

Article

Not peer-reviewed version

Vehicle Minor Fault Diagnosis Based on Multi-Residual Neural Networks and ER Rule

[Aoxiang Zhang](#), Lihong Tang, [Guanyu Hu](#)*

Posted Date: 8 December 2025

doi: 10.20944/preprints202512.0743.v1

Keywords: micro fault diagnosis; multi-residual neural network; ER Rule; diagnostic condition assessment



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Vehicle Minor Fault Diagnosis Based on Multi-Residual Neural Networks and ER Rule

Aoxiang Zhang¹, Lihong Tang² and Guanyu Hu^{2,*}

¹ Cavan Automotive, Beiqi Foton Motor Co., Ltd., Beijing, 102200, China; zhangaoxiang9469@163.com

² Key Laboratory of the Ministry of Education, Guilin University of Electronic Technology, Guilin, 541004, China; 24032306030@mails.guet.edu.cn

* Correspondence: huguanyu@guet.edu.cn

Abstract

Micro fault diagnosis of vehicle powertrain systems can significantly bring safety and economic benefits in preventing major accidents and extending equipment lifespan. However, under variable operating conditions, effectively capturing and diagnosing fault-related weak current fluctuation or high-frequency noise features, presents substantial technical challenges. Regarding this issues, this paper proposes a fault diagnosis model based on multi-residual neural networks (multi-ResNets) and Evidential Reasoning Rule (ER Rule). While the multi-ResNets are employed to extract subtle fault features and diagnostic reasoning, the ER Rule dynamically estimate diagnostic condition and conduct fault diagnosis through a sub-model real time credibility assessment mechanism. The experimental results indicate that compared traditional machine learning algorithms, the proposed multi-ResNets-ER Rule based model achieves higher diagnostic accuracy and result reliability for minor faults under variable operating conditions.

Keywords: micro fault diagnosis; multi-residual neural network; ER Rule; diagnostic condition assessment

1. Introduction

MICRO-faults as a typical fault type in complex vehicle systems, through early detected and diagnosed, it can bring substantial safety and economic benefits in aspects such as driving safety improvement, maintenance cost reduction, and equipment lifespan extension. However, due to the concealment and randomness of minor faults, their characteristics are often very difficult to detect. Moreover, fault signatures are more prone to being masked by interference or noise caused by vehicle operating conditions and load changes, making early monitoring and fault localization particularly challenging. These issues impose higher demands on diagnostic technologies in terms of feature extraction, diagnostic accuracy, and reliability. Micro-fault diagnosis is not only one of the core challenges in the current field of fault diagnosis, but also an urgent problem that needs to be solved. By organizing and analyzing current investigations, the researches on minor fault diagnosis mainly includes the following three types.

Knowledge-based approach: The diagnostic method based on qualitative knowledge is suitable for fields with complex system structures and extensive historical fault data, offering strong explainability and reasoning capabilities.

For the fault diagnosis and monitoring of gas turbine combustion chamber, literature [1] introduces the main working principle, conceptual design and development of gas turbine chamber expert system. For the problem of energy system fault diagnosis, literature [2] proposed to use the recombination of evolutionary algorithm and artificial random generation of selection operator to form the knowledge base rule of expert system. Aiming at the safety protection problem of process industry, literature [3] proposed a knowledge-based fault diagnosis method, and used fuzzy logic to reason expert knowledge and real-time data, and developed a computer-aided tool based on this method. For the problem of

automobile fault detection, literature [4] proposes to use Expert System (ES) to build the fault detection model of electric automobile, and develops the automobile fault diagnosis auxiliary system (CFMDAS) based on the expert system. For the problems of low generality and poor expansibility of existing expert systems, literature [5] proposes a task-based software architecture expert system, which allows a specific optimization based on a set of common rules, and allows users to add or modify rules in the human-machine dialog box to continue absorbing and improving professional knowledge. Literature [6] proposed a general fault diagnosis expert system framework with two basic characteristics: the first is fault diagnosis strategy, and the second is using the knowledge of fault classification and detection of the unit under test.

However, under variable operating conditions, the difficulty in extracting specific micro-fault features and expressing knowledge-based rules limits their diagnostic applicability in such scenarios.

Analytical model-based approach: The analytical diagnosis model is built upon a deep understanding of system fault evolution mechanisms, involving high-precision mathematical modeling of the system. It offers advantages such as high real-time performance and robustness in diagnosis.

In [7], a hierarchical multi-model fault diagnosis scheme based on UKF is proposed to detect and isolate actuator faults of robots. In [8], the OLA method in continuous systems is extended to fault diagnosis of nonlinear discrete systems, and the stability and robustness analysis of fault detection and compensation algorithms are given. Aiming at the fault of four-rotor actuator, a fault diagnosis system based on nonlinear fault detection estimator and nonlinear adaptive fault isolation estimator was designed in literature [9], which improved the robustness of the system and the sensitivity of the fault diagnosis algorithm. Based on robust navigation based on particle filtering, literature [10] designed a fault diagnosis system for underwater robot, which can effectively diagnose faults and has strong robustness.

For the diagnosis of micro-fault under variable operating conditions, a profound understanding of fault evolution mechanisms and precise mathematical modeling prove highly challenge. Moreover, the subtle nature of micro-fault makes them easily buried in noise, necessitating models with exceptional robustness and fault sensitivity.

Data-driven approach: Traditional data-driven fault diagnosis algorithms directly establish input-output mapping relationships by leveraging noise reduction and feature extraction algorithms. This enables them to model complex nonlinear systems.

In literature [11], discrete inverse wavelet transform is used to extract relevant frequency bands, and the current sequence is reconstructed or filtered under different fault states, and then trend fluctuation analysis is carried out to realize the effective differentiation of minor faults. Literature [12] reviews the development and application of continuous wavelet transform in the diagnosis of micro-fault of rolling bearing, and constructs a decision tree to select the optimal wavelet to improve the classification accuracy. Literature [13] combined wavelet transform with EMD, based on sample entropy and singular value decomposition, proposed a small fault diagnosis technology for auditory signals, reduced the signal-to-noise ratio, and improved the singular value decomposition. For the degradation of components and mechanical wear in the production system. Literature [14] uses ANN to detect and diagnose minor faults in the frequency domain through time-frequency analysis and technology. Literature [15] designed a new continuous decision function for SVM classifier, which can not only identify fault types, but also monitor the severity of faults.

The data-driven fault diagnosis model has achieved promising results in detecting minor faults. However, its generalization capability remains limited—when operating conditions (e.g., load or environmental variations) change, the model's performance may degrade significantly.

To address the challenges of mechanical micro-fault diagnosis under variable operating conditions, this paper proposes a fault diagnosis model based on a multi-residual neural network structure and evidential reasoning rules.

Firstly, multiple residual neural network structures are employed to extract micro-fault features and perform diagnosis reasoning under several variable operating conditions. Compared with

traditional artificial intelligence algorithms, deep learning [16–30] can pursue fundamental data characteristics and structural information at a deeper level.

Furthermore, the cross-entropy loss function is used to calculate the model divergence between predicted probability distributions and target probability distribution, and the loss values serves as belief degree for diagnostic condition assessment while also serves as the weights in subsequent diagnostic category assessment.

Case studies demonstrates that compared to traditional machine learning algorithms, the proposed fault diagnosis model achieves higher diagnostic accuracy and result reliability for micro-fault issues under varying operational conditions.

2. The Theoretical Basis

2.1. Information Transformation

Residual neural network model should utilizes the micro-fault images for feature extraction, as the time-frequency analysis tool, the short-time Fourier transform (STFT) can generate continuous spectrum through the time-frequency transformation of non-stationary signals:

$$G_f(\varepsilon, u) = \int f(t)g(t-u)e^{j\varepsilon t} dt \quad (1)$$

where, $f(t)$ is the original signal and $g(t-u)$ is the window function.

2.2. Training and Reasoning of Residual Neural Networks

Residual module as the basic building block of the residual neural network, in this section, a 4-layer residual module (see Figure 1) serves as an example to illustrate the forward and backward propagation processes, with the corresponding equations provided below:

$$F(x, \{W\}) = W_4\sigma(W_3\sigma(W_2\sigma(W_1x + b_1) + b_2) + b_3) \quad (2)$$

$$y = \sigma(F(x, \{W\}) + x) \quad (3)$$

where, x directly imports from the topmost layer; σ represents the activation function, W represents the convolution kernel weight, and b represents the bias. y is the output vector of the residual module.

For the output of residual neural network, the cross entropy loss function is adopted as follows:

$$\text{softmax}(y) = \frac{e^y}{\sum e^y} \quad (4)$$

$$\text{Loss} = - \sum \text{softmax}(y) \log y_{\text{label}} \quad (5)$$

Where, $\text{softmax}(y)$ is used for normalization; y_{label} represents the target output; Loss represents the loss value.

Based on the chain rule, the partial derivative of the loss function value can be obtained.

$$\frac{\partial \text{Loss}}{\partial y} = y - y_{\text{label}} \quad (6)$$

$$\frac{\partial \text{Loss}}{\partial W_4} = \frac{\partial \text{Loss}}{\partial y} \frac{\partial y}{\partial (F(x, \{W\}) + x)} \frac{\partial (F(x, \{W\}) + x)}{\partial W_4} \quad (7)$$

$$\frac{\partial \text{Loss}}{\partial W_3} = \frac{\partial \text{Loss}}{\partial y} \frac{\partial y}{\partial (F(x, \{W\}) + x)} \frac{\partial (F(x, \{W\}) + x)}{\partial y_3} \frac{\partial y_3}{\partial W_3} \quad (8)$$

$$\frac{\partial \text{Loss}}{\partial W_2} = \frac{\partial \text{Loss}}{\partial y} \frac{\partial y}{\partial (F(x, \{W\}) + x)} \frac{\partial (F(x, \{W\}) + x)}{\partial y_3} \frac{\partial y_3}{\partial y_2} \frac{\partial y_2}{\partial W_2} \quad (9)$$

$$\frac{\partial \text{Loss}}{\partial W_1} = \frac{\partial \text{Loss}}{\partial y} \frac{\partial y}{\partial (F(x, \{W\}) + x)} \frac{\partial (F(x, \{W\}) + x)}{\partial y_3} \frac{\partial y_3}{\partial y_1} \frac{\partial y_1}{\partial W_1} \quad (10)$$

$$W := W - \alpha \frac{\partial \text{Loss}}{\partial W} \quad (11)$$

Where, W represents the neural network original weight, α represents the learning rate, and W : represents the update weight.

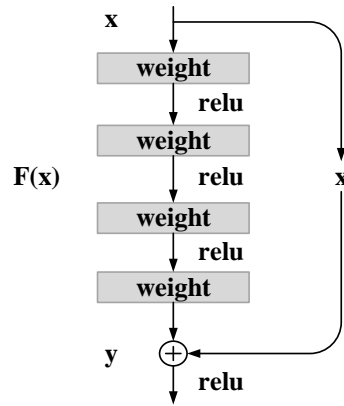


Figure 1. Residual module.

2.3. Diagnostic Condition Assessment

Cross entropy loss function reflect the deviation between the predicted distributions and the target distributions. In this subsection, the loss value is used to measure the relative credibility among the models and serves as belief degree $\beta_{n,i}$ for diagnostic condition assessment while also as the weight for subsequent diagnostic category assessment:

$$\text{Loss}_n = - \sum_{i=1}^C y_{n,i} \log(y'_{n,i}) \quad (12)$$

$$\beta_n = \frac{\log_{0.5} \text{Loss}_n}{\sum_n \log_{0.5} \text{Loss}_n}, \sum_n \beta_{n,i} = 1, 0 \leq \beta \leq 1 (n = 1, 2, \dots, N) \quad (13)$$

$$S(e) = \{(H_n, \beta_n), n = 1, 2, \dots, N\} \quad (14)$$

$$U(e) = \sum_{i=1}^N H_n \beta_n \quad (15)$$

Where, $y_{n,i}$ representing the target category in one-hot encoded form, $y'_{n,i}$ is the n th model's predicted probability of the i th category after normalization, β_n represents the belief degree of diagnostic condition e relative to the condition grade H_n , $S(e)$ represents the belief distribution, $U(e)$ represents the diagnostic condition utility value.

2.4. Diagnostic Category Assessment

The proposed model uses Evidence Reasoning Rule (ER Rule) as the diagnostic reasoning fusion theory. Firstly, the probability distributions of the diagnostic reasoning corresponding category need to be combined with the corresponding weights ω_i to obtain the basic probability masses.

$$m_{n,i} = \omega_i \beta_{n,i}, \sum_{i=1}^L \omega_i = 1, 0 \leq \omega_i \leq 1 \quad n = 1, 2, \dots, N \quad (16)$$

$$m_{H,i} = 1 - \sum_{n=1}^N m_{n,i} = 1 - \omega \sum_{n=1}^N \beta_{n,i} \quad (17)$$

where $m_{n,i}$ represents the basic probability masses. $m_{H,i}$ indicates the remaining basic probability masses which can be broken down into $\bar{m}_{H,i}$ and $\tilde{m}_{H,i}$ two main parts.

$$\bar{m}_{H,i} = 1 - \omega_i \quad (18)$$

$$\tilde{m}_{H,i} = \omega_i \left(1 - \sum_{n=1}^N \beta_{n,i} \right) \quad (19)$$

Secondly, by using the basic probability masses, the i diagnostic reasoning and the $(i + 1)$ th diagnostic reasoning can be aggregated into a new diagnostic reasoning, the aggregation process as follows:

$$\{H_n\} : m_{n,I(i+1)} = K_{I(i+1)} [m_{n,I(i)} m_{n,i+1} + m_{H,I(i)} m_{n,i+1} + m_{n,I(i)} m_{H,i+1}] \quad (20a)$$

$$m_{H,I(i)} = \tilde{m}_{H,I(i)} + \bar{m}_{H,I(i)}, \quad n = 1, 2, \dots, N \quad (20b)$$

$$\{H_n\} : \tilde{m}_{H,I(i+1)} = K_{I(i+1)} [\tilde{m}_{H,I(i)} \tilde{m}_{H,i+1} + \bar{m}_{H,I(i)} \tilde{m}_{H,i+1} + \tilde{m}_{H,I(i)} \bar{m}_{H,i+1}] \quad (20c)$$

$$\{H_n\} : \bar{m}_{H,I(i+1)} = K_{I(i+1)} [\bar{m}_{H,I(i)} \bar{m}_{H,i+1}] \quad (20d)$$

$$K_{I(i+1)} = \left[1 - \sum_{i=1}^N \sum_{j=1, j \neq i}^N m_{i,I(i)} m_{j,i+1} \right]^{-1}, \quad i = \{1, 2, \dots, L - 1\} \quad (20e)$$

where $m_{n,I(i+1)}$ is the combined probability mass generated by aggregating the first i diagnostic reasoning with the $(i + 1)$ th diagnostic reasoning; $\tilde{m}_{H,I(i+1)}$ the combined probability mass for due to the possible incompleteness, and $\bar{m}_{H,I(i+1)}$ for due to the combined relative importance.

Finally, after all L diagnostic reasoning have been aggregated, the combined probability distribution are generated by using the following normalization process:

$$\{H_n\} : \beta_n = \frac{m_{n,I(L)}}{1 - \bar{m}_{H,I(L)}}, \quad n = 1, 2, \dots, N \quad (21)$$

$$S(y) = \{(H_n, \beta_n), n = 1, 2, \dots, N\} \quad (22)$$

y is the diagnostic probability distribution which assessed to the fault category H with the belief degree $\beta_n (n = 1, 2, \dots, N)$.

3. Performance Verification and Result Analysis

3.1. Case Description and Model Construction

In this subsection, CWRU bearing dataset has been selected to verify the model ability. The dataset is one of the most widely used datasets in the field of mechanical fault diagnosis. It covers four load conditions of 0, 1, 2, and 3 horsepower, and includes three fault types: normal condition, inner race fault, outer race fault, and rolling element fault, with fault diameters categorized into three specifications: 7 mils, 14 mils, and 21 mils.

For the mechanical micro-fault diagnosis under variable operating conditions, the drive end bearing under the minimum defect size (7 mils) are selected as the diagnostic object, and the drive end data, sampling frequency in 48 kHz, is used to establish the dataset. For each bearing faults defect under each load conditions, including 800 training samples and 200 test samples. Table I gives the Drive end bearing faults defect under four load conditions. Figure 2 shows the sample's defect spectrum generated by STFT, which with the size of $224 \times 224 \times 3$.

Table 1. The bearing faults defect under four load conditions.

Motor Load (HP)	Motor Speed (rpm)	Normal	Inner Raceway	Ball	Outer Raceway Center
0	1797	-	7 mils	7 mils	7 mils
1	1772	-	7 mils	7 mils	7 mils
2	1750	-	7 mils	7 mils	7 mils
3	1730	-	7 mils	7 mils	7 mils

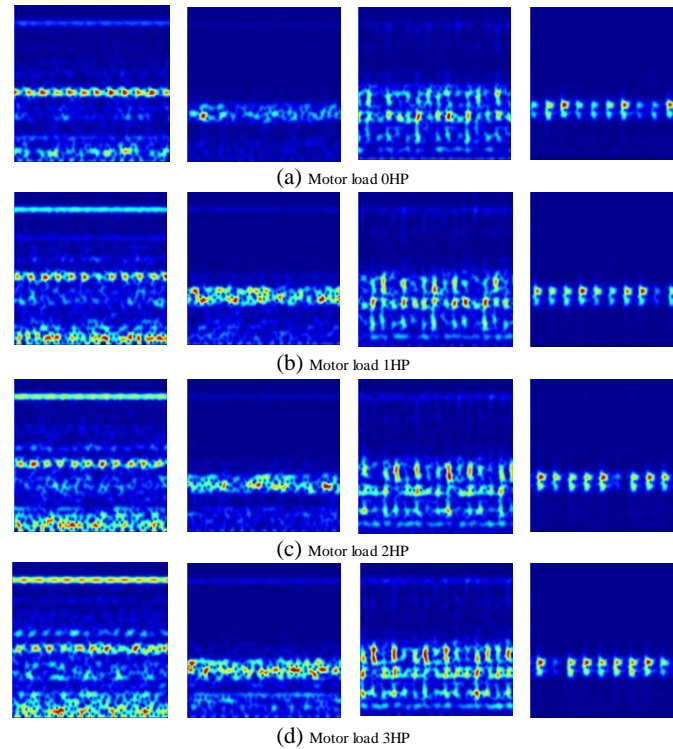


Figure 2. Defect spectrum.

For this case, a diagnostic model was established as shown in Figure 3, and the detailed structural parameters are given in Table 2. It has four 67 layer residual neural networks, for each load condition, an independent model is employed for training and diagnostic reasoning. Finally, the diagnostic condition assessment and diagnostic category assessment are conducted based on the ER rule.

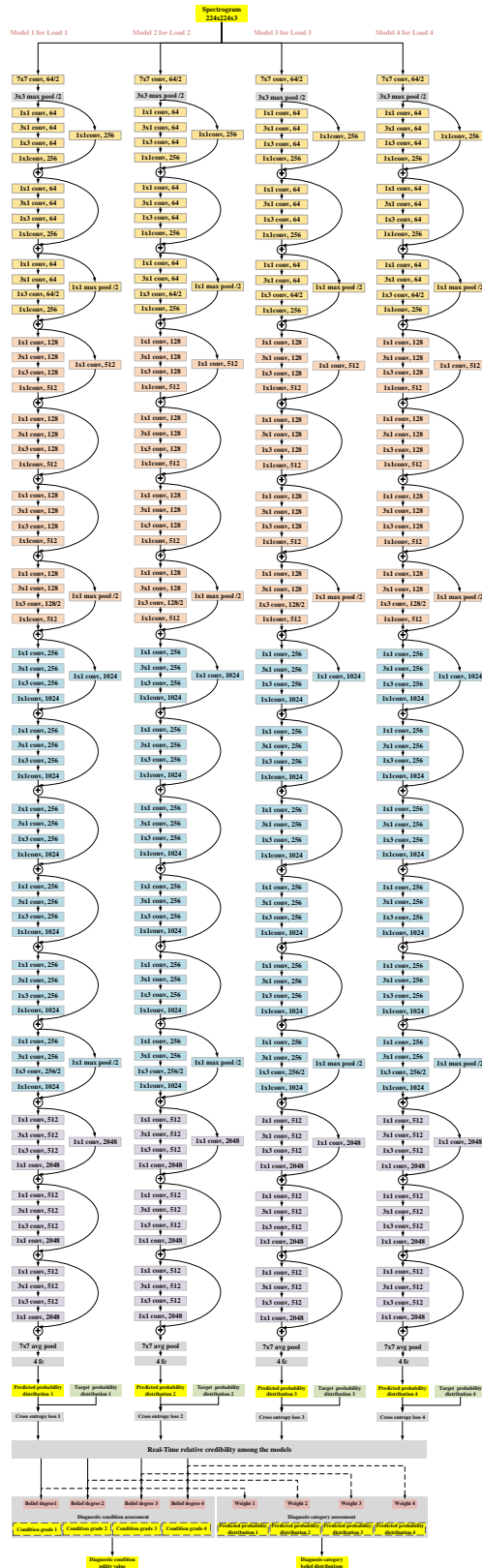


Figure 3. The structure of diagnostic model based on resnet network and ER rule.

Table 2. The parameters of the diagnostic model.

Residual Unit	Output Size	Network Layer Parameters	Unit Number	Sub-Model NUMBER
-	112×112	7×7 conv, 64/2	1	
-	56×56	3×3 max pool, 64/2	1	
unit_1	56×56	1×1, 64; 3×1, 64; 1×3, 64; 1×1, 256	2	
unit_1	56×56	1×1, 64; 3×1, 64; 1×3, 64/2; 1×1, 256	1	4
unit_2	28×28	1×1, 128; 3×1, 128; 1×3, 128; 1×1, 512	3	
unit_2	28×28	1×1, 128; 3×1, 128; 1×3, 128/2; 1×1, 512	1	
unit_3	14×14	1×1, 256; 3×1, 256; 1×3, 256; 1×1, 1024	5	
unit_3	14×14	1×1, 256; 3×1, 256; 1×3, 256/2; 1×1, 1024	1	
unit_4	7×7	1×1, 512; 3×1, 512; 1×3, 512; 1×1, 2048	2	
unit_4	7×7	1×1, 512; 3×1, 512; 1×3, 512/2; 1×1, 2048	1	
1×7×2	1×1	7×7 mean pool, 2048	1	
1×7×2	-	4 fc, Softmax	1	
ER Rule				1

3.2. Model Training and Result Analysis

In this subsection, the model conducted 2500-round iterative training, and the training results are shown in Figures 4 and 5. Wherein, Figure 4 shows the correct rate change of resnet-67 model-1~4 separately on the corresponding training set and test set. It can be seen from the change curve that the model has a high accuracy on the corresponding training set and test set, and there is no obvious over-fitting phenomenon.

Figure 5 shows the loss value change of resnet-67 model-1~4 separately on the corresponding training set, according to the change curve of loss value, it can be seen that model has the ability of rapid optimization.

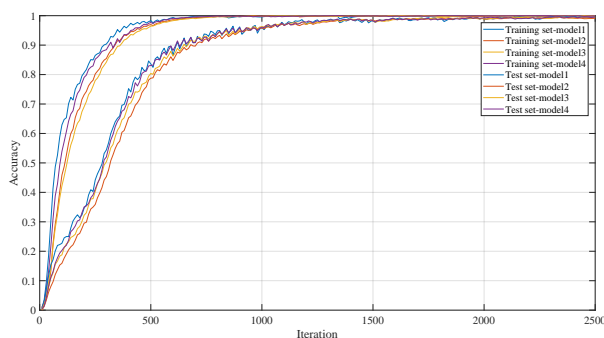


Figure 4. The correct rate of the resnet67 model -1~4.

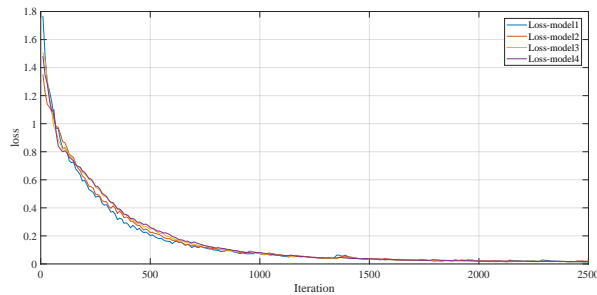


Figure 5. The loss value of the resnet67 model-1~4.

Figure 6 shows the diagnostic reasoning of model on the test set. It can be seen that the model has a high identification accuracy for all diagnostic category under four working conditions.

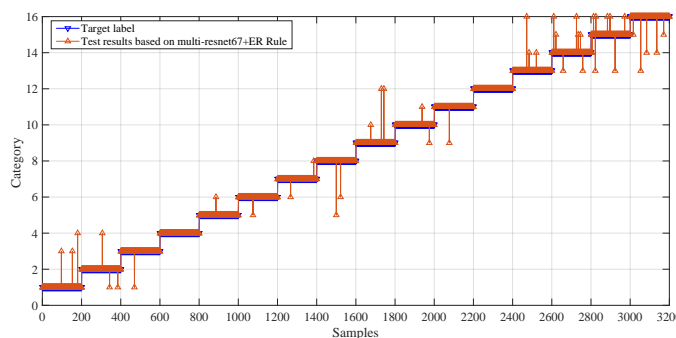


Figure 6. The diagnostic reasoning of diagnostic model on the test set.

3.3. Experimental Comparison

To evaluate the model’s performance, this subsection presents the traditional model based on the BP neural network [31], RBF neural network [32], and SVM [33]. The correct rate on test set between the proposed model and other models are shown in Figure 7 and summarized in Table 3.

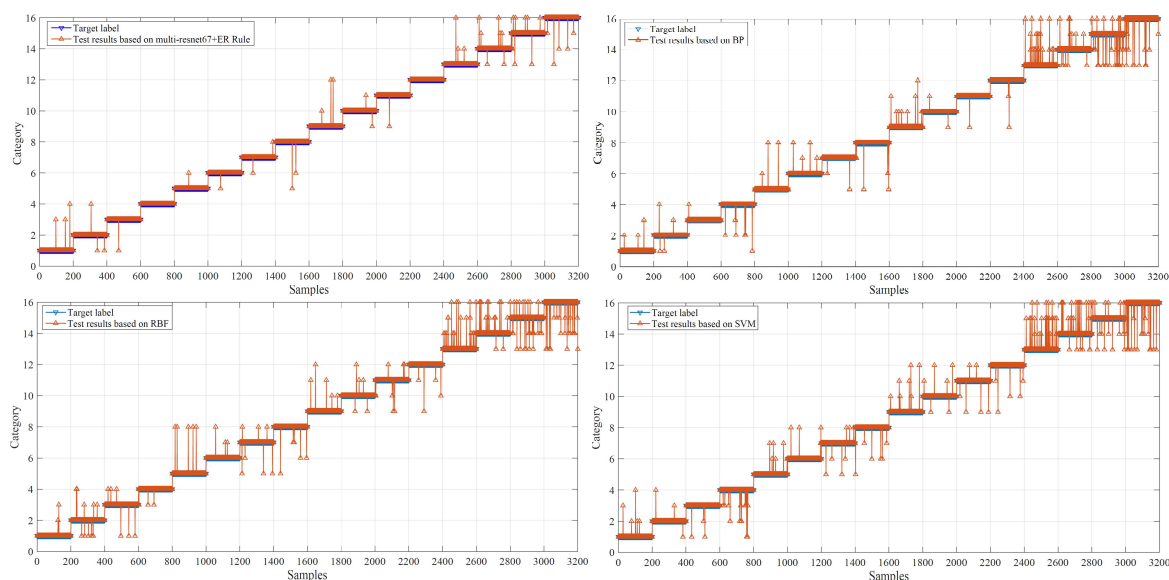


Figure 7. The diagnostic reasoning of multi-resnet67-ER rule model and other models.

Table 3. The correct rate of the model in the training set and test set.

Model	Training Set	Test Set
Multi-resnet-67+ER Rule	0.9916	0.9734
Wavelet packet+ BP	0.9772	0.9646
Wavelet packet+ RBF	0.9706	0.9553
Wavelet packet +SVM	0.9634	0.9521

The former presents the model diagnostic reasoning, while the latter displays the correct rates on the training set and test set. It is evident that the proposed model achieves a higher correct rate than both the BP network, RBF network and SVM models. These comparisons demonstrate the superior performance of the model.

4. Conclusions

To tackle the challenges of mechanical micro-fault diagnosis under variable operating conditions, this paper introduces a multi-residual neural network and evidential reasoning rule-based fault diagnosis model. Firstly, the multi-residual neural network architecture are utilized to extract micro-fault features and conduct diagnostic reasoning. Secondly, based on cross-entropy loss between predicted probability distributions and target probability distributions, the loss value is used to measure the relative credibility among the models and serves as belief degree for diagnostic condition assessment. Further, evidential reasoning rule is utilized to fuse the predicted probability distributions based on the weights which obtained from model belief degree.

Case studies have demonstrated that compared to traditional machine learning algorithms, the proposed model achieves higher diagnostic accuracy and result reliability for micro-fault issues under varying operational conditions.

The contributions of this study can be summarized as follows:

1. Multi-residual neural network structure was proposed to deep extract and classify the micro faults based on the spectrum diagram;
2. Diagnostic condition assessment and diagnostic category assessment mechanism were conducted by using ER Rule based on the models' relative credibility.

Author Contributions: Conceptualization, A.Z. and G.H.; methodology, A.Z. and G.H.; software, A.Z.; validation, A.Z.; formal analysis, A.Z. and L.T.; investigation, L.T.; resources, A.Z. and L.T.; data curation, L.T.; writing—original draft preparation, A.Z.; writing—review and editing, G.H.; visualization, L.T.; supervision, G.H.; project administration, G.H.; funding acquisition, G.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported partly by the National Natural Science Foundation of China (Grant No. 62273113), the Central Guiding Local Science and Technology Development fund project (Grant No. 2024ZYZX4027).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study were derived from the following resources available in the public domain: [Case Western Reserve University Bearing Data Center Website, URLs:<https://engineering.case.edu/bearingdatacenter/welcome>].

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ER Rule	Evidence Reasoning Rule
ResNet	Residual Network
STFT	Short-Time Fourier Transform
BP	Back Propagation network
RBF	Radial Basis Function network
SVM	Support Vector Machine
ES	Expert System
UKF	Unscented Kalman Filter
CFMDAS	Car Failure and Malfunction Diagnosis Assistance System
OLA	Online Approximator
ANN	Artificial Neural Network
EMD	Empirical Mode Decomposition
CWRU	Case Western Reserve University
HP	horsepower
RPM	Revolutions Per Minute

References

1. Afgan, N. H.; Carvalho, M. G.; Pilavachi, P. A. An expert system concept for diagnosis and monitoring of gas turbine combustion chambers. *Appl. Therm. Eng.* **2006**, *26*, 766–771.
2. Andrea, T.; Andrea, L. Energy system diagnosis by a fuzzy expert system with genetically evolved rules. *Int. J. Thermodyn.* **2008**, *11*.
3. Nan, C.; Khan, F.; Iqbal, M. T. Real-time fault diagnosis using knowledge-based expert system. *Process Saf. Environ. Protection* **2008**, *86*, 55–71.
4. Mostafa, S. A.; Ahmad, M. S.; Mohammed, M. A. Implementing an expert diagnostic assistance system for car failure and malfunction. *Int. J. Comput. Sci. Issues* **2012**, *9*.
5. Bo, M.; Zhi-Nong, J.; Zhong-Qing, W. Development of the task-based expert system for machine fault diagnosis. In Proceedings of the J. Phys.: Conf. Ser., 2012; pp. 012043.
6. Kodavade, D. V.; Apte, S. D. A universal object oriented expert system frame work for fault diagnosis. *Int. J. Intell. Sci.* **2012**, *2*, 8.
7. Hsiao, T.; Weng, M. C. A hierarchical multiple-model approach for detection and isolation of robotic actuator faults. *Robotics and Autonomous Systems* **2012**, *60*, 154–166.
8. Yang, Q. M.; Sun, Y. X. Automated fault accommodation for discrete-time systems using online approximators. In Proceedings of the 30th Chinese Control Conference, Shanghai, 2011; pp. 4264–4269.
9. Avram, R. C.; Zhang, X.; Muse, J. Quadrotor actuator fault diagnosis and accommodation using nonlinear adaptive estimators. *IEEE Trans. Control Syst. Technol.* **2017**, 1–8.
10. Zhao, B.; Skjetne, R.; Blanke, M. Particle filter for fault diagnosis and robust navigation of underwater robot. *IEEE Trans. Control Syst. Technol.* **2014**, *22*, 2399–2407.
11. Barman, D.; Sarkar, S.; Das, G.; Das, S.; Purkait, P. DFA and DWT based severity detection and discrimination of induction motor stator winding short circuit fault from incipient insulation failure. In Proceedings of the 2015 International Conference on Electrical, Electronics, Signals, Communication, and Optimization, Visakhapatnam, 2015; pp. 1–6.
12. Sugumaran, V.; Rao, A. V.; Ramachandran, K. I. A comprehensive study of fault diagnostics of roller bearings using continuous wavelet transform. *Int. J. Manuf. Syst. Des.* **2015**, *1*, 27–46.
13. Liang, J.; Yang, Z. X. A novel wavelet transform-empirical mode decomposition based sample entropy and SVD approach for acoustic signal fault diagnosis. In Proceedings of the 6th International Conference on Advances in Swarm and Computational Intelligence, Beijing, China, 2015; pp. 232–241.
14. Zhang, Z. Y.; Wang, Y.; Wang, K. S. Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *J. Intell. Manuf.* **2013**, *24*, 1213–1227.
15. Namdari, M.; Jazayeri-Rad, H. Incipient fault diagnosis using support vector machines based on monitoring continuous decision functions. *Eng. Applicat. Artif. Intell.* **2014**, *28*, 22–35.
16. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444.

17. Krizhevsky, A.; Sutskever, I.; Hinton, G. E. ImageNet classification with deep convolutional neural networks. In Proceedings of the NIPS, 2012.
18. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the CVPR, 2015.
19. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the CVPR, 2014.
20. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. In Proceedings of the NIPS, 2015.
21. Xu, K.; Ba, J.; Kiros, R.; Cho, K.; Courville, A.; Salakhutdinov, R.; Zemel, R. S.; Bengio, Y. Show, attend and tell: Neural image caption generation with visual attention. In Proceedings of the ICML, 2015.
22. Yao, L.; Torabi, A.; Cho, K.; Ballas, N.; Pal, C.; Larochelle, H.; Courville, A. Describing videos by exploiting temporal structure. In Proceedings of the ICCV, 2015.
23. Huval, B.; Wang, T.; Tandon, S.; Kiske, J.; Song, W.; Pazhayampallil, J.; Andriluka, M.; Rajpurkar, P.; Migimatsu, T.; Cheng-Yue, R.; others. An empirical evaluation of deep learning on highway driving. *arXiv* 2015, preprint at <https://arxiv.org/abs/1504.01716>.
24. Bojarski, M.; Del Testa, D.; Dworakowski, D.; Firner, B.; Flepp, B.; Goyal, P.; Jackel, L. D.; Monfort, M.; Muller, U.; Zhang, J.; others. End to end learning for self-driving cars. *arXiv* 2016, preprint at <https://arxiv.org/abs/1604.07316>.
25. Greenspan, H.; van Ginneken, B.; Summers, R. M. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Trans. Med. Imaging* **2016**, *35*, 1153–1159.
26. Ali, S. D. The impact of deep learning on investments: Exploring the implications one at a time. *Predictive Anal. Futur.* **2016**, *13*, 49–50.
27. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016; pp. 770–778.
28. Griffin, D. Signal estimation from modified short-time Fourier transform. *IEEE Trans. Acoust. Speech, Signal Process.* **1984**, *32*, 236–243.
29. Kwok, H. K.; Jones, D. L. Improved instantaneous frequency estimation using an adaptive short-time Fourier transform. *IEEE Trans. Signal Process.* **2000**, *48*, 2964–2972.
30. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016; pp. 2818–2826.
31. Sun, Y. J.; Zhang, S.; Miao, C. X. Improved BP neural network for transformer fault diagnosis. *J. China Univ. of Mining and Technol.* **2007**, *17*, 138–142.
32. Yi, Q.; Zhan-Ming, L.; Er-Chao, L. Fault detection and diagnosis for non-Gaussian stochastic distribution systems with time delays via RBF neural networks. *ISA Trans.* **2012**, *51*, 786–791.
33. Lv, G.; Cheng, H.; Zhai, H. Fault diagnosis of power transformer based on multi-layer SVM classifier. *Proceedings of the Chinese Society of Universities* **2005**, *75*, 9–15.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.