge correlation analysis and stochastic algorithm [6] is proposed to improve the fault detection rate. For non-linear problems, a kernel principal component analysis combining previous fault information is proposed in [9], [13], [14]. In [15]-[17], the multi-dimensional fault diagnosis of high-speed trains is transferred to the multi-dimensional convolutional neural network to improve the detection accuracy. However, these methods are not good at detecting slowly changing faults. Because the interference in the signal cannot be eliminated, it is difficult to detect slowly changing faults.

In order to avoid the slow-change faults evolved into permanent faults, timely fault detection has become a current hot research issue [18], [19]. In the research of [8], [20], a new data-driven method probability PCA was proposed, which is used in the electric drive of high-speed trains. This method effectively improves the fault detection rate and can also be used for non-Gaussian signals. In order to monitor bearing failures more effectively, a new method of on-line detection of initial failures of rolling bearings with adaptive depth feature matching is proposed in [21]. According to research [22], a parallel monitoring slow feature analysis is proposed to simultaneously monitor operating conditions and process abnormalities. Next, a new dynamic inner-PCA is proposed in [23], which use static methods to deal with dynamic problems by preprocessing the data.

In this study, we will first propose a novel dynamic SFA (DISFA) for slow-change faults. Immediately, a DISFA-based FD method is proposed and used for fault detection of running gears of high-speed trains. The main advantages and innovations of the proposed method are as follows:

1) Compared with the SFA method, the DISFA method is more sensitive to slow changes. It can not only effectively analyze fault-related data but also improve the detection efficiency of slowly changing faults.
2) It can increase the speed of slowly changing faults of the running gear and improves detection rates of the slowly changing faults.
3) The DISFA method takes both the static and dynamic characteristics of systems into account.
4) Its flexibility and realizability have great potential.

The main arrangements for the remaining chapters of this article are as follows. Section II of this article gives a brief introduction to the running gear system and slow-change failures. After that, the method proposed in this article is introduced. In Section III, the theoretical core of the proposed method and the detection of slowly changing faults are introduced in detail. Section IV discusses the test results of the gear experiment on high-speed trains. Finally, Section V summarizes and looks forward to this article.

II. PRELIMINARIES

In this section, a brief description of the running gear structure in high-speed trains is given, then the problem of slow-change faults is proposed on this basis, and corresponding solutions are proposed at the end.

A. RUNNING GEARS OF HIGH-SPEED TRAIN

In actual engineering, data collected from sensors mainly include speed, gearbox temperature, vibration and shock, etc. The health status of running gears can be judged through these data. In the collected data, temperature and vibration are the main reference standards for faults of running gear. The bogie mainly includes the following four parts: axle box bearing, motor bearing, motor driving end bearing and big gear.

In order to perform effective FD for slowly changing faults, an overall system model of the operating gear must be available. In order to make the analysis more in line with actual engineering, this article chooses to analyze the properties of the bogie-motor bearing subsystem.

B. FAULT DESCRIPTION AND OBJECTIVES

Considering the overall system model of running gears, the fault model is

\[ X = X^* + F + N \]  

where \( X^* \) represents normal data, \( F \) represents slowly changing faults, and \( N \) represents noise data. It is difficult to separate the slowly changing faults \( F \) from the noise data \( N \). The running gear is in a sub-healthy transition during long-term fatigue damage and damage accumulation. This state will not directly cause an accident, but a long-term sub-healthy state will cause serious accidents.

- The bogie is the running device of the high-speed electric multiple unit (EMU). It has important functions such as load-bearing, vibration reduction, guidance, traction and braking. It determines the operating speed and quality of trains. The bogie is prone to scratch and peel wheel tread, which will not affect to drive a certain extent. At this time, the bogie is in a sub-healthy transition. Minor abrasions do not affect the train operation, but with long-term driving, the fault changes slowly. If the fault is not found in a time range from sub-health to fault, it may cause a serious accidents.
- The EMU bogie bearing is generally a rolling bearing, which is the most severe part of EMU working conditions.
In operation of the EMU, it plays the role of bearing and transmitting loads. The bearing in the sub-health transition state changes with speed, friction loss, and the effect of transmitting load. The fault changes slowly until it becomes a fault.

As an indispensable universal component for connecting and transmitting torque in mechanical equipment, we must pay attention to the fatigue and vibration of the gear box. In the case of long-term high speed and high vibration, the gear failure in the sub-health state continues to change slowly. There is no impact from the running speed of the train. Once the rainy weather occurs, the anti-skid system is not well controlled and the speed cannot be reduced until the final accident.

On the basis of the above discussion, in order to effectively solve FD problems of slow-change faults of running gear system in the high-speed train, following problems need to be solved:

1) Establish a model that matches the system and no longer use mathematical models and expert knowledge to detect faults.
2) When the fault is in the sub-health state to the slow change of the fault state, the fault is detected in time.
3) Fully analyze the data to effectively detect the fault.

III. METHODOLOGY

In this part, a DISFA algorithm is proposed, then how to use the proposed method for FD is introduced in detail.

A. THE DYNAMIC INNER SFA

The purpose of this article is to obtain a matrix $X_f$ that is approximately equal to the slowly changing fault matrix $F$ by preprocessing the original data $M$, and then perform SFA on the matrix $X_f$ to perform FD on the process. Generally speaking, the dynamics among latent variables predicting past can be expressed as

$$t_i = \alpha_1 t_{i-1} + \cdots + \alpha_s t_{i-s} + r_i$$  \(2\)

The idea of dynamic inner model is a derivative of the moving average model, which is based on an autoregressive moving average model (ARMA). The method proposed in this article is extended on this basic idea.

Gives the data matrix $X$

$$X = [x_1, x_2, \cdots, x_{n+1}]^T$$

and the following matrix can be obtained by $X$

$$X_i = [x_i, x_{i+1}, \cdots, x_{n-i}]^T \quad \text{for} \quad i = 1, 2, \cdots, s + 1$$

$$\dot{M}_i = \dot{X}_i^T X_i \quad \text{for} \quad i = 1, 2, \cdots, s + 1$$

$$M_i = X_i^T \dot{X}_i \quad \text{for} \quad i = 1, 2, \cdots, s + 1$$

where $\dot{X}_i$ is the first derivative of $X_i$. The pure ARMA model can only make predictions data, and cannot guarantee that maximizing predicted data is similar to a fault-free data. Combination of the SFA idea and the ARMA model makes the predicted data close to fault-free data.

The SFA method is a development of independent principal component analysis method [22]. It is an algorithm that extracts slow feature signals from rapidly changing signals. Given a set of multi-dimensional input data $X_f$, the purpose of SFA is finding a set of transfer functions $k_1(x) k_2(x) \cdots k_m(x)$ to obtain output data $y_j = k_j(X_f)$, and the resulting components change slowly. The rate of changes in this article is expressed by the squared mean of the first derivative. Therefore, the specific framework of the optimization problem is as follows:

$$\min \Delta(y_j) = \langle \dot{y}_j^2 \rangle_k $$  \(3\)

Restrictions:

$$\langle y_j \rangle_k = 0 \quad \text{(4)}$$

$$\langle (\dot{y}_j)^2 \rangle_k = 1 \quad \text{(5)}$$

$$\langle y_j \dot{y}_j \rangle_k = 0 \quad \text{(6)}$$

Among them, $\dot{y}_j$ is the first derivative of $y_j$, the $\langle \rangle$ expression takes the mean in time, the constraints (4) and (5) avoid the situation of constant solution, and the constraint (6) avoid correlation. The correlation among components makes the information carried by each component different from each other.

The optimization problem of SFA can be transformed into the decomposition of generalized eigenvalues,

$${AW = BWA}$$  \(7\)

where $A = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is the diagonal matrix composed of generalized eigenvalues, $W = (w_1, w_2, \ldots, w_m)$ is eigenvector matrix of corresponding generalized, $A = (X_f^T X_f)_k$, $B = (X_f^T X_f)_k$. The slow feature $y_j$ can be easily obtained by $W$.

$$y = Rz = Wx$$  \(8\)

In general, solving the problem only requires two singular value decompositions (SVD) of equation (7). So the prediction of dynamic inner model can be expressed as

$$\max_{\alpha, \gamma} y^T (\alpha \otimes Z_s) y$$

s.t. $\|\alpha\| = 1$, $\|y\| = 1$  \(9\)

where $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_s]$, $\gamma$ is the feature weight vector, $Z_s = [(M_i)^{-1}M_s, \cdots, (M_1)^{-1}M_1]$ and $\alpha \otimes \gamma$ is the Hadamard product. In (9), the problem can be solved by using Lagrange multipliers. Define

$$J = y^T (\alpha \otimes Z_s) x + \frac{1}{2} c_{\alpha} (1 - \alpha^T \alpha) + \frac{1}{2} c_{\gamma} (1 - y^T y)$$  \(10\)

By analyzing the Lagrange formula $J$, the following relationship can be obtained. The specific solution process is given in Appendix.

$$c_{\alpha} = \gamma^T \left[ \sum_{i=1}^s (M_i)^{-1} M_i \right] \gamma$$
First suppose that s is deterministic, and then choose variable s and the number of cycles l need to be determined. Through Algorithm 1 and (17) can get the resulting number of cycles 95% of the autocovariance is capture by the first cycle. When \( s = 1 \), \( \Phi_i = T_{s+1} - \bar{T}_s \hat{\Xi} \) The sum of errors is \( \Delta = \Phi_1^2 + \Phi_2^2 + \ldots + \Phi_l^2 \), \( \Delta \) can rewrite as

\[
\Delta = (T_{s+1} - \bar{T}_s \hat{\Xi}^T)(T_{s+1} - \bar{T}_s \hat{\Xi})
\]

Let the first-order partial derivative of \( \Delta \) with respect to \( \hat{\Xi} \) be equal to 0, and find the minimum \( \Delta \) of \( \hat{\Xi} \)

\[
\frac{\partial \Delta}{\partial \hat{\Xi}} = -T_{s+1}^T \bar{T}_s - T_{s+1} \bar{T}_s + 2 \hat{\Xi} T_{s+1}^T \bar{T}_s = 0
\]

the estimated value is

\[
\hat{\Xi} = (T_{s+1}^T \bar{T}_s)^{-1} (T_{s+1}^T \bar{T}_s \bar{T}_{s+1})^{-1} T_{s+1} \hat{\Xi} \]

which can further be used to calculate the prediction of \( X_{s+1} \)

\[
\hat{X}_{s+1} = \hat{T}_{s+1} \hat{\Xi}
\]

where \( \hat{P} = [p_1 \ p_2 \ \ldots \ p_l] \) is loading matrix in each \( p_i \) defined in (12).

Forecast from past data, fault data matrix can be calculated.

\[
X_f = (M_{s+1} - \hat{T}_{s+1} \hat{\Xi})X
\]

Remark 2: The ARMA algorithm used in this article only uses one component, namely AR(s) model, to carry out a series of formula derivations and matrix operations.

Remark 3: The matrix \( X_f \) obtained by (17) approximately equal to the matrix \( f \), because the noise always exists when predicting \( X_{s+1} \) by DISFA model. In an ideal case, the noise can be exactly eliminated by making a difference in (17), and the residual matrix can be regarded as the initial fault matrix, that is, \( X_f = F \).

Performing SFA on the fault matrix \( X_f \) again and dividing all features into two parts, one part is slow feature and the other is fast feature. As follows

\[
m = m_d + m_e
\]

after \( m \) is divided into two parts, \( W \) is also divided into two parts

\[
W_2 = W_{2,d} + W_{2,e}
\]

where \( m \) represents all features, \( m_d \) represents slow features, \( m_e \) represents fast features, \( W_{2,d} \) is the feature matrix of slow features, and \( W_{2,e} \) a is the feature matrix of fast features.
Remark 4: \( m_d \) and \( m_e \) are determined by comparing the self-covariance of the features. The specific details can be found in the paper [22].

Therefore, the DISFA method proposed in this article can be expressed in Algorithm 2. The specific idea of the method proposed in this article is shown in Fig. 1.

Algorithm 2 DISFA Method
1: First normalize \( X \).
2: Obtain the slow-change fault data \( X_f \) through Algorithm 1 and (17).
3: By (8) we can get \( y \).
4: Decompose features into slow features and fast features as \( y_d, y_e \).

\[
Y_d = W_{2,d}X_f \\
Y_e = W_{2,e}X_f
\]

B. DISFA-BASED INCIPIENT FAULT DETECTION
According to the proposed DISFA, test statistics can be formulated as

\[
T^2_d = Y_d^T \Sigma_d^{-1} Y_d \\
T^2_e = Y_e^T \Sigma_e^{-1} Y_e
\]

and

\[
T^2_d \sim \chi^2_{m_d} \\
T^2_e \sim \chi^2_{m_e}
\]

where \( T^2_d \) and \( T^2_e \) are test statistics corresponding to the slow and fast features. \( \chi^2_{m_d} \) and \( \chi^2_{m_e} \) are the thresholds of \( T^2_d \) and \( T^2_e \) respectively. If only limited offline data can be used, the threshold of \( T^2_d \) and \( T^2_e \) statistics can be selected according to the following \( F \) distribution. As follows

\[
\chi^2_{m_d} = \frac{m_d((N + s)^2 - 1)}{(N + s)(N + s - m_d)}F_a(m_d, N + s - m_d)
\]

then, using (25) to determine whether the system is operating normally.

\[
\begin{align*}
H_0 : & \quad T^2_d \leq \chi^2_{m_d} \quad \text{and} \quad T^2_e \leq \chi^2_{m_e} \quad \Rightarrow \quad \text{fault free} \\
H_1 : & \quad \text{otherwise} \quad \Rightarrow \quad \text{fault and alarm}
\end{align*}
\]

IV. VERIFICATIONS
In order to test the credibility of the method proposed in this article, a dynamic and static parameter test bench developed by CRRC Changchun Railway Vehicle Co., Ltd., which is used to simulate the movement of running gears in high-speed train, as shown in Fig. 2. Taken the actual running train as an example, which is used to record the data of normal train operation and fault data.

\[
\text{FIGURE 1. Flow chart of DISFA algorithm.}
\]

\[
\text{FIGURE 2. Testing bench of high-speed trains.}
\]

A. EXPERIMENTAL VERIFICATION
1) INJECT SLOWLY CHANGING FAULTS
According to the train engine speed exceeding 1000r/min and running smoothly, the method is verified by collecting the signals of the four sensors of gearbox, bogie, motor and bearing. In order to obtain the verification data set, three initial failures were injected into running gear systems from 25 minutes as follows: pinion gear fault \( f_1 \), bogie fault \( f_2 \), and bearing fault \( f_3 \).

2) DETECT SLOWLY CHANGING FAULTS
On the basis of training data composed of fault data and normal data, DISFA algorithm given above is used for detection. Fig. 3 shows the detection picture when no fault is injected. In order to emphasize the effectiveness of the FD method proposed in this article, the monitoring results of the three methods of DISFA, dynamic PCA (DPCA) and SFA are compared and analyzed. For slowly changing faults \( f_1 \), the detection results of DISFA, DPCA, and SFA are shown in Fig. 4 to Fig. 6; for slowly changing faults \( f_2 \),
the detection results of DISFA, DPCA, and SFA are shown in Fig. 7 to Fig. 9; for slowly changing faults \( f_3 \), the detection results of DISFA, DPCA, and SFA are shown in Fig. 10 to Fig. 12.

3) COMPARISON AND EFFICIENCY ANALYSIS
The slowly changing fault \( f_1 \) is injected at 25 minutes, and the proposed method \( T_2^f \) detects the fault at 26 minutes. Before the fault injection, there are false alarms of individual values. Except for the extremely small points such
as 241, 261, which is below the threshold, this method is good for fault detection. DPCA and SFA detected faults at 35 min and 39 min respectively, and only a few points are below the threshold in the later period. In general, the proposed method shortens the time of first fault detection compared with DPCA and SFA.

The slowly changing fault \( f_2 \) is injected at 25 minutes, and the proposed method \( T_2^d \) detected the fault at 29 minutes. There is a short-term fluctuation in early stage of the fault detection, and individual values are below the threshold, which generally does not affect the detection result. DPCA and SFA detected faults at 41 min and 42 min respectively, then changing on the threshold. In general, three methods have obvious fluctuations after detecting the fault, but the proposed method is significantly better than DPCA and SFA in slowly changing faults.

The slowly changing fault \( f_3 \) is injected at 25 minutes, and the proposed method \( T_2^d \) detected the fault at 36 minutes. Compared with the previous two faults, the detection time of this fault has a certain delay. It may be that the bearing fault is slightly difficult to be detected. DPCA detecte the fault at 53min, SFA briefly fluctuate after the fault is first detected at 39min, and the fault could be completely detected after 53min. In general, the monitoring efficiency of DISFA is still better than DPCA and SFA.

### B. DISCUSSIONS

In order to emphasize the effectiveness of the FD method proposed in this article, the following discussion focuses on 1) comparing and analyzing the false alarm rate and missing alarm rate; 2) how the proposed method improves the efficiency of detection.

#### 1) COMPARISON ANALYSIS

In order to show that the proposed method is good for FD, false alarm rate (FAR) and missing alarm rate (MAR) of the three methods are shown in Table 1. It can be clearly seen from the table that for the three different failures, the MAR of the party mentioned in this article is lower than that of other methods. At the same time, the detection efficiency for \( f_2 \) and \( f_3 \) is better than other methods, and the detection efficiency for \( f_1 \) is not much different.
TABLE 1. Performance Comparisons for Running Gear.

| Method | f₁ | f₂ | f₃ |
|--------|----|----|----|
|        | MAR | FAR | MAR | FAR | MAR | FAR |
| SPA    | 8.73% | 3.97% | 6.35% | 1.08% | 5.16% | 5.42% |
| DPCA   | 17.20% | 5.09% | 6% | 3.64% | 3.6% | 6.18% |
| DISFA  | 2.78% | 4.29% | 2.38% | 0.36% | 2.78% | 0%

2) IMPROVE DETECTION EFFICIENCY

This article uses the ARMA model to predict the current time data from the past time data, and then obtain slowly changing fault data by difference. Although noise and interference suppression existed in the prediction, the final difference was made to cancel out the noise and interference data, and only the fault data remained. Compared with other methods, the proposed method extracts useful data from noise and initial data more effectively.

V. CONCLUSION

This article introduces the idea of dynamic inner to the multivariate statistical framework, this article proposes a new DISFA method to deal with slowly changing faults. By comparing with other methods, it can be seen that the method proposed in this article is more sensitive to slowly changing faults. Strict mathematical derivation and fully practiced experiments on actual operating devices have proved its effectiveness. It is worth mentioning that the classification work after slowly changing fault detection is worthy of further research in our future work.

APPENDIX

MATHMATICAL DERIVATIONS OF (11)

Proof: Find the partial derivatives of \( \gamma \) and \( \alpha \) for equation (J), and make their partial derivatives be zero, respectively.

We have

\[
\frac{\partial J}{\partial \gamma} = \alpha \gamma + (\alpha \gamma \gamma)^T - \gamma \gamma \gamma = 0
\]

\[
\frac{\partial J}{\partial \alpha} = \gamma \gamma \gamma - \gamma \gamma \alpha = 0
\]  

(A.1)

Let \( G_{\beta} = \alpha \gamma \), the above equation can be simplified as

\[
\frac{\partial J}{\partial \gamma} = G_{\beta} \gamma + G_{\beta}^T \gamma - \gamma \gamma \gamma = 0
\]  

(A.2)

\[
\frac{\partial J}{\partial \alpha} = \gamma \gamma \gamma - \gamma \gamma \alpha = 0
\]  

(A.3)

By multiplying \( \gamma \gamma \) on both sides of (A.2) and multiplying \( \alpha \gamma \) on both sides of (A.3), the following formula can be obtained

\[
c_{\gamma} \gamma = \gamma (G_{\alpha} + G_{\alpha}^T) \gamma = J
\]  

(A.4)

The above results show that the maximum value of \( J \) is equal to values of \( \lambda_{\beta} \) and \( \lambda_{\alpha} \), \( w \) is feature vector of \( (G_{\alpha} + G_{\alpha}^T) \)

corresponding to the largest eigenvalue. Defining the latent scores \( t \) as follows,

\[
i = \dot{X} \gamma
\]

\[
t = (X^{-1})^T \gamma
\]  

(A.5)

then the following results can be obtained

\[
c_{\alpha} \alpha = \gamma T \sum_{i=1}^T (M_i - 1)M_i \gamma
\]

\[
c_{\gamma} \gamma = \sum_{i=1}^T \alpha_i (M_i \gamma T + M_i^{-1} \gamma T)
\]  

(A.6)

Hence, (11) can be obtained.

REFERENCES

[1] Z. Wang, W. Zhang, Z. Yin, Y. Cheng, G. Huang, and H. Zou, “Effect of vehicle vibration environment of high-speed train on dynamic performance of axle box bearing,” Vehicle Syst. Dyn., vol. 57, no. 4, pp. 543–563, Apr. 2019.

[2] C. Cheng, X. Qiao, H. Luo, W. Meng, M. Gao, B. Zhang, and X. Yin, “A semi-quantitative information based fault diagnosis method for the running gears system of high-speed trains,” IEEE Access, vol. 7, pp. 38168–38178, 2019, doi: 10.1109/ACCESS.2019.2969796.

[3] D. Wang, Z. Peng, and L. Xi, “Theoretical and experimental investigations on spectral Lp/Lq norm ratio and spectral gini index for rotating machine health monitoring,” IEEE Trans. Autom. Sci. Eng., early access, May 29, 2020, doi: 10.1109/TAESIE.2020.2994741.

[4] H. Wang, Z. Liu, A. Nunez, and R. Drollevoet, “Entropy-based local irregularity detection for high-speed railway cateraries with frequent inspections,” IEEE Trans. Instrum. Meas., vol. 68, no. 10, pp. 3536–3547, Oct. 2019.

[5] Z. Liu, K. Liu, J. Zhong, Z. Han, and W. Zhang, “A high-precision positioning approach for catenary support components with multiscale difference,” IEEE Trans. Instrum. Meas., vol. 69, no. 3, pp. 700–711, Mar. 2020.

[6] Z. Chen, S. X. Ding, T. Peng, C. Yang, and W. Gui, “Fault detection for non-Gaussian processes using generalized canonical correlation analysis and randomized algorithms,” IEEE Trans. Ind. Electron., vol. 65, no. 2, pp. 1559–1567, Feb. 2018.

[7] W. Deng, S. Zhang, H. Zhao, and X. Yang, “A novel fault diagnosis method based on integrating empirical wavelet transform and fuzzy entropy for motor bearing,” IEEE Access, vol. 6, pp. 35042–35056, 2018, doi: 10.1109/ACCESS.2018.2834540.

[8] H. Chen, B. Jiang, N. Lu, and Z. Mao, “Deep PCA based real-time incipient fault detection and diagnosis methodology for electrical drive in high-speed trains,” IEEE Trans. Veh. Technol., vol. 67, no. 6, pp. 4819–4830, Jun. 2018.

[9] J. Ni, C. Zhang, L. Ren, and S. X. Yang, “Abrupt event monitoring for water environment system based on KPCA and SVM,” IEEE Trans. Instrum. Meas., vol. 61, no. 4, pp. 980–989, Apr. 2012.

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

[11] X. Ding, Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems. New York, NY, USA: Springer-Verlag, 2014.

[12] J. C. Hong, Z. P. Wang, W. Chen, and Y. Yao, “Synchronous multi-parameter prediction of battery systems on electric vehicles using long short-term memory networks,” Appl. Energy, vol. 254, no. 15, Nov. 2019, Art. no. 113648.

[13] X. Deng, X. Tian, S. Chen, and C. J. Harris, “Nonlinear process fault diagnosis based on serial principal component analysis,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 3, pp. 560–572, Mar. 2018.

[14] E. Tian, X. Wang, and C. Peng, “Probabilistic-constrained distributed filtering for a class of nonlinear stochastic systems subject to periodic DoS attacks,” IEEE Trans. Circuits Syst. I, Reg. Papers, early access, Jul. 16, 2020, doi: 10.1109/TCSI.2020.3007953.

[15] W. Yu and C. Zhao, “Broad convolutional neural network based industrial process fault diagnosis with incremental learning capability,” IEEE Trans. Ind. Electron., vol. 67, no. 6, pp. 5081–5091, Jun. 2020.
[16] L. Kou, Y. Qin, X. Zhao, and X. Chen, “A multi-dimension end-to-end CNN model for rotating devices fault diagnosis on high-speed train bogie,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2513–2524, Mar. 2020.

[17] T. Xue, M. Zhong, L. Li, and S. X. Ding, “An optimal data-driven approach to distribution independent fault detection,” *IEEE Trans. Ind. Informat.*, vol. 16, no. 11, pp. 6826–6836, Nov. 2020.

[18] H. Chen, B. Jiang, S. X. Ding, and B. Huang, “Data-driven fault diagnosis for traction systems in high-speed trains: A survey, challenges, and perspectives,” *IEEE Trans. Intell. Transp. Syst.*, vol. 66, no. 6, pp. 4716–4725, Jun. 2019.

[19] C. Yang, W. Gui, Z. Chen, J. Zhang, T. Peng, C. Yang, H. R. Karimi, and S. X. Ding, “Voltage difference residual-based open-circuit fault diagnosis approach for three-level converters in electric traction systems,” *IEEE Trans. Power Electron.*, vol. 35, no. 3, pp. 3012–3028, Mar. 2020.

[20] H. Chen, B. Jiang, W. Chen, and H. Yi, “Data-driven detection and diagnosis of incipient faults in electrical drives of high-speed trains,” *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4716–4725, Jun. 2019.

[21] W. Mao, J. Chen, X. Liang, and X. Zhang, “A new online detection approach for rolling bearing incipient fault via self-adaptive deep feature matching,” *IEEE Trans. Instrum. Meas.*, vol. 69, no. 2, pp. 443–456, Feb. 2020.

[22] C. Shang, F. Yang, X. Gao, X. Huang, J. A. K. Suykens, and D. Huang, “Concurrent monitoring of operating condition deviations and process dynamics anomalies with slow feature analysis,” *AIChE J.*, vol. 61, no. 11, pp. 3666–3682, Nov. 2015.

[23] Y. N. Dong and S. J. Qin, “A novel dynamic PCA algorithm for dynamic data modeling and process monitoring,” *J. Process Control.*, vol. 67, no. 5, pp. 1–11, Jul. 2018.

**YU SONG** received the B.S. degree from the Changchun University of Technology and the M.S. degree from the Shenyang University of Technology.

He is currently with the Changchun University of Technology. His research interests include fault detection, embedded systems, and UAVs.

**SHENG YANG** received the B.S. degree from Wuhan Donghu University. He is currently pursuing the degree with the Changchun University of Technology.

His current research interest includes the detection and diagnosis of the puzzle based on data-driven methods.

**PU XIE** received the Ph.D. degree from the Department of Mechanical Engineering, Tsinghua University, in 2016.

He joined CRRC Changchun Railway Vehicles Co., Ltd. (CRRC), Changchun, where he is operating with leading manufacturing technologies. He is currently the Team Leader of the Structural Integrity Group (14 Members). His current research interests include fault detection and diagnosis, including embedded sensors and stretchable sensor networks with built-in self-diagnostics, integrated diagnostics, damage tolerance, and the failure analysis of high-speed railway vehicles.

**CHAO CHENG** (Member, IEEE) received the B.S. degree from Dalian University for Nationalities, in 2007, and the Ph.D. degree from Jilin University, in 2014.

He is currently a Teacher and a Doctoral Supervisor of the Changchun University of technology. His current research interests include fault detection and diagnosis based on data-driven method, residual life prediction, and artificial intelligence.

**VOLUME 8, 2020**

**211379**