

## Research Article

# Fault Diagnosis Method for Bearing of High-Speed Train Based on Multitask Deep Learning

Jia Gu <sup>1,2</sup> and Ming Huang <sup>3</sup>

<sup>1</sup>Mechanical Engineering Institute, Dalian Jiaotong University, Dalian, China

<sup>2</sup>IT Department, CRRC Qingdao Sifang Co., Ltd., Qingdao, China

<sup>3</sup>Software Technology Institute, Dalian Jiaotong University, Dalian, China

Correspondence should be addressed to Jia Gu; [gujiajiafirst@163.com](mailto:gujiajiafirst@163.com)

Received 20 August 2020; Revised 13 November 2020; Accepted 24 November 2020; Published 14 December 2020

Academic Editor: Dejan Gjorgjevikj

Copyright © 2020 Jia Gu and Ming Huang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

High-speed trains often pass through tunnel, turnout, ramp, bridge, and other line features in the process of running. At the same time, the length of the operation time, weather conditions, changes in train running conditions, and other conditions will lead to the loss of the train. In view of the complexity of a high-speed train structure and operation environment, in order to effectively evaluate the health of the train in the operation process, this paper proposes a diagnosis method of bearing temperature anomaly of a high-speed train based on condition identification and multitask deep learning. In this paper, the important components of bogie axle box, gearbox, and traction motor are taken as the research object. Firstly, the operating condition parameters of the high-speed train are analyzed and identified, and the *K*-means algorithm is used to classify and identify the operating condition of the high-speed train. Then, based on the operating condition identification and multitask deep learning, the bearing temperature prediction model is constructed. In addition, according to statistical quality control theory, the difference between the value predicted by the model and the real value is used to diagnose the anomaly of the bearing temperature of the high-speed train. Finally, the accuracy and availability of the model are verified by an example. The model can judge whether the running train bearing temperature is in the normal range in real time and predict and alarm the abnormal bearing temperature.

## 1. Introduction

A high-speed train is a complex integrated system composed of tens of thousands of parts, with a complex structure and diverse working conditions. The high-speed EMU has a wide operating range, including the hot south, cold northeast, and strong sand northwest. The equipment of the EMU needs to withstand the test of various harsh environments such as high and low temperatures, wind, sand, and humidity. In addition, the conditions of the EMU operation line, such as line slope, line curve, turnout, and tunnel, have a very important influence on the operating state of the EMU. The high-speed EMU has a high frequency of use and complex operating environment. In order to effectively ensure the safe operation of EMU, real-time diagnosis of train failures

and assessment of the health status of EMU should be conducted.

The axle box, gear box, and motor stator are important components of the bogie, and it is very important to carry out in-depth research on it [1]. The temperature of the bearing on the train will become higher and higher with the loss of use under the same external conditions and gradually approach the limit temperature, so the bearing temperature under the same external conditions can also be used as one of the indicators to measure the performance of the bearing [2–5]. At present, scholars use complex evaluation methods based on mechanism, data model, failure mode, and impact and hazard analyses to diagnose and evaluate complex equipment failures and promote the development of equipment health assessment in the fields of aerospace, ship,

and power. However, due to the problems of high operating speed, large span of operating intervals, complicated operating environment, and long continuous running time, subjective factors have a great influence on the mechanism research and model construction, which makes it difficult to diagnose EMU failures. At present, the majority of EMU bogies judge the bearing health status by collecting the bearing temperature and comparing it with the set threshold value. However, due to the influence of wind force, ambient temperature, and track irregularity, the running condition of EMU bogie bearing is complex and changeable. The traditional health condition monitoring method based on preset threshold value cannot meet the requirements of condition monitoring. In addition, the former condition monitoring usually divides the operating conditions of EMU according to a running parameter (such as ambient temperature) and then defines the temperature monitoring threshold according to different operating conditions. If the operation condition of EMU is divided only according to a single parameter such as temperature, it cannot effectively describe the complex and changeable operation condition of EMU.

Machine learning-based methods, such as support vector machine (SVM) and BP neural network (BPNN) have become one of the hotspots in current research on fault diagnosis [6, 7]. The fault diagnosis effect of the method mainly depends on the quality of the extracted fault features. In the context of big data applications, its diagnostic and generalization capabilities are clearly insufficient. In recent years, as a new method in the field of machine learning, deep learning has powerful modeling and feature extraction capabilities and is widely used in image processing, speech recognition, and fault diagnosis. Deep learning can extract high-level features of faults through layer-by-layer feature extraction, which has unique advantages over traditional machine learning algorithms [8–11]. However, in practical applications, single-task deep learning algorithms often encounter insufficient training samples, which results in insufficient model training due to insufficient samples, resulting in underfitting problems. Moreover, the related information between tasks is not fully mined, which will affect the generalization ability of the model [12]. In view of the above problems, this paper proposes a high-speed train bearing fault diagnosis method based on working condition identification and multitask deep learning. Considering the variability of EMU operation environment and the complexity of state information, the uncertainty of the threshold setting of state characteristic parameters is avoided. Fault diagnosis is carried out under corresponding operation conditions, which improves the accuracy of fault diagnosis and adaptability to variable operating conditions.

## 2. Identification of Operating Conditions of High-Speed Trains

*2.1. Analysis of the Influence of Operating Parameters on Train Bearing Temperature.* High-speed trains are in various complex environments during long-term use and bear various stresses, such as different climates and line conditions, and the use status of trains will change according to

different conditions. The interaction between various parts of the train, such as time, mileage, route, and weather will cause the loss of the train. The reasonable selection of characteristic parameters of operation conditions is the premise of realizing the identification of operation conditions. The selected parameters should be those which have a direct or indirect influence on the characteristic parameters of the train operation state. In this paper, the Pearson correlation coefficient is used to calculate the correlation coefficient between axle box temperature and train operation data, acceleration, mileage, fresh air temperature, line characteristics, and axle box temperature for a certain type of EMU. Through correlation analysis, the key factors affecting axle box temperature are obtained.

From the correlation coefficient of gearbox temperature with train operation and line parameters in Figure 1, it can be seen that the fresh air temperature has the highest correlation with the gearbox, which has a very high influence weight. The higher the fresh air temperature is, the higher the gearbox temperature is. Speed and mileage have certain positive influence weight with gearbox temperature. The higher the speed and mileage, the higher the gearbox temperature.

From Figure 2, it can be concluded that the radius of the curve, bridge, tunnel, and slope also have certain influence on gearbox temperature.

In order to deeply analyze the influence of different line conditions and operating conditions on the axle temperature of the high-speed train, the data from January 18 to July 19 on Wuhan-Guangzhou railway line are selected, and the temperature of three bearing parts, i.e., the wheel side temperature of 1-axis large gearbox, the temperature of the 1-way 1-axle box, and the temperature of 1-axis motor stator are selected, and the average temperature value of each component is calculated. Figure 3 shows that the temperature of bearing in slope condition is higher than that in nonslope condition, which indicates that slope affects bearing temperature.

*2.2. Definition and Selection of Operating Condition Characteristic Parameters.* Due to the great changes of environmental temperature and line characteristics in actual operation, it is necessary to divide the operation conditions based on the condition identification before establishing the health status assessment model, so as to improve the accuracy of the evaluation model.

Based on the analysis in Section 2.1, this paper firstly divides the vehicle operation conditions into two categories. The speed grade and vehicle operation state are defined as the internal condition parameters affecting the train operation, and the line characteristic data and weather environment data are defined as the external condition parameters. Finally, the characteristic parameters of working conditions affecting train operation are selected as follows:

- (1) Running state: traction, braking, idling, and static
- (2) Speed classification: high-speed (speed greater than 200 km/h) and low-speed (speed less than 200 km/h)

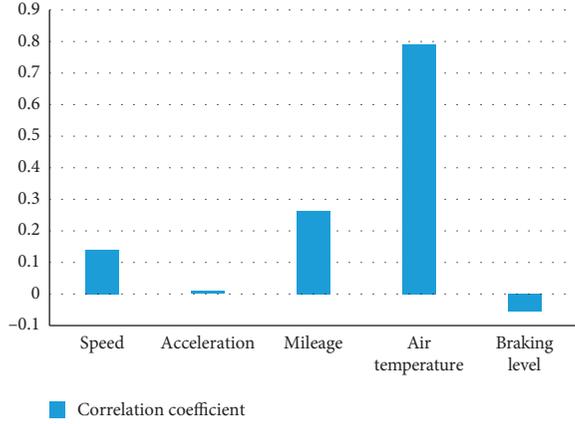


FIGURE 1: Correlation coefficient between gearbox temperature and train running state parameters.

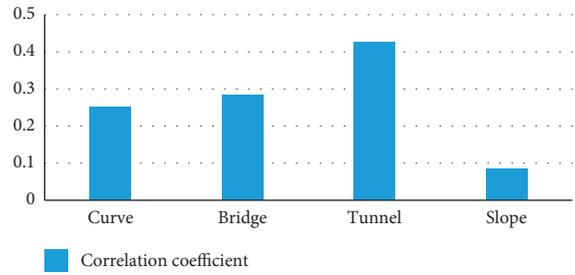


FIGURE 2: Correlation coefficient between gearbox temperature and line characteristics.

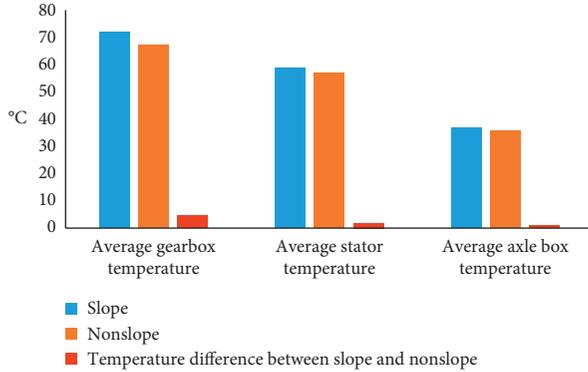


FIGURE 3: Comparison of bearing temperature on ramp.

- (3) Line features: bridges, tunnels, turns, and ramps
- (4) Weather environment: ambient temperature (fresh air temperature)

**2.3. Identification of High-Speed Train Operating Conditions Based on K-Means Algorithm.** The selected characteristic parameters of operating conditions are used to establish a feature set of operating conditions  $u$  (including speed, operation state, turning radius, whether it is on the bridge,

whether it is in the tunnel, slope gradient, and ambient temperature).

It is assumed that the operating condition space  $\Omega$  can be clustered into  $k$  operating condition subspaces using a certain clustering or space partitioning algorithm  $f(u)$ .

In this paper, the  $K$ -means clustering algorithm is used to identify the operating conditions of historical data and automatically classifies the categories by calculating the sample similarity, that is, the higher sample similarity is classified into one category. The  $K$ -means algorithm determines the method of class data  $k$  that minimizes the objective function  $E$  based on the randomly selected number  $k$  of clusters and an initial center point  $u_i$  ( $i = 1, 2, \dots, k$ ). Usually, the objective function  $E$  is the square error function as follows:

$$E = \sum_{i=1}^k \sum_{j=1}^{n_i} |x_j - u_i|^2, \quad (1)$$

where  $x_j$  is the sample data,  $u_i$  is the mean vector of the  $i$ -th cluster, and  $n_i$  is the number of samples in the  $i$ -th cluster [13].

This paper uses the Calinski–Harabaz index to evaluate the effect of clustering. The square sum of the distance between each sample in the class and the center point of the class is defined as the sample tightness, and the square sum of the distance between the class center point and the center point of the overall sample is defined as the separation degree of the sample. The CH score is the ratio of separation and tightness. The larger the CH score, the closer the data within the class, that is, the better the clustering effect:

$$s(k) = \frac{\text{tr}(\mathbf{B}_k)}{\text{tr}(\mathbf{W}_k)} \frac{n-k}{k-1}, \quad (2)$$

where  $s(k)$  is the CH score,  $n$  is the number of samples in the training set,  $k$  is the number of categories,  $\mathbf{B}_k$  is the covariance matrix between categories,  $\mathbf{W}_k$  is the covariance matrix of data within categories, and  $\text{tr}$  is the trace of the matrix. In other words, the smaller the covariance of the data within the category, the greater the covariance between the categories and the higher the Calinski–Harabasz score.

### 3. High-Speed Train Bearing Fault Diagnosis Method Based on Working Condition Identification and Multitask Deep Learning

**3.1. Multitask Deep Learning.** Multitask learning (MTL) is to set multiple training targets for training samples and to improve the generalization performance of model training through joint training and based on specific relevant information between targets [14]. Compared with single-task learning (STL) in traditional machine learning, MTL improves the prediction and generalization ability of main tasks by taking advantage of the interaction between multiple tasks. At present, MTL is mainly used in image recognition, text recognition, and other fields [15].

Compared with single-task deep neural network (ST-DNN), multitask deep neural network (MT-DNN) uses

multiple tasks to share the same architecture, for example, by sharing the same hidden layer, it can effectively extract common information between tasks and overcome insufficient model training due to insufficient training samples [16].

The multitask joint training process is as follows:  $T$  is the total number of tasks,  $(\mathbf{x}_i^t, \mathbf{y}_i^t)$  is the training sample data of the  $t$ -th task, where  $t \in (1, 2, \dots, T)$ ,  $i \in (1, 2, \dots, N)$ ,  $N$  is the total number of training samples,  $\mathbf{x}_i^t \in R$  is the feature vector, and  $\mathbf{y}_i^t \in R$  is the label of the  $t$ -th sample. The multitask objective function can be expressed as follows:

$$\arg \min_{\{\mathbf{w}^t\}_{t=1}^T} \sum_{t=2}^T \sum_{i=1}^N L(\mathbf{y}_i^t, f(\mathbf{x}_i^t; \mathbf{w}^t)) + \lambda \phi(\mathbf{w}^t), \quad (3)$$

where  $f(\mathbf{x}_i^t; \mathbf{w}^t)$  is the mapping function with the input feature vector  $\mathbf{x}_i^t$  and the weight parameter  $\mathbf{w}^t$ ,  $L(\cdot)$  is the loss function shown in equation (1),  $\phi(\mathbf{w}^t)$  is the regularization value of the weight parameter, and  $\lambda$  is the regularization coefficient factor [17].

**3.2. Construction of Bearing Temperature Prediction Model for High-Speed Trains.** MTL improves the ability of machine learning by training multiple related tasks at the same time. The bearing temperature prediction model (multitask learning-long short-term memory, MTL-LSTM) constructed in this paper is based on working condition identification and MTL as shown in Figure 4.

The steps for model construction are as follows:

- (1) An MTL-LSTM is constructed with 2 tasks and shared LSTM hidden layer. The network architecture is [MTL-LSTMnet] = feedforward  $(\mathbf{X}^{\text{main}}, \mathbf{X}^{\text{sub}}; H_{LSTM}; \mathbf{Y}^{\text{main}}, \mathbf{Y}^{\text{sub}}; \mathbf{W}, \mathbf{b})$   $\mathbf{X}^{\text{main}} = \{x_{1, \text{main}}, x_{2, \text{main}}, \dots, x_{N, \text{main}}\}$   $\mathbf{X}^{\text{sub}} = \{x_{i1, \text{sub}}, x_{i2, \text{sub}}, \dots, x_{id, \text{sub}}\}$ , which is the input sample of the high-speed train bearing temperature features;  $\mathbf{X}^{\text{sub}} = \{x_{1, \text{sub}}, x_{2, \text{sub}}, \dots, x_{N, \text{sub}}\}$   $\mathbf{X}_i^{\text{sub}} = \{x_{i1, \text{sub}}, x_{i2, \text{sub}}, \dots, x_{id, \text{sub}}\}$  is the input sample of working condition parameters that affect the operation of high-speed trains;  $\mathbf{Y}^{\text{main}} = \{y_1^{\text{main}}, y_2^{\text{main}}, \dots, y_N^{\text{main}}\}$  is the sample label of the main task,  $\mathbf{Y}^{\text{sub}} = \{y_1^{\text{sub}}, y_2^{\text{sub}}, \dots, y_N^{\text{sub}}\}$  is the sample label of the subtask for operating condition identification model, the expected value is 0, 1, 2, ...,  $K$ ;  $\mathbf{W}$  is the weight matrix, and  $\mathbf{b}$  is the bias vector;  $H_{LSTM}$  is the number of shared hidden layer LSTM neurons and the initialization parameters are  $\mathbf{W} = \text{random\_uniform0}$  and  $\mathbf{b} = \text{zero}()$ . The network architecture diagram is shown in Figure 5.

- (2) The main task of bearing temperature prediction and the subtask of train operating condition classification are constructed. The two tasks share the LSTM hidden layer features, where the main\_predict\_temperature\_output represents the main task of bearing temperature prediction, which is a regression problem. Subcondition\_classify\_output represents the subtask of train operating condition classification which is a

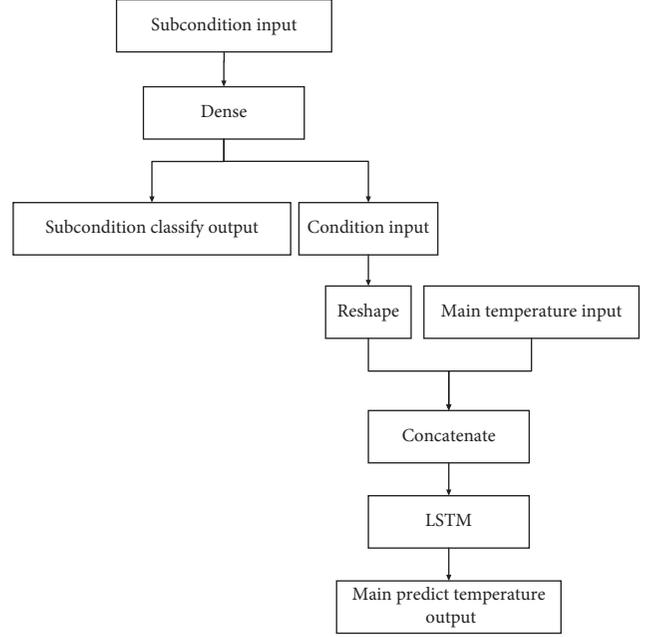


FIGURE 4: The bearing temperature prediction model.

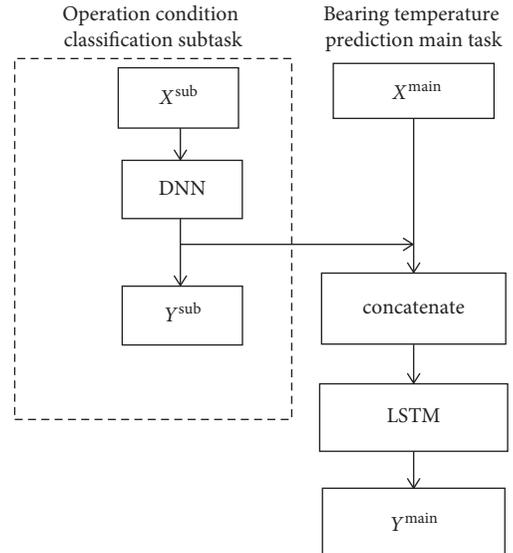


FIGURE 5: Multitask learning model.

classification problem. Among them, the main task loss function uses the mean-squared error (MSE), and the subtask loss function uses the categorical cross-entropy.

- (3) The total loss function of the model is calculated as follows:

$$J = \alpha J_{\text{main}} + \beta J_{\text{sub}} \quad (\alpha = 1, \beta = 1). \quad (4)$$

Among them,

$$\begin{aligned}
J_{\text{main}} &= \frac{1}{n} \sum_{i=1}^n (y_i - \overline{y_1})^2, \\
J_{\text{sub}} &= - \sum_{i=0}^n \overline{y_{i1}} \log y_{i1} + \overline{y_{i2}} \log y_{i2} + \dots + \overline{y_{ik}} \log y_{ik}, \\
\frac{\partial J_{\text{main}}}{\partial y} &= \frac{2}{n} \sum_{i=0}^n (y_i - \overline{y_1}), \\
\frac{\partial J_{\text{sub}}}{\partial y} &= - \sum_{i=0}^n \frac{\overline{y_{i1}}}{y_{i1}} - \sum_{i=0}^n \frac{\overline{y_{i2}}}{y_{i2}} - \dots - \sum_{i=0}^n \frac{\overline{y_{ik}}}{y_{ik}}.
\end{aligned} \tag{5}$$

- (4) Multitask model training: the joint loss function  $J$  of the above two tasks is established, the error function  $J$  is minimized, and the network parameters are updated using the gradient descent method. The formulae are as follows:

$$\begin{aligned}
\mathbf{W}_{jk}^{(l)} &= \mathbf{W}_{jk}^{(l)} - \eta \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{W}_{jk}^{(l)}}, \\
\mathbf{b}_j^{(l)} &= \mathbf{b}_j^{(l)} - \eta \frac{\partial J(\mathbf{W}, \mathbf{b})}{\partial \mathbf{b}_j^{(l)}}.
\end{aligned} \tag{6}$$

Among them,  $\eta > 0$  is the learning rate. In this paper, the stochastic gradient descent (SGD) algorithm is used to solve the minimum value of the total objective function.

**3.3. High-Speed Train Bearing Temperature Fault Diagnosis Method.** Based on the MTL-LSTM model constructed in Section 3.2, the difference between the predicted results and the real results is taken as the basis for anomaly judgment [18]. The specific judgment steps are as follows:

Step 1: the difference between the predicted value and the real result is calculated, and it is recorded as  $e$ .

Step 2: for  $e$ , weighted smoothing is performed according to the set sliding window interval, and it is recorded as  $e_s$ .

Step 3: statistical quality control theory is applied, and through statistical analysis of historical data,  $\mu + 2.5 * \sigma$  ( $\mu$  is the mean and  $\sigma$  is the standard deviation) is set as the outlier upper-line interval. Data greater than or equal to the upper-line interval will be recorded as outliers.

In summary, the real-time diagnosis process of bearing temperature anomalies for high-speed trains based on operating condition identification and MTL is shown in Figure 6.

## 4. Experimental Results and Simulation Analysis

**4.1. Experimental Data.** According to the above analysis, the health status of high-speed train bearings is closely related to the bearing temperature and the factors that affect the temperature such as environment temperature, vehicle speed, passenger capacity, operating status, bridge, tunnel, ramp, and other line characteristics which have a positive correlation to the bearing temperature. As one of the important high-speed railway lines, the Wuhan-Guangzhou high-speed railway line carries thousands of high-speed railway trains annually, with a total length of 965 kilometers [6]. This article selects some high-speed train status data from Wuhan to Guangzhou South as the analysis and statistical samples. The statistics of all 543 days of operation on the Wuhan-Guangzhou line from 2018 to 2019 are as follows: with a total of 7046 up and down lines, of which there are 3104 items in the upstream (Wuhan-Guangzhou South) and 3942 items in the downstream (Guangzhou South-Wuhan). The experimental data are extracted from the full amount of data, and part of the missing data is removed. The total retained data are as follows: the total number of vehicles participating in the analysis is 1700, accounting for 25% of the total 7046. A total of 125 unique train numbers participated in the analysis. During the analysis period, a total of 165 traveled on the Wuhan-Guangzhou line (the starting and ending stations were Wuhan and Guangzhou South). The number of trains in the analysis data accounted for 76% of the total. In some models, in addition, according to the real-time mileage of the train operation and GPS data, the line characteristic data of each train operation are correlated.

**4.2. Data Preprocessing.** According to the sorting and analysis results of EMU operation data, there are some dirty data such as missing value, duplicate value, and abnormal value in the original data of EMU. Through data cleaning, the problems of missing value, duplicate value, and noise data are solved.

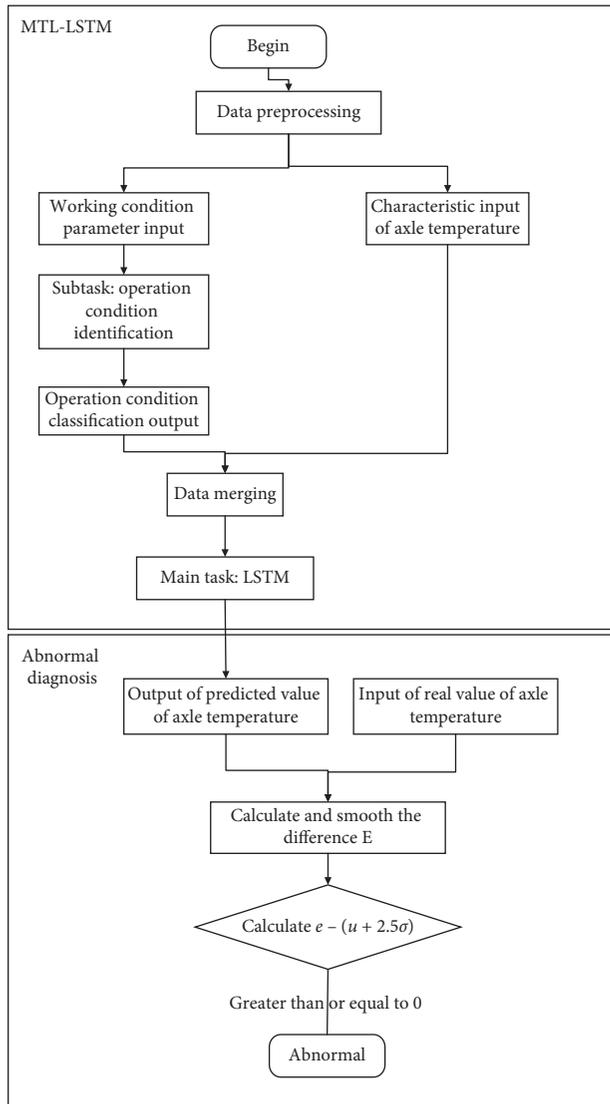


FIGURE 6: The flowchart of real-time diagnosis of bearing temperature abnormality.

The mileage is very important for the original data of real-time operation WTD of EMU. It is necessary to use the running mileage of each train to fit the line kilometer mark to further match the line characteristics of each train running position. Most of the mileage data collected on the vehicle are invalid or 0, which cannot be used. The accumulated mileage should be calculated according to the speed and time interval, and the missing value should be filled in. In addition, the environmental temperature collected by each compartment of the EMU is inconsistent. In this paper, the average value of the ambient temperature of multiple compartments is taken as the ambient temperature of the whole train.

The abnormal value will cause interference to the analysis results, and the bearing temperature data collected by EMU bearing temperature sensor may be abnormal in the process of data collection, so the abnormal bearing temperature data collected will be deleted.

**4.3. Classification of Train Operating Conditions.** A total of 434,100 pieces of historical operating data of trains involved in the analysis in Section 4.1 were identified using the  $K$ -means clustering algorithm. Input parameters include the speed, operating status (traction, braking, idle, and stop), turning radius, whether it is a bridge, a tunnel, or a ramp slope, and fresh air temperature.

The  $K$  value of the appropriate category number is selected, and the Calinski–Harabasz value from 2–20 traversal is calculated. Finally, when  $K = 7$ , the CH score is the largest, indicating that the greater the difference between the categories, the better the clustering effect.

$K = 7$  is brought into the model calculation to get the quantity under each category. Category 6 has very little data and belongs to anomalies, which can be ignored in the analysis. By calculating the center point, mean and median of the input data items of each category number, the maximum value and minimum values, after analysis, and the characteristics of each category can be summarized. For example, the average speed of category 0 is about 262 km/h, and the temperature of the fresh air is about 25 degree Celsius, and it belongs to a curve with a large radius and a small curvature. The average slope is in the uphill stage, and most of the data are concentrated on the bridge, in the nontunnel; the average speed of category 1 is about 254 km/h, and the fresh air temperature is about 28 degrees Celsius, mainly concentrated on noncurved and curved roads. On the tunnel, the average slope is in the uphill stage and is steeper than Category 0; Category 2 is concentrated in the downhill and tunnel; Category 3 operating environment has a low external temperature of about 15 degree Celsius and an average speed of 40 km/h; Category 4 has the highest average speed and belongs to high-speed operation. The speed for Category 5 is concentrated at 30 km/h, which is low-speed operation.

TSNE is used to reduce the dimensionality of the sampled data and it is displayed in three dimensions. As can be seen from Figure 7, there are differences between the categories, the data within the categories are closely connected, and so the train operating conditions on Wuhan-Guangzhou line can be classified into 6 categories.

**4.4. Model Training Results Based on Working Condition Identification and MTL.** The experimental data are the same as in Section 4.1. It is the operation status data of all trains running on Wuhan-Guangzhou line from 2018 to 2019, including bearing fault data caused by an abnormal increase of axle temperature. In the model, the main task input parameters are train axle box temperature, temperature rise rate (average value of temperature increase per minute), and working condition identification results. The output parameters are the temperature values corresponding to the input sample after 1 minute. The input parameters of the subtask are speed, running status, turning radius, whether it is a bridge, a tunnel, or ramp slope, and fresh air temperature. The output parameters of the subtasks are the results of the classification of the working conditions in Section 3.3; the ratios of 70% and 30% are used as training samples and test samples, that is, 303870 training sets and 130230 test

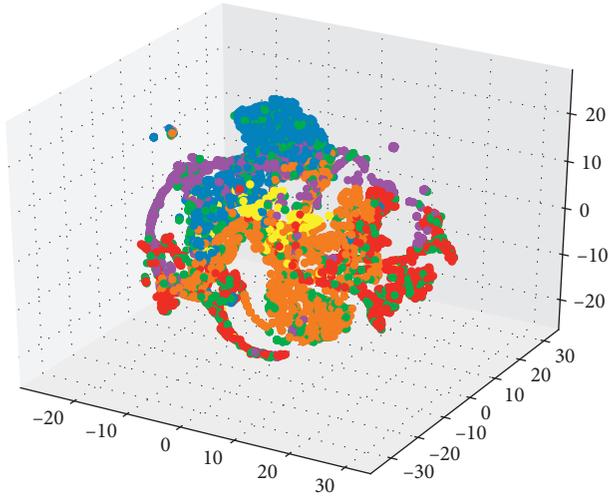


FIGURE 7: Classification results of train operation conditions.

sets. Both the training samples and test samples contain bearing temperature fault samples. Model initialization parameters are set as follows:

- (1) In the subtask, there are 7 neurons in the input layer, 50 neurons in the hidden layer, and 1 neuron in the output layer. In the main task, there are 2 neurons in input layer, 30 neurons in LSTM hidden layer, and 1 neuron in output layer.
- (2) The maximum number of iterations is set as  $\max = 300$ .
- (3) SGD is used in the optimizer, and the time-based attenuation scheme is used to adjust the learning rate. The initial learning rate is 0.1.
- (4) The weight  $w$  is initialized in a random manner.
- (5) The bias  $b$  is initialized to 0.

In this paper, the best network structure of the hidden layer is determined by the traversal method. The hidden layer neurons are from 10 to 200, and each structure repeats 30 times. When the LSTM hidden layer neuron 30 is the optimal, the average loss value is the lowest. Figure 8 shows the change of the loss of training set and verification set with the number of iterations.

In addition, in order to evaluate the performance of the algorithm, single-task learning LSTM (STL-LSTM) and MTL-LSTM in this paper are compared. The results are shown in Table 1. The model is evaluated by comparing the mean square error (MSE), mean absolute error (MAE), and coefficient of determination ( $R2\_score$ ). It can be seen from the results that the MTL-LSTM model achieves the best results whether it is mean square error, mean absolute error, or coefficient of determination, indicating that MTL can effectively improve the performance of fault diagnosis and better prediction the results of fault diagnosis.

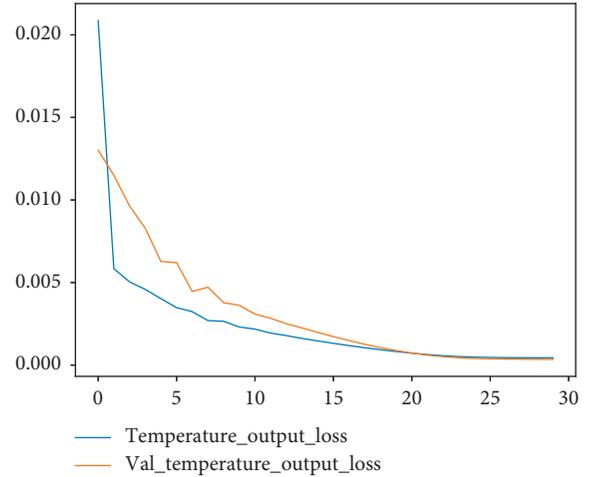


FIGURE 8: The change of iteration loss of training set and verification set.

TABLE 1: The performance comparison of fault diagnosis models.

Model	MSE	MAE	R2 score
STL-LSTM	0.000917	0.018564	0.938591
MTL-LSTM	0.000731	0.012677	0.951049

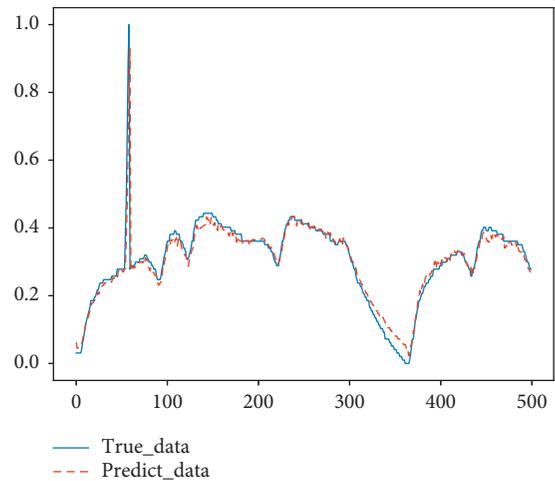


FIGURE 9: The comparison between the predicted value and true value.

**4.5. Fault Diagnosis and Verification.** The abnormal bearing temperature data of a train for abnormal diagnosis and verification are randomly selected. First, the predicted temperature and the real temperature are shown in Figure 9. The horizontal axis is the time series data of bearing temperature. The bearing temperature suddenly rises at about 50, using the abnormal diagnosis method of the difference between the predicted value and the real value in Section 3.3. The abnormal diagnosis result is shown in Figure 10.

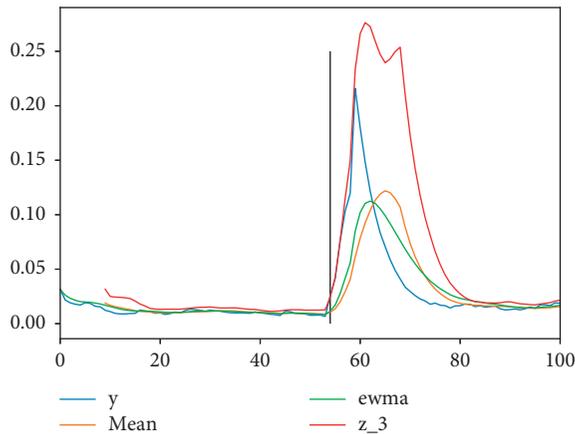


FIGURE 10: The diagnostic results of abnormal bearing temperature.

## 5. Conclusion

Based on the statistical analysis of the operating condition data, line characteristic data, and environmental data of high-speed trains on Wuhan-Guangzhou line for more than one year, it is preliminarily concluded that the factors strongly related to bearing temperature include environmental temperature, speed, and mileage at the same time. Line characteristics such as ramp, bend, bridge, and axle temperature also have a correlation with the shaft temperature; so for the effect of the high-speed train line operating conditions on the train bearing temperature, an abnormal diagnosis method for bearing temperature of the high-speed train is proposed based on operating condition identification and MTL.

Firstly, the operating condition parameters of high-speed trains are analyzed and selected, and the  $k$ -means algorithm is applied to identify the operating conditions of high-speed trains. Then, based on the condition identification and MTL, an MTL-LSTM model including the main task of bearing temperature prediction and the subtask of train operation condition classification is constructed. Finally, according to the statistical quality control theory, the difference between the predicted value and the real value of the MTL-LSTM model is used to diagnose the abnormal bearing temperature of the high-speed train.

In addition, the accuracy of this method is better than that of STL-LSTM. Therefore, a bearing prediction model based on condition identification and MTL is effective and feasible and can be used for bearing temperature anomaly diagnosis of high-speed trains. In the follow-up study, we will design a more effective network structure to improve the accuracy of fault diagnosis, try to use the dynamic weight method to improve the performance of the model, and further test the effect of the model.

## Data Availability

The data are extracted from the daily operation data of high-speed trains that cannot be disclosed.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] Z. Wu and L. Liu, "Study of computational fatigue life method of high speed railway bearings," *Journal of Machine Design*, vol. 32, no. 11, pp. 46–50, 2019.
- [2] O. Polach, "On non-linear methods of bogie stability assessment using computer simulations," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 220, pp. 13–27, 2006.
- [3] D. Liu, Q. Li, H. Lin, and F. Zhao, "Fatigue reliability model of bogie frame structure for high speed train," *China Railway Science*, vol. 38, no. 5, pp. 100–106, 2017.
- [4] W. Tang, M. Wang, and G. Chen, "Analysis on temperature distribution of failure axle box bearings of high speed train," *Journal of the China Railway Society*, vol. 38, no. 7, pp. 50–56, 2016.
- [5] P. Bagiński and G. Żywica, "The influence of temperature on dynamics of the rotor-foil bearing system," *Transactions of the Institute of Fluid-Flow Machinery*, vol. 134, pp. 11–28, 2016.
- [6] H. Pei, C. Hu, and X. Si, "Review of machine learning based remaining useful life prediction methods for equipment," *Journal of Mechanical Engineering*, vol. 55, no. 8, pp. 1–13, 2019.
- [7] P. Bangalore and L. Tjemberg, "An artificial neural network approach for early fault detection of gearbox bearings," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 980–987, 2017.
- [8] Q. Jiang, L. Shen, and W. Zhang, "Fallt diagnosis method based on Deep Learning," *Computer Simulation*, vol. 35, no. 7, pp. 409–413, 2018.
- [9] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical Systems and Signal Processing*, vol. 107, pp. 241–265, 2018.
- [10] N. Zeng, H. Li, Z. Wang, W. Liu, and X. Liu, "Deep-reinforcement-learning-based images segmentation for quantitative analysis of gold immunochromatographic strip," *Neurocomputing*, vol. 19, 2020.
- [11] X. Cao, B. Chen, and N. Zeng, "A deep domain adaption model with multi-task networks for planetary gearbox fault diagnosis," *Neurocomputing*, vol. 409, pp. 173–190, 2020.
- [12] N. Zeng, Z. Wang, H. Zhang, Y. Li, and X. Liu, "An improved particle filter with a novel hybrid proposal distribution for quantitative analysis of gold immunochromatographic strips," *IEEE Transactions on Nanotechnology*, vol. 18, no. 1, pp. 819–829, 2019.
- [13] J. Gao, D. Ren, and J. Xi, "Construction of vehicle driving cycle based on global K-means clustering algorithm," *Journal of Henan Polytechnic University*, vol. 30, no. 1, pp. 112–118, 2019.
- [14] R. Caruana, "Multitask Learning," *Machine Learning*, vol. 28, no. 1, pp. 41–75, 1997.
- [15] F. Cai and V. Cherkassky, "Generalized SMO algorithm for SVM—based multitask learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 6, pp. 997–1003, 2012.
- [16] J. Cheng and J. Wang, "Endpoint prediction method for steelmaking based on multi-task learning," *Journal of Computer Applications*, vol. 37, no. 3, pp. 889–895, 2017.
- [17] Z. Chen, V. Badrinarayanan, C. Lee, and A. Rabinovich, "Gradnorm: gradient normalization for adaptive loss balancing in deep multitask networks," *Computer Vision and Pattern Recognition*, vol. 80, pp. 793–802, 2018.
- [18] E. Viegas, A. Santin, and V. Abreu, "Enabling anomaly-based intrusion detection through model generalization," in *Proceedings of the 2018 IEEE Symposium on Computers and Communications*, arXiv:8538524, Fermo, Italy, November 2018.