IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 20, 2020, accepted May 31, 2020, date of publication June 4, 2020, date of current version June 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3000068

Intelligent Diagnosis for Railway Wheel Flat Using Frequency-Domain Gramian Angular Field and Transfer Learning Network

YONGLIANG BAI, JIANWEI YANG, JINHAI WANG, (Member, IEEE), AND QIANG LI
School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Jianwei Yang (yangjianwei@bucea.edu.cn)

This work was supported in part by the Key Project of the 2020 Science and Technology Plan of Beijing Municipal Commission of Education under Grant KZ202010016025, in part by the National Natural Science Foundation of China under Grant 51975038 and Grant 51605023, and in part by the Nature Science Foundation of Beijing under Grant 19L00001.

ABSTRACT The intelligent diagnosis of wheel flat based on vibration image classification is a promising research subject for performance maintenance of railway vehicles. However, the image representation method of vibration signal and classification network construction under small samples have become two obstacles to intelligent diagnosis of wheel flat. This paper presents a novel frequency-domain Gramian angular field (FDGAF) algorithm to encode the vibration signal of wheel flat to featured images. Furthermore, a modified transfer learning network is introduced to classify these featured images under small samples without any involvement of prior knowledge. The proposed FDGAF can calculate the Gramian angular matrix of axle box acceleration signal in frequency domain and assign frequency position dependence to the featured images to preserve original characteristic information. Then, these featured images can be intelligent classified by a transfer learning network under the condition of 30 sample without require of prior knowledge. To verify the efficiency of this proposed method, 12 cases of artificial wheel flats are processed on a scaled railway test rig, and their axle box acceleration signals are collected to obtain visual diagnosis results. The verfication proves that FDGAF is able to obtain accurate diagnostic results with high separability, for separability indexes of FDGAF reaches 10.8, 8.7, 14.9, and 5.8. We anticipate that this method will find use in the performance maintenance of railway vehicles and the improvement of industrial condition monitoring.

INDEX TERMS Intelligent transportation system, fault diagnosis, railway safety, wheels, frequency domain analysis, knowledge transfer.

I. INTRODUCTION

Railway vehicles are efficient due to their large capacities and high speeds. However, these advantages result in many wheel surface defects, especially wheel flats with certain numbers and sizes [1]–[3]. Wheel flat can increase contact force instantaneously and make severe noise [4]–[8]. Besides, rail crack propagation failure and wheel polygonization exacerbation can also be generated by the wheel flat [9]–[10]. These problems brought by wheel flat make it urgent to study intelligent diagnosis method. The current monitoring methods for railway wheel flat are to measure the length of flat on static vehicles, for example, the length limit of wheel flat in Europe is 60, and 40mm in China [11]. However, these measurements require high labor costs and comprehensive prior knowledge, which significantly reduces the intelligence of wheel flat monitoring. For better performance maintenance of the railway wheel, it is necessary to study the diagnosis method for wheel flat and increase its intelligence as much as possible.

The first thing for wheel flat diagnosis is to select a type of signal that can represent characterized wheel flat information. Some indirect approaches were studied based on strain, sound, electricity, and vibration [8]–[14]. Among them, vibration acquired from the axle box is the most widely used signal because of its convenient installation, insensitivity to vehicle speed, and low-latency acquisition [15], [16]. Generally, this kind of signal is abbreviated as ABA (axle box acceleration) [17]. After the vibration transmission, the ABA can be treated as a discrete-time series with

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

a lot of uncertainty [18]. There are some methods have been proposed based on ABA signals. For example, Yue and Chen [19] proposed a method for analyzing wheel vibration signals, which were carried out based on continuous wavelet transform. Skarlatos *et al.* [20] established a fuzzy logic method to classify healthy and damaged train wheels. Gibert *et al.* [21] proposed an intelligent detection method for rail fasteners using a deep convolutional neural network. Wang *et al.* [22] utilized torsional vibration to diagnose wheel flat. Yunguang *et al.* [23] constructed a data-driven method for estimating wheel condition. However, most of these existing methods for wheel flat diagnosis need certain prior knowledge to process the ABA signal, which brings extra labor cost for wheel flat diagnosis. Recently, Krummenacher [24] proposed a method called GAF, which can encode features of time series generated by wheel flat as images. However, this method is not suitable for ABA signals because of the sequence position dependence of time series. Based on the feature encoding ability of GAF and the frequency characteristics of wheel flat vibration, we proposed a novel frequency-domain Gramian angular field (FDGAF) algorithm to encode wheel flat vibration features as image. Also, a transfer learning network is introduced to realize the intelligent diagnosis for different types of railway wheel flat under four different vehicle velocities.

The main contributions of this paper are list below:

(1). To represent wheel falt vibration as images, we proposed a novel FDGAF algorithm. In the frequency domain, this algorithm can convert the ABA signal of wheel flat to a featured image by avoiding time independence of raw vibration. Based on the calculation of the Gramian matrix of frequency amplitude in polar coordinate, the feature information of wheel flat can be preserved, and the angular field of these frequency features can be represented as 2-D images for specific feature learning in a transfer learning network.

(2). To realize the intelligent classification of FDGAF images under small sample conditions, we introduced a modified transfer learning network to complete the task. The FDGAF images are utilized for the training of the fully connected layer, the softmax layer, and the classification layer of this network, while the ImageNet is introducted to pre-train the generic features. With this modified transfer learning network, FDGAF images with sample groups as small as 30 can also be used for intelligent classification.

(3). To verify the effectiveness of the proposed algorithm, we processed 3 types of artificial wheel flat on a scaled test rig and collected their ABA signals under four different velocities. After the signal processing by the proposed FDGAF, we obtained visual diagnostic results with clear boundaries under different velocity levels.

II. FREQUENCY FEATURE ANALYSIS OF WHEEL FLAT

A. WHEEL FLAT FORCE ANALYSIS DURING THE IMPACT

Wheel flat is a localized region of missing material caused by frequent braking and severe thermal stress on the wheel tread. It can cause fluctuations in the wheel-rail contact force and vehicle vibration. Generally, the shape of an actual wheel flat can be summarized as elliptical pits, as shown in Fig. 1.

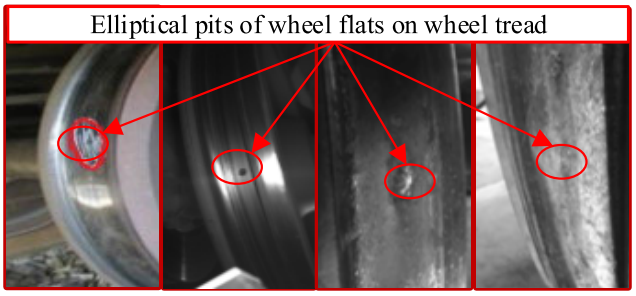


FIGURE 1. Actual shape of typical wheel flat.

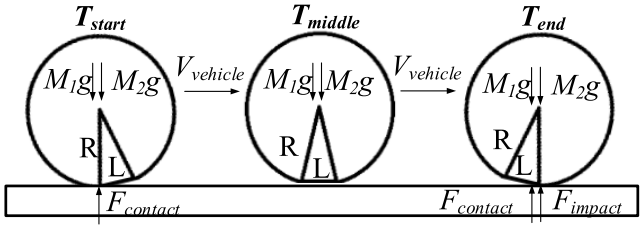


FIGURE 2. Contact model during the impact.

Based on the typical shape of the wheel flat shown in Fig.1, we use a standard wheel-rail contact model to analyze the fluctuation of the contact force. It is believed that the contact force only exists on the rolling circle of the wheelset, so the simplified model shown in Fig.2 can be used to analyze the contact status during the impact of wheel flat.

Based on the contact model in Fig.2, the impact force generated by the wheel flat can be divided into three phases: T_{start}, T_{middle} and T_{end}. In these three different phases, the contact force can be described as follows:

$$F_{wheel} = \begin{cases} M_1g + M_2g - F_{contact} & t \le T_{start} \\ M_1g + M_2g & T_{start} < t \le T_{middle} \\ M_1g + M_2g - F_{contact} - F_{impact} & T_{middle} < t \le T_{end} \\ M_1g + M_2g - F_{contact} & T_{end} < t \end{cases} \tag{1}$$

where M_1 is the sprung mass, M_2 is the unsprung mass, $F_{contact}$ is the supportive force caused by elastic deformation of rail, and F_{impact} is the transient impact force when a wheel flat hit the rail which can reach several times $F_{contact}$ during the impact [24].

It can be seen from Equation (1) that the contact force F_{wheel} changes nonlinearly with the change of contact phases. At the same time, with the rolling of wheelset, its corresponding contact force also shows periodicity related to the velocity $V_{vehicle}$. Therefore, the ABA signal generated by contact force fluctuation also has particular nonlinearity and periodicity related to wheel flat rotation frequency. It is necessary to analyze the ABA signals in time domain and frequency

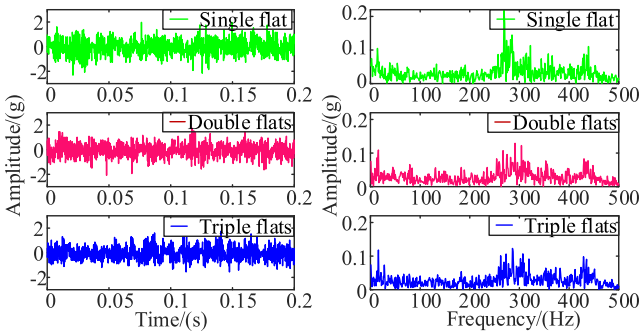


FIGURE 3. The waveform of wheel flat signals in the time domain and frequency domain.

domain to obtain a better understanding of the wheel flat feature.

B. COMPARISON OF FEATURE ANALYSIS IN TIME DOMAIN AND FREQUENCY DOMAIN

The ABA signal of wheel flat in field operation is complex because the existence of track irregularities, bogie system damping, and loading variations. To represent the characteristics of wheel flat vibration as real as possible, we collected ABA signals from a scaled wheel flat test rig and drew their waveforms both in time domain and frequency domain. Their waveforms were shown in Fig. 3. The velocity of wheel is 40 Km/h, the rotating frequency of wheel axle is 20.47Hz, and the acquisition time is 0.2 seconds.

From Fig.3, we can know that it is difficult to tell the difference between these three types of wheel flat in time domain because they have no significant feature difference. However, there are some distinct feature differences between the waveforms of three falt types in frequency domain, such as different concentrated frequency clusters, which appear at 250 to 350 Hz, 400 to 450 Hz, and 5 to 30 Hz. The highest peaks under three fault conditions are 269Hz, 291Hz, and 287Hz, respectively. It is noticed that in the case of double flats and triple flats, there is a severe frequency modulation near the highest frequency peak. Some features like frequency amplitudes, peak positions, and sparseness of frequency families that can represent wheel flat conditions show apparent differences in the frequency domain. Take the amplitude of the highest frequency peak as an example, it reaches 0.2g in the condition of single flat, while that value is 0.12g for double flats and 0.11g for triple flats. This phenomenon indicates that features in frequency domain can represent wheel defect type more vividly than the features in time domain.

III. FREQUENCY-DOMAIN GRAMIAN ANGULAR FIELD AND TRANSFER LEARNING NETWORK

A. TRADITIONAL IMAGE ENCODING FOR TIME SERIES

With the wheel flat features analyzed in time domain and frequency domain, it is crucial to encode the ABA signals as images with features information preserved. The Gramian angular field (GAF) method can encode time series as an image [24]. Given a time series $S = \{x_1, x_2, \cdots, x_N\}$ of N points, we rescale S to $\tilde{S}$ with min-max normalization to make sure that the values of $\tilde{s}$ fall in the interval $[-1\ 1]$. Then in the polar coordinates, the inverse cosine of $\tilde{s}(i)$ can be selected as the angle value, and the timestamp of x_i in $\tilde{S}$ can be treated as the radius value to conduct a polar transformation. Such a process can be described as:

$$\tilde{S}(x,y)\Big|_{x=t_i,y=\tilde{s}(i)} \xrightarrow{\textit{polar transform}} P(r,\varphi)|_{r=t_i/N,\varphi=\arccos(\tilde{s}(i))} \tag{2}$$

where t_i is the time stamp of x_i, and $P(r, \varphi)$ is the representation in polar coordinates. After obtaining $P(r, \varphi)$, the triangular sum between each point can be calculated and the time correlation in different time intervals can be identified. The corresponding GAF can be defined as follows:

$$\begin{aligned} GAF &= \tilde{S}^T \times \tilde{S} - \sqrt{I-\tilde{S}^2}^T \times \sqrt{I-\tilde{S}^2} \\ &= \begin{bmatrix} \cos(\varphi_1+\varphi_1) & \cdots & \cos(\varphi_1+\varphi_n) \\ \cos(\varphi_2+\varphi_1) & \cdots & \cos(\varphi_2+\varphi_n) \\ \cdots & \cdots & \cdots \\ \cos(\varphi_n+\varphi_1) & \cdots & \cos(\varphi_n+\varphi_n) \end{bmatrix} \end{aligned} \tag{3}$$

where I is the unit vector. By defining the inner product $< a, b >= a \cdot b - \sqrt{1-a^2} \cdot \sqrt{1-b^2}$, the GAF can be represented as a Gramian matrix. Then the representative feature map can be obtained when we use the value of this matrix as the pixel in an image.

Traditional GAF has two properties:

(1). For a time series, the GAF matrix produces only one bijective result in the polar coordinates because of the cosine mapping.

(2). Different from encoding mapping in Cartesian coordinates, the value of GAF is not only dependent on time stamp interval, but also influenced by absolute sequence position.

The property (1) means that the corresponding values will warp among different angular positions, so the dynamic change of amplitude in time series can be represented. However, when the GAF is utilized for periodic vibration signal processing, the image encoding results will be interfered with by the absolute starting point of the signal sample because of the property (2). For ABA signal of wheel flat vibration, this means the harmonics of sampling frequency/shaft-rate frequency, and it will bring intervention of prior knowledge and less intelligence in the fault diagnosis process. To solve this problem, we need to find a way to avoid this dependence on prior knowledge.

B. NOVEL FREQUENCY-DOMAIN GRAMIAN ANGULAR FIELD

Based on the analysis of wheel flat feature in frequency domain and description of GAF, we proposed a novel frequency-domain Gramian angular field method (FDGAF). To represent the wheel flat features with a discrete sequence, we transfer the original ABA signal from time domain to frequency domain using discrete Fourier transform.

Given a wheel flat vibration signal $S_{wf} = \{s_1, s_1, \cdots, s_N\}$, where N is the sampling point. Its frequency-domain sequence $S_{wf}(k) = \{s_{wf}(1), s_{wf}(2), \cdots, s_{wf}(k)\}$ can be described as

$$S_{wf}(k) = \sum_{i=0}^{N-1} S_{wf}(n) e^{-j\frac{2\pi}{N}ki} \tag{4}$$

where $k = 0, 1, \cdots, N-1$.

Then we convert the non-dimensional discrete sequence $S_{wf}(k)$ to the angular field. First, it should be rescaled to the interval [-1 1] to satisfy the amplitude principle of a cosine function. The min-max normalization is described as follow:

$$\tilde{s}_{wf}(k) = 2\frac{s_{wf}(k) - \min(S_{wf}(k))}{\max(S_{wf}(k)) - \min(S_{wf}(k))} - 1 \tag{5}$$

This normalized frequency sequence can be encoded in polar coordinates below:

$$\begin{cases} \tilde{r} = \dfrac{k}{N}, & k = 0, 1, \cdots, N-1 \\ \tilde{\varphi} = \arccos(\tilde{s}_{wf}(k)), & \varphi \in [-11] \end{cases} \tag{6}$$

Therefore, the frequency-domain Gramian angular field can be constructed as follow:

$$FDGAF = \begin{bmatrix} \cos(\hat{\varphi}_1 + \hat{\varphi}_1) & \cdots & \cos(\hat{\varphi}_1 + \hat{\varphi}_k) \\ \cos(\hat{\varphi}_2 + \hat{\varphi}_1) & \cdots & \cos(\hat{\varphi}_2 + \hat{\varphi}_k) \\ \vdots & \ddots & \vdots \\ \cos(\hat{\varphi}_k + \hat{\varphi}_1) & \cdots & \cos(\hat{\varphi}_k + \hat{\varphi}_k) \end{bmatrix} \tag{7}$$

There are three advantages to encode wheel flat vibration with FDGAF:

(1). The FDGAF uses frequency amplitude as angle value and frequency position as radius value to construct a matrix field. With this encoding process, the frequency features of wheel flat can be preserved graphically.

(2). The main diagonal of FDGAF contains the original wheel flat frequency information, which can be approximately reconstructed from high-level features learned by transfer learning.

(3). With image encoded in the frequency domain, the series position dependence in GAF can be avoided, and the intervention of fault-related prior knowledge can be eliminated by FDGAF.

To illustrate how the proposed FDGAF works for periodic signals, we constructed a modulated signal as an example:

$$\begin{cases} S1 = 0.7\sin(2\pi * 50t) + \sin(2\pi * 20t), & t \in [0, \frac{2}{3}T] \\ S2 = 0.7\sin(2\pi * 50\tilde{t}) + \sin(2\pi * 20\tilde{t}), & \tilde{t} \in [\frac{1}{3}T, T] \end{cases} \tag{8}$$

where $T = 1$s.

Based on Equation (2) to Equation (7), we performed GAF and FDGAF methods on S_1 and S_2 respectively, as shown in Fig.4. From Fig.4, we can know that these two periodic signals contain similar vibration information because they are segments with different starting points in one same signal

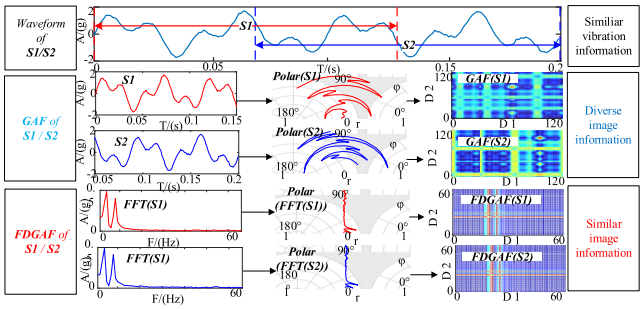


FIGURE 4. Comparison between GAF and FDGAF.

(S_1 start form 0s and S_2 start from $T/3$). However, GAF of S_1 and S_2 represented completely diverse image information, whereas the FDGAF of S_1 and S_2 represented similar featured image information.

C. THE TRANSFER LEARNING NETWORK FOR FDGAF IMAGES

After the image encoding by FDGAF, classification of these images becomes a critical issue in the intelligent wheel flat diagnosis. Deep convolution neural network (DCNN) has become a powerful tool for vibration-based fault diagnosis because of its reliable intelligence and powerful generalization [25]. Typical DCNN consists of an input layer, several hidden layers, and an output layer [26], [27]. The standard form of the input layer is an image. The hidden layer often contains a convolution layer, a pooling layer, an activation layer, and a fully connected layer to extract high dimensional features. The output layer is often used to predict probability.

For an input sample X, mostly a matrix of the input image, the convolution layer can extract some specific features. The m-th feature map of the l-th layer before the activation has the feature value $v^l_{m,}$, this convolution can be described as follow:

$$v^l_m = w^l_m * X + b^l_m \tag{9}$$

where w^l_m means the weight of m-th convolutional kernel of the l-th layer, b^l_m is the bias and $*$ means the convolutional operator.

After the activation function been involved, a large number of feature maps can be obtained through the convolution shown in Equation (9), then the pooling layer is used to reduce parameters of the deep network. The pooling process can be described as:

$$y^{l+1}_m = p(\beta^{l+1} D(z^l_m) + b^{l+1}) \tag{10}$$

where $p(\cdot)$ means the activation function, $D(\cdot)$ is the down-sampling function, and β^{l+1} denotes its corresponding weight.

The DCNN requires a large number of image samples to train the entire classification network at multiple levels. With a large number of samples, the classification results of DCNN can reach a high classification accuracy and strong robustness. However, in the actual diagnosis of wheel flat,

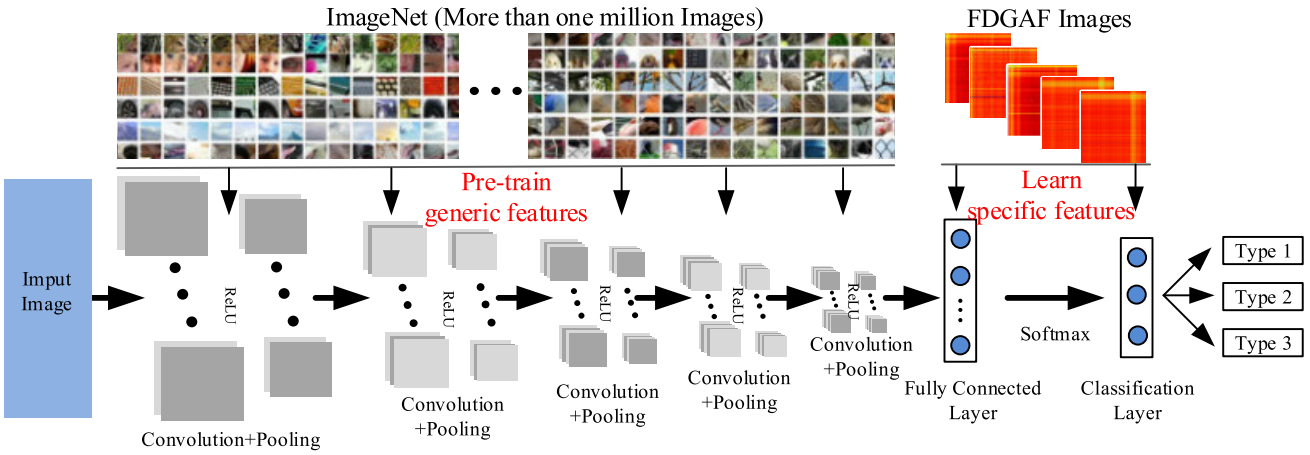


FIGURE 5. Structure of the transfer learning network for FDGAF images.

the number of fault samples is usually small, which can not provide enough dimensions and input samples to support the retraining of the whole DCNN network. Therefore, an intelligent network that can maintain high accuracy and classification robustness under the condition of small samples is needed to realize the intelligent diagnosis based on the FDGAF image. Research shows that in the image classification task, the convolution layer and pooling layers are supposed to learn generic features, like image edge, color, and image texture [28]. The specific feature learning that determines the classification characteristics is mainly trained in the fully connected layers, the softmax layers, and the output layers. For the FDGAF images of wheel flat, we use a transfer learning network to perform intelligent diagnosis. We utilized a typical structure of traditional CNN to construct our transfer learning network [29], [30]. The generic features of the top five convolutional and pooling layers were pre-trained by ImageNet [31], and the specific features related to wheel flat were learned by FDGAF samples. And then, we use a fully connected layer to get an intelligent classification result. With such a fine-tuning of transfer learning network parameters, we can realize the transfer learning from generic ImageNet to specific FDGAF images. Through such a transfer process, the generic features of the convolutional layers and the pooling layers that are the majority of the classification network can be maintained, and the FDGAF samples can only be used to retrain the last three layers, which account for a relatively small proportion of the network. Therefore, the proposed method can efficiently realize intelligent fault diagnosis under small samples while the high classification accuracy remained. The structure of the transfer learning network used for FDGAF images is shown in Fig.5.

D. INTELLIGENT DIAGNOSIS FOR WHEEL FLAT BASED ON FDGAF AND TRANSFER LEARNING NETWORK

With the FDGAF used to encode wheel flat features into images and the transfer learning network introduced to classify images intelligently, we constructed an intelligent method for wheel flat diagnosis. For each image encoded by FDGAF, the transfer learning network can obtain its corresponding attribution probability for the wheel flat type. The classification result with the most considerable attribution probability is taken as the attribution of the input image, which is

$$Case_{FDGAF(i)} = \max[p(FDGAF(i)) \in (Case_1, \cdots, Case_M)] \tag{11}$$

where M is the number of wheel flat types, $FDGAF(i)$ denotes the i-th FDGAF encoding image.

To further illustrate the classification result of the proposed method, we use t-distributed stochastic neighbor embedding (t-SNE) to visualize the classification [32]. The flow chart of the intelligent diagnosis method based on FDGAF and transfer learning network is described in Fig. 6.

IV. TEST VERIFICATION

To verify the proposed method in this paper, we used a scaled test rig for simulation of railway wheel flat. This test rig is mainly composed of one suspension system, one drive motor, four axle boxes, two wheelsets, and one circular rail. It is manufactured at a scale ratio of 1:5 relative to a field railway vehicle. The radius of the wheel's rolling circle is 86.42 mm, and the velocity could be adjusted from 0 km/h to 120 km/h.

A. DESCRIPTION OF TEST

Based on the actual wheel flat length shown in Fig. 1, we processed artificial wheel flats with sizes of 10 mm × 8 mm on the test rig in equal proportions. As shown in Fig. 7, we machined three types of wheel flats with a high-speed grinder. Type 1 is a single wheel flat, Type 2 was double wheel flats with the flats separated by 90°, and Type 3 was triple wheel flats with the flats separated by 45°. To verify the applicability of this intelligent method at different velocities, we set the vehicle velocities to 20Km/h, 40Km/h, 60Km/h, and 80Km/h. The specific conditions of the test were shown in TABLE 1.

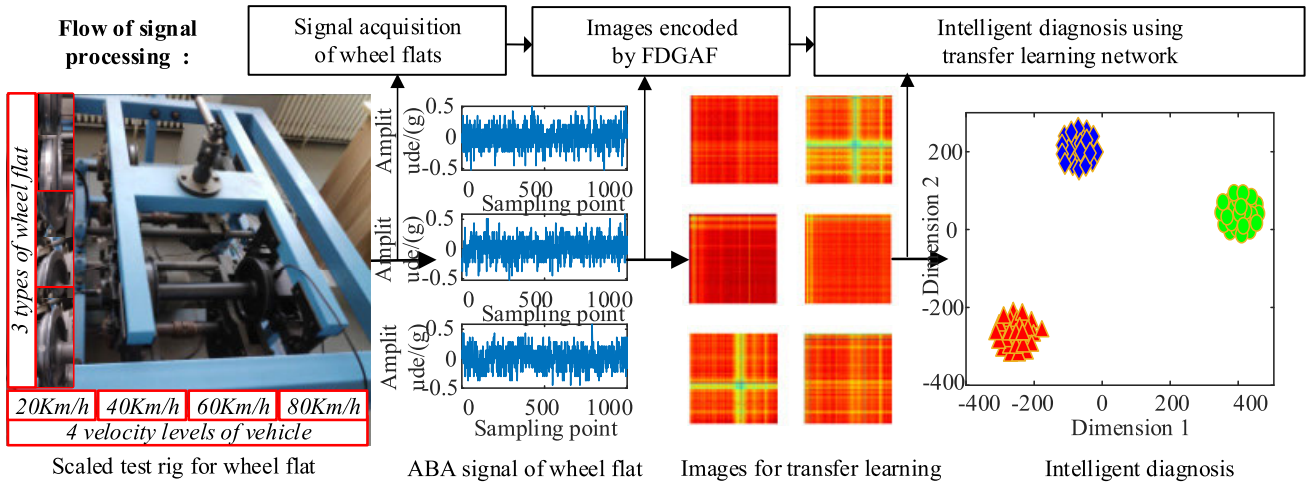


FIGURE 6. Flow of wheel flat intelligent diagnosis based on FDGAF.

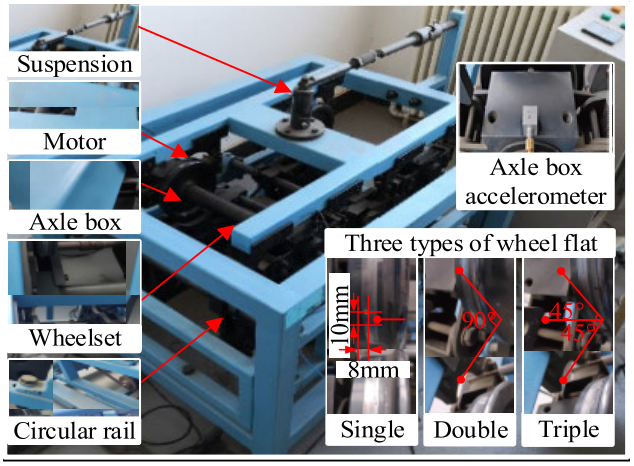


FIGURE 7. Scaled test rig of wheel flat.

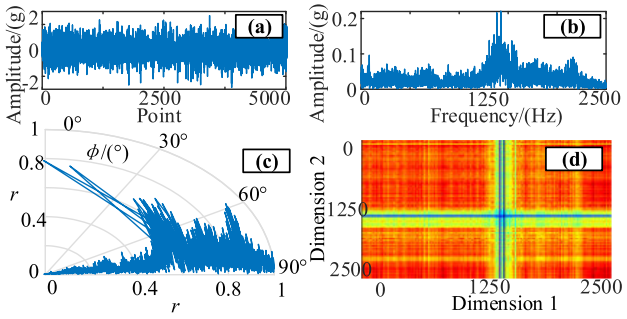


FIGURE 8. (a) Time-domain waveform. (b) Frequency-domain waveform. (c) Transform in polar coordinates. (d) Encoded image of FDGAF.

TABLE 1. Detailed condition of the test.

Type of wheel flat	Vehicle velocity			
	20 Km/h	*40 Km/h*	*60 Km/h*	*80 Km/h*
Single flat	$Case_1$	$Case_4$	$Case_7$	$Case_{10}$
Double flats	$Case_2$	$Case_5$	$Case_8$	$Case_{11}$
Triple flats	$Case_3$	$Case_6$	$Case_9$	$Case_{12}$

B. DESCRIPTION OF TEST

To collect the ABA signals, we mounted an accelerometer on the axle box vertically. The sampling frequency was set to 20KHz to obtain reliable vibration details. The test rig is shown in Fig. 7.

C. VERIFICATION UNDER DIFFERENT CASES

The verification can be divided into three steps:

(1). Signal acquisition and segmentation

The ABA signals of three different wheel flat were collected and segmented. Each signal is divided into 30 segments.

(2). The FDGAF encoding of segmented samples

Taking one sample in $Case_1$ as an example, the image encoding process using FDGAF is shown in Fig. 8.

The time-domain waveform of the first sample in $Case_1$ is shown in Fig. 8(a), and its frequency-domain waveform is drawn in Fig. 8(b). To obtain the wheel flat vibration features in frequency-domain, we convert the waveform from Cartesian coordinates to polar coordinates, as shown in Fig. 8(c). With the polar representation, the encoded image of FDGAF shown in Fig. 8(d) can be obtained.

(3). Perform intelligent diagnosis of wheel flat

After obtaining the FDGAF images for all the cases, the transfer learning network can be applied to diagnose the wheel flat intelligently. All FDGAF image samples are divided into four groups according to their velocity level, and each group contains three different types of wheel flats. To fully utilize the intelligent training and verification ability of the transfer learning network, 70% of the FDGAF samples were used for training and 30% for testing. We froze the top five convolutional layers and pooling layers to preserve the generic features of the network pre-trained by ImageNet, and

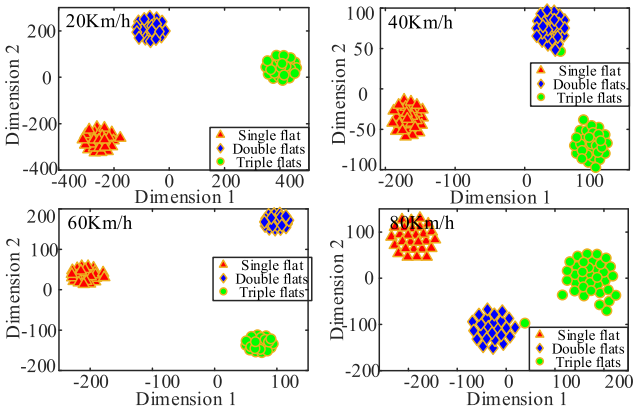


FIGURE 9. Intelligent diagnosis of wheel flats based on FDGAF.

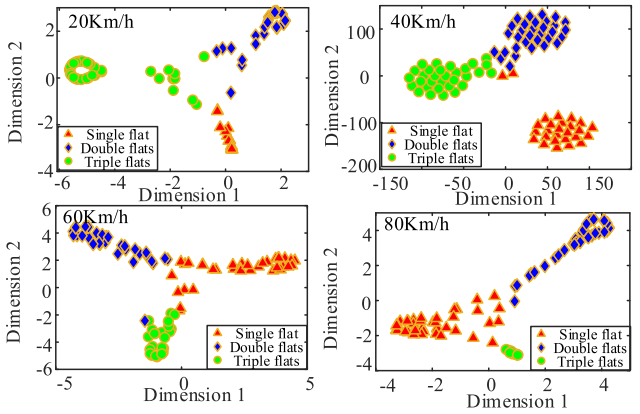


FIGURE 10. Intelligent diagnosis of wheel flats based on GAF.

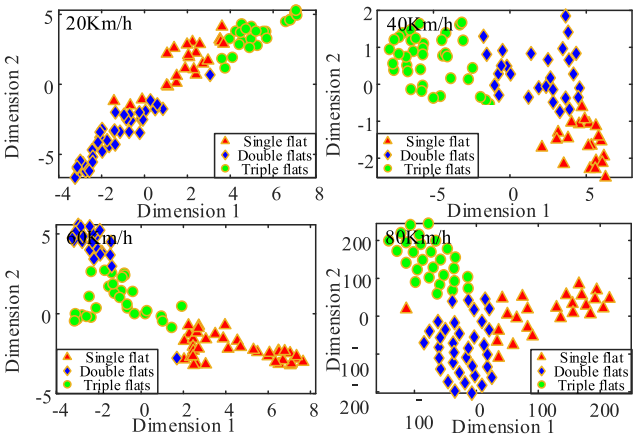


FIGURE 11. Intelligent diagnosis of wheel flats based on STFT.

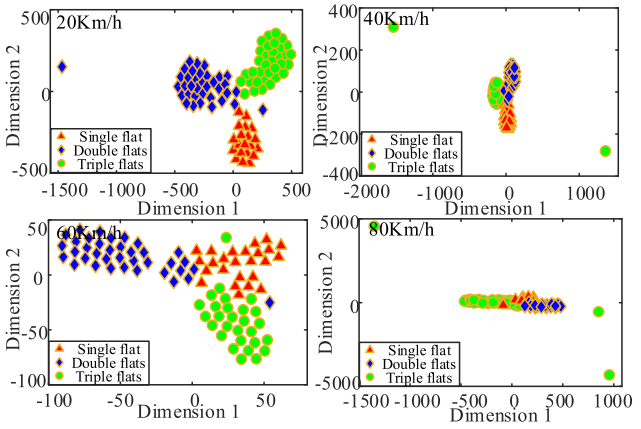


FIGURE 12. Intelligent diagnosis of wheel flats based on WVD.

train the last three fully connected layers and softmax layers with FDGAF images to learn specific features that related to wheel flat characteristics. Then an output layer is constructed to obtain the probability of each sample belonging to the wheel flat type.

To illustrate the superiority of the proposed FDGAF, we conducted the intelligent classification process based on FDGAF, and obtained its visual diagnosis result. Its diagnosis results are shown in Fig. 9.

As shown in Fig. 9, for FDGAF-based diagnosis, the classification results at four different velocities have distinct classification boundaries. Moreover, the cluster for single flat, double flats, and triple flats are centralized within the classes and scattered among the classes, which means that the diagnosis based on FDGAF can realize outstanding classification.

In contrast, we also conducted the intelligent classification based on traditional GAF, typical time-frequency representation method STFT, and typical time-frequency representation method WVD. In order to keep the consistency of the method in the aspect of no prior knowledge, we choose the classical parameter setting of STFT and WVD, which may lead to some small differences in the results. The results of these three typical method are shown in Fig.10 to Fig.12.

As can be seen from Fig.10 to Fig.12, all classification results at four velocities are not separable, there are a many points existed in wrong classification, and their classification boundary are instinct. The poor performancd of thse three methods were probably caused by the time dependence of GAF, the time-frequency resolution conflict of STFT, and the serious cross term of WVD.

For a more accurate quantitative evaluation of the classification result, we calculated the Euclidean class distance between FDGAF clusters, GAF clusters, STFT clusters, and WVD clusters under four velocities. The calculation results are shown in Table 2.

Comparison of the numbers in Table 2 indicates that there are apparent disparities between intra-class distance and inter-class distance for FDGAF under all four velocities, the former ones are generally 5 to 10 times higher than the latter ones, which means that FDGAF can classify the samples with same attributes correctly and distinguish the samples with different attributes effectively. In contrast, the multiple numbers between intra-class distance and inter-class distance for GAF, STFT, and WVD vary from 1.8 to 5.8, which are relatively small compared with the numbers for FDGAF. This result shows that FDGAF can obtain classification clusters with

TABLE 2. Class distance between clusters.

Cluster type		Intra-class diatance	Inter-class distance
20Km/h	*FDGAF*	582.9	53.9
	GAF	5.4	1.1
	STFT	6.7	20.6
	WVD	583.9	1686.5
40Km/h	*FDGAF*	216.2	24.7
	GAF	173.8	45.3
	STFT	6.4	20.9
	WVD	367.6	758.6
60Km/h	*FDGAF*	321.7	21.5
	GAF	6.2	1.7
	STFT	6.4	19.9
	WVD	98.2	243.6
80Km/h	*FDGAF*	285.5	48.8
	GAF	5.6	1.1
	STFT	259.6	705.7
	WVD	952.03	1685.0

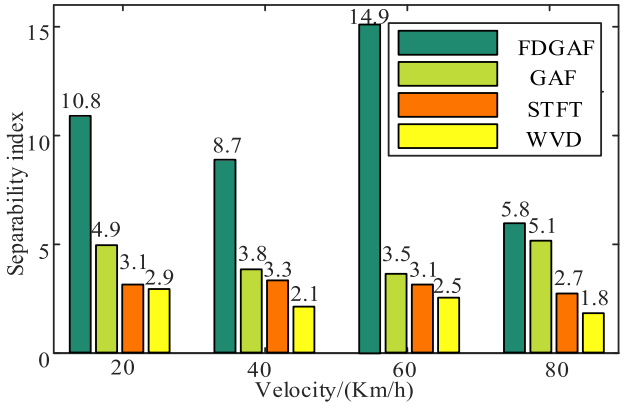


FIGURE 13. Comparison of separability between FDGAF and GAF.

clear boundaries, while other image representation method lack this kind of classification ability.

To further illustrate the superiority of proposed FDGAF, we define the separability index as inter-class distance divided by intra-class distance and draw the separability indexes of FDGAF, GAF, STFT, and WVD at four velocities in Fig.13.

It can be seen from Fig.13 that the diagnostic separability indexes of FDGAF are all higher than the separability indexes of other methods. In the case of 20Km/h, 40Km/h, and 60Km/h, the indexes of FDGAF are more than two times higher than the other indexes, which can fully demonstrate the superiority of the intelligent diagnosis method using FDGAF.

V. CONCLUSION

In this paper, we presented an intelligent wheel flat diagnosis method using FDGAF and transfer learning network. Moreover, we proved its superiority with vibration signals from a scaled test rig for railway wheel flat. Through the comparison of wheel flat vibration in time-domain and frequency-domain, we concluded that the wheel flat features are mainly represented in frequency domain. By mapping the vibration signals into frequency domain, we proposed the FDGAF method, which can construct a Gramian field in polar coordinates for the frequency-domain signals, thus discarding the wheel flat-related prior knowledge in the process of vibration signal image encoding. Then, a transfer learning network is used to realize the intelligent diagnosis of FDGAF images. In this network, the generic features of the top five layers are pre-trained by the ImageNet, and FDGAF learns the specific features of the last three layers. We verified the effectiveness of the proposed method for the wheel flat signals obtained on a scaled test rig under twelve cases and obtained four intelligent diagnosis results with distinct classifications. We calculated the classification separability indexes based on Euclidean class distance, and the results indicated that the separate indexes of FDGAF are much higher than GAF, STFT, and WVD.

This method is developed for different types of wheel flats under different constant velocities. In practical wheel flat diagnosis, the condition of the rail, including track irregularities and track joints, should be considered as a factor that may affect the classification performance. Future work will include an analysis of the influence of track irregularities and track joint impacts, an examination of the stabilization of wheel flat vibration signals under variable velocities, and field experiments for further verification of the method.

REFERENCES

[1] A. M. Remennikov and S. Kaewunruen, "A review of loading conditions for railway track structures due to train and track vertical interaction," *Struct. Control Health Monitor.*, vol. 15, no. 2, pp. 207–234, Mar. 2008.

[2] M. J. M. M. Steenbergen, "The role of the contact geometry in wheel–rail impact due to wheel flats: Part II," *Vehicle Syst. Dyn.*, vol. 46, no. 8, pp. 713–737, Aug. 2008.

[3] A. Zhu, S. Yang, Q. Li, J. Yang, C. Fu, J. Zhang, and D. Yao, "Research on prediction of metro wheel wear based on integrated data-model-driven approach," *IEEE Access*, vol. 7, pp. 178153–178166, 2019.

[4] Y. Li, M. J. Zuo, J. Lin, and J. Liu, "Fault detection method for railway wheel flat using an adaptive multiscale morphological filter," *Mech. Syst. Signal Process.*, vol. 84, pp. 642–658, Feb. 2017.

[5] S. Jia and M. Dhanasekar, "Detection of rail wheel flats using wavelet approaches," *Struct. Health Monit., Int. J.*, vol. 6, no. 2, pp. 121–131, Jun. 2007.

[6] N. A. Thakkar, J. A. Steel, and R. L. Reuben, "Rail–wheel contact stress assessment using acoustic emission: A laboratory study of the effects of wheel flats," *Proc. Inst. Mech. Eng. F, J. Rail Rapid Transit.*, vol. 226, no. 1, pp. 3–13, Jan. 2012.

[7] L. J. Feng, W. Xiaohao, and C. Kang, "Safety region estimation and fault diagnosis of wheels based on least squares support vector machine and probabilistic neural networks," *Mach. Building Autom.*, vol. 46, no. 1, pp. 141–145, 2017.

[8] M. Aktas, E. H. Gunel, P. Yilmazer, and T. Akgun, "Detection of wheel flatten defect on the moving train with acoustic emission sensor," in *Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2018, pp. 386–390.

[9] Y. Ye, D. Shi, P. Krause, Q. Tian, and M. Hecht, "Wheel flat can cause or exacerbate wheel polygonization," *Vehicle Syst. Dyn.*, early access, pp. 1–30, Jun. 2019, doi: 10.1080/00423114.2019.1636098.

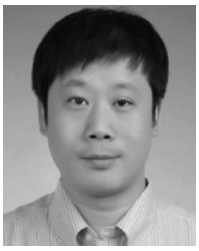
[10] J. Yang, Y. Zhao, J. Wang, Y. Bai, and C. Liu, "Investigation on impact response feature of railway vehicles with wheel flat fault under variable speed conditions," *J. Vibrat. Acoust.*, vol. 142, no. 3, pp. 1–22, Jan. 2020.

[11] N. Bosso, A. Gugliotta, and N. Zampieri, "Wheel flat detection algorithm for onboard diagnostic," *Measurement*, vol. 123, pp. 193–202, Jul. 2018.

[12] R. S. Dwyer-Joyce, C. Yao, R. Lewis, and H. Brunskill, "An ultrasonic sensor for monitoring wheel flange/rail gauge corner contact," *Proc. Inst. Mech. Eng. F, J. Rail Rapid Transit*, vol. 227, no. 2, pp. 188–195, Mar. 2013.
[13] G. Kouroussis, C. Caucheteur, D. Kinet, G. Alexandrou, O. Verlinden, and V. Moeyaert, "Review of trackside monitoring solutions: From strain gages to optical fibre sensors," *Sensors*, vol. 15, no. 8, pp. 20115–20139, Aug. 2015.
[14] A. Nunez, A. Jamshidi, and H. Wang, "Pareto-based maintenance decisions for regional railways with uncertain weld conditions using the Hilbert spectrum of axle box acceleration," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1496–1507, Mar. 2019.
[15] J. Yang, C. Liu, J. Wang, R. Sun, and H. Zhu, "Research on dynamic of urban rail vehicle of variable speed and long-wave excitation conditions," *J. Mech. Eng.*, vol. 55, no. 4, p. 118, Jan. 2019.
[16] M. Molodova, Z. Li, and R. Dollevoet, "Axle box acceleration: Measurement and simulation for detection of short track defects," *Wear*, vol. 271, nos. 1–2, pp. 349–356, May 2011.
[17] J. Yang, Y. Bai, J. Wang, and Y. Zhao, "Tri-axial vibration information fusion model and its application to gear fault diagnosis in variable working conditions," *Meas. Sci. Technol.*, vol. 30, no. 9, Sep. 2019, Art. no. 095009.
[18] J. Xiong, X.-H. Chang, and X. Yi, "Design of robust nonfragile fault detection filter for uncertain dynamic systems with quantization," *Appl. Math. Comput.*, vol. 338, pp. 774–788, Dec. 2018.
[19] Y. Jianhai, Q. Zhengding, and C. Boshi, "Application of wavelet transform to defect detection of wheelflats of railway wheels," in *Proc. 6th Int. Conf. Signal Process.*, vol. 1, 2002, pp. 29–32.
[20] D. Skarlatos, K. Karakasis, and A. Trochidis, "Railway wheel fault diagnosis using a fuzzy-logic method," *Appl. Acoust.*, vol. 65, no. 10, pp. 951–966, Oct. 2004.
[21] X. Gibert, V. M. Patel, and R. Chellappa, "Deep multitask learning for railway track inspection," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 153–164, Jan. 2017.
[22] Z. Wang, P. Allen, J. Wang, G. Huang, S. Xu, and W. Zhang, "Torsional vibration as a method of diagnostic tool for wheel flatness," in *Proc. Adv. Dyn. Vehicles Roads Tracks*. Cham, Switzerland: Springer, 2020, pp. 821–826, doi: 10.1007/978-3-030-38077-9_95.
[23] Y. Ye, D. Shi, P. Krause, and M. Hecht, "A data-driven method for estimating wheel flat length," *Vehicle Syst. Dyn.*, early access, pp. 1–19, May 2019, doi: 10.1080/00423114.2019.1620956.
[24] G. Krummenacher, C. S. Ong, S. Koller, S. Kobayashi, and J. M. Buhmann, "Wheel defect detection with machine learning," *IEEE Trans. Intell. Transport. Syst.*, vol. 19, no. 4, pp. 1176–1187, Apr. 2018.
[25] Z. Shi, H. Hao, M. Zhao, Y. Feng, L. He, Y. Wang, and K. Suzuki, "A deep CNN based transfer learning method for false positive reduction," *Multimedia Tools Appl.*, vol. 78, no. 1, pp. 1017–1033, Jan. 2019.
[26] M. Ha, Y. Byun, J. Kim, J. Lee, Y. Lee, and S. Lee, "Selective deep convolutional neural network for low cost distorted image classification," *IEEE Access*, vol. 7, pp. 133030–133042, 2019.
[27] S. Xia, Y. Xia, H. Yu, Q. Liu, Y. Luo, G. Wang, and Z. Chen, "Transferring ensemble representations using deep convolutional neural networks for small-scale image classification," *IEEE Access*, vol. 7, pp. 168175–168186, 2019.
[28] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" presented at the 27th Int. Conf. Neural Inf. Process. Syst., Montreal, QC, Canada, vol. 2, 2014.
[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
[30] P. Ma, H. Zhang, W. Fan, C. Wang, G. Wen, and X. Zhang, "A novel bearing fault diagnosis method based on 2D image representation and transfer learning-convolutional neural network," *Meas. Sci. Technol.*, vol. 30, no. 5, May 2019, Art. no. 055402.
[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and F. F. Li, "ImageNet: A Large-Scale Hierarchical Image Database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
[32] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.

YONGLIANG BAI received the M.Eng. degree in control engineering from the School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing, in 2017. He is currently pursuing the Ph.D. degree in vehicle operation engineering with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, China.

His research interests include signal processing and fault diagnosis.



JIANWEI YANG received the Ph.D. degree in vehicle operation engineering from the China Academy of Railway Sciences, Beijing, China, in 2006.

He is currently a part-time Professor with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing. He is also a Full Professor with the School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing. His research interests include fault diagnosis and vehicle system dynamics and control.

JINHAI WANG (Member, IEEE) received the B.Eng. and M.Eng. degrees in mechanical engineering and vehicle operation engineering from the Beijing University of Civil Engineering and Architecture, Beijing, China. He is currently pursuing the Ph.D. degree in vehicle operation engineering with Beijing Jiaotong University.

In summer 2011, he was a Visiting Student with the Gyeonggi College of Science and Technology, Sinhyeon, South Korea. His research interests include mechanical dynamics, system reliability engineering, and machinery fault diagnosis.

QIANG LI received the Ph.D. degree in mechanical engineering from Northern Jiaotong University, Beijing, in 1995.

He is currently a Full Professor with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing. His research interests include fatigue strength and reliability of structures.