

# Intelligent fault diagnosis for elevator door system based on discriminative multi-scale residual network

Journal of  
Intelligent  
Manufacturing  
and Special  
Equipment

175

Jiefeng Li

Key Laboratory of High-Parameter Elevator Intelligent Operation and Maintenance,  
State Administration for Market Regulation,  
Special Equipment Safety Supervision Inspection Institute, Nanjing, China

He Ren and Qiuyu Song

School of Mechanical Engineering and Rail Transit, Changzhou University,  
Changzhou, China, and

Liang Ye and Gongning Li

Key Laboratory of Intelligent Operation and Maintenance of High-parameter Lifts,  
State Administration for Market Regulation, Suzhou, China and  
Special Equipment Safety Supervision Inspection Institute of Jiangsu Province,  
Nanjing, China

Received 12 November 2025

Revised 29 January 2026

Accepted 22 February 2026

## Abstract

**Purpose** – This study addresses the challenge of extracting effective state features for intelligent fault diagnosis in elevator door systems, aiming to enhance model representational capability for reliable decision-making and accident prevention in real-world engineering.

**Design/methodology/approach** – The proposed discriminative multi-scale residual network (DMSRN) uses a one-dimensional multi-scale architecture to capture rich and hierarchical features, combined with a novel feature discriminative enhancement strategy that enhances intra-class compactness and inter-class separability, thereby improving feature discriminability.

**Findings** – Comprehensive experimental results validate the superiority of the proposed approach. The DMSRN achieves an average diagnostic accuracy that exceeds the best baseline method by more than 18.16%, demonstrating a substantial competitive advantage in fault diagnosis for elevator door systems.

**Originality/value** – The key contribution of this research lies in the novel integration of multi-scale feature learning with a specialized feature discriminative enhancement mechanism. This combination offers significant practical value by improving diagnostic reliability in safety-critical elevator applications.

**Keywords** Elevator door system, Intelligent fault diagnosis, Multi-scale residual network, Metric learning

**Paper type** Research article

## 1. Introduction

The proliferation of high-rise structures in contemporary urban settings has established elevators as a critical component of vertical transit systems (Ren *et al.*, 2024). Nevertheless, failures in these systems, including irregular door movement, passenger entrapment,

© Jiefeng Li, He Ren, Qiuyu Song, Liang Ye and Gongning Li. Published in *Journal of Intelligent Manufacturing and Special Equipment*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at [Link to the terms of the CC BY 4.0 licence](#).

**Funding:** This work was supported in part by the National Key Research and Development Program of China (award number: 2023YFC3010400), and the Open Project of the Key Laboratory of Intelligent Operation and Maintenance of High-parameter Lifts, State Administration for Market Regulation (award number: JSTJ-IOMHL-202504).



Journal of Intelligent Manufacturing and  
Special Equipment  
Vol. 7 No. 2, 2026  
pp. 175-189

Emerald Publishing Limited

e-ISSN: 2633-660X

p-ISSN: 2633-6596

DOI 10.1108/JIMSE-11-2025-0027

operational noises and sudden impacts, may lead to serious outcomes, compromising both passenger safety and operational integrity. Statistics indicate that door system faults account for over 80% of elevator-related accidents (Wang *et al.*, 2018), highlighting the pressing demand for reliable and precise fault diagnosis techniques to facilitate timely maintenance of elevator door mechanisms.

Traditional fault diagnosis approaches, such as rule-based expert systems, rely heavily on domain knowledge, requiring experts to meticulously define fault signatures and develop corresponding diagnostic procedures. Moreover, periodic shutdown maintenance often entails substantial time, labor and financial resources. In recent years, data-driven diagnostic techniques have gained increasing attention in the field of prognostics and health management for their potential to enhance asset reliability (Lu *et al.*, 2025a). For instance, Bai *et al.* (2021) developed a fault prediction model based on an optimized particle swarm optimization-back propagation algorithm. Chen *et al.* (2019) provided a comprehensive review of various artificial intelligence techniques applied in elevator fault diagnosis, including back propagation neural networks, radial basis function networks, K-means clustering and support vector machines (SVMs). A key limitation of these approaches is their inadequate capability in learning discriminative features, for which deep learning presents a promising alternative.

In contrast to traditional machine learning methods, deep learning-based diagnosis has seen rapid development due to its capability to autonomously extract hierarchical features in a cost-effective manner. For example, Kim *et al.* (2023) utilized a variational autoencoder with a restricted latent space and Bayesian optimization to construct a margin-maximized hyperspace for accurate elevator door fault classification. Similarly, Wang *et al.* (2021) implemented a distributed deep learning model that incorporates geolocation information by processing features through hidden parameters of the network. Chen *et al.* (2025) introduces a principal component analysis-long short-term memory method for fault diagnosis of an elevator. Pan *et al.* (2024) conducted research on fault prediction method of elevator door system based on techniques including graph neural networks, long-short-term memory networks and so on. Wan *et al.* (2025) combined variational mode decomposition and multi-scale convolutional neural network (CNN) and applied them to intelligent fault diagnosis of the center-opening elevator door system. Lu *et al.* (2025b) proposed a multi-scale convolution capsule network with data augmentation and attention mechanisms and realized positive elevator fault diagnosis.

Nevertheless, traditional convolutional neural networks often exhibit limited feature representation capacity, owing to their single-scale receptive fields and susceptibility to gradient vanishing, which can compromise the reliability of intelligent diagnostic systems. Moreover, the discriminative feature extraction for elevator door systems remains an understudied area. To address these limitations, this article proposes a discriminative multi-scale residual network (DMSRN), which integrates a one-dimensional multi-scale residual architecture with metric learning. This design optimizes the structure of the feature space, enabling the extracted deep features to be not only separable but also semantically discriminative. As a result, the proposed model enhances the reliability of intelligent fault diagnosis. The main contributions of this work are summarized as follows.

- (1) A feature extractor with one-dimensional multi-scale residual mechanism is designed to enhance the quality of the extracted features. By integrating the powerful concepts of multi-scale parallel processing and fast identity mapping, it addresses the core pain points of traditional CNN in terms of depth, multi-scale processing, and training stability.
- (2) A feature discriminative enhancement strategy that applies metric learning to optimize the feature space, explicitly improves semantic discriminability through enforced intra-class compactness and inter-class separation.

- (3) A novel and simple framework called DMSRN is proposed for automatically and accurately diagnosing faults of elevator door systems. Experimental findings, including accuracy, confusion matrices and feature visualization obtained from the laboratory datasets, have demonstrated the superior diagnosis performance of the DMSRN method.

The remainder of this article is structured as follows. [Section 2](#) details the proposed DMSRN methodology. [Section 3](#) presents and discusses the experimental setup and results. Finally, [Section 4](#) concludes the study and [Section 5](#) outlines potential directions for future research.

## 2. Related work

### 2.1 Residual networks

Residual networks are widely applied in various fields, including rolling element bearing fault diagnosis ([Ni et al., 2023](#)), bird song recognition ([Hu et al., 2023](#)), classification of power quality disturbances ([Pan et al., 2025](#)) and so on. The fundamental advantage of residual networks over conventional networks lies in their ability to overcome the degradation problem, where increasing the depth of a neural network leads to higher training error rather than better performance. By introducing skip connections, residual networks allow the input to bypass one or more layers, effectively teaching the network to learn a residual function rather than the full underlying mapping. This architecture mitigates the vanishing gradient problem by providing a gradient highway that enables smoother backpropagation to earlier layers. Consequently, Residual networks can be scaled to hundreds or even thousands of layers, achieving significantly higher accuracy and faster convergence while remaining easier to optimize than their single-path counterparts.

### 2.2 Multi-scale convolutional architectures

Compared to single-scale designs, the key advantage of multi-scale convolutional architectures lies in their superior ability to handle scale variation and capture richer hierarchical features. While single-scale models employ a fixed kernel size and thus a static receptive field at each layer, multi-scale architectures process input data through multiple parallel pathways or hierarchical levels with varying spatial resolutions and kernel sizes. This enables the network to simultaneously extract fine-grained local details like edges and textures, as well as broad global context, including object shapes and spatial relationships. By integrating these complementary perspectives, multi-scale models become more robust in detecting objects of vastly different sizes, ranging from small distant instances to large foreground objects. Moreover, this approach often enhances computational efficiency; rather than simply increasing network depth to enlarge the receptive field, multi-scale designs decompose convolutions into factored operations, achieving stronger representational capacity with fewer parameters. As a result, they consistently deliver improved accuracy in challenging tasks such as medical image segmentation ([Yin et al., 2023](#)) and Electroencephalogram-based motor imagery decoding ([Liu et al., 2022](#)).

### 2.3 Metric-learning

Metric learning is widely utilized in tasks such as facial recognition and image retrieval, as it enables models to learn an effective embedding space where semantically similar samples are mapped to adjacent positions ([Li et al., 2023](#)). Depending on the learning paradigm, current mainstream metric learning methods are generally categorized into comparison-based ([Wu et al., 2025](#)), proxy-based ([Chan et al., 2023](#)), and center-based losses ([Dong and Lam, 2024](#)). Among these, center-based loss is increasingly favored for its superior optimization stability and rapid convergence. Specifically, center loss functions by simultaneously learning a centroid for each class and penalizing the distances between individual samples and their corresponding centers. While the traditional Softmax loss ensures inter-class separability, it

does not explicitly encourage intra-class compactness. Center loss addresses this by directly minimizing intra-class variation, thereby producing tighter feature clusters. Furthermore, unlike triplet loss (Zhou et al., 2024), which requires computationally expensive hard-sample mining, center loss is highly efficient as it only involves updating class centers. Consequently, center loss is typically employed as a joint supervision signal with Softmax, i.e. the latter ensures global separability while the former enhances local compactness. This dual-objective approach results in more stable training and faster convergence compared to purely pair-based methods.

### 3. Methodology

This section presents the proposed framework for enhancing the accuracy and reliability of intelligent fault diagnosis in elevator door systems. The framework integrates a one-dimensional MSRN with a metric learning-based feature discriminative enhancement strategy, enabling effective extraction of discriminative state features and accurate fault identification.

#### 3.1 Problem definition

Some notations and definitions should be firstly described. There are  $n_{train}$  labeled samples  $\mathbf{x}$  in training dataset  $D_{train} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n_{train}}$  and  $n_{test}$  unlabeled samples in test dataset  $D_{test} = \{\mathbf{x}_i\}_{i=1}^{n_{test}}$ , where  $y$  denotes the labels of  $M$  health states of elevator door systems. The goal is to learn powerful fault diagnosis models using labeled fault data. That is to say, the category prediction function  $f_s(\cdot)$  established on the basis of the training dataset can be well applied to prediction tasks on test datasets that have not participated in model training.

#### 3.2 Overview of the proposed DMSRN method

Figure 1 displays the overview of the proposed DMSRN. Its framework includes a feature extractor  $G$  based on MSRN and a classifier  $C$  based on fully connected layers. By integrating the multi-scale one-dimensional residual architecture with metric learning, DMSRN forms an effective fault diagnosis framework for elevator door systems.

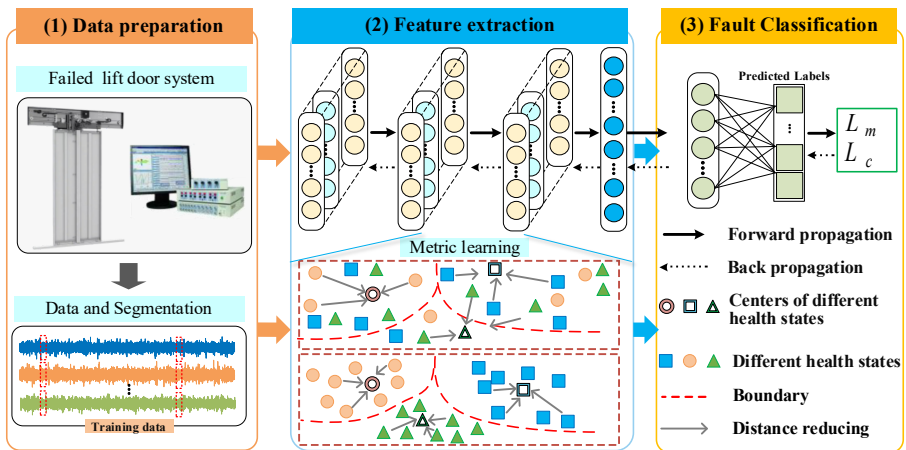


Figure 1. Structure of the proposed DMSRN method. Source(s): Created by authors

### 3.3 One-dimensional MSRN

The one-dimensional MSRN represents an advanced architecture tailored for processing sequential one-dimensional data, such as time-series signals, through the incorporation of residual learning across multiple scales. Its fundamental concept aims to mitigate the vanishing gradient problem while improving feature propagation via residual blocks equipped with skip connections, thereby facilitating more stable training and enhanced hierarchical feature extraction.

The basic residual block in a one-dimensional residual network can be mathematically represented as Eq. (1)

$$r = g(z, \{W\}) + z, \quad (1)$$

where  $z$  represents the input to the residual block,  $g(z, \{W\})$  is the residual function (e.g. a series of one-dimensional convolutional layers with weights  $W$ ), and  $r$  is the output of the block. The skip connection (the addition of  $z$ ) allows the network to learn residual mappings, which are often easier to optimize than the original mappings.

In a multi-scale configuration, the network processes input data across varying scales or resolutions. This is accomplished through parallel convolutional layers with diverse kernel sizes or dilation rates, enabling the extraction of features at distinct temporal granularities. The outputs from these parallel pathways are subsequently fused, typically via concatenation or summation, to construct a comprehensive feature representation.

As illustrated in Figure 2, a typical building block of the one-dimensional MSRN comprises three parallel convolutional layers with kernel sizes of 3, 5 and 7. The outputs from these layers are integrated and processed through an activation function before being propagated to subsequent residual blocks or layers.

The overall architecture generally consists of multiple such residual blocks stacked in sequence, with potential downsampling operations interspersed between blocks to reduce temporal dimensionality and expand the receptive field. This hierarchical design facilitates effective learning of multi-level representations, capturing both localized patterns and broader contextual features within the input data.

### 3.4 Feature discriminative enhancement strategy

The model incorporates metric learning to improve feature discrimination by increasing separability between classes. To this end, the center loss function is utilized, which learns a unique class center for each category and constrains the distance between feature representations and their respective centers. This approach enhances feature representation by jointly enforcing intra-class compactness and inter-class separability, as illustrated in Figure 3.

The center loss consists of two processes, i.e. (1) Learning the class centers of deep features; (2) Using the learned class centers to constrain the features belonging to that class. For the training set  $\{(\mathbf{x}_i, y_i)\}_{i=1}^{n_{train}}$ , the center loss  $L_{center}$  is defined as:

$$L_{center} = \frac{1}{2} \sum_{i=1}^{n_{train}} \|\mathbf{x}_i - \mathbf{c}_{y_i}\|_2^2 \quad (2)$$

where  $\mathbf{c}_{y_i}$  is the  $y_i$ th class center of the features.

Since  $\mathbf{c}_{y_i}$  cannot be directly obtained, it is first randomly initialized and then gradually optimized during model training. The network parameters are updated in each small batch of sample training. Meanwhile, the gradient of  $\mathbf{c}_{y_i}$  is calculated to update the position of the center. The calculation process of the gradient of  $L_{center}$  with respect to  $\mathbf{x}_i$  and  $\mathbf{c}_{y_i}$  is as follows:

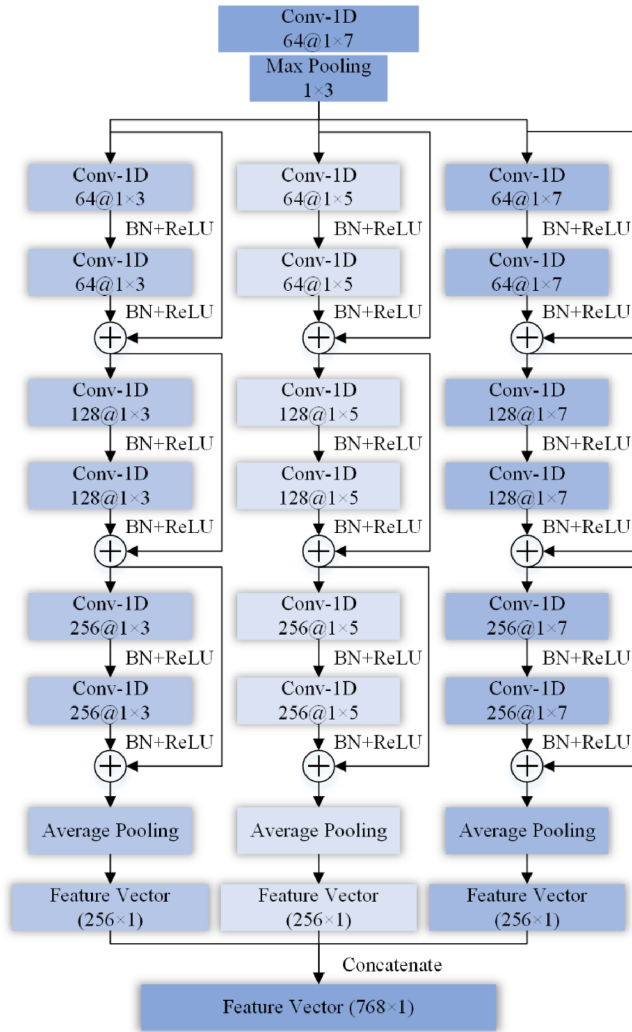
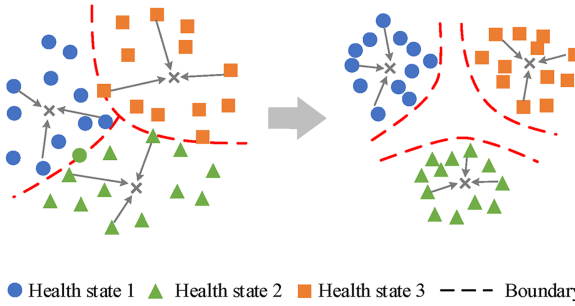


Figure 2. Details of the one-dimensional MSRN. Source(s): Created by authors

$$\frac{\partial L_{center}}{\partial \mathbf{x}_i} = \mathbf{x}_i - \mathbf{c}_{y_i}, \quad (3)$$

$$\Delta \mathbf{c}_j = \frac{\sum_{i=1}^m \delta(y_i = j) (\mathbf{c}_j - \mathbf{x}_i)}{1 + \sum_{i=1}^m \delta(y_i = j)}, \quad (4)$$

where, if  $y_i = j$ , then  $\delta(y_i = j) = 1$ ; otherwise,  $\delta(y_i = j) = 0$ . For the center  $\mathbf{c}_{y_i}$  of  $j$ , it is updated in the following way:



**Figure 3.** Schematic diagram of central loss. **Source(s):** Created by authors

$$\mathbf{c}_j^{t+1} = \mathbf{c}_j^t - \alpha \Delta \mathbf{c}_j^t, \quad (5)$$

where  $\mathbf{c}_j^t$  represents the feature center at time  $t$  and needs to be continuously updated during the model training process.  $\alpha$  is a scalar used to control the learning rate of the center to avoid disturbances caused by some mislabeled samples. In this article, the default value of  $\alpha$  is 0.5.

The center loss promotes the formation of tight clusters around class centers in the feature space, which significantly clarifies inter-class boundaries and enhances feature discriminability.

### 3.5 Objective function optimization

In conjunction with the center loss, the standard cross-entropy loss is employed to jointly train the feature extractor and the fault classifier for the classification task.

$$L_c = - \sum_{i=1}^{n_{train}} y_i \log C(G(\mathbf{x}_i)) \quad (6)$$

The overall training object of DMSRN can be summarized as follows:

$$L_{all} = L_c + \lambda L_{center} \quad (7)$$

where  $\lambda$  is the weighting coefficient of different losses. With Adam optimizer, the computation of parameters at each training epoch  $q$  is formulated in Eqs. (8) and (9).

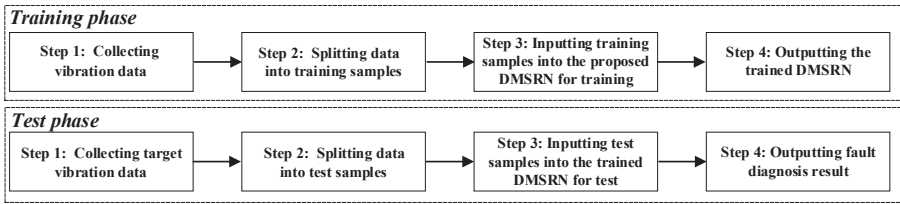
$$\theta_G^{(q)} \leftarrow \theta_G^{(q)} - \mu \left( \frac{\partial L_c}{\partial \theta_G^{(q)}} + \lambda \frac{\partial L_{center}}{\partial \theta_G^{(q)}} \right), \quad (8)$$

$$\theta_C^{(q)} \leftarrow \theta_C^{(q)} - \mu \frac{\partial L_c}{\partial \theta_C^{(q)}} \quad (9)$$

where  $\mu$  denotes the learning rate.

### 3.6 Diagnosis process using proposed DMSRN method

Figure 4 introduces the diagnosis process using proposed DMSRN method step by step.



**Figure 4.** Diagnosis process using proposed DMSRN method. **Source(s):** Created by authors

#### 4. Experimental verification

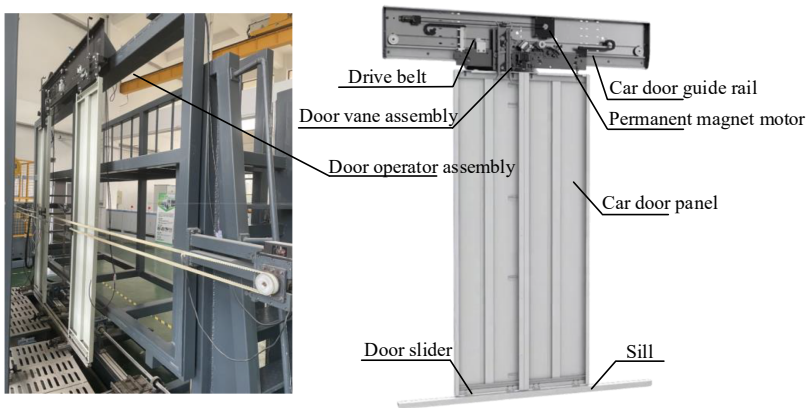
This section presents a comprehensive evaluation of the proposed DMSRN’s diagnostic performance and practical utility, conducted on a custom-built laboratory dataset for elevator door systems.

##### 4.1 Datasets description

Figure 5 illustrates the test bench used for evaluating the elevator door system in a laboratory setting. A vibration sensor was mounted on the inner surface near the outer edge of the car door to collect real-time operational data at a sampling frequency of 10 Hz. The system was subjected to four health conditions: normal operation (labeled as 0), missing one threshold slide block (labeled as 1), missing two threshold slide blocks (labeled as 2) and debris presence in the threshold (labeled as 3). From the collected data, samples with a length of 1,024 points were extracted, resulting in 720 samples per fault class.

##### 4.2 Implementation details

To evaluate the effectiveness and superiority of the proposed DMSRN, comparative experiments are conducted against one-dimensional CNN, SVM, MSRN and intra-domain adversarial network (IDAN) (Huang *et al.*, 2025). Implementation details of these comparison methods are introduced as follows. Specifically, the architecture of CNN used in this experiment is introduced in Table 1. The architecture of MSRN is the same as that of DMSRN presented in Figure 2. The fault classifier of CNN, MSRN and DMSRN are both composed of two fully connected layers. CNN, MSRN and DMSRN employ the same experimental settings. The Adam optimizer with betas of (0.9, 0.99) is adopted for model training, with the



**Figure 5.** Test bench of elevator door system. **Source(s):** Created by authors

**Table 1.** Architecture of CNN in this experiment

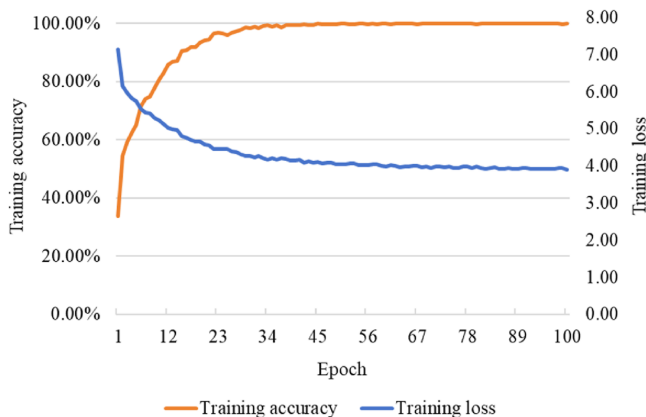
Layer	Operation and parameters
1-Conv	Kernel $8-15 \times 1$ , stride 2, BN, ReLU
2-Conv	Kernel $16-15 \times 1$ , stride 4, BN, ReLU
3-Conv	Kernel $32-15 \times 1$ , stride 4, BN, ReLU
4-Conv	Kernel $64-15 \times 1$ , stride 4, BN, ReLU
5-AveragePool	Kernel $256-8 \times 1$ , stride 1

learning rate  $\mu$  empirically set to 0.001. IDAN is a state-of-the-art fault diagnosis method which leverages multi-scale branches, an improved adversarial learning mechanism and a perspective sharing strategy to extract generalized fault representations. The hyperparameters and architectures of IDAN are adopted from those of the reference of [Huang et al. \(2025\)](#). To ensure the robustness and stability of the model, a sixfold cross-validation strategy was employed. In this approach, the total dataset was partitioned into six disjoint subsets. In each iteration, 5 folds were used for training the model, while the remaining fold served as the test set to evaluate performance. The training uses 100 epochs with a batch size of 10. Additionally, SVM employs a linear kernel, with the regularization coefficient varying across the set  $\{0.01, 0.1, 1, 10, 100, 1,000\}$ , and the results are recorded at each traversal.

#### 4.3 Results and analysis

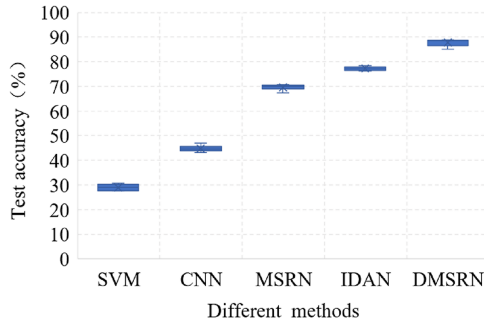
[Figure 6](#) provides the iterative curves of training accuracy and training loss of the proposed DMSRN in the whole training phase, the convergence of which shows the optimization of the model.

As for the test results based on sixfold cross-validation strategy, the fault diagnostic accuracies of the elevator door system using different methods are exhibited in [Table 2](#) and [Figure 7](#). The proposed DMSRN demonstrates a significant improvement of 58.85%, 43.06%, 18.16% and 10.60% in accuracy compared to CNN, SVM, MSRN and IDAN, respectively. For a more detailed and intuitive comparison, confusion matrices of different methods are presented in [Figure 8](#). Compared with the proposed DMSRN, CNN, SVM, MSRN and IDAN exhibit inferior performance in identifying health states of the elevator door system. The results verify the advantage of the proposed DMSRN over other comparison methods.

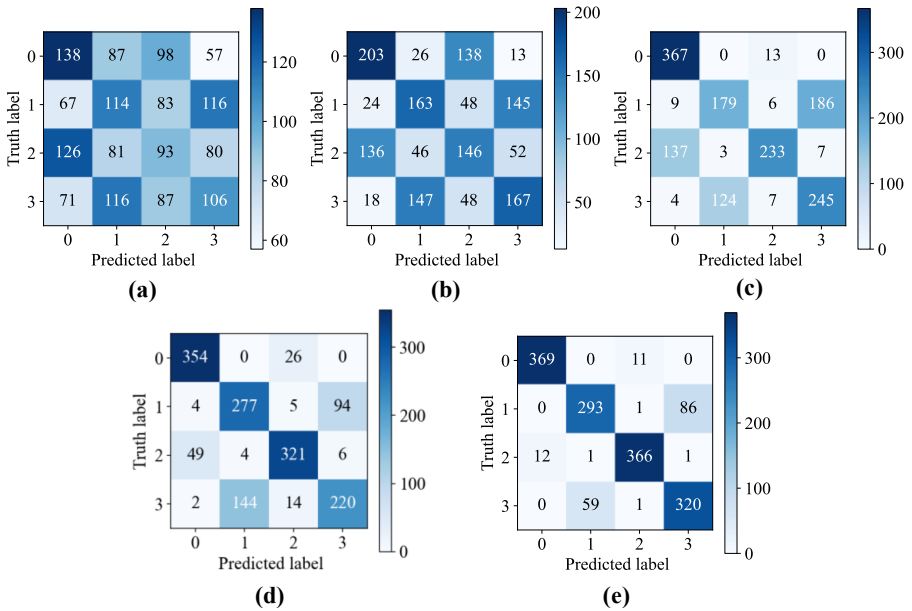
**Figure 6.** Iterative curves during training phase. **Source(s):** Created by authors

**Table 2.** Diagnostic accuracy of different methods

Methods	1st time	2nd time	3rd time	4th time	5th time	6th time	Average	Standard deviation
SVM	29.67	30.72	28.3	27.5	30.00	27.5	28.95	1.37
CNN	44.47	44.93	46.97	44.21	43.16	44.67	44.74	1.25
MSRN	69.54	70.26	69.67	70.79	70.2	67.37	69.64	1.20
IDAN	76.72	77.76	76.13	78.37	76.42	77.78	77.20	0.89
DMSRN	88.82	85.07	88.55	87.04	88.62	88.68	87.80	1.49



**Figure 7.** Box plot of diagnostic accuracy and standard deviation of different methods. **Source(s):** Created by authors



**Figure 8.** Confusion matrices of different methods: (a) SVM, (b) CNN, (c) MSRN, (d) IDAN and (e) DMSRN. **Source(s):** Created by authors

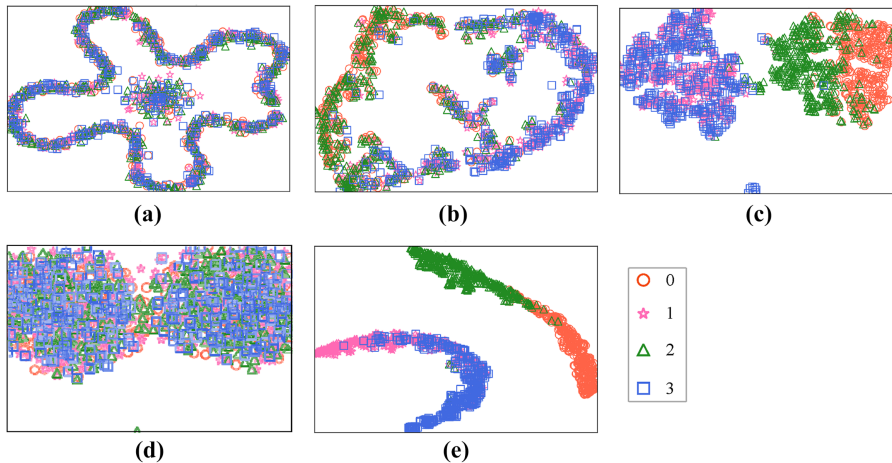
To visually assess the robustness of the extracted features, t-distribution stochastic neighbor embedding is applied to project the high-dimensional features of the elevator door system dataset into a two-dimensional space. As shown in Figure 9, the features learned by CNN, SVM, IDAN and MSRN exhibit less distinct clustering, indicating limited discriminative power. In contrast, those extracted by the proposed DMSRN show well-separated clusters corresponding to different health states, demonstrating qualitatively stronger feature discrimination. Based on the abovementioned analysis, the experimental results attest to the superior diagnostic capability of DMSRN over all comparative methods in intelligent fault diagnosis for elevator door systems.

#### 4.4 Ablation experiments on multi-scale branches

To evaluate the effectiveness of the proposed multi-scale residual architecture in DMSRN, two sets of ablation experiments were conducted. As illustrated in Table 3, the diagnosis accuracies of multi-scale branches significantly outperform any single-scale branch. Moreover, the influence of the number of residual blocks and channel multiplier, i.e. depth and width, on diagnostic accuracy is investigated and the results are summarized in Table 4. Consider minimizing network complexity as much as possible, (Depth, width)=(3,3), which brings the best result is set as the optimal network settings.

#### 4.5 Discussion on computational burden

To assess the computational complexity of the proposed DMSRN, Table 5 summarizes the floating-point operations (FLOPs), number of parameters and storage demands. While the multi-scale branches and metric learning modules introduce additional computational cost, this is offset by the enhanced feature extraction capacity of DMSRN. Moreover, as model training is performed offline, the overall computational burden remains acceptable. During testing, DMSRN achieves inference speeds comparable to those of state-of-the-art methods, preserving time efficiency. These findings underscore the model's potential for real-time fault diagnosis in practical engineering scenarios.



**Figure 9.** Feature visualization of different methods: (a) SVM, (b) CNN, (c) MSRN, (d) IDAN and (e) DMSRN. **Source(s):** Created by authors

**Table 3.** Diagnostic accuracy of different branches

Methods	Branch of kernel size 3	Branch of kernel size 5	Branch of kernel size 7	Multi-scale branches
Accuracy	83.75	83.22	79.61	87.76

**Table 4.** Diagnostic accuracy of different network settings

(Depth, width)	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
Accuracy	43.20	45.78	56.80	54.30	68.41	70.64	69.60	77.21	87.71

**Table 5.** The computational burden of different methods

Evaluation criterion	SVM	CNN	MSRN	IDAN	DMSRN
FLOPs/(M)	0.02	2.29	5.97	10.2	7.89
Parameters Number/(M)	0.01	0.13	0.37	0.51	0.42
Storage Occupancy/(MB)	0.01	0.67	1.98	5.11	3.20
Training Time/(s)	1.5	66	99.5	123	102.3
Test Time/(s)	1.0	1.3	1.4	1.5	1.3

#### 4.6 Sensitivity analysis for weighting coefficient

To conduct sensitivity analysis, a grid search was implemented on weighting coefficient  $\lambda$  for center loss, which is drawn from the range of {0.01, 0.1, 0.5,1}. As reported in Table 6, diagnostic accuracy is positively correlated with values of the weighting coefficient because  $\lambda$  determines the contribution of center loss in feature learning. The value of  $\lambda = 1$  bringing about the highest accuracy is recommended in this study.

#### 4.7 Analysis of diagnostic performance with different noise in test conditions

In real-world scenarios, the test data may be affected by noise. To investigate the diagnostic performance of already-trained DMSRN as the additional stochastic noises in the test samples vary, we utilized the models to diagnose samples with different signal-to-noise ratios (SNRs) in the task. The method of adding noise to test samples according to SNRs is shown in Eq. (10). The training samples were not added with noise. Table 7 lists the accuracies with different test SNRs by different methods. A smaller SNR indicates stronger noise is applied, which mostly results in lower classification accuracies. It can be observed that although noise significantly destroys the performance of models, the general diagnostic performance of the proposed DMSRN is superior to comparison methods.

$$\tilde{\mathbf{x}} = \mathbf{x} + \text{torch.randn\_like}(\mathbf{x}) \cdot \sqrt{\frac{\text{mean}(\mathbf{x}^2)}{10^{\text{SNR}/10}}} \quad (10)$$

**Table 6.** Diagnostic accuracy of different values for weighting coefficient

Value	0.01	0.1	0.5	1
Accuracy	70.02	72.39	82.65	87.14

**Table 7.** Accuracy (%) with different testing SNRs by different methods

SNR/(dB)	20	10	0	-10
SVM	30.07	29.87	26.91	26.25
CNN	42.96	37.04	25.99	25.13
MSRN	65.39	44.74	27.30	25.66
IDAN	69.11	48.53	32.64	26.87
DMSRN	78.29	61.64	50.20	27.79

where  $\tilde{\mathbf{x}}$  and  $\mathbf{x}$  represent the noisy and original signals, respectively; *SNR* represents signal to noise ratio; *torch.randn\_like* is the function for generating Gaussian white noise in Pytorch; *mean* ( $\cdot$ ) represents the operation of calculating the mean.

## 5. Conclusion

The intelligent fault diagnosis of elevator door systems is critical to operational safety, yet remains challenging due to the difficulty in extracting discriminative state features via deep learning. To overcome this limitation, a DMSRN method is proposed for enhanced intelligent fault diagnosis of elevator door systems. The model leverages a one-dimensional multi-scale architecture to extract representative features, combined with a metric learning-based feature discriminative enhancement strategy that enhances feature discriminability by promoting intra-class compactness and inter-class separation. Experimental validation shows that DMSRN achieves a significant performance advantage, improving average diagnostic accuracy by over 18.16% compared to the best baseline method, thereby substantially increasing the reliability of fault identification in elevator door systems.

## 6. Future work

Future work will focus on the following aspects: (1) optimizing the model architecture to further improve fault recognition accuracy and inference efficiency; (2) enhancing the model's generalization capability to enable reliable diagnosis under diverse operating conditions and environments and (3) investigating the relationship between the physical mechanisms underlying typical elevator door faults and the representations learned by the artificial intelligence model.

## References

- Bai, D., An, Z., Wang, N., Liu, S. and Yu, X. (2021), "The prediction of the elevator fault based on improved PSO-BP algorithm", *Journal of Physics: Conference Series*, Vol. 1906 No. 1, 012017, doi: [10.1088/1742-6596/1906/1/012017](https://doi.org/10.1088/1742-6596/1906/1/012017).
- Chan, P., Li, S., Deng, J., Yeung, S. and Daniel (2023), "Multi-proxy based deep metric learning", *Information Sciences*, Vol. 643, 119120, doi: [10.1016/j.ins.2023.119120](https://doi.org/10.1016/j.ins.2023.119120).
- Chen, L., Lan, S. and Jiang, S. (2019), "Elevators Fault diagnosis based on artificial intelligence", *Journal of Physics: Conference Series*, Vol. 1345, 042024, doi: [10.1088/1742-6596/1345/4/042024](https://doi.org/10.1088/1742-6596/1345/4/042024).
- Chen, C., Ren, X. and Cheng, G. (2025), "Research on distributed fault diagnosis model of elevator based on PCA-LSTM", *Algorithms*, Vol. 17 No. 6, 250, doi: [10.3390/a17060250](https://doi.org/10.3390/a17060250).
- Dong, R. and Lam, K. (2024), "Bi-center loss for compound facial expression recognition", *IEEE Signal Processing Letters*, Vol. 31, pp. 641-645, doi: [10.1109/LSP.2024.3364055](https://doi.org/10.1109/LSP.2024.3364055).

- Hu, S., Chu, Y., Tang, L., Zhou, G., Chen, A. and Sun, Y. (2023), "A lightweight multi-sensory field-based dual-feature fusion residual network for bird song recognition", *Applied Soft Computing*, Vol. 146, 110678, doi: [10.1016/j.asoc.2023.110678](https://doi.org/10.1016/j.asoc.2023.110678).
- Huang, K., Ren, Z., Zhu, L., Lin, T., Zhu, Y., Zeng, L. and Wan, J. (2025), "Intra-domain self generalization network for intelligent fault diagnosis of bearings under unseen working conditions", *Advanced Engineering Informatics*, Vol. 64, 102997, doi: [10.1016/j.aei.2024.102997](https://doi.org/10.1016/j.aei.2024.102997).
- Kim, M., Son, S. and Oh, K.-Y. (2023), "Margin-maximized hyperspace for fault detection and prediction: a case study with an elevator door", *IEEE Access*, Vol. 11, pp. 128580-128595, doi: [10.1109/ACCESS.2023.3330137](https://doi.org/10.1109/ACCESS.2023.3330137).
- Li, X., Yang, X., Ma, Z. and Xue, J. (2023), "Deep metric learning for few-shot image classification: a review of recent developments", *Pattern Recognition*, Vol. 138, 109381, doi: [10.1016/j.patcog.2023.109381](https://doi.org/10.1016/j.patcog.2023.109381).
- Liu, K., Yang, M., Yu, Z., Wang, G. and Wu, W. (2022), "FBMSNet: a filter-bank multi-scale convolutional neural network for EEG-based motor imagery decoding", *IEEE Transactions on Biomedical Engineering*, Vol. 70 No. 2, pp. 436-445, doi: [10.1109/TBME.2022.3193277](https://doi.org/10.1109/TBME.2022.3193277).
- Lu, J., Chen, H., Chen, J., Xiao, Z., Li, R., X. G. and Wang, Q., (2025a), "Temporal knowledge graph fusion with neural ordinary differential equations for the predictive maintenance of electromechanical equipment", *Knowledge-Based Systems*, Vol. 317, 113450, doi: [10.1016/j.knosys.2025.113450](https://doi.org/10.1016/j.knosys.2025.113450).
- Lu, J., Zhang, W., Lu, C., Xiao, G. and Wang, Q. (2025b), "A multi-scale convolution capsule network with data augmentation and attention mechanisms for elevator fault diagnosis", *ISA Transactions*, Vol. 167, pp. 1873-1887, doi: [10.1016/j.isatra.2025.09.041](https://doi.org/10.1016/j.isatra.2025.09.041).
- Ni, Q., Ji, J., Halkon, B., Feng, K. and Nandi, A.K. (2023), "Physics-Informed Residual Network (PIResNet) for rolling element bearing fault diagnostics", *Mechanical Systems and Signal Processing*, Vol. 200, 110544, doi: [10.1016/j.ymsp.2023.110544](https://doi.org/10.1016/j.ymsp.2023.110544).
- Pan, J., Shao, C., Dai, Y., Wei, Y., Chen, W. and Lin, Z. (2024), "Research on fault prediction method of elevator door system based on transfer learning", *Sensors*, Vol. 24 No. 7, 2135, doi: [10.3390/s24072135](https://doi.org/10.3390/s24072135).
- Pan, S., Nie, X., Zhai, X., He, C. and Ding, Z. (2025), "Classification of power quality disturbances using residual networks with channel attention mechanism", *Engineering Applications of Artificial Intelligence*, Vol. 151, 110641, doi: [10.1016/j.engappai.2025.110641](https://doi.org/10.1016/j.engappai.2025.110641).
- Ren, Y., Li, R., Ru, X. and Niu, Y. (2024), "Suppression of horizontal vibrations in high-speed elevators using active shock absorber to assist traditional damping systems", *Journal of Intelligent Manufacturing and Special Equipment*, Vol. 5 No. 1, pp. 170-189, doi: [10.1108/JIMSE-09-2023-0006](https://doi.org/10.1108/JIMSE-09-2023-0006).
- Wan, A., Tong, X., AL-Bukhaiti, K., Zhou, Z., Su, Y. and Cheng, X. (2025), "Intelligent fault diagnosis for elevator door systems using variational mode decomposition and multi-scale convolutional networks", *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, Vol. 47 No. 10, 508, doi: [10.1007/s40430-025-05816-2](https://doi.org/10.1007/s40430-025-05816-2).
- Wang, Q., Leng, Y., Li, D., Zhang, X., Li, R., Zhu, H. and Zhang, H. (2018), "MCU system-based intelligent high-speed elevator door operator fault analysis and research", *IOP Conference Series: Materials Science and Engineering*, Vol. 428 No. 1, 012028, doi: [10.1088/1757-899X/428/1/012028](https://doi.org/10.1088/1757-899X/428/1/012028).
- Wang, H., Liu, C., Jiang, D. and Jiang, Z. (2021), "Collaborative deep learning framework for fault diagnosis in distributed complex systems", *Mechanical Systems and Signal Processing*, Vol. 156, 107650, doi: [10.1016/j.ymsp.2021.107650](https://doi.org/10.1016/j.ymsp.2021.107650).
- Wu, C., Cheng, X. and Wang, H. (2025), "A contrastive clustering loss function increases class-balanced in time series classification", *Expert Systems with Applications*, Vol. 283, 127493, doi: [10.1016/j.eswa.2025.127493](https://doi.org/10.1016/j.eswa.2025.127493).

---

Yin, Y., Han, Z., Jian, M., Wang, G., Chen, L. and Wang, R. (2023), "AMSUnet: a neural network using atrous multi-scale convolution for medical image segmentation", *Computers in Biology and Medicine*, Vol. 162, 107120, doi: [10.1016/j.combiomed.2023.107120](https://doi.org/10.1016/j.combiomed.2023.107120).

Zhou, H., Qin, Q., Hou, J., Dai, J., Huang, L. and Zhang, W. (2024), "Deep global semantic structure-preserving hashing via corrective triplet loss for remote sensing image retrieval", *Expert Systems with Applications*, Vol. 238, 122105, doi: [10.1016/j.eswa.2023.122105](https://doi.org/10.1016/j.eswa.2023.122105).

**Corresponding author**

He Ren can be contacted at: [hren@cczu.edu.cn](mailto:hren@cczu.edu.cn)