# A Machine Learning-Based Approach for Elevator Door System Fault Diagnosis

Taiwang Liang, Chong Chen, Tao Wang, Ao Zhang, and Jian Qin

Abstract—The door system is the core part of the elevator. An accurate diagnosis of the door system can aid engineers in troubleshooting and reduce maintenance costs. However, the research of fault diagnosis based on elevator operation and maintenance data is still in its infancy. With the development of the industrial Internet-of-things, real-time monitoring data of elevator can be collected and used for fault diagnosis modeling. This paper investigates a machine learning-based approach to achieve accurate elevator door fault diagnosis. An experimental study was conducted based on the monitoring data collected from the real-world elevator door system. The experimental results revealed that XGBoost algorithm can accurately identify the fault type of the elevator door.

# I. INTRODUCTION

With the acceleration of urban development, the elevator has become indispensable equipment in daily life. However, in the actual operation of the elevator, the failure of the elevator door system often leads to the phenomenon of stopping the elevator and trapped people, which seriously endangers the life safety of passengers [1, 2]. Therefore, modeling and analysis of elevator operation and maintenance data through big data analysis technology is of great significance for elevator maintenance. When the elevator fails or is about to fail, it can accurately and quickly locate the fault area and find out the cause of the failure. Traditional fault diagnosis is generally based on expert system, which combines the knowledge of domain experts and uses knowledge representation and knowledge reasoning techniques in artificial intelligence to simulate complex problems that can only be solved by the domain experts [3, 4]. However, expert system needs to be dominated by domain experts and design a large number of rules, and it has no learning ability and poor portability. This dependence is not conducive to the popularization of methods, but also affects the accuracy of diagnosis results. Nowadays, with the rapid development of industrial Internet and artificial intelligence technology, realtime monitoring data of elevator equipment can be collected and analyzed [5]. However, it is of great significance to mine the characteristic information related to faults and analyze the types of faults from massive monitoring data.

In the context of industrial big data, fault diagnosis is gradually becoming intelligent with the rapid development of artificial intelligence and machine learning [6]. In the field of fault diagnosis, machine learning has attracted extensive attention because of its powerful learning ability and training

\*Our work is supported by multiple funds in China, including the Key Program of NSFC-Guangdong Joint Funds (U2001201), the Natural Science Foundation of Guangdong Province (2020A1515010890). Our work is also supported by Guangdong Provincial Key Laboratory of Cyber-Physical System (2020B1212060069).

mode [7]. It can learn the complex relationship between input and output automatically, and achieves end-to-end modeling from original data to predicted results, reducing the need for expert knowledge and improving the accuracy and efficiency of fault diagnosis[8]. However, the research of fault diagnosis based on elevator operation and maintenance data is still in its infancy. The fault diagnosis of elevator door system is mainly based on the classification and modeling of faults, but there is no in-depth study on different fault modes classification tasks [9].

To address the above issues, this paper proposed a machine learning-based approach for elevator door system fault diagnosis based on elevator operation and maintenance data, and realizes accurate classification of common fault modes of the door system. This approach is suitable for multi-feature fault detection of the elevator door system. A fault diagnosis model can be trained using XGBoost algorithm, which can identify the failure mode of the door system. It can reduce the workload of maintenance personnel and the cost of elevator operation and maintenance, which is of great significance to improving the intelligent operation and maintenance level of the elevator. The rest of this paper is organized as follows: Chapter 2 overview. Chapter 3 introduces the methods. Chapter 4 summarizes the experimental results and analysis. Chapter 5 Concludes.

# II. LITERATURE REVIEW

Fault diagnosis refers to the detection of the running status of the device to find exceptions and analyze the exceptions to maintain the device to ensure that the device can work properly [10-13]. With the increasing importance of fault diagnosis in equipment maintenance, more machine learning algorithms are introduced into the field of fault diagnosis [14-16].

In the field of artificial intelligence, classical machine learning algorithm has been widely used in various fault diagnosis scenarios due to their lightweight property and strong analytical performance [17]. Liu et al. [18] used Bayesian network model of elevator fault diagnosis, and reasoning mechanism based on Monte Carlo study the model parameters and structure, but a Bayesian network for the expression form of the input data is extremely sensitive, and the model of classification effect is heavily dependent on the correlation between input attributes, when attribute correlation is larger, the classification of the model effect is very bad. Liu et al. [19] used the decision tree classification algorithm to

Taiwang Liang, Chong Chen, Tao Wang, and Ao Zhang are with Guangdong Provincial Key Laboratory of Cyber-Physical System, Guangdong University of Technology, Guangzhou, China. Jian Qin is with Welding Engineering and Laser Processing Centre, School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, UK (corresponding author e-mail: Chenc2021@gdut.edu.cn).

diagnose elevator faults, but the decision tree model had high requirements for training data and could not effectively train unbalanced data sets. Ye et al. [20] proposed a solution to improve the accuracy of rolling bearing fault detection. The method is based on variational mode decomposition (VMD), multiscale permutation entropy (MPE) and the particle swarm optimization-based support vector machine (PSO-SVM). Wu et al. [21] proposed a support vector machine (MPSO-SVM) transformer diagnosis method based on improved particle swarm optimization, which overcomes the shortcomings of PSO-SVM, such as slow convergence speed and easy to fall into local optimum.

With the development of artificial intelligence, deep learning has been applied to all aspects of industrial production with its powerful learning ability [22]. In the field of bearing fault diagnosis, Chen et al. [23] proposed an inverted residual convolution neural network model for bearing fault diagnosis, which reduced the feature loss in the process of down sampling, and adopted the concept of lightweight to reduce the calculation of the model. Gu et al. [24] proposed a multi-sensor fault diagnosis model based on long and short-term memory network (LSTM), which used the LSTM network to extract detailed fault information in frequency domain features after wavelet transformation for fault analysis. Chen et al. [25] proposed an integrated model to predict the automobile failure time using a merged-LSTM algorithm. In the field of gearbox fault diagnosis, Shi et al. [26] proposed a long and short-term memory network model based on two-way convolution to automatically extract the space-time characteristics of vibration and speed data, so as to detect gearbox fault categories. Li et al. [27] proposed an adaptive neural network model integrating instantaneous speed to carry out end-to-end learning guided by speed information. By combining instantaneous speed information, fault detection of planetary gearbox can be realized under different working conditions. In the field of elevator fault diagnosis, Bai et al. [28] used BP neural network model for elevator fault diagnosis, but bp neural network convergence speed is slow, poor convergence, and easy to fall into local extremums, often converging to different local minima, resulting in poor classification effect.

Although traditional machine learning and deep learning have rich research results on fault diagnosis in industrial scenarios, the research on fault diagnosis of elevator equipment is very limited. And the research towards the elevator door machine system fault diagnosis algorithm for intelligent operations has the important significance of the elevator, this paper puts forward a fault diagnosis algorithm based on XGBoost elevator door system, help operations engineers in the identification of elevator door machine system health, reduce the operational cost of the lift, effectively enhance the level of the elevator of the intelligent operations.

#### III. METHODOLOGY

The elevator door system mainly consists of a door machine, car door, layer door and its related components. With the development of elevator manufacturing technology and big data technology, elevator door systems can sense and collect more and more data, including door motor current, door motor command speed, door motor feedback speed, car door position,

switch threshold signal, opening and closing time, opening and closing state and so on.

The condition monitoring data of the elevator door system is collected from the door control system, which is a type of time series data. This paper presents an XGBoost-based door machine fault diagnosis method. Feature variables closely related to faults were found through mechanism analysis, and the maximum value, minimum value, average value, standard deviation, skewness and kurtosis in each fixed time window were statistically analyzed by sliding window approach, and then the XGBoost model was introduced for training. Figure 2 shows the sliding window approach process with a step size of 1s and window length of 4s. The meanings of the six statistical variables are shown in Table 1. The flow chart of fault diagnosis is shown in Figure 1, and the steps are as follows.

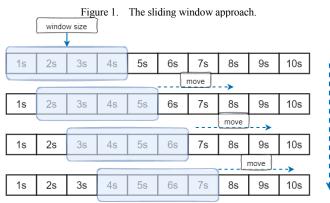


TABLE I. MEANINGS OF STATISTICAL VARIABLES

Index	Statistical variables	Descrition	
1	Maximum value	The maximum value of the instances in a time window	
2	Minimum value	The minimum value of the instances in a time window	
3	Mean value	The mean value of the instances in a time window	
4	Standard deviation	The Standard deviation of the instances in a time window	
5	Skewness	Direction and degree of skew of statistical data distribution in time window	
6	Kurtosis	The number of features representing the peak value at the mean of the probability density distribution curve in a given time window	

- Through mechanism analysis, determine which sensor signals are strongly related to the fault to be collected.
- Data sets containing multiple common man-made fault samples are collected through the elevator big data platform, and each sample contains a multichannel feature item and label item corresponding to the fault type, and the data sets are divided into training sets and test sets according to a certain proportion. Among them, the training set is used to train the model, and the test set is used to test the generalization ability of the model after training.
- The sliding window approach was used to calculate six statistical features of each data in each time

window, including maximum value, minimum value, mean value, standard deviation, skewness and kurtosis.

- By means of the grid search, the parameters in the XGBoost model are tuned to find the best parameter combination.
- After the training, save the model and input the test set into the model.

#### A. XGBoost

XGBoost algorithm, proposed by Professor Tianqi Chen in 2016, uses the negative gradient of the loss function as the residual value of the current fitting to achieve an accurate classification effect. XGBoost improves the traditional GBDT objective function by adding regular terms based on the original function, reducing the possibility of over-fitting and speeding up the convergence speed. The most important feature of the algorithm is that it supports multi-thread computation and uses regularization enhancement technology to reduce over-fitting, so as to ensure the robustness of the model. At the same time, loss function can be customized, sparse feature processing, allowing missing values, etc., which has the advantages of flexibility, fast computing speed, not easy to be interfered with by outliers and good robustness. Objective function:

$$Obj^{t} = \sum_{i=1}^{n} l\left(y_{i}, y_{i}^{\wedge (t-1)} + f_{t}\left(x_{i}\right)\right) + \Omega\left(f_{t}\right) + c \tag{1}$$

Taylor expansion:

$$f(x + \Delta x) \cong f(x) + f'(x) * \Delta x + 1/2 f''(x) * \Delta x^2$$
 (2)

Model complexity calculation formula:

$$\Omega(f_t) = \gamma T + 1/2\lambda \|w\|^2 = \gamma T + 1/2\lambda \sum_{j=1}^{T} w_j^2$$
 (3)

Optimization process:

$$g_i = \frac{dL(y_i, y_i^{\wedge (t-1)})}{\partial y_i^{\wedge (t-1)}} \tag{4}$$

$$h_i = \frac{\partial^2 L(y_i, y_i^{\wedge (t-1)})}{\partial y_i^{\wedge (t-1)}} \tag{5}$$

Substitute formula (2), (3), (4) and (5) into Formula (1), and the final result is:

$$Obj^{t} = \sum_{j=1}^{T} \left[ (\sum g_i) w_j + \frac{1}{2} (\sum h_i + \lambda) w_j^2 \right] + \gamma T \quad (6)$$

$$G_{j} = \sum g_{i}, H_{j} = \sum h_{i} :$$

$$Obj^{t} = \sum_{j=1}^{T} \left[ G_{j} w_{j} + \frac{1}{2} (H_{j} + \lambda) w_{j}^{2} \right] + \gamma T$$
(7)

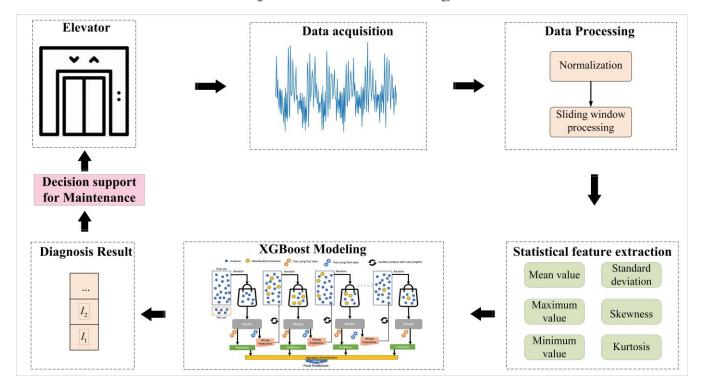
So:

$$w_j^* = -\frac{G_j}{H_j + \lambda} \tag{8}$$

Target results:

$$Obj = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} + \gamma YT$$
 (9)

Figure 2. The overall flow of the methodology.



#### IV. EXPERIMENTAL STUDY

# A. Experimental Setup and Data

In order to verify the validity of the proposed model, the experimental data of a civil elevator in a residential area were collected through a fault injection experiment and analyzed by simulation. Through the statistical analysis of the maintenance records of the maintenance personnel, it is found that the main human factors leading to the failure of the elevator door system include: blocking door by hand (BDBH), object to block the door (OBTD), door groove blocked by an object (DGBO). BDBH usually causes the elevator doors to not close the first time, and then return to normal the next time the passenger enters the elevator. OBTD can cause the elevator door to remain closed for a long time until the object is removed. DGBO will not cause the elevator door to be unable to close. but it will affect the speed of the elevator door closing in the process of the elevator door closing so that the elevator door cannot close normally within the specified time. The identification of these faults can provide information for the operation and maintenance schedule. Therefore, in the injection failure experiment, artificial hands, with a box to block the door, the foreign body into the elevator door, and then let the elevator door open and close operation. For each type of fault experiment, 10 groups were collected, each group lasted for 30 s, and the collection frequency was every 40ms. The sampled data included the position of the door engine, given speed, feedback speed and working current. The malfunction injection experiment categories are shown in Table 2.

The collected original data were processed by sliding window, and six statistical features of maximum value, minimum value, mean value, standard deviation, skewness and kurtosis in each sampling signal time window were counted. 24 features consistent with model input were obtained after data processing. At the same time, data sets with different window lengths and steps of 1s were made for the data, including 1s, 2s, 4S, 6s, 8s, 10s and 12s.

TABLE II. THE MALFUNCTION INJECTION EXPERIMENT CATEGORIES.

Index	Experimental class	Abbreviations	
1	normal opening and closing doo	NOCD	
2	blocking door by hand	BDBH	
3	door groove blocked by object	DGBO	
4	object to block the door	OBTD	

In order to evaluate the classification efficiency of the model, accuracy, precision, recall and F1 values were used to comprehensively evaluate the model. The specific formula is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
 
$$Pr \ e \ cision = \frac{TP}{TP + FP}$$
 
$$recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{10}$$

where, TP represents the number of correctly identified categories; FP represents the number of other categories incorrectly identified as such; FN represents the number of categories misidentified as other categories.

In this experiment, an XGBoost algorithm was designed to do the elevator door system fault diagnosis model. After several experiments, the model parameters were finally adjusted as table 3.

TABLE III. KEY PARAMETERS OF XGBOOST.

Parameter	Value	Parameter	Value
Learning_rate	0.2	Min_child_weight	5
Max_depth	10	Gamma	0
Subsample	0.8	Colsample_bytree	0.8

In order to verify the validity of the XGBoost gate system fault diagnosis model, the following three experiments were designed. All experiments were conducted on Intel I5-6500 3.20Ghz PC with 2\*8GB memory and sklean, a machine learning library.

In order to find the appropriate time window size, the door system fault diagnosis model based on XGBoost was established. The model was trained with 7 data sets made and verified by the ten-fold cross-validation method, so as to avoid the influence of data set division on the model classification results. The results of Accuracy Precision, Recall and F1 values in each different time window were analyzed.

Compare the performance of different algorithms. In this paper, the advantages of the fault diagnosis model based on the XGBoost gate system are analyzed by comparing the support vector machine (SVM), random forest (RF) and k nearest Neighbor method (k-NN). Among them, SVM, RF and k-NN are implemented using sklean machine learning library and use default parameters. The experimental data set uses a time window of 4s data set. The results of accuracy, precision, recall and F1 values of gate system fault diagnosis under different algorithm models were compared.

## B. Experimental Results

The results of experiment 1 are shown in the following table. Figure 3 shows the accuracy, precision, recall and F1 values of fault classification of 7 data sets divided by different window sizes.

It can be seen from Figure 3 that when the time window is short, the data set contains insufficient information due to too short window time, and the four performance indicators are all low. When the time window reaches 10s, the data set contains sufficient information due to its window, and each indicator exceeds 95%. In order to better display the classification effect of the model on various faults in the data set, the diagnosis results are visualized through the confusion matrix, as shown in Figure 4.

It can be seen from the confusion matrix that the classification effect of DGBO and NOCD is very good, but the

judgment accuracy of BDBH and OBTD is relatively poor. Since both BDBH and OBTD may completely hinder the operation of car doors and floor doors, there is a certain similarity in data. The results of experiment 2 are shown in Figure 5, which shows the results of different models on the accuracy, precision, recall and F1 values of the data set with a time window of 4S. As can be seen from the figure, the XGBoost algorithm proposed in this paper has improved in accuracy, precision, recall and F1 values compared with traditional SVM, RF and k-NN.

Figure 3. Comparison in different time Windows.

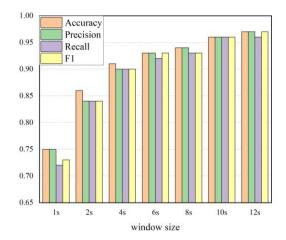


Figure 4. The Confusion matrix of XGBoost modelling.

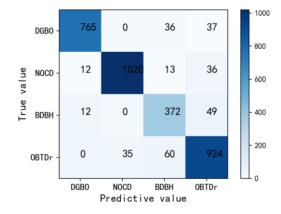
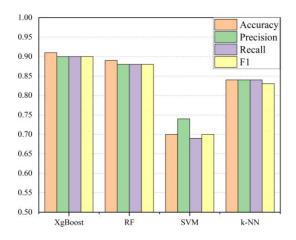


Figure 5. Comparison of the results of different algorithms.



#### V.DISCUSSION

The experimental results have demonstrated that XGBoost can achieve good performance in terms of accuracy, precision, recall and F1 values. With the increase of the window size, the performance of XGBoost tends to be converged. In the actual elevator operation scenario, the large window size will cause a heavy burden to the data transmission unit, especially for the elevator service company which monitors thousands of elevators. Hence, the window size which is 4s or 6s is suggested to be taken for the real-world application. In the future, this work will be further extended with the development of edge computing. The current fault diagnosis strategy is transmitting the condition monitoring data in the customer end to the operation centre for fault diagnosis, which is high in data transmission cost and latency. Edge computing that enables the algorithms to be deployed in the edge device can bring tangible benefits to the elevator fault diagnosis.

## VI. CONCLUSION

When the elevator door system fails, the traditional expert system has low fault diagnosis efficiency due to its difficult acquisition and poor learning ability. In this paper, a fault diagnosis model of elevator door systems based on XGBoost is proposed, which can effectively identify whether and what kind of faults occur in elevator door system from massive detection data. For fault diagnosis modelling, the key features within the sliding window are first extracted and then fed into the XGBoost algorithm. Accuracy, precision, recall and F1 values were used as evaluation indexes to compare with SVM, RF and *k*-NN. The results show that the proposed method can achieve a more accurate fault diagnosis of elevator door systems, which can help engineers to identify the health status of elevator door machine systems, and reduce the elevator operation and maintenance costs.

#### REFERENCES

- Ziliang, A., B. Dingsong, H. Yuxin, N. Wang, D. Yuji, G. Ningli, and L. Shaofeng, *Building Elevator Safety Monitoring System Based on the BIM Technology* Journal of Physics: Conference Series, 2021. 1939(1).
- [2] Guo, B., L. Zhang, and Y. Li, Research on the path of residents' willingness to upgrade by installing elevators in old residential quarters based on safety precautions Safety Science, 2019. 118.
- [3] Tan, C.N., C.F. Tan, and M.A. Abdullah, A Fault Diagnosis Expert System for Building Chillers International Review of Mechanical Engineering IREME, 2017. 11(4).
- [4] Ti, L.C., C.W. Yuan, and I. Jutarporn, A Framework for Improving Fault Localization Effectiveness Based on Fuzzy Expert System IEEE ACCESS, 2021. 9.
- [5] Qingchao, J., Y. Shifu, C. Hui, and Y. Xuefeng, Local-Global Modeling and Distributed Computing Framework for Nonlinear Plant-Wide Process Monitoring With Industrial Big Data. IEEE transactions on neural networks and learning systems, 2020.
- [6] Sunday, O. and S. Mahmood, Performance Metrics for Artificial Intelligence (AI) Algorithms Adopted in Prognostics and Health Management (PHM) of Mechanical Systems. Journal of Physics: Conference Series, 2021. 1828(1).
- [7] Anyi, L., Y. Xiaohui, D. Huanyu, X. Zihao, and Y. Chunsheng, Machine Learning-Based Sensor Data Modeling Methods for Power Transformer PHM. Sensors (Basel, Switzerland), 2018. 18(12).
- [8] S, M. and D. K, Fault diagnosis of various rotating equipment using machine learning approaches A review. Proceedings of the

- Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, 2021. 235(2).
- [9] Liu, C., X. Zhang, X. Liu, and C. Chen, The Research Of Elevator Fault Diagnosis Method Based On Decision Tree Algorithm. Proceedings of the 2017 2nd Joint International Information Technology, Mechanical and Electronic Engineering Conference (JIMEC 2017), 2017.
- [10] He, C., T. Wu, C. Liu, and T. Chen, A novel method of composite multiscale weighted permutation entropy and machine learning for fault complex system fault diagnosis. Measurement, 2020. 158(prepublish).
- [11] Yonghao, M., Z. Boyao, L. Jing, Z. Ming, L. Hanyang, L. Zongyang, and L. Hao, A review on the application of blind deconvolution in machinery fault diagnosis. Mechanical Systems and Signal Processing, 2022. 163.
- [12] Yongxiang, L., W. Tianzheng, Y. Hui, Y. Gang, and Z. Na, Study on Application of Infrared Spectrum Technology in Power Equipment. MATEC Web of Conferences, 2019. 260.
- [13] Li, Quankun, Liao, Mingfu, Jing, and Xingjian, Transmissibility function-based fault diagnosis methods for beam-like engineering structures: a review of theory and properties. Nonlinear Dynamics, 2021. 106(3).
- [14] Chen, X., B.Y. Wei, and X. Wen, Study of support vector machine based faults diagnosis for satellite's actuators. Chinese Space Science and Technology, 2018. 38(2): p. 47-55.
- [15] Liu, L.H., Q. Zhang, H. Wang, G. Li, H. Wu, Z.P. Wang, B.Z. Guo, and J.Y. Zhang, Fault diagnosis for momentum wheel bearing based on spectral kurtosis entropy and hierarchical extreme learning machine. Chinese Space Science and Technology, 2021. 41(3): p. 97-104
- [16] Zhang, Y.F., B. Ma, J.L. Zhang, M. Chen, Y.H. Fan, and W.C. Li, Study on Fault Diagnosis of Power-Shift Steering Transmission Based on Spectrometric Analysis and SVM. Spectroscopy and Spectral Analysis, 2010. 30(6): p. 1586-1590.
- [17] MingHong, H., W. Yaman, H. Yunfeng, and W. Yumin, A Fault Diagnosis Method based on Improved Synthetic Minority Oversampling Technique and SVM for Unbalanced Data IOP Conference Series: Materials Science and Engineering, 2021. 1043(5).
- [18] Liu, C., X.Z. Zhang, X.D. Liu, C. Chen, and Iop. The research of elevator health diagnosis method based on Bayesian network. in 2nd International Conference on Materials Science, Energy Technology and Environmental Engineering (MSETEE). 2017. Zhuhai, PEOPLES R CHINA: Iop Publishing Ltd.
- [19] Liu, C., X.Z. Zhang, X.D. Liu, and C. Chen. The Research Of Elevator Fault Diagnosis Method Based On Decision Tree Algorithm. in 2nd Joint International Information Technology, Mechanical and Electronic Engineering Conference (JIMEC). 2017. Chongqing, PEOPLES R CHINA: Atlantis Press.
- [20] Maoyou, Y., Y. Xiaoan, and J. Minping, Rolling Bearing Fault Diagnosis Based on VMD-MPE and PSO-SVM Entropy, 2021. 23(6).
- [21] Yuhan, W., S. Xianbo, Y. Pengfei, and W. Zhihao, Transformer fault diagnosis based on Improved Particle Swarm Optimization to support Vector Machine Journal of Physics: Conference Series, 2021. 1750(1).
- [22] Wen, C.L. and F.Y. Lu, Review on Deep Learning Based Fault Diagnosis. Journal of Electronics & Information Technology, 2020. 42(1): p. 234-248.
- [23] Deqiang, H., L. Chenyu, C. Yanjun, J. Zhenzhen, L. Xianwang, and S. Sheng, A rolling bearing fault diagnosis method using novel lightweight neural network. Measurement Science and Technology, 2021. 32(12).
- [24] Gu, K., Y. Zhang, X.B. Liu, H. Li, and M.F. Ren, DWT-LSTM-Based Fault Diagnosis of Rolling Bearings with Multi-Sensors. Electronics, 2021. 10(17): p. 13.
- [25] Chen, C., Y. Liu, X. Sun, C. Di Cairano-Gilfedder, and S. Titmus, An integrated deep learning-based approach for automobile maintenance prediction with GIS data. Reliability Engineering & System Safety, 2021. 216: p. 107919.
- [26] Junchuan, S., P. Dikang, P. Zhongxiao, Z. Ziyang, G. Kai, and W. Dazhong, Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks Mechanical Systems and Signal Processing, 2022. 162.
- [27] Peng, C., L. Yu, W. Kesheng, and Z.M. J., An automatic speed adaption neural network model for planetary gearbox fault diagnosis Measurement, 2021. 171.

[28] Dingsong, B., A. Ziliang, W. Ning, L. Shaofeng, and Y. Xintong, The Prediction of the Elevator Fault Based on Improved PSO-BP Algorithm Journal of Physics: Conference Series, 2021. 1906(1).