

## Research Article

# Deep Sparse Autoencoder for Feature Extraction and Diagnosis of Locomotive Adhesion Status

Changfan Zhang <sup>1</sup>, Xiang Cheng,<sup>1,2</sup> Jianhua Liu,<sup>1</sup> Jing He <sup>1</sup> and Guangwei Liu<sup>1</sup>

<sup>1</sup>College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou 412000, China

<sup>2</sup>CHNV Technology Co., Ltd., Dongguan 523000, China

Correspondence should be addressed to Jing He; [hejing@263.net](mailto:hejing@263.net)

Received 6 January 2018; Revised 30 May 2018; Accepted 27 June 2018; Published 15 July 2018

Academic Editor: Qiang Liu

Copyright © 2018 Changfan Zhang et al. 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.

The model is difficult to establish because the principle of the locomotive adhesion process is complex. This paper presents a data-driven adhesion status fault diagnosis method based on deep learning theory. The adhesion coefficient and creep speed of a locomotive constitute the characteristic vector. The sparse autoencoder unsupervised learning network studies the input vector, and the single-layer network is superimposed to form a deep neural network. Finally, a small amount of labeled data is used to fine-tune training the entire deep neural network, and the locomotive adhesion state fault diagnosis model is established. Experimental results show that the proposed method can achieve a 99.3% locomotive adhesion state diagnosis accuracy and satisfy actual engineering monitoring requirements.

## 1. Introduction

Precise diagnosis of the wheel-rail adhesion state is an important prerequisite for adhesion control. Currently, the wheel-rail adhesion state of a locomotive is mostly diagnosed based on the detection and analysis of relevant parameters to determine the type of adhesion state and the degree of adhesion [1]. For the diagnosis of adhesion states, a sampling eigenvector should be generated based on the creeping speed of the driving wheel and the wheel-rail cohesion coefficient, a sample feature should be extracted, and the feature should be coded; then, various intelligent algorithms should be used to classify the eigenvector [2]. Many studies on the use of neural networks in the adhesion field have been reported. For example, Castillo et al. [3] used a neural network to estimate the adhesion state in an ABS system. Castillo [4] trained an artificial neural network to calculate the best creep operating point for each road on the basis of traffic information collected by a vehicle sensor. Li Ningzhou [5] studied the adhesion feature of the air brake of a locomotive and used the optimized recursive neural network to optimize the parameters of the adhesive controller and improve the

utilization rate of locomotive adhesion, thereby obtaining a good experimental result.

These methods are more convenient and intelligent than the general mechanism analysis method. However, these methods still belong to the supervised learning area [6]. Thus, they require sufficient data for feature extraction. Meanwhile, extracting the right features is often relatively complex and difficult. To obtain labels, experiments and rich professional knowledge are required. With the artificial participation factors, the uncertainty of feature extraction and optimization is greatly increased, thereby making the diagnosis of the right adhesion state difficult. Furthermore, a traditional neural network essentially uses hidden-layer neurons for nonlinear transformation [7]. It can learn potential features from a given sample and fit out an approximation function [8]. Taking the classical BP neural network as an example, obtaining high-precision features becomes more difficult while the layers are few. If the number of layers is excessive, then the gradient may disappear, and the local optimal solution is another defect that is difficult to overcome [9].

Sample feature extraction is a key step in determining the accuracy of fault diagnosis [10]. The change in adhesion

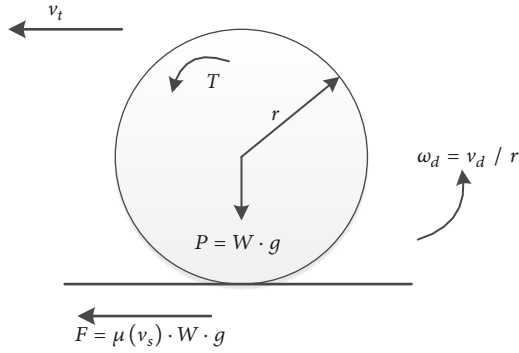


FIGURE 1: Schematic of adhesion.

state is a complex process that is affected by multiple factors, producing a complex nonlinear relation between factors and outcomes. Fault prediction and analysis are particularly challenging. The introduction of deep learning [11] has made a breakthrough in research on high-precision feature extraction. As an unsupervised learning algorithm, deep neural network not only has an excellent feature extraction ability but can also overcome the common problem of obtaining sample labels [12]. Thus, deep neural network has become a popular research area in the field of fault diagnosis [13–15]. This paper proposes a sparse autoencoder deep neural network with dropout to diagnose the wheel-rail adhesion state of a locomotive. This deep neural network can significantly reduce the adverse effect of overfitting, making the learned features more conducive to classification and identification.

The rest of this paper is organized as follows: Section 2 describes the adhesion principle and characteristics. Section 3 describes the principle and process of the deep neural network algorithm. Section 4 discusses the comparative experimental research and result analysis. Section 5 presents the conclusions.

## 2. Description of Adhesion Status

Adhesion is the ultimate manifestation of locomotive driving force in the wheel-rail relationship and the fundamental motive force for locomotives [16]. The wheel pair rolls forward when subjected to a tangential traction, and the rolling pressure causes deformation between the wheel and the rail. Simultaneously, the gravity of the car body imposed on the rail keeps the contact surface between the wheel and the rail relatively stable. This phenomenon is called adhesion. As shown in Figure 1, the contact point between the wheel and the rail is elastically deformed under the action of the wheel load ( $P$ ). The wheel rolls forward under the action of the driving torque ( $T$ ), the original contact surface deformation develops into a new elliptical deformation, and the tractive effort at the wheel rim ( $F$ ) is generated.

Adhesion coefficient  $\mu$  is typically defined as the ratio of traction to axle weight (see (1)), where  $W$  is the axle weight (kg) and  $g$  is the gravitational acceleration ( $\text{m/s}^2$ ).

$$\mu = \frac{F}{W \cdot g} \quad (1)$$

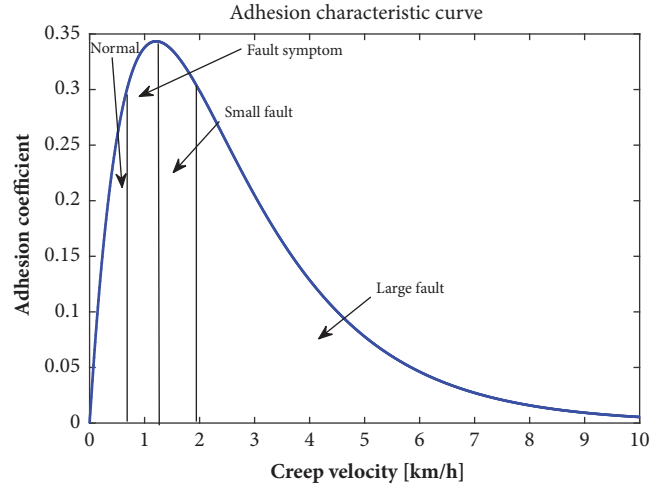


FIGURE 2: Refined adhesion characteristic curve.

In the process of normal movement, the train body speed ( $v_t$ ) is always less than the wheel speed ( $v_d$ ) due to the wheel-rail microsliding generated by the deformation. This phenomenon is called creep, and the speed difference between them is defined as creep speed  $v_s$ .

$$v_s = v_d - v_t \quad (2)$$

Creep is a slight wheel-spin phenomenon produced by the locomotive drive system. The adhesion coefficient of the rail contact surface rises constantly with the creep speed within a certain range [17], and the locomotive has a great available traction. Once the range is exceeded, the wheel-rail adhesion coefficient drops sharply with the increase in the creep speed.

Figure 2 shows the adhesion characteristic curve of the locomotive. The adhesion peak point is taken as the boundary in which the left side is called the creep region and the right side is called the slid region [4]. However, when the adhesion state is divided into two categories, abnormal adhesion can be identified, but the predicted foundation of potential creep failure cannot be provided. To this end, this paper further refines the adhesion features: normal (N0), fault symptom (N1), small fault (F1), and large fault (F2).

The adhesion state is divided into four categories. When a minor fault is encountered, fault tolerant control methods [18] can be adopted to prevent serious system performance deterioration [1].

## 3. Deep Neural Network

Unsupervised learning can be used to automatically learn potential features from the samples without labels [19, 20]. This method has a significant advantage when addressing complex problems, such as adhesion state recognition. The sparse autoencoder is an unsupervised algorithm, and this deep neural network can effectively extract the characteristics that reflect the adhesion state [21, 22].

**3.1. Sparse Autoencoder.** From the structural point of view, the autoencoder is an axisymmetric single hidden-layer

neural network [23]. The autoencoder encodes the input sensor data by using the hidden layer, approximates the minimum error, and obtains the best-feature hidden-layer expression [24]. The concept of the autoencoder comes from the unsupervised computational simulation of human perceptual learning [25], which itself has some functional flaws. For example, the autoencoder does not learn any practical feature through copying and inputting memory into implicit layers, although it can reconstruct input data with high precision. The sparse autoencoder inherits the idea of the autoencoder and introduces the sparse penalty term, adding constraints to feature learning for a concise expression of the input data [26, 27].

For the adhesion state identification of locomotive,  $k$  sets of monitoring data  $\{x_1, x_2, x_3, \dots, x_n\}$  exist, which are reconstructed into a  $N \times M$  data set  $\{x(1), x(2), x(3), \dots, x(N)\}$ ,  $x(i) \in R^M$ . These data are used as input matrix  $X$ . The input data encoded by the automatic encoder are used to construct a mapping relationship. In this paper, the activation function of the autoencoder is sigmoid, which is designed to obtain a better representation of input data:  $h(X, W, b) = \sigma(WX + b)$ . A sparse penalty term is added to the sparse autoencoder cost function to limit the average activation value of the hidden-layer neuron. Normally, when the output value of a neuron is 1, it is active, and the neuron is inactive when its output value is 0. The purpose of enforcing sparsity is to limit the undesired activation.  $a_j(x)$  is set as the  $j$ th activation value. In the process of feature learning, the activation value of the hidden-layer neuron is usually expressed as  $a = \text{sigmoid}(WX + b)$ , where  $W$  is the weight matrix and  $b$  is the deviation matrix. The mean activation value of the  $j$ th neuron in the hidden layer is defined as

$$\rho_j = \frac{1}{n} \sum_{i=1}^n [a_j(x(i))]. \quad (3)$$

The hidden layer is kept at a lower value to ensure that the average activation value of the sparse parameter is defined as  $\rho$ , and the penalty term is used to prevent  $\rho_j$  from deviating from parameter  $\rho$ . The Kullback–Leibler (KL) divergence [28] is used in this study as the basis of punishment. The mathematical expression of KL divergence is as follows:

$$KL(\rho \parallel \rho_j) = \rho \ln \frac{\rho}{\rho_j} + (1 - \rho) \ln \frac{1 - \rho}{1 - \rho_j}. \quad (4)$$

When  $\rho_j$  does not deviate from parameter  $\rho$ , the KL divergence value is 0; otherwise, the KL divergence value will gradually increase with the deviation. The cost function of the neural network is set as  $C(W, b)$ . Then, the cost function of adding the sparse penalty term is

$$C_{\text{sparse}} = C(W, b) + \beta \sum_{j=1}^{S_2} KL(\rho \parallel \rho_j) \quad (5)$$

where  $S_2$  is the number of neurons in the implicit layer and  $\beta$  is the weight of the sparse penalty term. The training essence of a neural network is to find the appropriate weight and threshold parameter ( $W, b$ ). After the sparse penalty term is defined, the sparse expression can be obtained by minimizing the sparse cost function.

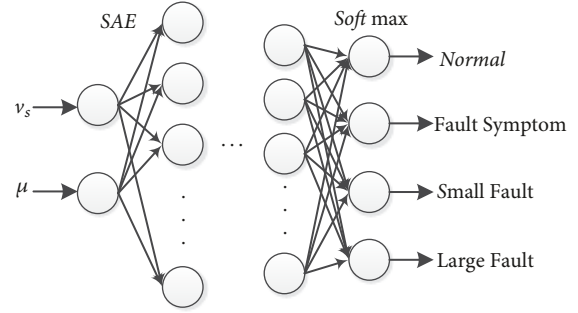


FIGURE 3: Schematic of the network structure.

**3.2. Softmax Regression.** The sparse autoencoder can form the deep network structure through the multilayer stack, which can be used for feature learning and clustering of the adhesion data collected by the sensor. However, this autoencoder has no ability to classify. Therefore, this paper presents a deep neural network architecture that combines the stacked sparse self-encoder and softmax regression. The schematic of the network structure is shown in Figure 3.

Softmax regression is an extension of the logistic regression model on multiple classifications [29]. The category tag of the logistic regression can only take two values, whereas the softmax tag can take on multiple values [30]. Let us suppose  $m$  training samples of adhesion state  $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$ ,  $y^{(i)} \in \{1, 2, \dots, n\}$ . The hypothesis function is used to estimate the probability value  $p(y = j | x)$  for each category  $j$ . The softmax output is defined as follows:

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix} \quad (6)$$

$$= \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix}$$

where  $\theta_1, \theta_2, \dots, \theta_k \in \mathfrak{R}^{n+1}$  is the model parameter.  $1 / \sum_{j=1}^k e^{\theta_j^T x^{(i)}}$  normalizes the probability distribution such that the sum of all probabilities is 1.

**3.3. Overfitting and L2 Regularization.** L2 regularization is a way of effectively reducing the neural network overfitting [31]. In this study, this method is used to avoid overlearning on features caused by synergies. The basic principles of L2 regularization are as follows:

$$C'_{\text{sparse}} = C_{\text{sparse}} + \alpha \|w\|^2 \quad (7)$$

TABLE 1: Description of adhesion state of locomotive wheel rail.

Different types of States	State representation
Normal	Running with a low adhesion
Fault symptom	Near the peak adhesion coefficient
Small fault	Across the peak adhesion point
Large fault	Serious slip running unsafe already

where  $C_{sparse}$  is the cost function of the neural network,  $\alpha$  is coefficient, and  $\|w\|^2$  is the penalty term. The greater the coefficient, the deeper the weight attenuation.

*3.4. Framework of the Algorithm.* The feature learning ability of the single sparse autoencoder is limited. To construct a model with improved feature extraction capacity, we stacked the sparse autoencoders into a deep structure (SAE). In this process, the output of the upper layer of the encoder is taken as the input of the next layer to achieve a multilearning sample feature. The flowchart of the deep neural network algorithm is shown in Figure 4 and subsequently described.

*(1) The Creep Speed ( $v_s$ ) and Adhesion Coefficient  $\mu$  Monitored by the Sensor Are Used to Train the Sparse Autoencoder*

- (1) The parameters, such as the network learning rate and the dropout parameter, are set; weights  $W$  and thresholds  $b$  are initialized.
- (2) The number of iterations is set, and the mean activation value  $\rho_j$  and the sparse cost function are calculated according to (3)–(5); the network parameters are updated based on the backpropagation (BP) algorithm.

*(2) The Deep Neural Network Is Fine-Tuned with a Small Number of Labeled Samples*

- (1) With the abovementioned step, the parameters of threshold  $b$  and weight  $W$  learned from the network are saved.
- (2) The L2 regularization and learning rates are set, and the mean square error is calculated.
- (3) The BP algorithm is used to update the weights of the network and fine-tune the entire network.

*(3) The Performance of the Identification Model Is Tested*

- (1) The sample size is usually 30% of the total number of samples.
- (2) According to the L2 coefficient, the weight of neural network is attenuated while performing the front-propagation algorithm.
- (3) The output of DNN is compared with the sample labels, and comparative statistics are made.

## 4. Experimental Research and Analysis

For performance comparison with this dropout-based deep neural network, a BP neural network and an optimized BP

TABLE 2: State type.

Different types of states	State encoding
Normal (N0)	000
Fault symptom (N1)	001
Small fault (F1)	010
Large fault (F2)	100

neural network based on a genetic algorithm (GA-BP) are used as state recognition models. Taking the creep speed ( $v_s$ ) and the adhesion coefficient ( $\mu$ ) of the locomotive as characteristic signals, the adhesion features include normal operation zone, wheel-spin warning zone, slight wheel-spin zone, and serious wheel-spin. Detailed information is shown in Table 1. The test sample added a Gaussian noise with a mean of 0 and a variance of 0.02 for an improved fit construction.

The changes in the state of adhesion directly affect the running safety of a locomotive. These changes are reflected in the sensor monitoring data. In this section, the adhesion state of the locomotive is identified and diagnosed according to the inherent characteristics of the sample data of deep neural network based on the sparse autoencoder. In general, the original data samples of sensors are divided into training and test sets according to a 7:3 ratio. A total of 700 training samples and 300 test samples are used in this experiment.

*4.1. Simulation of the BP Neural Network.* We need to set parameters before the experiment. The number of hidden-layer nodes is set according to the empirical formula  $l = 2n + 1$ , where  $n$  is the input dimension and  $l$  is the number of hidden-layer nodes. In this study, the creep speed ( $v_s$ ) and the adhesion coefficient ( $\mu$ ) are used as inputs, so that  $l$  should be 5, and the adhesion state is expressed in binary form, as shown in Table 2. This BP neural network should be a  $2 \times 5 \times 3$  structure (Figure 5).  $W$  is the weight and  $\beta$  is the bias of the BP neural network. The mean square error curve is shown in Figure 6.

To demonstrate the advantages of the proposed algorithm, a genetic algorithm- (GA-) optimized BP neural network is used as a contrast experiment. The crossover operator uses a single point crossover, the crossover probability is 0.7, and the mutation probability is 0.01. The evolutionary process of the GA is shown in Figure 7, and the error descent curve of the BP neural network that is optimized by the GA is shown in Figure 8.

*4.2. Simulation of Sparse Autoencoder Deep Neural Network.* The visualization of the target classification is shown in Figure 9 to provide a clear analysis of the classification ability of the proposed algorithm. The plane between the yellow and blue modules is the desired classification plane. In this experiment, 1,000 sets of monitoring data are selected as experimental samples, and the 7:3 ratio is used to divide the samples into training and test samples. The actual results are shown in Figure 10, which shows that the actual classification plane is basically consistent with the expected one.

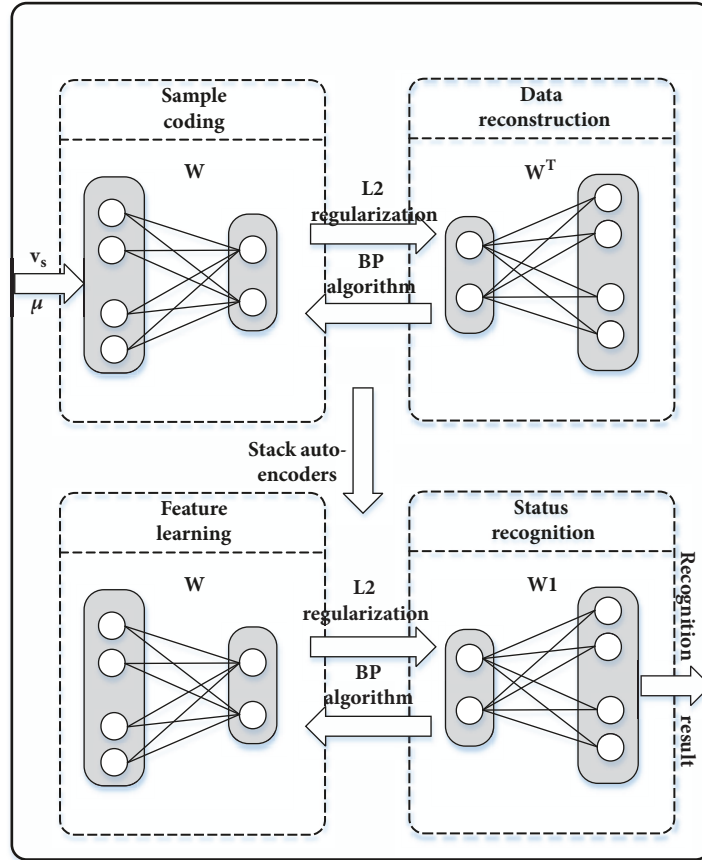


FIGURE 4: Flowchart of deep neural network algorithm.

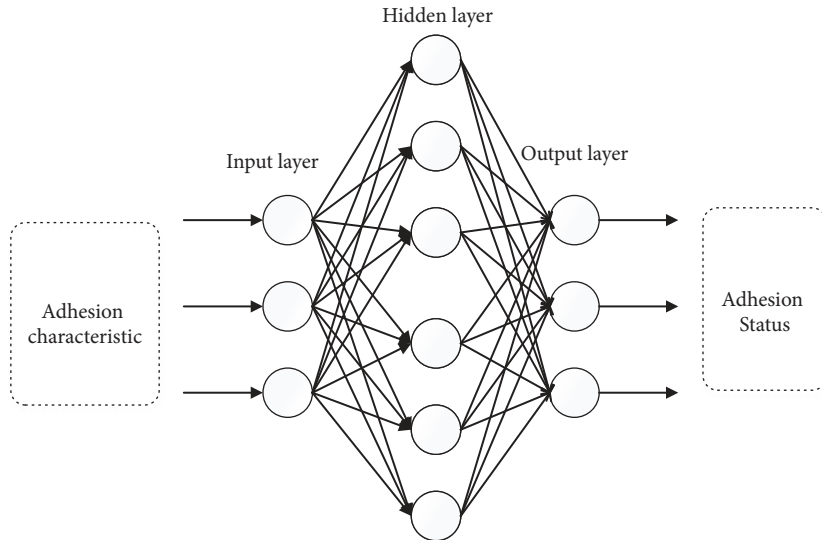


FIGURE 5: Schematic of the structure of the BP neural network.

Figure 9 shows the visualization of the classification target of adhesion status. In general, it is necessary to divide the adhesion status into four different statuses. Three classification planes are needed to achieve this (yellow and blue junction in the figure). The requirement of training

neural network to accurately classify the adhesion state of locomotives is to make the test dataset also present a clear classification plane. The actual state division result is shown in Figure 10. The classifying plane of the adhesion state is clear.



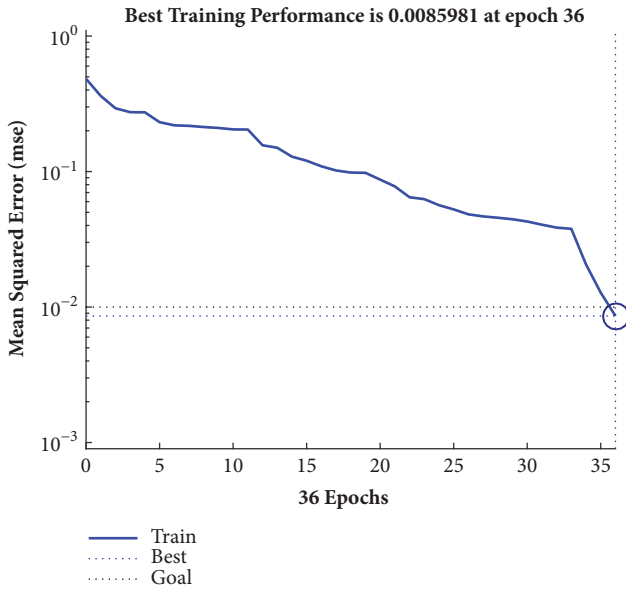


FIGURE 6: Simulation error curves of BP neural network.

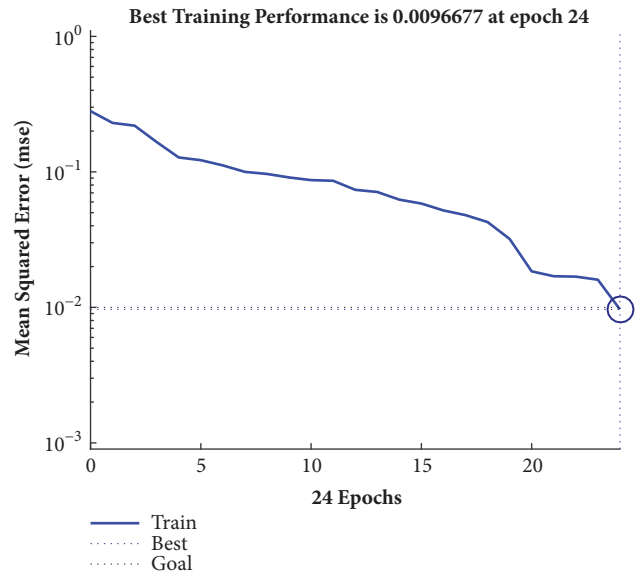


FIGURE 8: Simulation error curves of the GA-BP neural network.

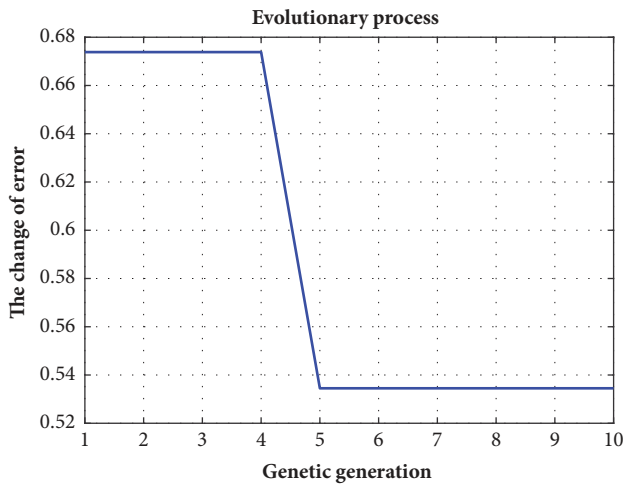


FIGURE 7: Evolutionary process.

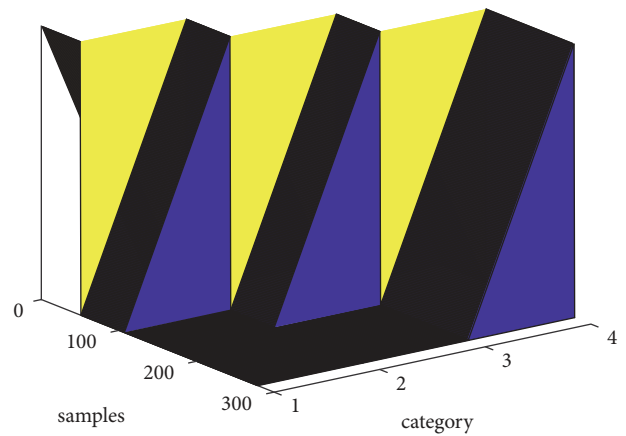


FIGURE 9: Classification target visualization.

From the error histogram in Figure 11, the error distribution of the deep neural network in this chapter is basically in line with the normal distribution, which meets the needs of practical application. The accuracy of adhesion state recognition is shown in Figures 12 and 13. The horizontal axis represents the desired target category, the vertical axis represents the experimentally predicted adhesion state category, and the gray block shows the exact percentage of prediction and expectation.

Figure 12 shows that the accuracy of the deep neural network adhesion state recognition is 96.1%. The overfitting of the neural network generally appears as the trained neural network does not accurately identify the test samples. The adhesion state is a continuously changing process; in order to ensure the safety of driving, the identification of the adhesion state must be as accurate as possible. Since the recognition accuracy rate does not reach the ideal state and there are

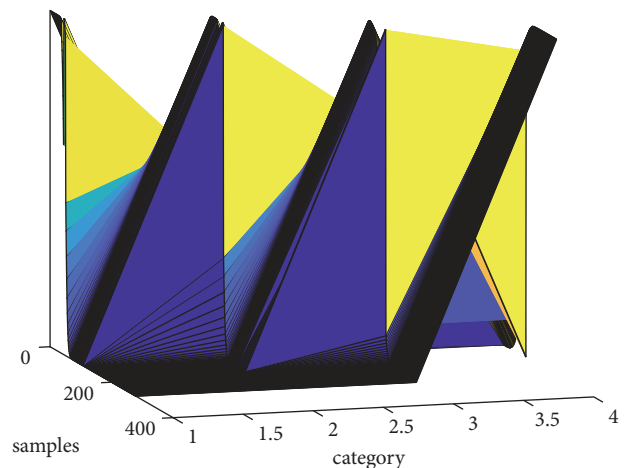


FIGURE 10: Contrast experiment with sparse coefficient 0.08.

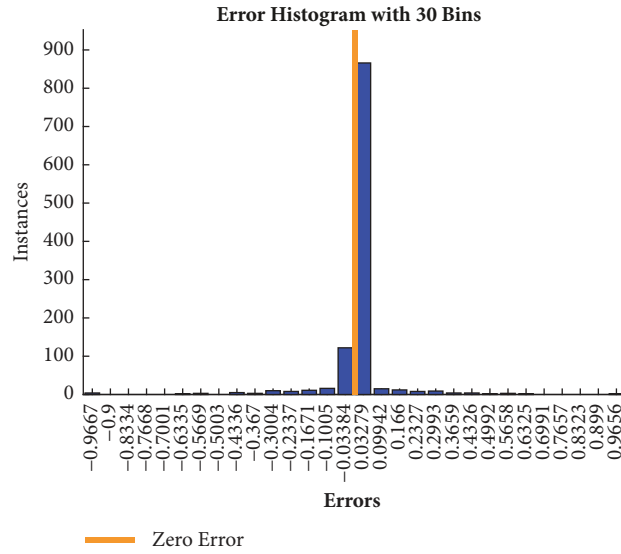


FIGURE 11: Error histogram.

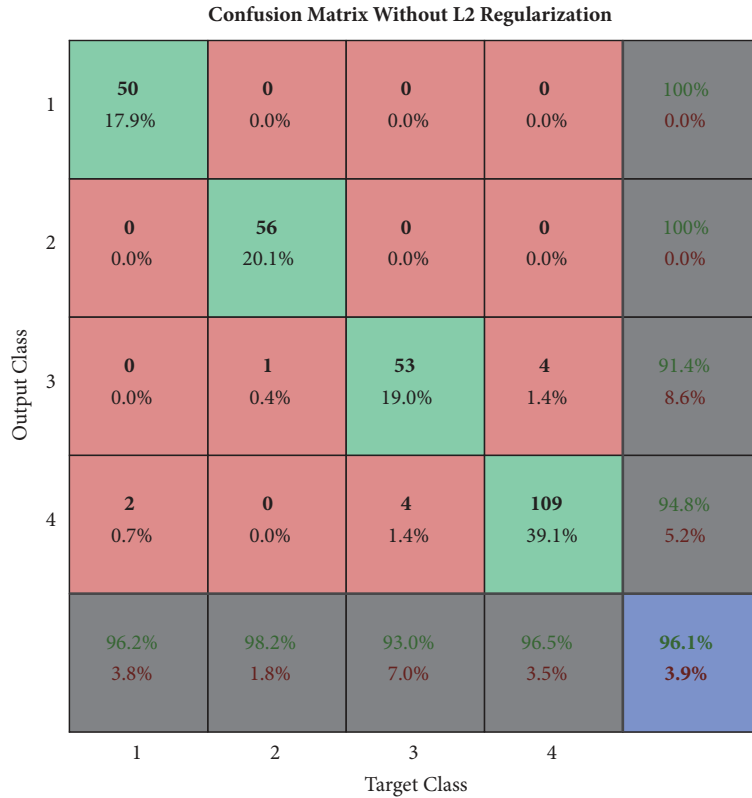


FIGURE 12: Confusion matrix without L2 regularization.

significantly more training samples than test samples, there is ample reason to speculate that the deep neural network used in this section has been overfitted. To improve this phenomenon, the L2 regularization method was used to attenuate the weights of the deep neural network. Figure 13 shows the results of the deep neural network adhesion state recognition after L2 regularization. The accuracy rate reaches

99.6%. The accuracy rate of the neural network for the adhesion state test set is improved, and the proposed L2 regularization can improve the overfitting phenomenon that may occur in the adhesion state recognition based on deep neural network.

The experimental results show that the SAE-based locomotive adhesion diagnosis can meet the requirements of

**Confusion Matrix**

Output Class	1	50 17.9%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	2	0 0.0%	57 20.4%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	0 0.0%	57 20.4%	0 0.0%	100% 0.0%
	4	2 0.7%	0 0.0%	0 0.0%	113 40.5%	98.3% 1.7%
		96.2% 3.8%	100% 0.0%	100% 0.0%	100% 0.0%	99.3% 0.7%
	1	2	3	4	<b>Target Class</b>	

FIGURE 13: Confusion matrix.

TABLE 3: Test accuracy of the four methods.

Method	Type of adhesion state			
	N0	N1	F1	F2
BP	83.33	57.14	77.78	95.45
GA-BP	84.23	55.49	88.62	97.66
DNN	96.2	100	100	99.3

high-accuracy recognition under a reasonable error distribution. Table 3 presents the comparison of the three algorithms in this paper with the traditional BP neural network and GA-optimized neural network performance.

## 5. Conclusions

In this paper, an adhesion state fault diagnose method based on SAE is proposed. The effectiveness of the proposed method is validated by computer simulation. The conclusions are elaborated as follows:

(1) The adhesion state is divided into four categories, which provide a strong basis for wheel skid warning.

(2) The sparse automatic encoder can extract data features effectively, make classification easier, and extract more robust data features.

(3) Compared with the traditional BP neural network, the deep neural network of the sparse autoencoders can ensure effective fault diagnosis of the locomotive adhesion condition.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Authors' Contributions

Changfan Zhang and Xiang Cheng conceived and designed the research; Changfan Zhang and Xiang Cheng performed the research; Changfan Zhang, Xiang Cheng, Jing He, Jianhua Liu, and Guangwei Liu wrote the paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. 61773159, 61473117), Hunan



Provincial Natural Science Foundation of China (nos. 2018JJ2093, 2017JJ4031), Key Laboratory for Electric Drive Control and Intelligent Equipment of Hunan Province (no. 2016TP1018), and Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province.

## References

- [1] C. Zhang, J. Sun, J. He, and L. Liu, "Online Estimation of the Adhesion Coefficient and Its Derivative Based on the Cascading SMC Observer," *Journal of Sensors*, vol. 2017, Article ID 8419295, 11 pages, 2017.
- [2] C. Zhang, X. Cheng, J. He, and G. Liu, "Automatic recognition of adhesion states using an extreme learning machine," *International Journal of Robotics and Automation*, vol. 32, no. 2, pp. 194–200, 2017.
- [3] J. J. Castillo, J. A. Cabrera, A. J. Guerra, and A. Simón, "A Novel Electrohydraulic Brake System with Tire-Road Friction Estimation and Continuous Brake Pressure Control," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 3, pp. 1863–1875, 2016.
- [4] J. J. C. Aguilar, J. A. C. Carrillo, A. J. G. Fernández, and E. C. Acosta, "Robust road condition detection system using in-vehicle standard sensors," *Sensors*, vol. 15, no. 12, pp. 32056–32078, 2015.
- [5] N. Li, X. Feng, and X. Wei, "Optimized adhesion control of locomotive airbrake based on GSA-RNN," in *Proceedings of the 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2015*, pp. 157–161, China, August 2015.
- [6] A. Stanescu and D. Caragea, "An empirical study of ensemble-based semi-supervised learning approaches for imbalanced splice site datasets," *BMC Systems Biology*, vol. 9, no. 5, article no. S1, 2015.
- [7] M. Madhjarasan and S. N. Deepa, "Comparative analysis on hidden neurons estimation in multi layer perceptron neural networks for wind speed forecasting," *Artificial Intelligence Review*, vol. 48, no. 4, pp. 449–471, 2017.
- [8] N. J. Guliyev and V. E. Ismailov, "A single hidden layer feed-forward network with only one neuron in the hidden layer can approximate any univariate function," *Neural Computation*, vol. 28, no. 7, pp. 1289–1304, 2016.
- [9] M. Cai and J. Liu, "Maxout neurons for deep convolutional and LSTM neural networks in speech recognition," *Speech Communication*, vol. 77, pp. 53–64, 2016.
- [10] Y. Liu, Y. Li, X. Ma, and R. Song, "Facial expression recognition with fusion features extracted from salient facial areas," *Sensors*, vol. 17, no. 4, 2017.
- [11] G. E. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [12] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction," *Sensors*, vol. 17, no. 4, 2017.
- [13] Xiao He, Zidong Wang, Gang Li, Zhijie Zhou, and Youqing Wang, "Fault Diagnosis and Application to Modern Systems," *Journal of Control Science and Engineering*, vol. 2017, pp. 1–3, 2017.
- [14] C. Li, R.-V. Sánchez, G. Zurita, M. Cerrada, and D. Cabrera, "Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning," *Sensors*, vol. 16, no. 6, article no. 895, 2016.
- [15] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, 2017.
- [16] T. Ishrat, G. Ledwich, M. Vilathgamuwa, and P. Borghesani, "Wheel slip control based on traction force estimation of electric locomotives," in *Proceedings of the Australasian Universities Power Engineering Conference, (AUPEC '16)*, Australia, 2016.
- [17] M. Spiryagin, P. Wolfs, C. Cole et al., "Theoretical investigation of the effect of rail cleaning by wheels on locomotive tractive effort," in *Core*, p. V001T10A003, 2016.
- [18] Xiao He, Yamei Ju, Yang Liu, and Bangcheng Zhang, "Cloud-Based Fault Tolerant Control for a DC Motor System," *Journal of Control Science and Engineering*, vol. 2017, pp. 1–10, 2017.
- [19] H. Elghazel and A. Aussem, "Unsupervised feature selection with ensemble learning," *Machine Learning*, vol. 98, no. 1–2, pp. 157–180, 2013.
- [20] F. Anselmi, J. Z. Leibo, L. Rosasco, J. Mutch, A. Tacchetti, and T. Poggio, "Unsupervised learning of invariant representations," *Theoretical Computer Science*, vol. 633, pp. 112–121, 2016.
- [21] W. Zhou, Z. Shao, C. Diao, and Q. Cheng, "High-resolution remote-sensing imagery retrieval using sparse features by auto-encoder," *Remote Sensing Letters*, vol. 6, no. 10, pp. 775–783, 2015.
- [22] O. Tsinalis, P. M. Matthews, and Y. Guo, "Automatic Sleep Stage Scoring Using Time-Frequency Analysis and Stacked Sparse Autoencoders," *Annals of Biomedical Engineering*, vol. 44, no. 5, pp. 1587–1597, 2016.
- [23] M. Kang, K. Ji, X. Leng, X. Xing, and H. Zou, "Synthetic aperture radar target recognition with feature fusion based on a stacked autoencoder," *Sensors*, vol. 17, no. 1, 2017.
- [24] J. Leng and P. Jiang, "A deep learning approach for relationship extraction from interaction context in social manufacturing paradigm," *Knowledge-Based Systems*, vol. 100, no. C, pp. 188–199, 2016.
- [25] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.
- [26] D. Yang, J. Lai, and L. Mei, "Deep representations based on sparse auto-encoder networks for face spoofing detection," *Springer International Publishing*, 2016.
- [27] J. Xu, L. Xiang, Q. Liu et al., "Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 119, 2016.
- [28] A. Youssef, C. Delpha, and D. Diallo, "An optimal fault detection threshold for early detection using Kullback-Leibler Divergence for unknown distribution data," *Signal Processing*, vol. 120, pp. 266–279, 2016.
- [29] J. Yang, Y. Bai, G. Li, M. Liu, and X. Liu, "A novel method of diagnosing premature ventricular contraction based on sparse auto-encoder and softmax regression," *Bio-Medical Materials and Engineering*, vol. 26, pp. S1549–S1558, 2015.
- [30] M. M. A. Rahhal, Y. Bazi, H. Alhichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Information Sciences*, vol. 345, pp. 340–354, 2016.
- [31] E. Phaisangittisagul, "An Analysis of the Regularization Between L2 and Dropout in Single Hidden Layer Neural Network," in *Proceedings of the 7th International Conference on Intelligent Systems, Modelling and Simulation, ISMS 2016*, pp. 174–179, Thailand, January 2016.



Hindawi

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

