

Research Article

Elevator Running Fault Monitoring Method Based on Vibration Signal

Mingxing Jia , Xiongfei Gao , Hongru Li, and Hali Pang

College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

Correspondence should be addressed to Mingxing Jia; jiamingxing@ise.neu.edu.cn

Received 3 July 2021; Accepted 24 September 2021; Published 7 October 2021

Academic Editor: Changqing Shen

Copyright © 2021 Mingxing Jia 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.

According to the one-dimensional characteristics of the vibration signal, this paper proposes an elevator operation fault monitoring method based on one-dimensional convolutional neural network (1-DCNN). It can solve the problems of traditional elevator fault monitoring methods that require complex feature extraction processes and a large amount of diagnostic experience. Because the elevator fault monitoring field has less fault information, it is different from the large sample situation in the field of face recognition. Aiming at the problem of small samples, this paper first preprocesses elevator vibration signals through singular value decomposition (SVD) and wavelet transform, then uses wavelet transform to extract wavelet energy features of the original vibration signals, and then use PCA to reduce the feature data to the dimension with a cumulative contribution rate of greater than 85%. When reducing the dimensionality, the original characteristics of the features are preserved as much as possible. When designing the 1-CNN, the K-fold cross-validation method is added to obtain as many abnormalities from the sample set as possible. The information is finally trained using the 1-CNN and classified by softmax regression. In order to verify the performance of the algorithm, the original vibration signal was used as the input of the 1-CNN, and the wavelet energy feature without PCA dimensionality reduction was used as the input of the 1-CNN. The experimental results showed that the 1-DCNN model with PCA dimension-reduced feature data as input can effectively extract and identify the features of normal and abnormal states and has high fault identification accuracy, and good results have been obtained.

1. Introduction

The elevator is essential special equipment in modern life. In order to ensure the safe and reliable operation of elevator equipment, it is of great practical significance to carry out elevator condition monitoring research, which can provide technical support for elevator safety risk identification and evaluation [1–3]. Since the vibration signal is easier to reflect the operating status of the elevator, the most frequently used monitoring method in the field of elevator fault monitoring is to use the vibration signal as the original signal. The fault diagnosis method based on signal processing mainly extracts the vibration signal in time domain and frequency domain to obtain the characteristic information, which can represent the elevator running state. The traditional time-domain and frequency-domain features can directly reflect the characteristic information of elevator operation state. The time-

domain features mainly include slope, kurtosis, peak value, mean value, and variance. The frequency-domain features mainly include frequency slope, frequency kurtosis, mean square frequency, root mean square frequency, stability factor, and amplitude corresponding to the second harmonic frequency.

In order to complete the extraction of time-domain and frequency-domain features, scholars at home and abroad have studied many traditional methods: wavelet transform, fast Fourier transform, empirical mode decomposition, and so on. Xu et al. [4] used wavelet transform and empirical mode decomposition (EMD) vibration signal extraction methods to monitor and extract features of vibration signals of elevator traction gearbox under normal and abnormal noise conditions, so as to realize the state monitoring and identification of potential fault sources of elevator mechanical components without dismantling the machine. The

frequency spectrum obtained by Zhang and Wang [5] after using fast Fourier transform can be used to analyze the abnormal vibration fault of elevator traction machine caused by bearing problem, and the error caused by spectrum aliasing can be avoided by order analysis. Li et al. [6] proposed an adaptive time-frequency signal processing method, and Dragomiretskiy and Zosso [7] proposed a completely nonrecursive variational mode decomposition model and used the alternating direction method of multiplier method to optimize the variational model, which has stronger robustness.

With the rise of machine learning [8] and deep learning [9], more and more fault monitoring researches are combined with them. The fault diagnosis technology based on deep learning can be divided into four categories: fault diagnosis method based on trestle self-coding (SAE), fault diagnosis method based on deep confidence network (DBN), fault diagnosis method based on convolutional neural network (CNN), and fault diagnosis method based on recurrent neural network (RNN) [10]. Shi et al. [11] used the fusion of multiple superimposed sparse self-encoders to monitor the tool condition in ultraprecision machining, and the classification accuracy of this method for ultraprecision machining cases reached more than 96%. Chen et al. [12] proposed a fault diagnosis method based on improved depth confidence network, which optimized the network feature extraction ability, to improve the ability of network learning and classification to reduce the dependence of network training on data. The performance of the improved deep confidence network model is tested by using the open network dataset, and the improved network model is applied to the fault data set of small samples. Janssens et al. [13] used a single-layer convolutional layer for bearing fault diagnosis and achieved an accuracy of 93.16%. Shaowu et al. [14] proposed to use convolutional neural network to classify and identify hydraulic pump time-frequency images with different volumetric efficiencies to realize the health monitoring of hydraulic pumps. First, collect the vibration signal data of the hydraulic pump. The time-frequency diagram has short-time Fourier transform, wavelet transform, and Wigner-Will distribution. Then, the generated time-frequency graph is divided into training set and test set. The training set is used to train the convolutional neural network, and the test set is used to verify the recognition results, and the recognition rate reaches 99%. Xia [15] proposed a method to extract features by discrete wavelet transform and then establish a cyclic neural network model. Cui et al. [16] used FFT to transform time-domain signals into frequency-domain signals to extract features and train the recurrent neural network model.

In this paper, the research object is focused on the elevator, which has more extensive application and more complex vibration fault causes. First, the original signal is preprocessed by singular value decomposition and wavelet transform and then is trained by 1-CNN to extract features automatically. Finally, the classification is carried out by softmax regression. Aiming at the problem of high dimensionality and high complexity of feature data in the process of feature extraction, in this paper, principal

component analysis (PCA) method is used to optimize the dimensionality of feature data. Full use of PCA can effectively reduce the complexity of feature data, identify the most important features, and effectively improve the recognition rate.

2. Pretreatment of Elevator Vibration Signal

2.1. Singular Value Decomposition for Noise Reduction. Because people move around in the elevator car, the vibration signal of the elevator will produce more noise components in the measurement process. At the same time, the acceleration sensor will produce more noise components in the measurement process. Therefore, before using the vibration signal for fault monitoring, the original signal needs to be denoised. In this paper, the singular value decomposition (SVD) method is used. SVD is a signal processing method with good numerical robustness and adaptability, which can effectively identify noise component and fault feature component by singular value decomposition. In the field of signal processing, SVD is a classical orthogonal transformation method. This method transforms an orthogonal matrix into a singular value matrix (diagonal matrix) by multiplying an orthogonal matrix by its left and right, respectively. The number of singular values obtained reflects the number of independent row (column) vectors in the original matrix, and its size represents the intrinsic properties of the signal. The algorithm has good stability and invariance and can be used for denoising, feature extraction and weak signal separation. According to SVD theory, for a matrix $A \in R_{m \times n}$, there must be an orthogonal matrix U , V such that

$$A = U_{m \times m} S V_{n \times n}^T \quad (1)$$

where diagonal matrix $S = [\text{diag}[\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_q] 0]$ or its transposition. The singular value matrix S_{m-1} is obtained by SVD decomposition of Hankel. The singular value curvature spectrum method and the principle of singular value contribution rate are used to adaptively eliminate noise singular values and smooth signal singular values (singular value is set to zero) to retain the effective singular value of abrupt signal. Then, the singular value matrix of sudden change signal is reconstructed by SVD inverse operation method, and the mutation signal $x'(t)$ is obtained. The specific content of SVD is shown in [17]. The original vibration signal waveform and SVD filtered waveform are shown in Figure 1. The waveforms of the elevator during normal operation, the waveforms during abnormal operation (small jitter amplitude), and the waveforms when there is severe jitter are listed, respectively.

2.2. Wavelet Energy Extraction. When an elevator traction machine fails, its transfer function will change accordingly, so the amplitude frequency and phase frequency characteristics of different frequency signals will change. The enhancement and inhibition of each frequency component of the output signal changed obviously. At this time, compared with the output of normal elevator traction machine, the

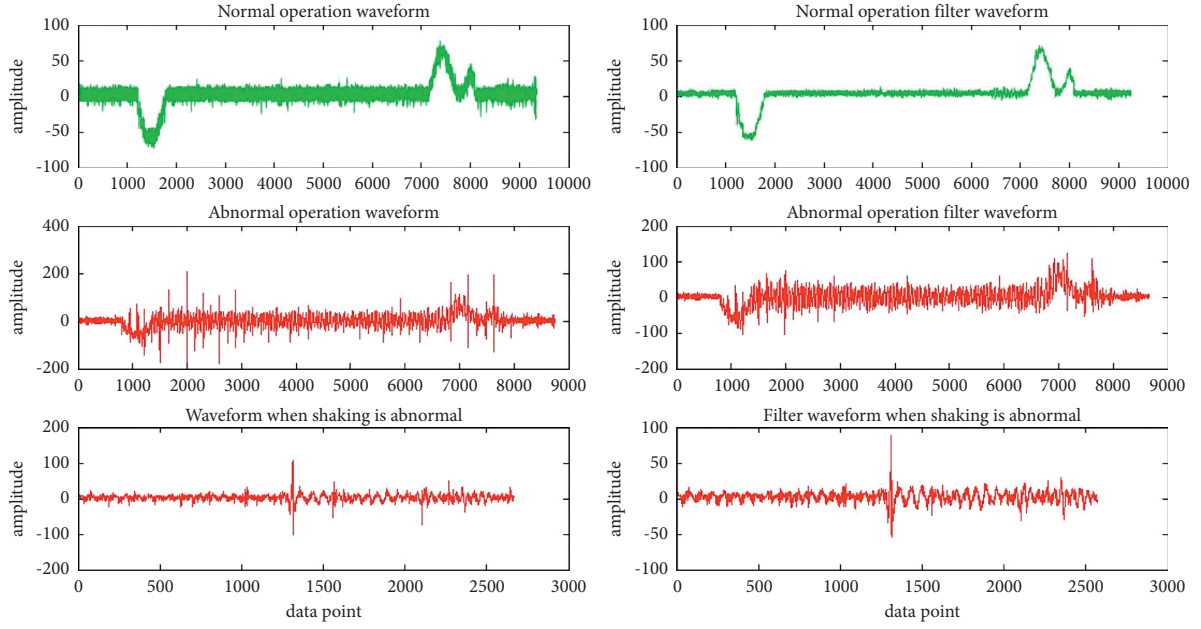


FIGURE 1: Original waveform and filtered waveform.

signal energy with the same frequency band will be greatly different. Wavelet packet can decompose different components of the signal into independent frequency bands without omission and redundancy. The energy of signals in these subbands provides important information for elevator operation condition monitoring and fault diagnosis. Because the energy of each frequency of elevator car vibration contains rich fault information, if one or more frequency energy changes, it means that a fault occurs. Therefore, compared with the traditional spectrum analysis, wavelet packet is suitable for energy detection according to frequency band. Taking the vibration signal $W(t)$ of an elevator car as an example, the specific steps are as follows:

- (1) *Decomposition of Vibration Signal.* The collected elevator car vibration signal $W(t)$ is decomposed by

$$S = AAA3 + DA A3 + A DA 3 + D D A3 + AA D3 + DA D3 + A DD 3 + D D D3. \quad (2)$$

- (3) *Calculation of the Energy of Each Subband.* Let E_j be the total energy of each subband signal S_j ; then,

$$\begin{aligned} E_j &= \int |S_j(t)|^2 dt \\ &= \sum |x_{jk}|. \end{aligned} \quad (3)$$

- (4) The feature vector is constructed with energy as the element. When the elevator car system fails, the signal energy in each frequency band will change greatly:

$$R = [E'_0, E'_1, \dots, E'_i, \dots, E'_j], \quad (4)$$

wavelet packet in time domain and frequency domain, and the characteristic frequency bands in each range are extracted. Using a three-layer wavelet packet decomposition, the decomposition structure is shown in Figure 2, where S is the original signal, A is the high-frequency signal, and D is the low-frequency signal.

- (2) *Wavelet Packet Decomposition Coefficient Reconstruction.* The decomposition coefficients of each layer after wavelet packet decomposition are reconstructed, and each subband signal is extracted. The reconstructed signal is represented by $AAA3$, and the total signal reconstruction is as follows:

where E_i represents the feature vector with prominent energy changes in each subband energy E_j .

- (5) The normal signal is compared with the fault signal to determine the frequency band position of the fault signal, so as to diagnose the possible fault.

By extracting the signal energy of each frequency band of characteristic quantity and comparing with the normal signal of corresponding frequency band, the characteristic difference value of different frequency band can be obtained. Combining with the characteristics of different parts of elevator, the factors related to elevator car vibration can be analyzed. Elevator car vibration is often not only a single fault, but also a combination of multiple faults.

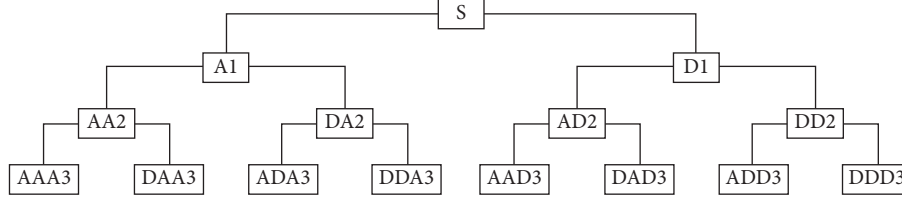


FIGURE 2: Wavelet packet decomposition diagram.

The above-mentioned is the energy calculation method of wavelet packet decomposition. The $x'(t)$ after SVD denoising is taken as the input signal of wavelet analysis. Through wavelet analysis, the energy of each subband is obtained, and an array $R = [E'_0, E'_1, \dots, E'_p, \dots, E'_j]$ is formed. The energy array will form the dataset of the later 1-CNN.

2.3. PCA Dimensionality Reduction Processing

2.3.1. Normalization. The core of PCA dimensionality reduction is orthogonal decomposition. The purpose is to select new mutually orthogonal spatial basis vectors to express the original data. Based on a few key principal component vectors, data dimensionality reduction is achieved by reconstructing the original data. PCA dimensionality reduction can be understood as looking for one or more vectors, obtaining a new vector space, and then mapping the samples that need dimensionality reduction to the new vector space.

2.3.2. Calculating the Dimensionality Reduction Matrix. First, calculate the covariance matrix Σ of the sample features, as in formula (5). In the actual calculation, it can be considered that the covariance processing and analysis process of the wavelet energy matrix is actually the process of principal component analysis of the wavelet energy matrix [18]. Then, use singular value decomposition to find the eigenvalue λ_i of Σ and the corresponding orthogonalized unit eigenvector a_i . After dimensionality reduction, the dimensionality reduction matrix is obtained, and the sample can be mapped to a low-dimensional space through the dimensionality reduction matrix.

$$\Sigma = (S_{ij})_{p \times p}. \quad (5)$$

Among them, $S_{ij} = (1/n - 1) \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$, $i, j = 1, 2, \dots, p$, and x_{ki} and x_{kj} in this article represent the energy values of the rows and columns in the wavelet energy matrix, while \bar{x}_i and \bar{x}_j are the average values of the i -th row and j -th column in the wavelet energy matrix, respectively.

2.3.3. Contribution Rate. Calculate the contribution rate, as in formula (6). m is the number of eigenvectors in the selected dimension reduction matrix U . The larger the m is, the more the feature vectors in U are and the smaller the dimensionality reduction error is, that is, the more the original features are preserved. The value of m can be determined by the cumulative contribution rate number $G(m)$, as in formula (7). When the

cumulative contribution rate is greater than 85%, it can sufficiently reflect the information of the original variable, and the corresponding m is the first m extracted principal ingredient.

$$\alpha_i = \frac{\lambda_i}{\sum_{i=1}^m \lambda_i}, \quad (6)$$

$$G(m) = \frac{\sum_{i=1}^m \lambda_i}{\sum_{j=1}^p \lambda_j}. \quad (7)$$

3. Design of Elevator Operation State Monitoring Model Based on 1-CNN

3.1. 1-CNN Model Design

3.1.1. Convolution Layer. The core structure of convolution neural network is convolution layer, and the weight parameter matrix in convolution layer is also called convolution kernel. The function of convolution kernel is equivalent to different filter functions, which perform different filtering operations on the input signal in order to extract the characteristics of the input signal. There are many kinds of convolution kernels, such as 1, 3, 5, and 7. This paper selects 64 convolution kernels and 3 convolution kernels. Different sizes of convolution kernels can connect different numbers of neurons; that is to say, different sizes of visual field can be sensed from input signals, and different features can be extracted.

3.1.2. Pooling Layer. Pooling layer is actually a form of downsampling. The role of pooling layer is to retain the main features while reducing the amount of calculation of network parameters, which can improve the generalization ability of the model for different input data and prevent the overfitting phenomenon. Different from the convolution layer, the pool layer does not need parameters and is easier to optimize. The pooling layer is usually used with the convolution layer; that is, the output of the convolution layer is used as the input of the pooling layer. There are many different forms of nonlinear pooling functions. There are two common pooling layers: maximum pooling and mean pooling. The maximum pooling is used in this article.

3.1.3. Activation Function. The basic requirement of activation function is nonlinear. Adding activation function to convolutional neural network can increase the non-linearity of the network and increase the expression ability of the network, so as to better fit the target. Common

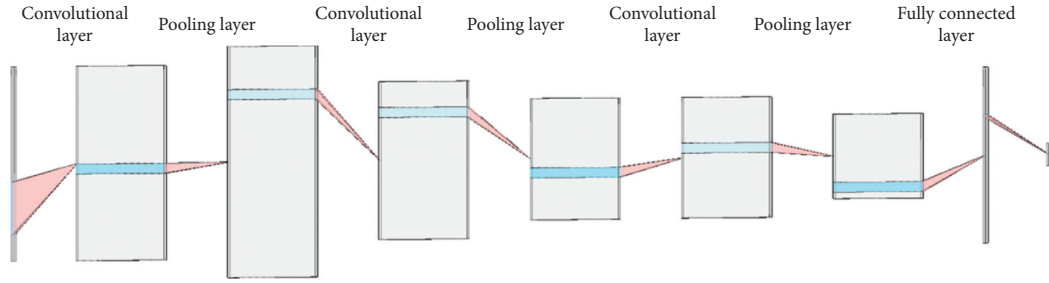


FIGURE 3: One-dimensional convolution neural network model.

TABLE 1: Design parameters of model structure.

Parameter name	Parameter value
Convolution level	3
Pool layer number	3
Convolution kernel size	64 * 1, 3 * 1
Pool size	16, 2
Convolution kernel layer number of convolution kernels	16, 32, 64
Input dimension	370 * 1, 256 * 1

TABLE 2: Data sets.

	Elevator 1		Elevator 2	
	Dataset I	Dataset II	Dataset III	Dataset IV
Number of groups	172	172	272	272
Data points per group	370	256	370	256

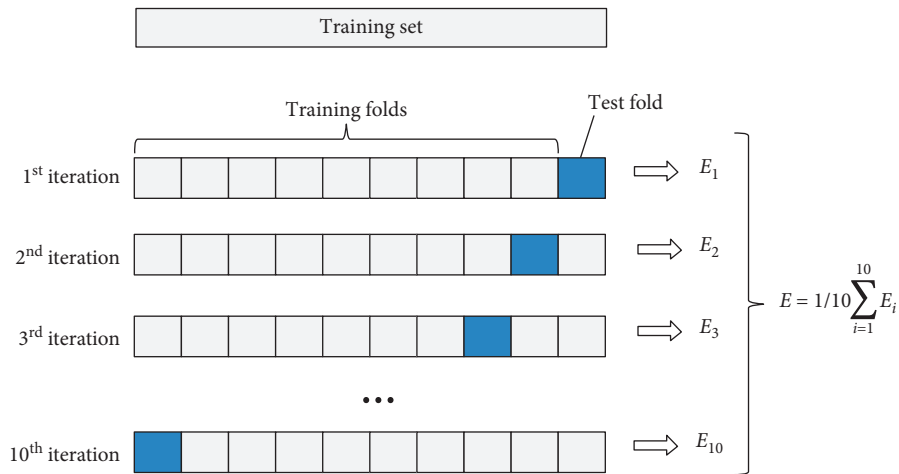


FIGURE 4: K-fold cross-validation process.

activation functions include sigmoid function, tanh function, ReLU function, and SReLU function. It is found that the tanh function is more suitable for the case that the original signal is vibration signal, and the tanh function is selected as the activation function in this paper.

3.1.4. Full Connection Layer. After passing through multiple groups of convolution layers and pooling layers, a fully connected layer will be connected. It can be applied to different classification models. The full connection layer is followed by a hidden layer, and the softmax regression layer completes the

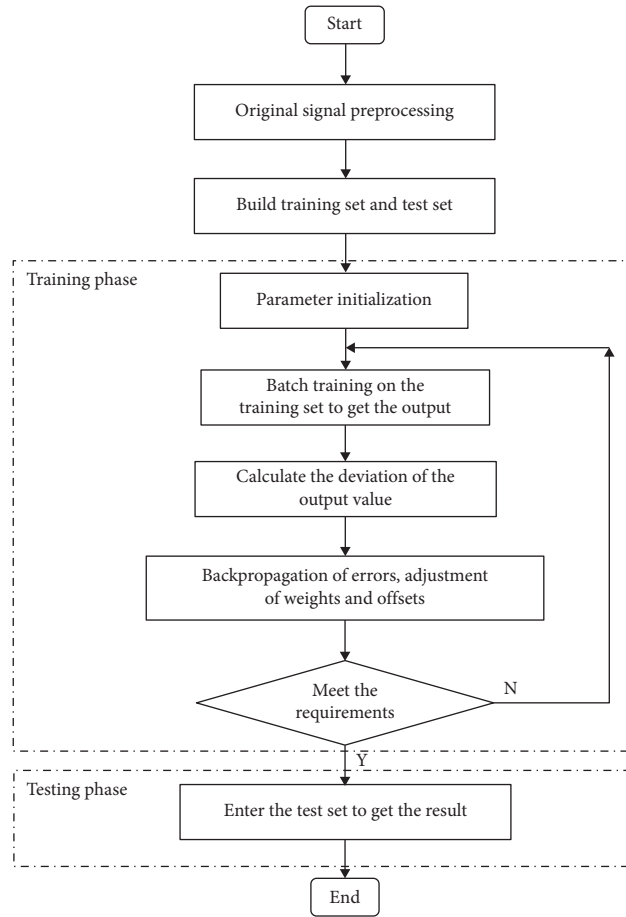


FIGURE 5: Modeling flowchart.

classification.

3.1.5. The Structure and Parameter Design of the Elevator Operation State Monitoring Model. Because the dimensions of the vibration signal and the picture are different, the one-dimensional signal is best solved with a one-dimensional model. Therefore, a one-dimensional convolution neural network model as shown in Figure 3 is established in this paper. The model consists of input layer, feature extraction layer (composed of multiple convolution layers and pooling layers alternately), full connection layer, and output layer. In this pool, the number of convolution layers and the number of convolution layers in the pool can be reduced. The first set of convolution kernels is $64 * 1$ in order to enlarge the receptive field of the input signal and avoid the input signal whose value is limited to a small range. The size of the last two convolution kernels is $3 * 1$, because $3 * 1$ is the most frequently used convolution kernel. The specific design parameters of the network structure are shown in Table 1.

3.2. Construction of Input Dataset of 1D Convolutional Neural Network. The data used in this article comes from two traction elevators with a long service life and intermittent abnormal operation in the school complex, a total of 15 floors. It mainly collected the vibration signal in the car and collected it during

the time period when many people took the elevator and no one did. This work lasted for a month. The original vibration data of the two elevators are segmented, and noise reduction operation is carried out. The size of each segment is 370, and the datasets I and III are obtained. Each group of data in datasets I and III is decomposed by wavelet transform “db3” in eight levels. Two hundred and fifty-six components are obtained in the eighth layer, and the energy value of the components is calculated to obtain datasets II and IV, as shown in Table 2.

3.3. K-Fold Cross Validation. In order to improve the accuracy of training results and make the results more convincing, the K-fold cross-validation method is introduced.

The steps of this method are as follows:

- (1) First, all samples are divided into k subsets of equal size, as shown in Figure 4.
- (2) The K -subset is traversed in turn. Each time, the current subset is used as the verification set, and all other samples are used as the training set to train and evaluate the model.
- (3) Finally, the average value of K times evaluation index is taken as the final evaluation index. In practical experiments, K is usually taken as 10. At this time, the training set becomes $k * D$ (d represents the number of data samples contained in each copy).

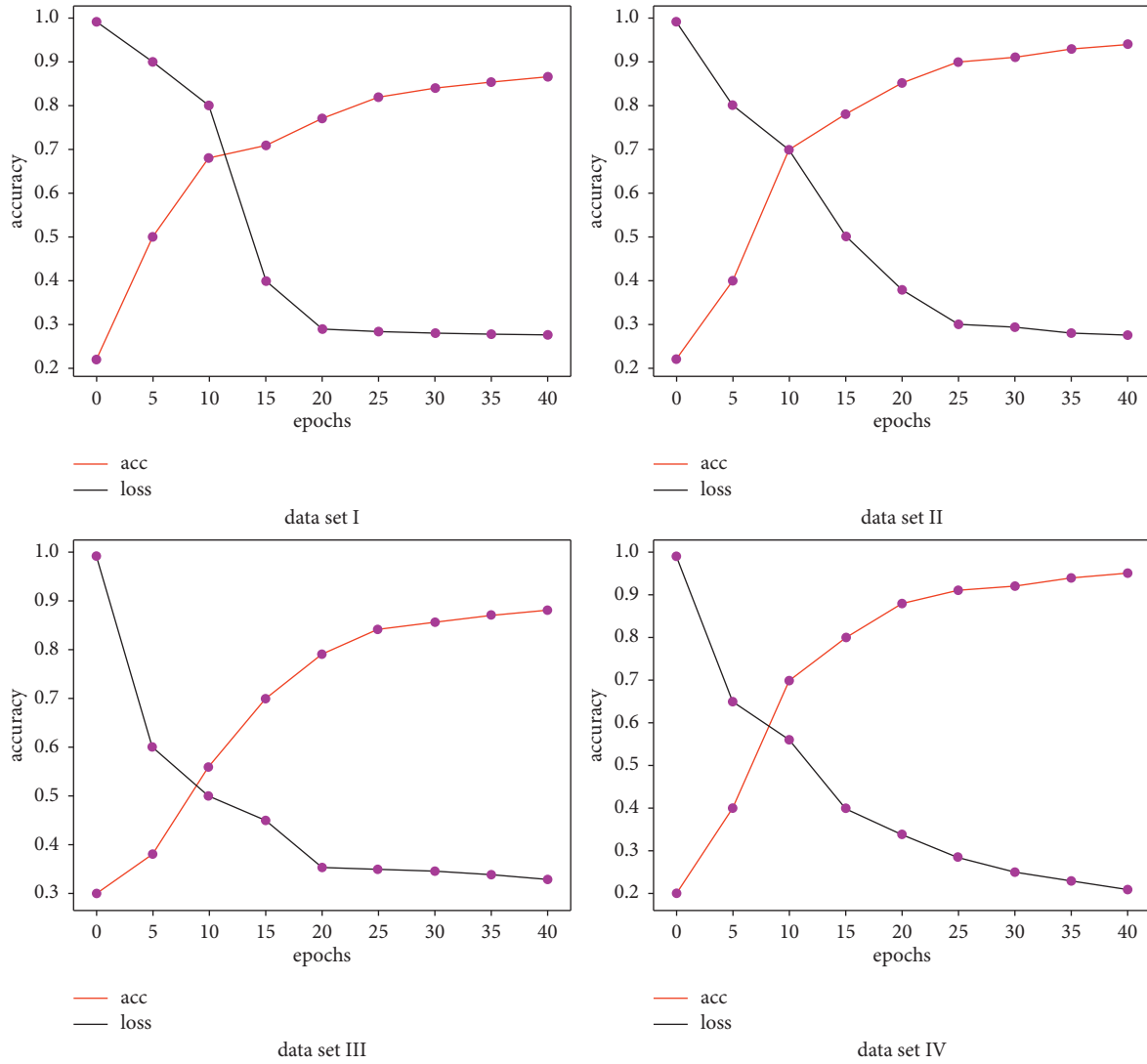


FIGURE 6: Variation curve of accuracy and loss function.

TABLE 3: Training results.

	Training set		Test set	
	Classification accuracy (%)	Loss function value	Classification accuracy (%)	Loss function value
Dataset I	88.15	0.0256	86.45	0.1063
Dataset II	96.26	0.0230	94.22	0.0956
Dataset III	90.12	0.0322	88.15	0.1131
Dataset IV	97.65	0.0206	95.20	0.1041

4. Experiment

4.1. Comparison Experiment before and after Wavelet Energy Feature Extraction. The process of establishing one-dimensional neural network model is shown in Figure 5. According to the design parameters of the model structure in Table 1, the network structure is built. The training samples in the training set are input into the network to calculate the output, and the loss function will calculate the difference between the real value and the output value. Taking the minimum loss function as the objective

function, the backpropagation algorithm can calculate the gradient of all parameters in the network and use the set optimization algorithm Adam to update the weight and bias, and the learning rate is set to 0.001 until the objective function value (loss function value) meets the requirements or the number of training rounds reaches the preset value. When testing, only the test set is input into the network, and the result obtained is the final monitoring result.

The datasets I, II, III, and IV are trained and tested by the constructed one-dimensional convolution neural

TABLE 4: Wavelet energy value.

No.	1	2	3	...	510	511	512
1	34.25	1.63	0.13	...	0.03	0.02	0.03
2	54.17	2.77	0.53	...	0.004	0.01	0.003
3	61.12	3.16	0.79	...	0.01	0.002	0.01
4	64.96	3.35	1.19	...	0.03	0.01	0.01
5	49.21	2.91	1.50	...	0.01	0.02	0.01
...
268	61.07	2.20	3.46	...	0.01	0.001	0.002
269	58.89	1.75	2.52	...	0.01	0.003	0.004
270	19.64	5.10	2.28	...	0.11	0.01	0.02
271	82.77	3.74	1.17	...	0.001	0.005	0.001
272	79.15	4.52	0.59	...	0.002	0.02	0.002

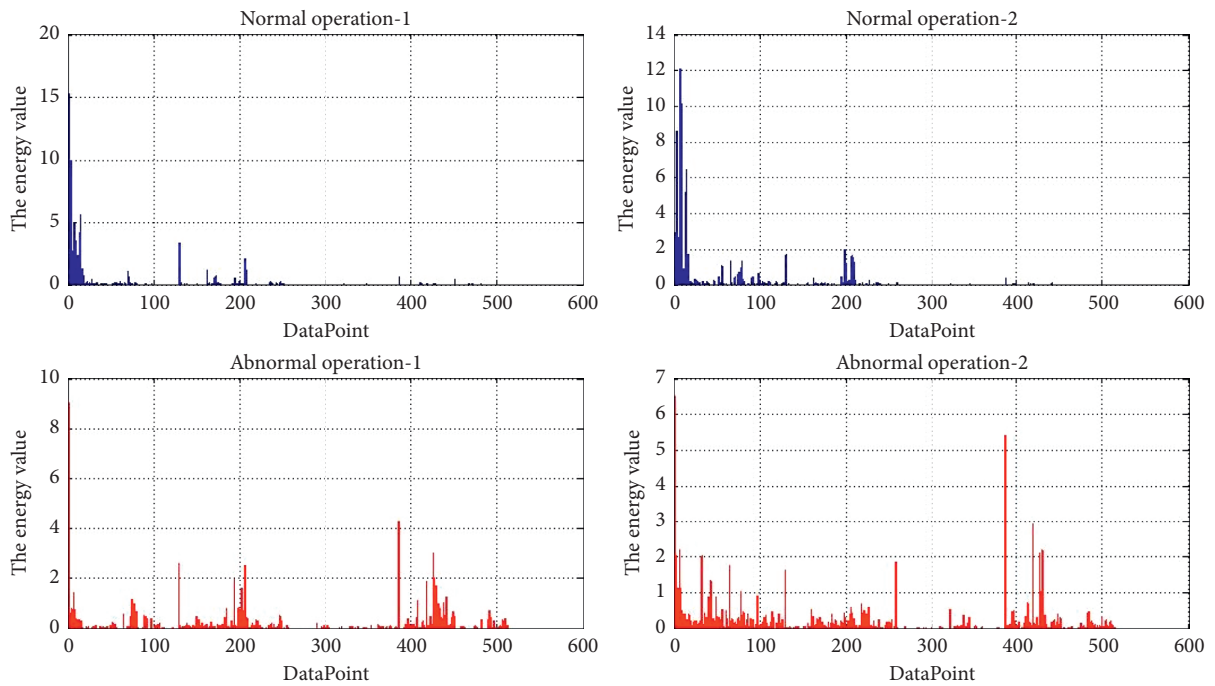


FIGURE 7: Comparison of wavelet energy during normal operation and abnormal operation.

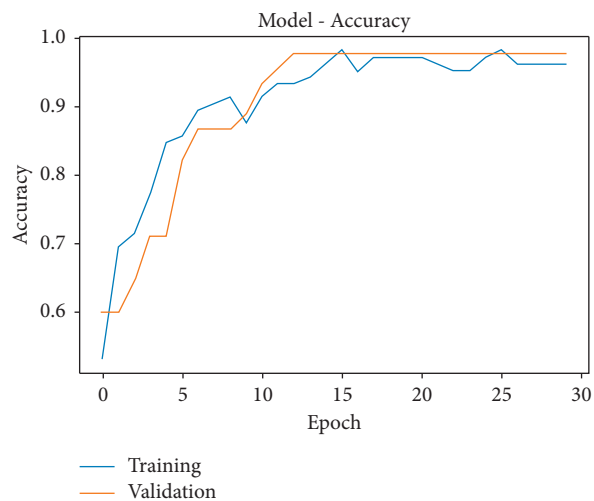


FIGURE 8: Before PCA dimensionality reduction.

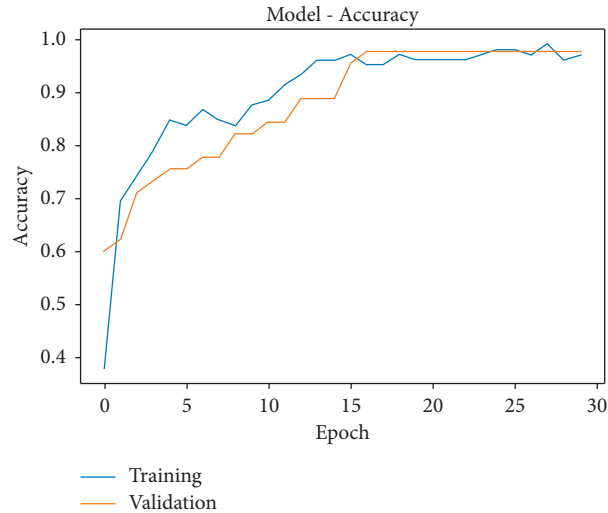


FIGURE 9: After PCA dimensionality reduction.

TABLE 5: Comparison of training results before and after PCA dimensionality reduction.

	Training set		Test set	
	Classification accuracy (%)	Loss function value	Classification accuracy (%)	Loss function value
Before PCA dimensionality reduction	90.12	0.0322	88.15	0.1131
After PCA dimensionality reduction	96.57	0.0295	96.11	0.0341

network, and the training results are shown in Figure 6. Datasets I and III are original vibration signals, and the accuracy rate is about 88%. Datasets II and IV are datasets composed of original vibration signals and energy features extracted by wavelet transform. The accuracy rate is significantly improved to 95%. The training results are shown in Table 3.

4.2. Comparison Experiment before and after PCA Dimensionality Reduction Optimization. Since there are 512 feature components in each data sample in dataset III, the number of features is a bit large. PCA dimensionality reduction can effectively reduce the complexity of the data and identify the most important features, so you can first use the 512-dimensional feature sample. Perform PCA dimensionality reduction processing to reduce the complexity of data samples.

First, the 272 groups of original vibration signal data samples in dataset III are, respectively, decomposed in nine layers using wavelet transform “db3.” and 512 components are obtained in the ninth layer, and the energy values of the components are calculated. The energy values are as shown in Table 4 and normal. The comparison of wavelet energy during operation and abnormal operation is shown in Figure 7.

Next, perform PCA dimensionality reduction on the 512-dimensional wavelet energy feature data to obtain a 272×128 two-dimensional matrix. The results are as follows.

Finally, the 272×128 two-dimensional matrix is used as the input of the 1-CNN for training, and the test is compared

with the results without PCA dimensionality reduction. As shown in Figures 8 and 9, the accuracy of PCA before dimensionality reduction is about 88%, the accuracy of PCA after dimensionality reduction can reach 96%. The training results are shown in Table 5.

5. Conclusion

This paper proposes an elevator operation fault monitoring method based on 1-CNN. Before training the model, the original vibration is decomposed by wavelet packet, and the energy features are extracted, and the energy is used as the input of the network. Comparing the original signal as the input and the energy as the input, the accuracy is significantly improved. According to the time-domain vibration signals collected under different monitoring and working conditions, the method proposed in this paper is used for verification, and the accuracy rate is improved. In addition, this paper also compares the training results before and after PCA dimensionality reduction. The results show that the extracted wavelet energy features are first subjected to PCA dimensionality reduction and then recognized, which can reduce the complexity of data samples, optimize input features, and further improve the recognition rate. This method also has some notable and improved problems. For example, the data used in the test is not comprehensive enough, and the selection of various layer parameters during network construction requires multiple attempts to adjust to obtain better results.

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 that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (no. 2019YFF0302203).

References

- [1] Y. Lin, Y. Wu, T. Wang, and W. Su, "Research on elevator health real-time monitoring system based on information terminal," *Electronic measurement technology*, vol. 39, no. 8, pp. 159–162, 2016.
- [2] H. Zhao, Y. Bu, and X. He, "On line condition monitoring and fault diagnosis system for rotating machinery," *Electromechanical engineering*, vol. 23, no. 8, pp. 15–17, 2006.
- [3] J. Xu, "Research on risk identification and improvement of radar electromechanical control system based on PFMEA," *Electromechanical engineering*, vol. 35, no. 5, pp. 512–516, 2018.
- [4] J. Xu, L. Xu, H. Wang, B. Zheng, and B. Tang, "Research on elevator condition monitoring technology based on vibration analysis," *Electromechanical engineering*, vol. 36, no. 3, pp. 279–283, 2019.
- [5] C. Zhang and G. Wang, "Abnormal analysis of elevator traction machine bearing based on fast Fourier transform," *China elevator*, vol. 28, no. 7, pp. 7–9, 2017.
- [6] H. Li, X. Wang, L. Chen, and E. Li, "Denoising and R-peak detection of electrocardiogram signal based on EMD and improved approximate e," *Circuits, Systems, and Signal Processing*, vol. 33, no. 4, pp. 1261–1276, 2014.
- [7] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.
- [8] K. P. Murphy, "Machine learning: a probabilistic perspective," *Chance*, vol. 27, no. 2, pp. 62–63, 2012.
- [9] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [10] C. Wen and F. Lv, "A review of fault diagnosis methods based on deep learning," *Acta Electronica Sinica*, vol. 42, no. 1, pp. 234–248, 2020.
- [11] C. Shi, G. Panoutsos, B. Luo, H. Liu, B. Li, and X. Lin, "Using multiple-feature-spaces-based deep learning for tool condition monitoring in um," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 5, pp. 3794–3803, 2019.
- [12] Z. Chen, W. Jiao, H. Sun, and Y. Li, "Fault diagnosis method based on improved deep confidence network," *Journal of Detection and Control*, vol. 42, no. 4, pp. 98–105, 2020.
- [13] O. Janssens, V. Slavkovikj, B. Vervisch et al., "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331–345, 2016.
- [14] S. Shaowu, Z. Sheng, J. Wanlu, and L. Zhenbao, "Study on the health condition monitoring method of hydraulic pump based on convolutional neural network," in *Proceedings of the 2020 12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pp. 149–153, Phuket, Thailand, February 2020.
- [15] Z. Xia, "Bearing fault detection method of Brushless DC motor based on DWT and RNN ej3," *Micro special motor*, vol. 45, no. 6, p. 17, 2017.
- [16] Q. Cui, Z. Li, J. Yang, and B. Liang, "Rolling bearing fault prognosis using recurrent neural network," in *Proceedings of the 2017 29th Chinese Control And Decision Conference (CCDC) Melbourne*, pp. 1196–1201, IEEE, Chongqing, China, 28–30 May 2017.
- [17] X. Zhao, B. ye, and Y. Lin, "Extraction of amplitude modulation feature information from bearing vibration signal by singular value decomposition," *Journal of Beijing University of Technology*, vol. 31, no. 5, pp. 572–577, 2011.
- [18] R. Sharifi and R. Langari, "Nonlinear sensor fault diagnosis using mixture of probabilistic PCA models," *Mechanical Systems and Signal Processing*, vol. 85, pp. 638–650, 2017.