

The Recognition of Rail Surface State Based on Improved ResNet-50 Deep Learning Network

Lili Liu¹, Jihong Zuo^{2*}, Ping Ouyang³, Chuanyin Yang¹, Qingwan Tang¹, Tianle Kuang¹

¹College of Intelligent Control, Hunan Railway Professional Technology College, Zhuzhou, China

²College of Rail Transit Locomotive and Rolling Stock, Hunan Railway Professional Technology College, Zhuzhou, China

³College of Rail Transit, Hunan University of Technology, Zhuzhou, China

Email: liulili2023@outlook.com, *zuojihong2006@outlook.com

How to cite this paper: Liu, L.L., Zuo, J.H., Ouyang, P., Yang, C.Y., Tang, Q.W. and Kuang, T.L. (2025) The Recognition of Rail Surface State Based on Improved ResNet-50 Deep Learning Network. *Open Journal of Applied Sciences*, 15, 735-746.
<https://doi.org/10.4236/ojapps.2025.153048>

Received: February 20, 2025

Accepted: March 22, 2025

Published: March 25, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The traction/braking performance of rail transit vehicles depends on the wheel rail contact condition. However the running performance of vehicles is quite different under different rail surface conditions, such as dry, wet and greasy. In view of the problems of large subjectivity and serious lag in traditional artificial experience based rail surface state recognition, a rail surface state recognition method based on an improved ResNet-50 deep learning network is proposed. Firstly, the ResNet-50 network is used to build the rail surface state recognition model. Secondly, transfer learning is introduced to improve the structure and parameters of the ResNet-50 network, and the improved ResNet-50 network is used to train the image data of the rail surface state. Finally, a model that can be used for classification is obtained to recognize the rail surface state. The results show that the optimized ResNet-50 model is more effective in identifying the rail surface state compared with the traditional ResNet-50 model. The model performance is better, and the identification accuracy can reach 92.75%.

Keywords

Rail Surface State, Deep Learning Network, ResNet-50, Transfer Learning

1. Introduction

The effective exertion of traction/braking performance of rail transit vehicles depends on the adhesion utilization when the wheel set and the track contact each other [1]-[3]. The important factor that causes the change of wheel-rail adhesion characteristics is the sudden change of rail surface state. Therefore, effective recognition of different rail surface states is the prerequisite for accurate detection of

wheel-rail adhesion characteristics. The wheel-rail adhesion characteristics under different rail surface conditions are quite different. For example, the adhesion coefficient of snow-covered rail surfaces is much smaller than that of dry rail surfaces. Due to the complicated operation areas of rail transit vehicles, sudden changes in rail surface state occur from time to time during the operation. It is of great significance to realize the effective recognition of rail surface state to improve the utilization rate of wheel-rail adhesion and the operation efficiency of rail vehicles [4] [5].

In order to further explore the influence of rail surface state on wheel-rail adhesion, domestic and foreign scholars have done a lot of research work by means of tests and numerical simulation. In literature [6] [7], the state of the rail surface was changed by applying the “third medium”, such as water, oil, ice and snow. The adhesion change laws under different rail surface states were given in combination with experimental data. However, most of the existing literature simply analyzes the influence of water, oil, ice, snow and other rail surface states on wheel-rail adhesion, but doesn't describe and characterize the information on rail surface states.

With the development of machine vision technology, the visual feature extraction method of rail surface state has received attention [8]-[13]. Xiong and Li *et al.* realized visual inspection of rail surface defects by using a laser to measure the contour [14]. Yu and Li *et al.* constructed a multi-level model for rail surface defect detection by using computer vision technology [15]. Li and Ren *et al.* focused on the image enhancement and automatic thresholding algorithm, aiming at the problems of irregular reflection often encountered in the visual inspection system of track surface defects [16]. Zhang and Jin *et al.* proposed a robust Gaussian mixture model based on Markov random field to rapidly segment rail surface defects [17]. Hu and Zhu *et al.* proposed a track surface defect detection algorithm based on visual saliency in view of the influence of uneven illumination, stray light, rail surface smoothness change and other factors [18]. Gan and Wang *et al.* proposed a background-oriented defect inspector (BODI) to improve defect detection by considering specified characteristics of the track during inspection [19]. However, most of the current rail surface state feature extraction methods are based on the feature difference between the target area and the background area. They use the obvious characteristics of the edge gradient information to obtain information such as rail surface defects [20]. Less consideration is given to the adhesion difference caused by the rail surface state of water, oil and other media. This is not conducive to the improvement of traction/braking performance of rail transit vehicles.

Therefore, this paper intends to combine the visual deep learning and transfer learning methods, and combine the unique characteristics of the rail surface state, and realize the effective recognition of the rail surface state by improving the ResNet-50 deep learning network. The basic structure of the paper is as follows: the recognition network architecture of the rail surface state is given in Section 2, and

the specific functions and functional analysis of each network module are described. The proposed method is verified and analyzed in combination with actual data in Section 3. The conclusion is given in Section 4.

2. The Recognition Network of Rail Surface State

2.1. Overall Frame

Combined with the actual experimental conditions in a certain area, the rail surface state is divided into three types: dry rail surface without pollutants, wet rail surface with water, and oily rail surface with mixed lubricating oil and water. The deep transfer learning network architecture of the rail surface state is shown in Figure 1.

Figure 1.

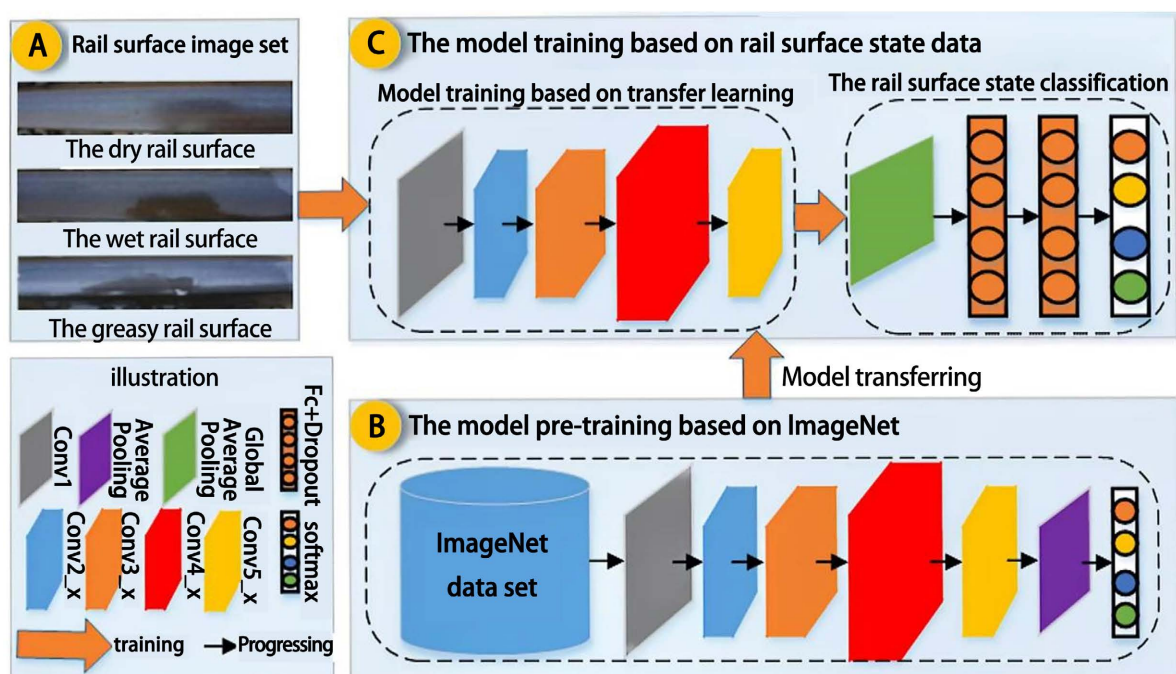


Figure 1. The deep transfer learning network architecture of the rail surface state.

As shown in Figure 1, firstly, the model pre-training network is used to train the ImageNet data set until its classification accuracy converges. Its model parameters are saved. Secondly, the pre-training network model will be removed from the classification layer. The remaining network model will be used as the rail surface image feature extractor of the rail surface state image database, and a custom classification layer will be designed. Then, the network model parameters are fine tuned until the classification accuracy of the rail surface state image converges. The network model and parameters trained by the ImageNet data set are transferred to the rail surface state recognition network model.

2.2. ResNet-50 Deep Learning Model

The basic structure of ResNet-50 is shown in Figure 2. ResNet-50 deep learning

network structure consists of five parts: an input layer (Input), five convolution groups (Conv1 - Conv5_x), an average pool layer, a full connection layer (FC) and an output layer (Output). Each convolution group can generally contain one or more basic convolution operations and one down-sampling operation, which can shorten the feature map to half of the original. The down-sampling can be realized by two methods: the first is to pool the maximum value, and the step size is usually 2. This method is only applicable to the second convolution group (Conv2). The second method uses convolution, and the step size is also 2. This method is applicable to the remaining four convolution groups (Conv1, Conv3, Conv4, Conv5) except the second convolution group.

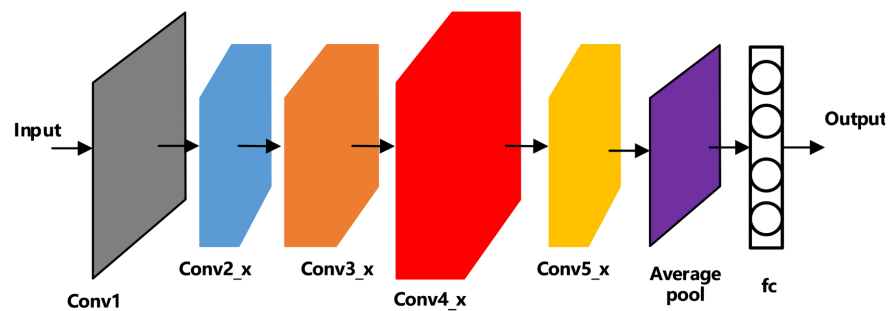


Figure 2. The basic structure of ResNet-50.

The input data is subject to feature learning through five convolution groups (conv1 - conv5_x). The first convolution group (Conv1) contains only one convolution calculation. The second to fifth convolution groups (conv2_x - conv5_x) contain multiple identical residual units. After extracting the features, the main features are retained by the compression of the average pooling layer and transmitted to the full connection layer for classification and recognition, the output is what we want.

The ResNet-50 network structure is complex and the deep network is difficult to train. The requirements for computer hardware resources are very high, and its recognition effect depends on the number of data samples. If the training data samples are insufficient, the deep learning network cannot learn the required sample characteristics. Even if the training data samples are sufficient, after the timeliness of the data samples, the data samples cannot be better applied to new tasks. Therefore, the introduction of migration learning method can reduce the dependence of deep learning network on the number of training set sample data in the target area. Through this method, the problem of insufficient training data samples can be solved, the pressure of computer hardware can be reduced, and the time of training can be shortened.

2.3. Model Optimization Based on Transfer Learning

The optimization process of ResNet-50 based on transfer learning is divided into two steps. The first step is pre-training, ResNet-50 network model is trained in

ImageNet data set until its classification accuracy converges, and then its model parameters are saved. The second step is fine-tuning, the last classification layer is removed from the pre-training ResNet-50 model. The remaining ResNet-50 model is used as the rail surface image feature extractor of the rail surface state image database. Then the designed custom classification layer is connected, and the new ResNet-50 network model is fine tuned until the classification accuracy of the rail surface state image converges. The model and parameters of ResNet-50 trained by ImageNet data set are transferred to the self-built rail surface data set.

1) Pre-training

When using deep learning to deal with problems, the deeper the network is selected, the better the learning ability of the model, and the model will need more data samples for training. If a supervised learning method is adopted, a large number of data samples need to be labeled, but the labeled samples are difficult to obtain. Too few training samples will easily lead to overfitting. When the deep network features extracted by the model are more, there will be a lot of multi-feature problems, such as multi data sample fusion and feature selection. The problem of multi-layer neural network parameter optimization is actually a high-order nonconvex optimization problem, which often results in local solutions with relatively poor convergence. As the network is too deep, the problem of gradient diffusion will occur. The gradient calculated by BP algorithm will obviously decrease with the deepening of the network, which will make use of the previous network parameters very few and the speed of updating the network is very slow.

Therefore, the built network model is pre-trained for the classification of specific images. Firstly, the parameters of the built network model are initialized, and then the built network model is trained by the prepared image data sample set. During the training process, the model is constantly adjusted to make the damage function of the model gradually stable and tend to zero. The model parameters will change continuously at the beginning of training until the classification accuracy converges. The model parameters will be saved to facilitate the same operation on other images next time.

The pre-training designed in this paper is to train ResNet-50 neural network with ImageNet data set until its classification accuracy converges. ImageNet data set is often used as a training set because its image data sample is very large, including 1.2 million images. It is also conducive to the training of universal models. For other images other than the ImageNet data set, the pre-trained deep learning network can also show good generalization ability.

2) Fine-tuning

The parameters of ResNet-50 model saved after pre-training are used as initialization parameters for rail surface state recognition. Then some modifications are made continuously according to the results during training. The specific steps are as follows:

a) Firstly, only the input layer and the first five convolution layers are reserved in the ResNet-50 model after pre-training. These five convolution layers are used

as the feature extractors of the rail surface image data.

b) Since the average pooling layer and the full connection layer are removed from the ResNet-50 model after pre-training, a custom classification layer needs to be designed. As shown in **Figure 3**, the custom classification layer designed in this paper consists of the Global Average Pooling layer and the full connection layer (FC). Global Average Pooling can average the entire Feature Maps and regularize the entire ResNet-50 network structurally to prevent overfitting. So the original Average Pooling layer is replaced. While FC can play the “firewall” function in the process of model transfer learning, so three-layer FC is added after Global Average Pooling. In order to effectively prevent too many parameter coupling of the first two layers of FC, dropout is introduced to effectively suppress overfitting. The last layer FC connects a function of the softmax logic partition regressor to classify the rail surface state.

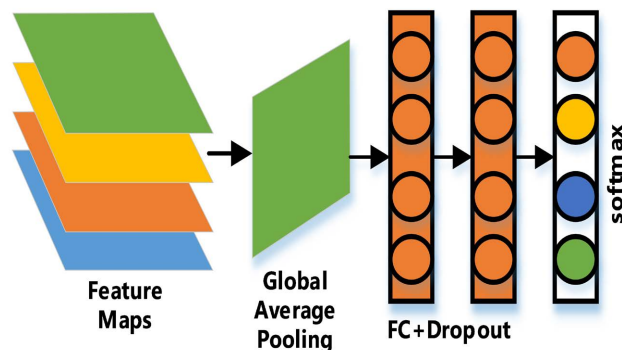


Figure 3. The custom classification layer.

c) Finally, the new ResNet-50 model is trained by inputting the rail surface state image data. When using the optimized ResNet-50 model to train the rail surface state image, the parameters in the ResNet-50 network layer that have not been replaced should be frozen in the first step of training. The optimized ResNet-50 model performs forward calculation, but the parameters will not be updated during reverse transfer. Only the FC parameters in the customized classification layer will be updated during the training of the optimized ResNet-50 model. Then, a relatively small learning rate is selected for training, such as 0.001. Finally, when the FC parameters are learned almost, the ResNet-50 network layer that has not been replaced can be unfrozen. Then the entire ResNet-50 network can be trained.

3. Data Validation and Result Analysis

3.1. Analysis of Experimental Results with Different Activation Functions

The custom classification layer adds a full connection layer, different activation functions will affect the accuracy of rail surface image recognition. In order to obtain better classification effect, three activation functions of Sigmoid, Tanh, ReLU are used in this experiment to identify the rail surface image respectively.

The average recognition accuracy of the three activation functions on the rail surface image is shown in **Table 1**. The images of the training set and the test set are divided into the rail surface image database according to 7:2. Adam is selected as the learner, with a learning rate of 0.001, Dropout = 0.5, Epoch = 50, batch_size = 6.

Table 1. The average recognition accuracy of three activation functions for rail surface image.

Activation function	Average recognition accuracy
Sigmoid	92.75%
Tanh	47.25%
ReLU	73.91%

It can be seen from **Table 1** that under the same conditions, the activation function of Sigmoid is used to identify the rail surface image, and the highest accuracy is 92.75%. Using the activation function of ReLU to identify the rail surface image is poor, and the recognition accuracy is 73.91%. The recognition effect of Tanh's activation function is the worst, and the recognition accuracy is 47.25%. Therefore, Sigmoid is selected as the activation function of the custom layer. The experimental parameters are set as learner = Adam, learning rate = 0.001, dropout = 0.5, epoch = 50, batch_size = 6.

3.2. Performance Analysis of Rail Surface Recognition in Different States

In order to verify the feasibility of the optimized ResNet rail surface recognition model, the model is used to identify the dry, wet and oily rail surfaces. The experimental parameters are set according to Section 3.1. For better visual recognition effect, the confusion matrix is used to represent the recognition results and recognition accuracy of the test set. The results are shown in **Figure 4** and **Figure 5** respectively.

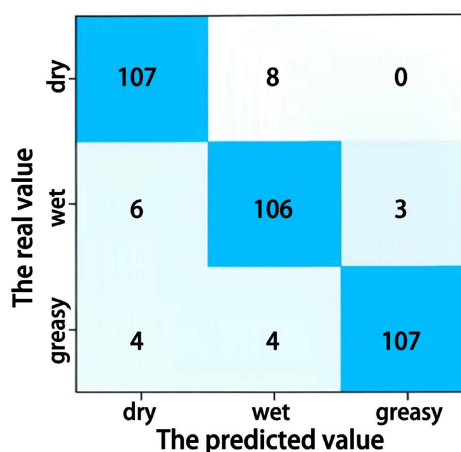


Figure 4. The recognition results of the test set.

Figure 4 is the confusion matrix of the recognition results for the test set. The horizontal axis represents the predicted rail surface. The vertical axis represents the real rail surface. The sum of the elements for each row of the matrix represents the total number of samples. The elements on the diagonal of the matrix represent the correct number of dry, wet and greasy rail surface identifications respectively. It can be seen from **Figure 4** that there are 115 images of dry, wet and greasy rail surfaces respectively. Among them, the number of correct recognition of dry rail surfaces is 107, the number of misidentifications as wet rail surfaces is 8, and the number of misidentifications as greasy rail surfaces is 0. The number of correct recognition of wet rail surfaces is 106, the number of misidentifications as dry rail surfaces is 6, and the number of misidentifications as greasy rail surfaces is 3. The number of correct recognition of greasy rail surfaces is 107, the number of misidentifications as dry rail surfaces is 4, and the number of misidentifications as wet rail surfaces is 4. It can be seen that the optimized ResNet-50 rail surface state identification model can effectively identify different rail surface states, but the error rate is low.

The real value	dry	0.9145	0.0678	0
	wet	0.0513	0.8983	0.0273
	greasy	0.0452	0.0339	0.9727
		dry	wet	greasy
		The predicted value		

Figure 5. The recognition accuracy of test set.

Figure 5 is the recognition accuracy confusion matrix of the test set. The elements on the diagonal of the matrix represent the correct recognition accuracy of dry, wet and greasy rail surfaces respectively. It can be observed from **Figure 5** that the recognition accuracy rate of dry rail surface is 0.9145, that of wet rail surface is 0.8983, and that of greasy rail surface is 0.9727. It can be seen that the recognition accuracy of the optimized ResNet-50 rail surface state recognition model is above 0.8983. The model has high recognition accuracy and excellent performance.

3.3. Methods Comparative Analysis

The optimized ResNet-50 model and ResNet-50 model are respectively used to identify the rail surface images in this experiment. The experimental parameters are also set according to Section 3.1. The recognition accuracy of the two models

on the test set is shown in Table 2. The accuracy and loss distribution of the optimized ResNet-50 model and ResNet-50 model with the increase of epoch from 1 to 50 are shown in Figure 6 and Figure 7 respectively.

It can be observed from Table 2 that the accuracy of the optimized ResNet-50 rail surface identification model is higher. The recognition accuracy rate reaches 92.75%. The recognition accuracy rate of the ResNet-50 rail surface recognition model is only 70.14%.

Table 2. The recognition accuracy of two models for test sets.

Model	Recognition accuracy
ResNet-50	70.14%
Optimized ResNet-50	92.75%

Figure 6 shows the accuracy-Epoch curve of the optimized ResNet-50 model and the ResNet-50 model training set. It can be seen that with the increase of epoch, the accuracy of the optimized ResNet-50 model gradually approaches 1. But the accuracy of the ResNet-50 model gradually approaches 0.8. It can be seen that the optimized ResNet-50 model has a better recognition effect and better performance.

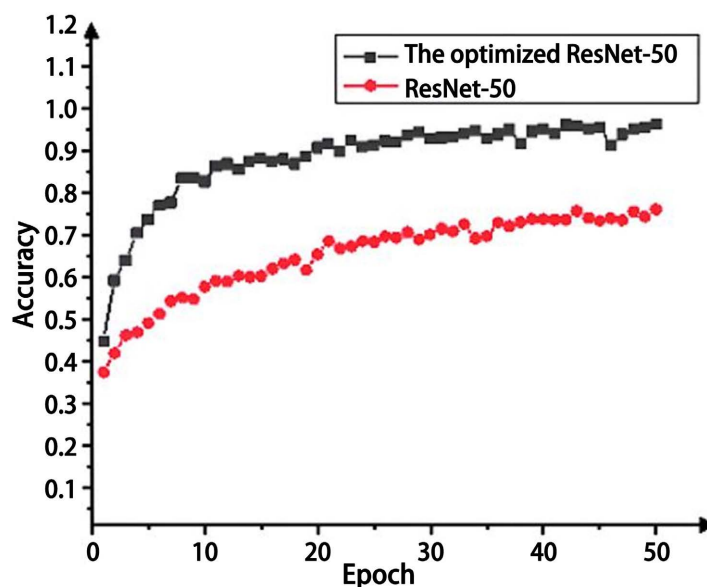


Figure 6. The accuracy-Epoch diagram of ResNet-50 model and the optimized ResNet-50 model training set.

Figure 7 shows the loss-Epoch curve of the optimized ResNet-50 model and the ResNet-50 model training set. It can be observed that the loss convergence of the optimized ResNet-50 model is faster. With the increase of epoch, the loss shows a fluctuating downward trend and gradually approaches 0. While the loss of the ResNet-50 model converges slowly with the increase of epoch, the convergence

time also needs to be longer. It can be seen that the optimized ResNet-50 model converges more quickly and has a shorter convergence time in the same Epoch.

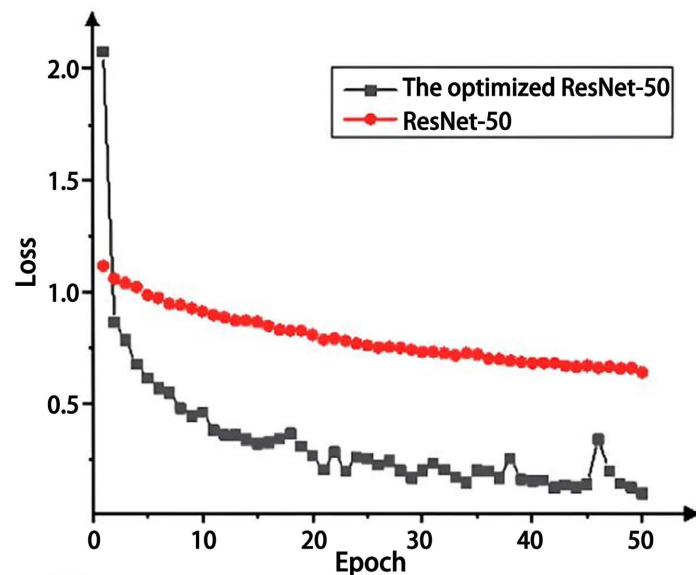


Figure 7. The loss-Epoch diagram of ResNet-50 model and the optimized ResNet-50 model training set.

In summary, the optimized ResNet-50 rail surface identification model adopted in this paper is more effective in identifying rail surface status. The model performance is better, and the recognition accuracy can reach 92.75%.

4. Conclusions

1) An improved ResNet-50 deep learning network model for rail surface state recognition is proposed. The model uses the residual network ResNet-50 as the backbone network. The network structure and parameters are improved based on transfer learning. The rail surface state image data is trained based on the improved ResNet-50 network. The results show that this method reduces the number of parameters and improves the network training speed on the premise of ensuring accuracy.

2) Compared with the traditional ResNet-50 model, the improved ResNet-50 rail surface recognition model has better performance. The accuracy of rail surface recognition can reach 92.75%.

Acknowledgements

The research is funded by the Scientific research project of Hunan Provincial Department of Education, China (Wheel-Rail Adhesion Depth Modeling Based on Rail Surface Condition Identification, under Grant 22B1013).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Liu, J., Liu, L., He, J., Zhang, C. and Zhao, K. (2020) Wheel/Rail Adhesion State Identification of Heavy-Haul Locomotive Based on Particle Swarm Optimization and Kernel Extreme Learning Machine. *Journal of Advanced Transportation*, **2020**, Article ID: 8136939. <https://doi.org/10.1155/2020/8136939>
- [2] Zou, R., Ma, W. and Luo, S. (2018) Influence of the Wheel Diameter Difference on the Wheel/Rail Dynamic Contact Relationship of the Heavy Haul Locomotive. *Australian Journal of Mechanical Engineering*, **16**, 98-108. <https://doi.org/10.1080/14484846.2018.1456787>
- [3] Pichlik, P. and Bauer, J. (2021) Adhesion Characteristic Slope Estimation for Wheel Slip Control Purpose Based on UKF. *IEEE Transactions on Vehicular Technology*, **70**, 4303-4311. <https://doi.org/10.1109/tvt.2021.3072484>
- [4] Wu, B., Wu, T., Wen, Z. and Jin, X. (2016) Numerical Analysis of High-Speed Wheel/Rail Adhesion under Interfacial Liquid Contamination Using an Elastic-Plastic Asperity Contact Model. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, **231**, 63-74. <https://doi.org/10.1177/1350650116645025>
- [5] He, J., Liu, G., Liu, J., Zhang, C. and Cheng, X. (2018) Identification of a Nonlinear Wheel/Rail Adhesion Model for Heavy-Duty Locomotives. *IEEE Access*, **6**, 50424-50432. <https://doi.org/10.1109/access.2018.2868177>
- [6] Chang, C., Chen, B., Cai, Y. and Wang, J. (2019) An Experimental Study of High Speed Wheel-Rail Adhesion Characteristics in Wet Condition on Full Scale Roller Rig. *Wear*, **440**, Article 203092. <https://doi.org/10.1016/j.wear.2019.203092>
- [7] Liu, X., Xiao, C. and Meehan, P.A. (2019) The Effect of Rolling Speed on Lateral Adhesion at Wheel/Rail Interface under Dry and Wet Condition. *Wear*, **438**, Article 203073. <https://doi.org/10.1016/j.wear.2019.203073>
- [8] Feng, H., Jiang, Z., Xie, F., Yang, P., Shi, J. and Chen, L. (2014) Automatic Fastener Classification and Defect Detection in Vision-Based Railway Inspection Systems. *IEEE Transactions on Instrumentation and Measurement*, **63**, 877-888. <https://doi.org/10.1109/tim.2013.2283741>
- [9] Torabi, M., Mousavi, S.G.M. and Younesian, D. (2018) A High Accuracy Imaging and Measurement System for Wheel Diameter Inspection of Railroad Vehicles. *IEEE Transactions on Industrial Electronics*, **65**, 8239-8249. <https://doi.org/10.1109/tie.2018.2803780>
- [10] Chandra, B.S., Sastry, C.S. and Jana, S. (2019) Robust Heartbeat Detection from Multimodal Data via CNN-Based Generalizable Information Fusion. *IEEE Transactions on Biomedical Engineering*, **66**, 710-717. <https://doi.org/10.1109/tbme.2018.2854899>
- [11] He, J., Yin, L., Liu, J., Zhang, C. and Yang, H. (2022) A Fault Diagnosis Method for Unbalanced Data Based on a Deep Cost Sensitive Convolutional Neural Network. *IFAC-PapersOnLine*, **55**, 43-48. <https://doi.org/10.1016/j.ifacol.2022.05.008>
- [12] Liu, J., Yang, H., He, J., Sheng, Z. and Chen, S. (2022) Unbalanced Fault Diagnosis Based on an Invariant Temporal-Spatial Attention Fusion Network. *Computational Intelligence and Neuroscience*, **2022**, Article ID: 1875011. <https://doi.org/10.1155/2022/1875011>
- [13] Yu, Y., Sun, W., Liu, J. and Zhang, C. (2022) Traffic Flow Prediction Based on Depth-wise Separable Convolution Fusion Network. *Journal of Big Data*, **9**, Article No. 83. <https://doi.org/10.1186/s40537-022-00637-9>
- [14] Xiong, Z., Li, Q., Mao, Q. and Zou, Q. (2017) A 3D Laser Profiling System for Rail

Surface Defect Detection. *Sensors*, **17**, Article 1791.

<https://doi.org/10.3390/s17081791>

- [15] Yu, H., Li, Q., Tan, Y., Gan, J., Wang, J., Geng, Y., et al. (2019) A Coarse-to-Fine Model for Rail Surface Defect Detection. *IEEE Transactions on Instrumentation and Measurement*, **68**, 656-666. <https://doi.org/10.1109/tim.2018.2853958>
- [16] Li, Q. and Ren, S. (2012) A Visual Detection System for Rail Surface Defects. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **42**, 1531-1542. <https://doi.org/10.1109/tsmcc.2012.2198814>
- [17] Zhang, H., Jin, X., Wu, Q.M.J., Wang, Y., He, Z. and Yang, Y. (2018) Automatic Visual Detection System of Railway Surface Defects with Curvature Filter and Improved Gaussian Mixture Model. *IEEE Transactions on Instrumentation and Measurement*, **67**, 1593-1608. <https://doi.org/10.1109/tim.2018.2803830>
- [18] Hu, Z., Zhu, H., Hu, M. and Ma, Y. (2018) Rail Surface Spalling Detection Based on Visual Saliency. *IEEJ Transactions on Electrical and Electronic Engineering*, **13**, 505-509. <https://doi.org/10.1002/tee.22594>
- [19] Gan, J., Wang, J., Yu, H., Li, Q. and Shi, Z. (2020) Online Rail Surface Inspection Utilizing Spatial Consistency and Continuity. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **50**, 2741-2751. <https://doi.org/10.1109/tsmc.2018.2827937>
- [20] Gibert, X., Patel, V.M. and Chellappa, R. (2017) Deep Multitask Learning for Railway Track Inspection. *IEEE Transactions on Intelligent Transportation Systems*, **18**, 153-164. <https://doi.org/10.1109/tits.2016.2568758>