IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 July 2024, accepted 2 August 2024, date of publication 8 August 2024, date of current version 19 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3440649

RESEARCH ARTICLE

Real-Time Prewarning Method of Subway Turnout Jamming Failure Based on Short-Term Load Prediction and Ensemble LVQ Learning

HAO WEN[1], JIE XIAO[1], AND GUANGXIANG XIE[2]
[1]School of Rail Communication & Signals and Electrification, Wuhan Railway Vocational College of Technology, Wuhan, Hubei 430205, China
[2]Wuhan Metro Operation Company Ltd., Wuhan, Hubei 430030, China

Corresponding author: Jie Xiao (20150012@wru.edu.cn)

This work was supported in part by the Scientific Research Project of Hubei Provincial Department of Education under Grant B2022527, and in part by the Scientific Research Team of Wuhan Railway Vocational College of Technology under Grant CXTD201901.

ABSTRACT Turnout is the key equipment for realizing the turnback of subway trains. Frequent movements and environmental changes often lead to abnormal increases in resistance during turnout conversion, resulting in jamming failures that directly affect traffic safety and efficiency. In order to effectively foresee the risk of failure during operation and minimize the adverse effects of failure, a real-time turnout jamming failure prewarning method is proposed. Firstly, a weighted grey prediction machine using PSO for weight optimization (PSO-WGPM) is proposed to predict the short-term action load index, and the predicted index series is used to characterize the future changes in the resistance state of the turnout; Secondly, the predicted index and the generated index are cascaded into a risk identification index series, and a multi-dimensional hybrid prewarning feature set with time continuity are constructed by time-domain characteristics and overload statistical characteristics of the risk identification index series; Then, the prewarning feature set is fed into the learning vector quantization(LVQ) network for prewarning discrimination, and an ensemble learning based on a hybrid voting strategy is designed to obtain the final prewarning result, in order to fit the learning environment of unbalanced small-scale samples; After giving an prewarning, the occurrence time range of jamming failure is inferred based on the overload rate of the predicted index series. The experimental results show that: the proposed method can improve the prewarning success rate and effectively control false alarms, with good correctness; it can infer the range time of fault occurrence, achieving prewarning accuracy; the prewarning calculation time is much shorter than the failure advance time, meeting the timeliness requirements for applications.

INDEX TERMS Subway, turnout, jamming failure, real-time prewarning, load prediction, ensemble LVQ learning.

I. INTRODUCTION

Turnout (including the point machine, the same below) is the most basic and critical driving equipment in the current rail transit system, used to convert the direction of track opening to achieve train turnback and change the direction of travel. Once the turnout fails, it may lead to the interruption of train operation, or even the derailment and rollover of the train, causing adverse social impact. Subway turnback turnouts operate very frequently. Owing to some factors such as continuous mechanism wear and train vibration, they often enter the overload working state due to abnormal increase of conversion resistance during operation, resulting in mechanical failures. Compared to sudden electrical failures, turnout mechanical failures such as jamming have more complex causes and higher incidence, but their formation is generally

The associate editor coordinating the review of this manuscript and approving it for publication was Gongbo Zhou.

gradual and predictable [1]. If the risk can be predicted before the failure, the impact of the failure can be minimized or even avoided by carrying out emergency repairs or changing the turnback line in advance. Therefore, it is of great significance to study the real-time prewarning method of turnout jamming failure to ensure the subway traffic safety. From a macro view, the PHM (Prediction and Health Management) method for rail transit turnout is usually driven by data [2], [3], and then implemented through specific techniques such as condition safety assessment, action anomaly identification, and potential risk mining. The representative methods are summarized in the following.

Eker et al. [4] proposed a strategy for collecting equipment parameters that can simulate the failure development process, and established a simple state evaluation method based on Hidden Markov process, which can predict the remaining useful life (RUL) of turnouts based on the degree of oil shortage of the slide plate. Work [5] improved the shortage for ignoring the state duration in [4], and achieved better practical application effect. Chen et al. [6] describes the state transition and failure evolution of the equipment degradation process using Hidden Semi-Markov model (HSMM), thereby establishing the failure degradation identification and state prediction model of turnout point machine, and using SA-CPSO to optimize HSMM parameters to improve the accuracy of the model. Methods [7], [8] takes the time-frequency characteristics of turnout action parameters as the carrier, and established the RUL prediction model for turnout based on SVM joint spectral analysis forecast and CNN-BiLSTM, respectively. The nonlinear mapping ability of machine learning is used to better characterize and predict the process of turnout state degradation. A feature fusion method that can effectively capture the relationship between the RUL and the appropriate health indicator (HI) is proposed in [9], and realized the prediction of turnout failure state based on genetic programming. Methods [10], [11] takes turnout mechanical parameters as the carrier, and established comprehensive evaluation model for turnout conditions based on granular computing and fuzzy hierarchical evaluation to measure the health of equipment. These models can provide condition grading prewarning based on abnormal evaluation results and make preventive maintenance decisions. In addition, work [12] has realized the prediction of gap failure risk level of turnout point machine based on discretization method and random forest. The main purpose of the above methods [4], [5], [6], [7], [8], [9], [10], [11], [12] is to predict the RUL or failure risk period, focusing on fitting the equipment state degradation failure process throughout the whole cycle, which can provide a reference for daily preventive maintenance of turnouts. In fact, the action environment of the turnout is complex and changeable, the pre-established RUL prediction model is uncertain, and many static parameters cannot be obtained in real-time, so this kind of method often does not have the timeliness of short-term abnormal state perception during operation.

To achieve failure prediction and prewarning during operation, real-time turnout action monitoring data must be used as the basis. The relevant methods take the action power/current curve [13], [14], [15], [16], sound signal [17] or vibration signal [18], [19] generated with the turnout action as the carrier, and establish feature matching or learning models based on DTW, SVDD, deep forest, fuzzy clustering, FWPDE-BPSO-SVM, KPCA-SVM, GAN to realize health prewarning, failure prediction and diagnosis. These methods have the characteristics of real-time processing, but it is more one-sided to predict the occurrence of failures and give prewarning only based on the monitoring data generated by a single action at the current time, which is easy to cause misjudgments, as failures are often the result of qualitative changes caused by multiple anomalies. A method of failure criticality discrimination based on statistical strategy was proposed in [20], which can alleviate the one- sidedness of failure prediction based on single abnormal identification. However, the problem remains that the generated action data is used as the basis for judgment, and when prewarning is issued, it often indicates that the failure is approaching, which can easily lead to delayed or even missed prewarning. García et al. [21] first adopted a two-step process of parameter prediction and fault prejudge to perform failure prewarning, which involves first predicting the action parameters in real-time, and then directly using the predicted values to make failure discrimination. The recently proposed methods [22], [23], [24] also adopt a similar processing framework, using discrete gray model (DGM), gated recurrent units (GRU), autoregressive moving average to predict the turnout action parameters in real-time, and then establish the failure prediction model. Such methods [21], [22], [23], [24] are beneficial for early detection of failure symptoms; however, these methods only use predictive features for failure discrimination, which can easily ignore the continuity of predictive features with current state features, making the effectiveness of the method highly dependent on the accuracy of feature prediction and lacking robustness, especially prone to misjudgments during non-failure periods. Excessive false alarm will adversely affect the normal order of operation, so reducing false alarms is also an important part of improving the correctness of prewarning. Finally, the existing methods are main still in the stage of ''fuzzy prewarning'' to judge whether a failure has occurred, and there are some limitations in the accuracy of prewarning; If we can further infer the approximate time-period of failure based on the correct prewarning, it will be more beneficial for on-site emergency measures to be taken. This is an important direction for methodological progress.

Considering the shortcomings of existing technologies, three core issues need to be addressed for the practical application of real-time failure prewarning methods: first, the prewarning correctness requires to be further improved; second, the prewarning accuracy needs a breakthrough; and third, the prewarning timeliness must be guaranteed. In summary, this article focuses on the three core issues of failure

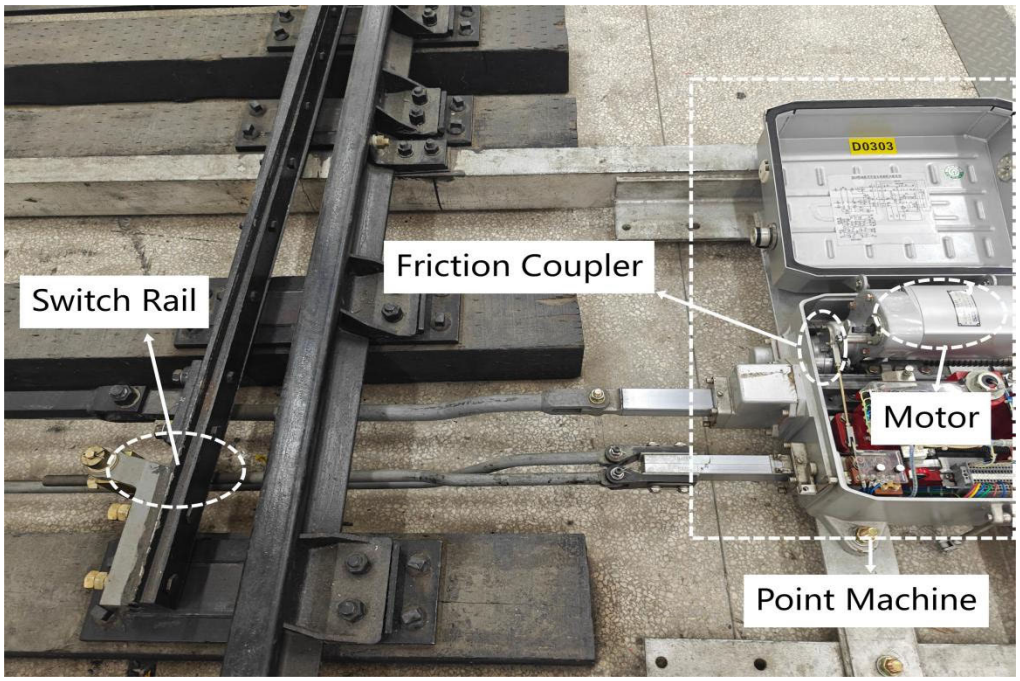

FIGURE 1. A basic turnout system.

prewarning, and proposes a real-time failure prewarning method for the most common turnout jamming failure in subway operation. This method adopts a four-step process: process1 (parameter prediction), process2 (feature calculation), process3 (prewarning discrimination), and process4 (failure time inference). The specific innovations and contributions focus on each of the above processes, which are summarized as follows:

1. Established a PSO-WGPM model to realize the short-term precise prediction of action parameter series, which can also achieve good prediction accuracy for the fluctuation series, and lay the foundation for the subsequent prewarning process.

2. Proposed a multi-dimensional hybrid prewarning feature calculation strategy with spatio-temporal continuity, which can comprehensively characterize the action state changes of the turnout from the past to the future, effectively capture the signs of failure, and improve the robustness of prewarning.

3. Designed a LVQ ensemble learning model based on hybrid voting strategy, which combines voting rate and confidence score for prewarning discrimination, improving the adaptability and accuracy of the method in unbalanced small-scale sample learning environments.

4. It is realized to infer the occurrence time of the turnout jamming failure based on the predicted action parameters and prewarning features, and the prewarning accuracy is obtained while ensuring the timeliness, which improves the practical application value of the method.

II. PROBLEM FORMULATION AND PRELIMINARIES

This chapter first provides a detailed description of the objects to be studied in this article, then elaborates on the connotation of the issues to be addressed and gives definitions of relevant evaluation indicators, and finally introduces the overall implementation idea of the method in this article.

A. PRINCIPLE AND PARAMETER CHARACTERIZATION OF TURNOUT JAMMING FAILURE

The equipment composition of a basic turnout system is shown in FIGURE 1.

Based on the system in FIGURE 1, an overview for the main process of turnout failure caused by jamming is given: the switch rail of turnout suffers abnormal resistance during the conversion process, causing the point machine with traction switch rail in overload working state; When the load exceeds a certain limit, the friction coupler in the point machine will fail, so as to actively disconnect the switch rail from the point machine to prevent the motor from burning out. At this time, the turnout will be unable to continue to operate properly. Therefore, selecting the appropriate load index to characterize the action state of turnouts is an important foundation for determining the risk of jamming failure. From the physical meaning, the turnout action load can refer to the total power taken by the point machine to overcome the resistance at a certain action time, so the power-time curve generated in real-time with the turnout action is commonly used in the actual site to reflect its resistance state. Next, with a single conversion movement as a cycle, the power curves of normal action, critical jamming action and jamming failure are drawn in a display window for comparison, as shown in FIGURE 2.

From FIGURE 2, The power value at the potential failure time will rise abnormally before the critical jamming failure occurs, and the overall action time will also increase slightly.

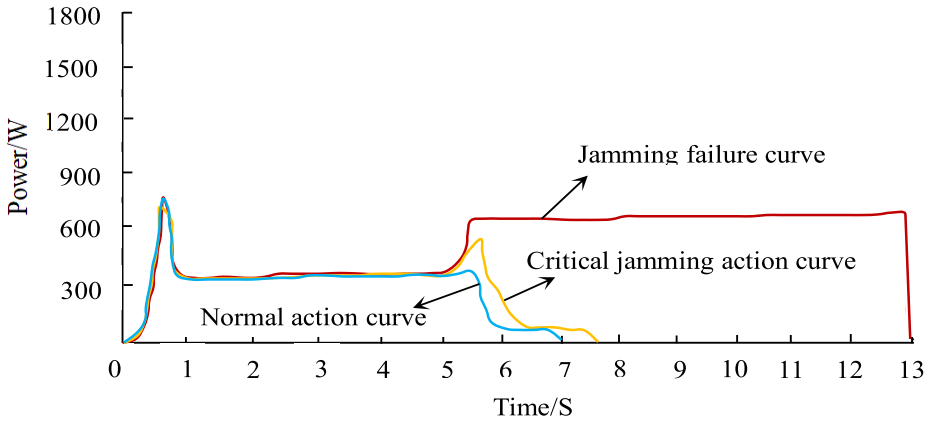

FIGURE 2. Comparison of power curves under different states.

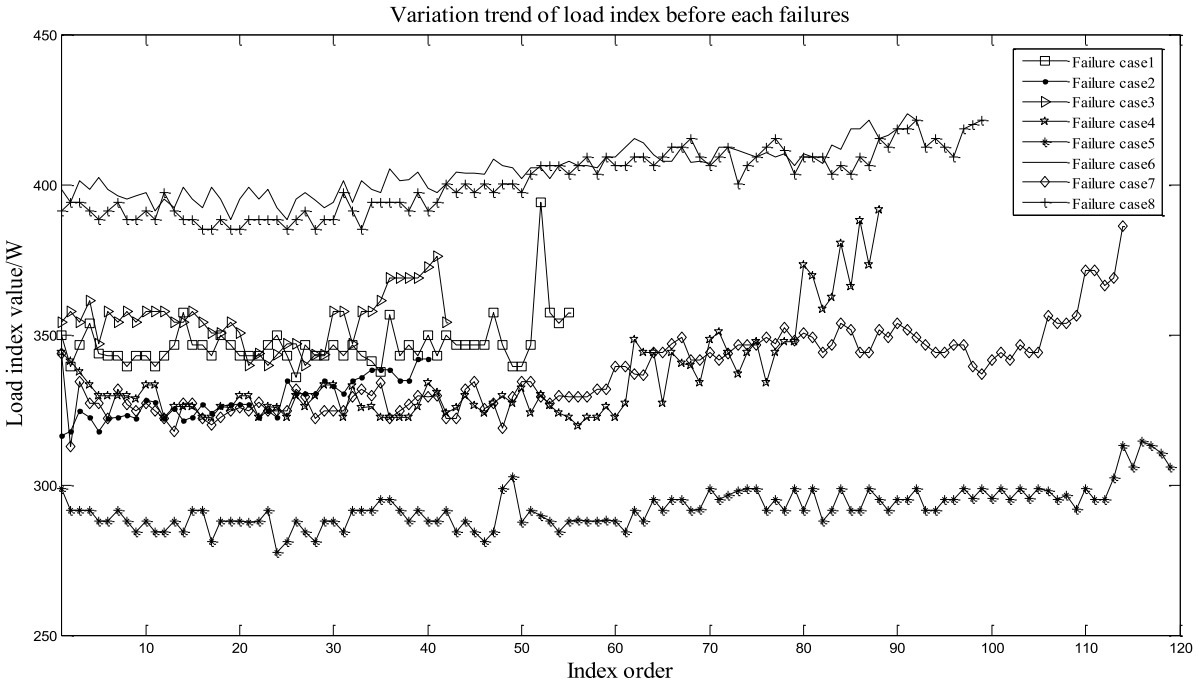

FIGURE 3. Variation trend of load index before failure in each case.

Therefore, the main action parameter that can effectively characterize the jam problem is the point machine power, followed by the action time.

Since a single action of a subway turnout in the same direction requires about 6-7 seconds, we calculate the root mean square (RMS) value for fixed sampling point of power curve during the period from turnout unlocking to locking as the load index, which can highlight the change of power and reflect the length of time. Then, the turnout load index N within one action cycle can be expressed as:

$$N = \sqrt{\frac{\sum_{h=1}^{H} Xh^2}{H}} \tag{1}$$

where Xh is the power value corresponding to the h-th sampling point of the power curve. To verify the correlation between the calculated load index change and the turnout jamming failure, FIGURE 3 shows the change trend of the load index before eight real turnout jamming failures occurred in a subway (Note: the collection object of the verified case here is basically homologous with the case in research [20]). These cases occurred in seven sets of Turnouts at different stations on multiple subway lines, with failure case 8 being a secondary occurrence of failure case 6, which has both diversity and universality.

As shown in FIGURE 3, although there are some fluctuations in the local variations of the load index before each failure case, they all show a clear upward trend overall, which

confirms that the calculated load index can reflect the gradual process of abnormal increase in switching resistance before the turnout jamming failure, and has a prewarning prompting function. Moreover, the characteristics of this chaotic change also suggest that the load index series has short-term predictability.

Based on this, this article continuously calculates the load index of power curves for the previous limited number of actions in the same direction based on the current turnout action completion time, and obtains the generated load index series $\{Ni|i = 1, 2, 3 \ldots\}$ as input for subsequent short-term prediction and prewarning discrimination.

B. THE CONNOTATION AND EVALUATION OF CORE ISSUES TO BE ADDRESSED

As previously pointed out, this article focuses on the three core issues (correctness, accuracy and timeliness) of real-time failure prewarning. The following describes the connotations of these three core issues.

1) ISSUE 1: PREWARNING CORRECTNESS

The correctness of prewarning discrimination is the foundation of the application of the method, which is embodied in the prewarning success rate (PSR) and the false alarm rate (FAR). Next, we give the definitions of the PSR and the FAR among this article.

Definition 1: Taking the operation day as the statistical unit, the ratio of all failure cases (failure days) that can successfully give prewarning before the failure on that day is called PSR.

Definition 2: Taking the operation day as the statistical unit, the ratio of all non-failure cases (non-failure days) that erroneous give prewarning on that day is called FAR.

2) ISSUE 2: PREWARNING ACCURACY

The prewarning accuracy refers to the ability to effectively infer the occurrence time of failure based on the successful prewarning, and further enhance the practical application value of real-time failure prewarning. The definition and calculation approach of inferred failure occurrence time are given below:

Definition 3: From the completion time of the current turnout action, if it is estimated that the failure will occur during the subsequent S-th action, and the current train headway is T minutes, it can be inferred that the failure occurs at $T \cdot S$ minutes after the current time. Considering the uncertainty of the estimated action sequence of failure occurrence, given a penalty term $\varepsilon(\varepsilon \in \mathrm{N}^{+}\&1 \leq \varepsilon \leq S)$ here, so the estimated action sequence range of failure occurrence is $S \pm \varepsilon$. Then, the actual inferred failure occurrence time can be expressed by a range $\tilde{T}$:

$$\tilde{T} = [T\,(S - \varepsilon)\,, T\,(S + \varepsilon)] \tag{2}$$

According to equation (2), it is easy to get that the inferred time range of the failure occurrence spans $2 \cdot T \cdot \varepsilon$.

3) ISSUE 3: PREWARNING TIMELINESS

The prewarning timeliness means that prewarning can be given in time within a limited time range to avoid the meaninglessness of prewarning due to untimely. For the real-time failure prewarning method, if prewarning calculation is carried out every time the turnout completes action, then the time required for a single prewarning calculation $\bar{T}$ must be much less than the prewarning lead time $\widehat{T}$ to ensure the timeliness of prewarning. The following provides the definition and calculation approach of $\widehat{T}$.

Definition 4: Assuming that a prewarning is given after the current turnout action completed, and the actual failure occurs in the subsequent Sa-th turnout action, recorded the current train headway as T minutes, then $\widehat{T}$ is calculated by equation (3):

$$\widehat{T} = T \cdot Sa \tag{3}$$

In summary, this article constructs a prewarning performance evaluation system based on PSR, FAR, $\tilde{T}$, $\bar{T}$ and $\widehat{T}$. Obviously, we expect that the method can give prewarning in time, improve PSR while reducing FAR, and the inferred failure occurrence time is as accurate as possible.

C. OVERALL FRAMEWORK OF THIS METHOD

According to the core issues and expectations to be addressed in this article, FIGURE 4 shows the overall framework of the method implementation.

Compared with the existing two-step process of parameter prediction and prewarning discrimination, this method adds two processes of prewarning feature calculation and failure occurrence time inference, hoping to improve the prewarning correctness, obtain the prewarning accuracy and ensure the prewarning timelines.

III. PROPOSED METHOD

We have constructed a specific real-time prewarning model for turnout jamming failure based on the four-step process of the method framework in FIGURE 4. This chapter first gives the implementation process of the proposed prewarning model, and then introduces the way of each process in detail.

A. PREWARNING IMPLEMENTATION PROCESS DURING OPERATION

Taking the operation period of each day as a unit, the detailed implementation process of the proposed real-time prewarning method for subway turnout jamming failure in daily is show in FIGURE 5.

In FIGURE 5, the content in the dotted line box needs to be handled in advance, while the content in the solid line box is the specific process to be executed. The whole prewarning process starts from the operation of the day to the end of the operation. If a prewarning is given, the process is interrupted and appropriate emergency response measures are taken according to the prewarning information. Next, the core processes in the flow chart: load index series prediction,

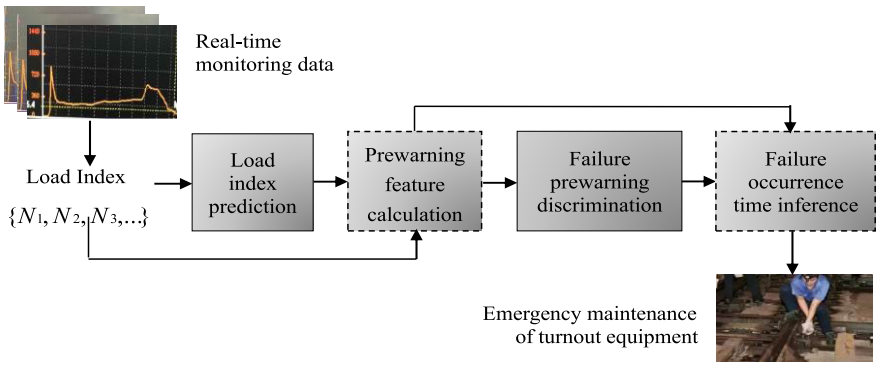

FIGURE 4. Overall framework of the method.

prewarning feature calculation, failure prewarning discrimination, and failure occurrence time inference are introduced in detail.

B. LOAD INDEX SERIES PREDICTION BASED ON PSO-WGPM

Under the idea of this method, short-term prediction of load index involves predicting a finite-length output series from the finite-length input series, which is used to characterize the changes of turnout conversion resistance in the short-term in future. To achieve good prediction accuracy and effectively reflect the prediction trend and fluctuation degree, a PSO-WGPM model is established to improve the short-term prediction effect of load index series here. The specific implementation process is shown in FIGURE 6.

According to FIGURE 6, the core of the entire prediction process is the selection and optimization for base predictor, series sampling weighting, and weight optimization. The details are as follows:

1) THE SELECTION AND OPTIMIZATION FOR BASE PREDICTOR

The base predictor makes the model have prediction ability, and any time series prediction model can be applied. Considering the calculation efficiency and accuracy of short-term prediction, this method also uses the DGM as the base predictor in method [22], to facilitate testing and comparison. Here, the DGM standard structure is improved and optimized, namely:

Introducing weight parameter e and translation parameter c to weight and translate the original series pair-wise to obtain an intermediate series, and then accumulating to generate a prediction input series. Improving prediction performance from the perspectives of series smoothing and adjusting the stepwise ratio. Let $e \in \Omega$ and $c \in \Phi$, use each group (e, c) in the value space to predict the original series in advance, respectively calculate the root mean square error(RMSE) and Spearman distance(SD, namely one minus the sample Spearman's rank correlation coefficient) between the original series and the corresponding simulated prediction series, and then establish an error-correlation constraint equation:

$$\ell(e, c) = RMSE(e, c) + SD(e, c) \tag{4}$$

Traverse the value of (e, c) to obtain the parameter group that minimizes the objective $\ell(e, c)$:

$$(e0, c0) = \underset{e\in\Omega \& c\in\Phi}{\arg\min}\ \ell(e, c) \tag{5}$$

The optimized prediction can be realized by replacing the parameter group $(e0, c0)$ obtained in advance with the prediction model.

2) DESIGN PRINCIPLE OF WEIGHTED PREDICTION

Let the length of the input series $\{Ni\}$ be $2l$, down sample it to get two sets of odd and even complementary series $\{N1, N3, \ldots, N2l-1\}$ and $\{N2, N4, \ldots, N2l\}$. Then use the base predictor to predict $\{N1, N2, \ldots, N2l\}$, $\{N1, N3, \ldots, N2l-1\}$ and $\{N2, N4, \ldots, N2l\}$ separately to obtain $\{ni|i = 1, 2, \ldots, l\}$, $\{\bar{n}i|i = 1, 2, \ldots, l/2\}$ and $\{\tilde{n}i|i = 1, 2, \ldots, l/2\}$. Given the weight series $\{b\}$, the odd-order prediction value $\widehat{no}$ and even-order prediction value $\widehat{ne}$ of the final prediction series $\{\widehat{ni}|i = 1, 2, \ldots, l\}$ can be obtained by interleaving weighting $\{ni\}$ with the corresponding order's $\{\bar{n}i\}$ and $\{\tilde{n}i\}$ respectively. Namely:

$$\widehat{no} = [n + r \cdot \mathrm{sgn}(n - \bar{n})] \cdot b + [\bar{n} + r \cdot \mathrm{sgn}(\bar{n} - n)](1 - b) \tag{6}$$

$$\widehat{ne} = [n + r \cdot \mathrm{sgn}(n - \tilde{n})] \cdot b + [\tilde{n} + r \cdot \mathrm{sgn}(\tilde{n} - n)](1 - b) \tag{7}$$

In the above formulas, r is a floating parameter, whose size is determined by the relative error rate obtained from the prior prediction of the base predictor, and its role is to reasonably expand the range of predicted values so as to better fit the fluctuation situation. Since the action interval of subway turnouts is short and relatively fixed, the action load index series can be regarded as a chaotic time series. Therefore, using equations (6) and (7) to obtain the final prediction series is meaningful in combining multiple time-dimensional information to regulate and predict trends and reflect fluctuations.

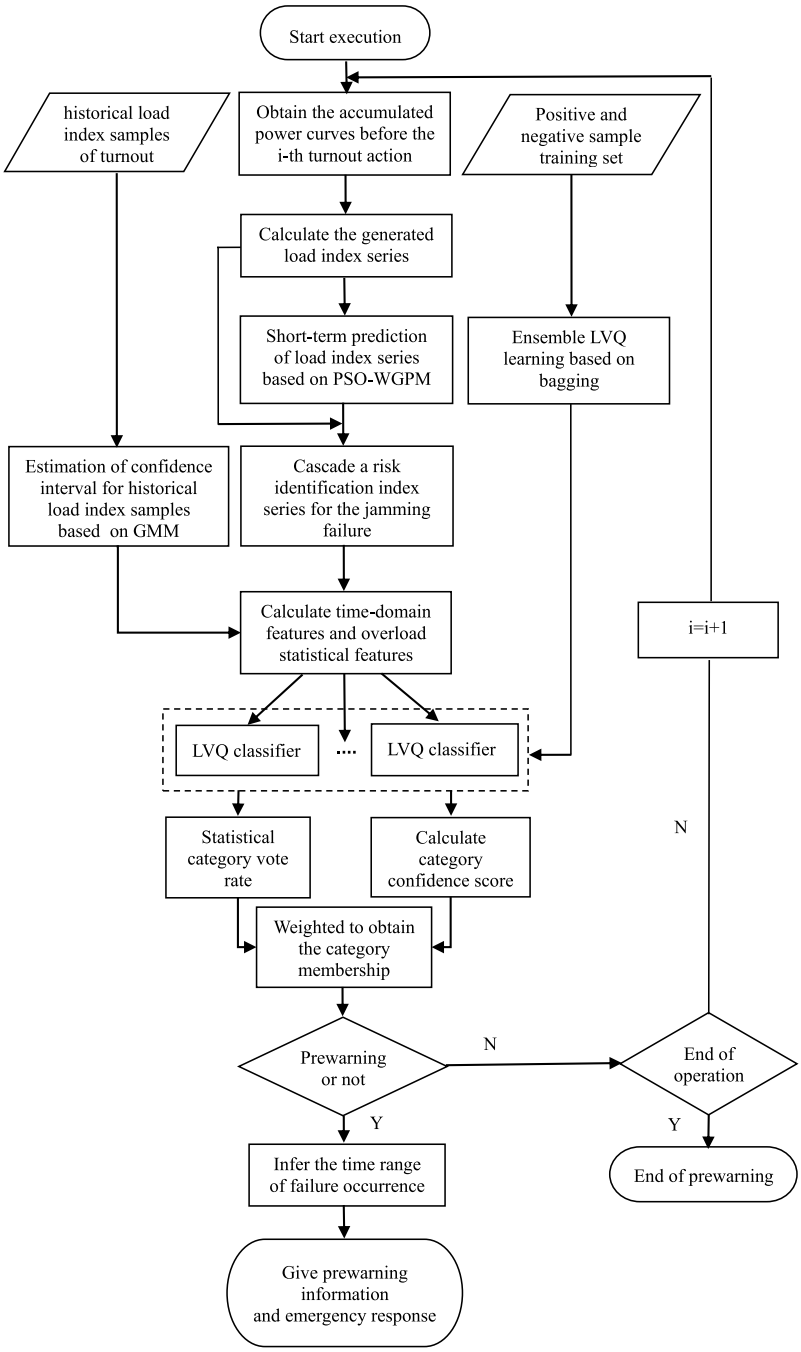

FIGURE 5. Implementation process of real-time failure prewarning during operation.

3) WEIGHT LEARNING OPTIMIZATION

Taking the weighted prediction series $\{\widehat{n}\}$ as the objectives to be optimized, a set of optimal weights is trained by using the known series sample group to make the prediction value of each step closer to the true value. Given the weight value range B, and set the true value corresponding to the single-point prediction value at each step as V, the prediction loss *loss* of the prediction series under the weight sequence $\{b1, b2, \ldots, bl\}$ is also defined based on RMSE and SD:

$$loss = \sqrt{\frac{\sum_{i=1}^{l}\left(\widehat{ni} - Vi\right)^2}{l}} + \lambda \cdot \left(\frac{6 \cdot \sum_{i=1}^{l}\left[\mathrm{rg}(\widehat{ni}) - \mathrm{rg}(Vi)\right]^2}{l \cdot \left(l^2 - 1\right)}\right) \tag{8}$$

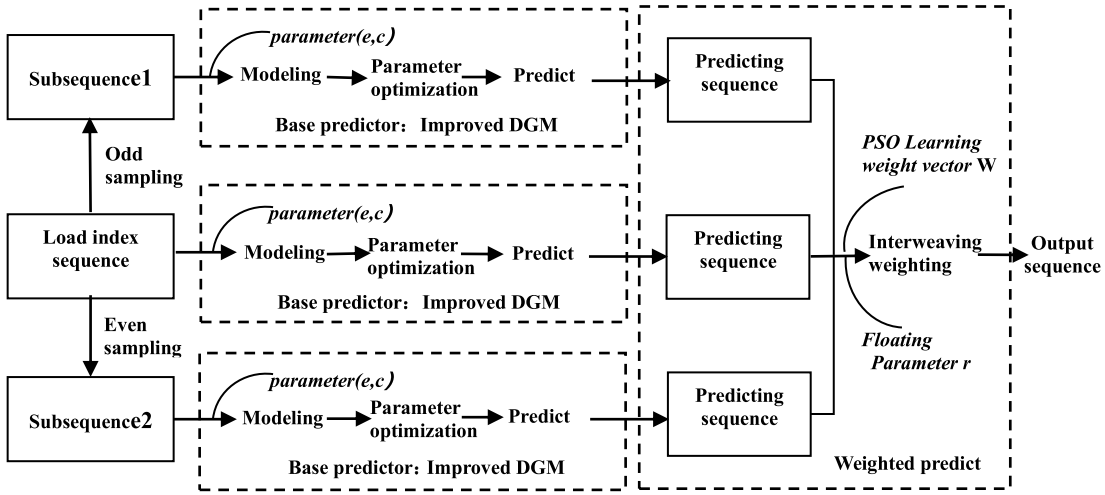

FIGURE 6. Sequence prediction process based on PSO-WGPM.

where λ is the regularization parameter and $\mathrm{rg}(\cdot)$ is used to calculate the rank. Further covering Q samples, the learning objective is to obtain a weight vector $\mathbf{W}_{1\times l}$ that minimizes the total loss, which can be expressed as:

$$\mathbf{W}_{1\times l} = \arg\min_{B} \frac{1}{Q} \sum_{Q} loss_{(\mathbf{W}_{1\times l})} \tag{9}$$

For solving the learning objective, here we use particle swarm optimization (PSO) to optimize the weights after limiting the changes range of weight. The implementation process of standard PSO can be referred to [25].

C. MULTI-DIMENSIONAL HYBRID PREWARNING FEATURE EXTRACTION

Due to the limited length and inevitable error of the load index series obtained from short-term prediction, it is not suitable to directly use the series prediction value itself to characterize the turnout action state. Based on the completion time of the current turnout action, the predicted load index series $\{\widehat{ni}|i = 1, 2, 3, \ldots, l\}$ and the generated load index series $\{Ni|i = 1, 2, 3 \ldots, l\}$ are cascaded into a risk identification index series $I = N||\widehat{n}$, which is used to characterize the change in the resistance to conversion of the turnout from the past to the future. Then the time-domain characteristics and overload statistical characteristics of $\{I\}$ are extracted for subsequent prewarning discrimination.

1) TIME-DOMAIN FEATURES EXTRACTION

The time-domain features are mainly used to reflect the fluctuation and change trend of the series itself. Referring to methods [20], [26], this article calculates four parameters as time-domain characteristics, including peak to peak $F1$, standard deviation $F2$, acceleration peak factor $F3$, and Mann-Kendall trend test value $F4$. The calculation method is as follows:

$$\left.\begin{aligned} F1 &= \max(\{Ii\}) - \min(\{Ii\}) \\ F2 &= \sqrt{\tfrac{1}{2l-1}\textstyle\sum_{i=1}^{2l}(Ii - \bar{I})^2} \\ F3 &= \max(\{Ii\}) \Big/ \sqrt{\tfrac{1}{2l}\textstyle\sum_{i=1}^{2l} Ii^2} \\ F4 &= Func\_MK(\{Ii\}) \end{aligned}\right\} \tag{10}$$

where, $\bar{I}$ denotes the average value of $\{I\}$; $Func\_MK(\cdot)$ is employed to compute the Mann-Kendall trend test value, and the computation process is omitted.

2) OVERLOAD STATISTICAL FEATURES CALCULATION

Overload statistical features refer to the statistics that the series value of risk identification index exceeds the normal upper limit of load, which can describe the overload working state of turnout from a statistical perspective. Since the action load limit of the turnout is closely related to the equipment usage and work environment, there is no unified standard. Therefore, the action load limit of each turnout is determined by calculating the confidence interval of the historical load index samples. In case of sufficient samples, this article uses the Gaussian mixture model (GMM) [27] with good approximation performance to estimate the overall value confidence interval of the historical load index samples. The main implementation steps are as follows:

Step1. Collect a large number of power curves generated by the action in the same direction on the non- failure day of turnout to calculate the load index, and establish the sample set $\{x\}$ of historical load index.

Step2. The one-dimensional probability density estimation expression $\rho(x)$ of $\{x\}$ is established based on GMM, namely:

$$\begin{aligned} \rho(x) &= \sum_{k=1}^{K} \omega k \cdot \phi\left(x|\mu k, \sigma k\right) \\ &= \sum_{k=1}^{K} \omega k \cdot \frac{1}{\sqrt{2\pi k}\delta k} \exp\left(-\frac{(x-\mu k)^2}{2\sigma k^2}\right) \end{aligned} \tag{11}$$

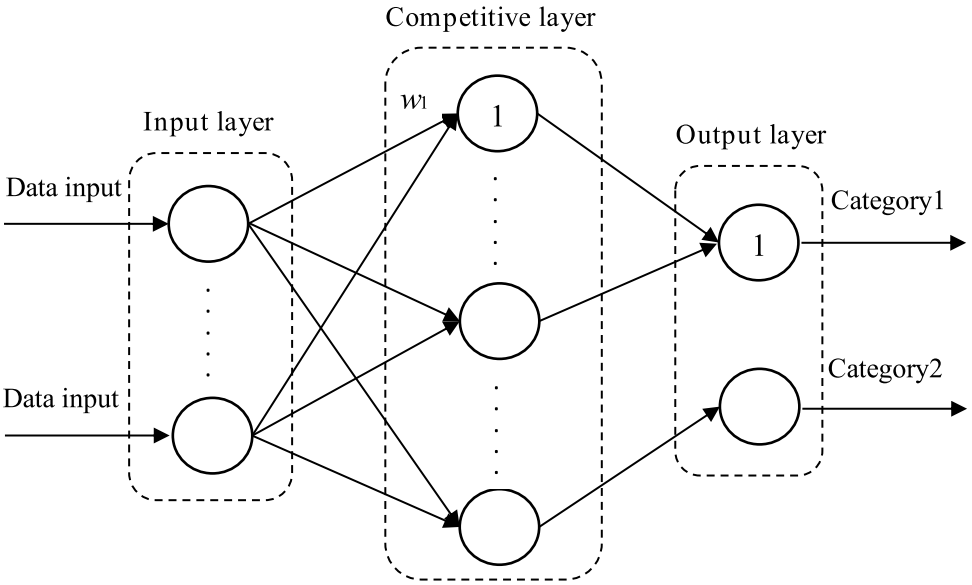

FIGURE 7. Basic binary LVQ network topology.

In formula (11), $\phi(x|\mu k, \sigma k)$ represents the k-th one-dimensional Gaussian distribution used for superimposed mixing, and the superimposed weight ωk satisfies $\omega 1 + \omega 2 + \ldots + \omega K = 1$. μk and σk represent the mean and standard deviation of $\phi(x|\mu k, \sigma k)$ respectively, and K represents the mixing order.

Step3. Use the expectation maximization (EM)algorithm to determine ωk, μk and σk for $\rho(x)$. its process and parameter selection refer to the paper [28].

Step4. Given the 95% confidence level, the calculation relationship between $\rho(x)$ and the overall value confidence interval $[Al, Ah]$ of $\{x\}$ is as follows:

$$\int_{Al}^{Ah} \rho(x)\, dx = 0.95 \tag{12}$$

According to equation (12), the upper bound Ah of the confidence interval at 95% confidence level is calculated as the lower limit to measure whether the load index exceeds the standard. Then, for the elements in $\{I\}$, count the number $F5$ of exceeding Ah and calculate the cumulative sum $F6$ of exceeding Ah as overload statistical characteristic parameters:

$$\left.\begin{array}{l} F5 = accout(Ii \geq Ah)\ i = 1, 2, 3, \ldots, 2l \\ F6 = acsum(Ii - Ah)\ onlyif\ \ Ii \geq Ch \end{array}\right\} \tag{13}$$

where, $accout(\cdot)$ and $acsum(\cdot)$ are conditional counting function and cumulative summation function, respectively.

In summary, this article extracts parameters $F1 \sim F6$, to form a multi-dimensional hybrid prewarning feature set with time continuity, which is used for the subsequent judgment of jamming failure prewarning.

D. FAILURE PREWARNING DISCRIMINATION BASED ON ENSEMBLE LVQ

After extracting the prewarning feature parameters of the risk identification index series, the classifier is trained for prewarning discrimination. It is clearly that this is a binary classification problem of ''prewarning or not''. We select LVQ neural network as the base classifier, and then an ensemble learning strategy is designed to enhance the classification performance.

1) LVQ NETWORK TRAINING

LVQ is essentially a prototype supervised clustering algorithm, belonging to the self-organizing competitive neural network [29], and has high accuracy and learning efficiency for classification with low parameter dimensions and few recognition types. In particularly, LVQ has certain noise suppression ability, which can help to reduce the FAR in this research. Its efficiency, robustness and generalization ability are suitable for the classification environment here. A basic binary LVQ network topology is shown in FIGURE 7.

It can be seen that the input layer and the competition layer of LVQ neural network are fully connected, the competition layer and the output layer are partially connected, and the value of the output neuron is only 0 or 1.

The core idea of its network training is that each competing neuron represents a central point, and the input data belongs to the category connected to which neuron is close to the central point(that is, when the competition is successful, set the central point position to 1, as $w1$ shown in the FIGURE 7, and the output connected to it set to 1 too). Next, the basic training steps of LVQ network are given below:

Step1. Initialize the weight wpj and learning rate η ($\eta > 0$) between the input layer and competition layer.

Step2. Send the input vector value $\mathbf{v} = (v1, v2, \ldots, vR)^{\mathrm{T}}$ to the input layer, and calculate the distance d between the competition layer neurons and the input vector:

$$dp = \sqrt{\sum_{j=1}^{R} (vj - wpj)^2} \quad p = 1, 2, \ldots, U \tag{14}$$

where, R is the dimension of the input vector and U is the number of competing neurons.

Step3. Select the competing neuron with the smallest distance from the input vector. If dp is the smallest, record the class label of the linear output layer neuron connected to it as Cp

Step4. Note that the class label corresponding to the input vector is Cv. if $Cp = Cv$, then use the following equation (15) to adjust the weight:

$$w\_new = w\_old + \eta \cdot (v - w\_old) \tag{15}$$

Otherwise, update the weight according to the following equation (16):

$$w\_new = w\_old - \eta \cdot (v - w\_old) \tag{16}$$

After reaching the specified iterations, the training is completed and can be used for test classification.

2) ENSEMBLE LEARNING DECISION BASED ON HYBRID VOTING

In the subway site, the number of turnout action failure is very few compared with the number of normal actions, so the turnout failure prediction and diagnosis methods based on machine learning belong to the small-scale sample learning tasks with positive and negative proportion imbalance [2], [30]. In this case, using a single LVQ classifier for prewarning discrimination is unstable. Inspired by method [31], this paper uses ensemble learning to improve the discrimination performance. According to the error analysis theory of machine learning [32], in the learning environment with insufficient data, the generalization error is greatly affected by variance. Therefore, an ensemble learning idea based on bagging [33] is used to improve the classification variance of LVQ. The specific implementation steps are as follows:

Step1. Calculate the load index series of $2l$ actions before the occurrence of each turnout jamming failure to form a positive sample resource pool, with the category label $m = 2$(the category of ''prewarning required'').Then, the load index series of any continuous $2l$ actions during the non-failure period of each turnout is calculated to form a negative sample resource pool, with the category label $m = 1$(the category of ''no prewarning required'').

Step2. Based on bootstrap sampling, D index series sets with the same capacity are extracted from the positive and negative sample resource pools respectively, and combined to form D training sets of positive and negative samples 1:1.

Step3. Firstly, the D training sets obtained in the previous step are used to train D LVQ networks to form a base classifier group; Then, the positive samples collected in Step 1 are used as the test group to perform prewarning discrimination tests on each base classifier, and statistical test results regard as a priori prewarning rate f; Finally, the prior prewarning rate of the base classifier is ranked from large to small, and the first g base classifiers are selected for the final integrated decision.

Step4. For the prewarning feature set to be identified, use the g base classifiers obtained in the previous step to simultaneously perform prewarning discrimination, and then calculate the number of votes $Votes(m)$ for the two categories, then the vote rate Xm of category m is:

$$Xm = Votes(m)/g \tag{17}$$

In addition to calculating Xm, the results of the base classifier with a higher PSR should be given a greater weight, so we further calculate the confidence score Ym corresponding to category m:

$$Ym = \sum_{t=1}^{g} \frac{ft}{\sum\limits_{1 \to g} f} ||Lt - m| - 1| \tag{18}$$

where, ft corresponds to the prior prewarning rate of the t-th base classifier, and Lt refers to the category label of the t-th base classifier for prewarning judgment of the identified object.

Step5. Weighting the vote rate Xm and the confidence score Ym to obtain the final category membership degree Zm, which is given by:

$$Zm = \tau \cdot Xm + (1 - \tau) \cdot Ym \tag{19}$$

where, τ is the weight. Obviously, the corresponding category with the largest membership degree is output as the final prewarning discrimination result. If two categories have the same membership degree, the default output category is ''prewarning required''.

E. INFERENCE OF FAILURE OCCURRENCE TIME

According to definition 3, the time range of failure occurrence can be determined by inferring the turnout action sequence at failure and combining with the train headway. Considering that the inference of the action sequence at failure is fault tolerant, it is transformed into a simple probability problem for uncertainty treatment as follows:

First, we assume that the turnout is prone to jamming failure when the action overload rate reaches $\bar{O}$ in l actions ($\bar{O}$ can be obtained by using known failure samples easily).Then, starting from the action sequence when the prewarning is given at first time, the turnout action sequence S of subsequent failure occurrence can be estimated by formula (20):

$$S \approx \left\lceil \frac{l \cdot \bar{O} - 1}{\tilde{O}} \right\rceil \tag{20}$$

In the formula(20), $\lceil \cdot \rceil$ represents the upward rounding, and $\tilde{O}$ is the overload rate of the elements in the series sequence $\{I\}$ at the first prewarning. According to the joint formula (2)

TABLE 1. Main parameters setting.

Parameter type	Specific requirements for indicators
Index sequence prediction parameters	①Prediction index series length $l=10$,and input series length $2l=20$ (Remark: There are abnormal changes in the index within the 10 actions before the failure, generally)
	②Weighted floating parameter $r=0.015$ (Remark: The value is determined by the average relative error rate of the improved DGM in this article)
Parameters for PSO optimization	③The value range of weight b is from 0.1 to 0.9 in steps of 0.1. Set the population to 25 and the evolution times to 80
Main parameter of GMM	Mixed order K is set to 2 (Remark: Refer to the literature [34,35], $K=2$ can accurately fit the non-Gaussian probability distribution)
Parameters of LVQ network training	Set the number of competitive neurons to 20, the learning rate to 0.01, and the number of iterations to 1000
Parameters of ensemble learning	D is set to 500 and g is set to 15.

TABLE 2. Statistical test and comparison of load index prediction.

Group prediction / Evaluating indicator	Group1（0≤sd＜3）		Group2（3≤sd＜6）		Group3（6≤sd＜9）		Group4（sd≥9）	
	PSO-WGPM	DGM	PSO-WGPM	DGM	PSO-WGPM	DGM	PSO-WGPM	DGM
RMSE-ave	**2.81**	2.84	**4.92**	6.18	**6.56**	9.28	**9.61**	14.87
MAPE-ave	**0.67%**	0.72%	**1.22%**	1.61%	**1.59%**	2.51%	**2.44%**	3.62%
SD-ave	**0.52**	0.59	**0.77**	1.1	**0.74**	0.96	**0.79**	1.23

and (20), it can be inferred that the time range of future failure occurrence as:

$$\tilde{T}=\left[T\left(\left\lceil\frac{l\cdot\bar{O}-1}{\tilde{O}}\right\rceil-\varepsilon\right),T\left(\left\lceil\frac{l\cdot\bar{O}-1}{\tilde{O}}\right\rceil+\varepsilon\right)\right] \tag{21}$$

IV. EXPERIMENTS

This chapter first introduces the setting of some main parameters of the model in this article, and then conducts the load index prediction test, prewarning correctness test, prewarning accuracy test and prewarning timeliness test. Through analysis for the test results, it demonstrates the advantages of the method in alleviating the three core issues of prewarning.

A. TEST CASES AND MAIN PARAMETERS SETTING DESCRIPTION

All the instance data in the experimental test are from the real jamming failure monitoring data of multiple turnouts in a subway, and the monitoring data for the 5 non-failure days before and after the failure. The setting requirements of main parameters are summarized in TABLE 1.

B. SHORT-TERM LOAD INDEX PREDICITON TEST

Some turnouts are selected as objects, and a large number of historical load index series with length 30 are collected to form a sample set, which is divided into two parts: training set and test set. Among them, the training set is used for PSO-WGPM weight optimization, while the test set is divided into four test groups with different fluctuation degrees according to standard deviation (sd). Then, proposed PSO-WGPM and DGM in [22] are used to test the prediction effect of the four test groups respectively, and the statistical average values of the three evaluation indicators, namely RMSE, SD, and mean absolute percentage error (MAPE), are recorded in TABLE 2 to evaluate the prediction effect.

Furthermore, one test instance is extracted from each of the four test groups to visually demonstrate the changes of the predicted value and the real value. The results are shown in FIGURE 8.

According to the test results in TABLE 2 and FIGURE 8, for group1 with minimal fluctuation, the prediction accuracy of PSO-WGPM and DGM is comparable, both achieving good prediction results. However, as the fluctuation of the prediction object increases, the prediction accuracy of PSO-WGPM becomes more dominant compared to DGM. For group4 with the highest fluctuation, the three

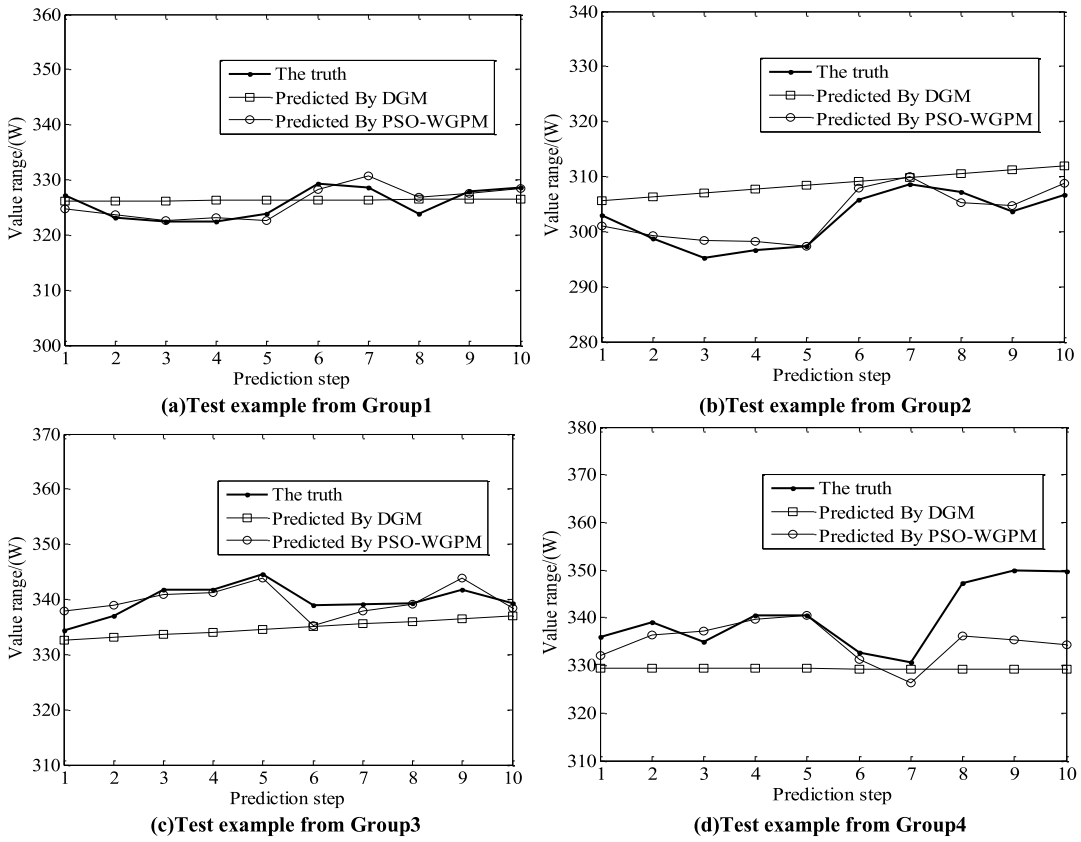

FIGURE 8. Load index series prediction instances.

evaluation indicators RMSE, MAPE and SD obtained using PSO-WGPM decreased by 35.4%, 32.6% and 35.8% respectively compared with DGM. Further analysis of the test results for the two methods: DGM usually accumulates the original series without obvious rules to generate an intermediate series with increasing trend for fitting prediction, so the prediction series can reflect the overall variation trend, but it is difficult to effectively fit the local fluctuation changes, so the prediction error will increase significantly as the fluctuation of the measured series intensifies. The proposed PSO-WGPM uses the improved DGM as a base predictor and designs a composite prediction mode based on interleaving weighting, it improves the prediction accuracy and generalization performance by optimizing the weights, so that the prediction results are not only closer to the real value, but also can effectively reflect the fluctuation degree and local variation trend.

Due to the obvious fluctuations and upward trend in load index before turnout jamming failures, it is more meaningful to improve the prediction performance of sequences with large fluctuations. Overall, the proposed prediction model can meet the short-term prediction and prewarning discrimination requirements in this article.

C. PREWARNING CORRECTNESS TEST AND ANALYSIS

We use the real cases data to review the operation process of the day, and then conduct prewarning discrimination testing based on the implementation process shown in FIGURE 5, and the test is suspended after the prewarning is given at first time. In addition, since the existing similar methods have their own specific application scenarios and the parameters are not all disclosed, it is difficult to directly compare them. Therefore, ablation experiments are designed to test the method and conduct longitudinal comparison.

1) ABLATION EXPERIMENT SETTING(i)

Maintain the load prediction and prewarning discrimination framework unchanged, refer to the types of existing methods, derive two special situations by changing the structure of the load index series. The derivative method 1 analogy methods [13], [14], [20], and only uses the generated load index series for prewarning discrimination. The derivative method 2 analogy methods [21], [22], [23], [24], and directly uses the predicted load index series (the prediction length is doubled) for prewarning discrimination. Then use the proposed method and the two derived methods to perform prewarning discrimination for all cases. The partial results of the turnout action sequence at the first prewarning (abbreviated as 'TASFP') of each case are recorded in TABLE 3. Among them, cases 1 to 8 are the failure cases shown in FIGURE 3, and cases 9 to 16 are the non-failure cases at the day before or after the failure day corresponding to cases 1 to 8, respectively.

TABLE 3. Comparison of prewarning discrimination replay tests.

Case number	Turnout action sequence when the actual failure occurrence	TASFP of this method	TASFP of derivative method 1	TASFP of derivative method 2
1	56	47	52	36
2	41	33	**N/A**	25
3	43	35	41	33
4	89	80	82	80
5	120	114	116	49
6	94	87	93	63
7	115	109	111	107
8	100	90	98	68
9	No failure occurred	N/A	N/A	N/A
10	No failure occurred	N/A	N/A	N/A
11	No failure occurred	N/A	N/A	N/A
12	No failure occurred	N/A	N/A	N/A
13	No failure occurred	N/A	N/A	**57**
14	No failure occurred	N/A	N/A	N/A
15	No failure occurred	N/A	N/A	N/A
16	No failure occurred	N/A	N/A	N/A

The ''N/A'' in TABLE 3 indicates that the event has not occurred and there is no data record. In overall view to the test results in TABLE 3, there are only one missed prewarning and one false alarm in the three methods, indicating that the proposed prewarning feature calculation strategy and prewarning discrimination method are effective. Further detailed analysis of the test results for each method as follows:

The advantage of derivative method 1 is that there is no false alarm, but also has problem of prewarning untimely or even missed. This is because if only the generated load index series is used for prewarning discrimination, the prewarning will not be given until the turnout action load index deteriorates to a certain extent. Therefore, this method is not prone to false alarm due to the sporadic overload; but for cases 3, 6, and 8 with relatively gentle deterioration process, prewarning are only given near the failures, resulting in prewarning untimely; In case 2, not only the failure progressed smoothly, but also the overload degree before the failure is not enough, which directly led to prewarning missed.

The advantage of the derivative method 2 is that it has no prewarning missed and can catch failure signs early, but also has problem of prone to false alarm. This is because method 2 only uses predicted load index series for prewarning discrimination, and the predicted values are often very sensitive to abnormal increases in current load index, and cumulative errors can occur as the number of prediction steps increases, resulting in poor robustness. For instance, the power of failure case 8 began to rise abnormally at the 68-th action. The method is very sensitive to catch the symptom of the failure and gave prewarning; While non-failure case 13 is generally normal during operation, but sporadic abnormal power increase occurred during the 57-th action, resulting in false alarm. In fact, the working conditions of the turnout system are complex and changeable, and many factors such as environmental changes, power fluctuations, motor characteristics changes may all lead to occasional abnormal power rise, resulting in an increase of FAR.

Considering the advantages and disadvantages of derivation methods 1 and 2, this article adopts a composite load index series structure that can combine the two situations to form a complementary approach to enhance overall robustness. According to the test results, each case of failure was successfully prewarning at the 6th-10th turnout action before the failure, and there were no false alarms in non-failure cases. It can be concluded that this method can not only ensure the PSR, but also effectively control FAR, and can avoid the problem of prewarning untimely, with better comprehensive performance.

2) ABLATION EXPERIMENT SETTING(ii)

Keeping the load index series structure and prediction framework unchanged, we test the effect of ensemble learning and non-ensemble learning for prewarning discrimination respectively, and add the SVM used in methods [7], [17], [18] as the base classifier for comparison. This prewarning discrimination test uses all failure samples and non-failure samples collected, and then calculate PSR and FAR of various methods. Moreover, we also add the most commonly used binary classification evaluation metric F1-Score under imbalanced data to comprehensively reflect the correctness of prewarning (the definition and computing of F1-score refer to [36]). TABLE 4 shows the test results of ensemble LVQ (i.e. used in this article), single LVQ, ensemble SVM and single SVM, in which the single classifier is trained by positive and negative samples 1:1.

TABLE 4. Prewarning effect of different discrimination methods.

Discrimination method / Evaluating indicator	Ensemble LVQ	Single LVQ	Ensemble SVM	Single SVM
PSR	100%	95%	100%	100%
FAR	1.43%	1.43%	5.71%	4.29%
F1-Score	0.98	0.95	0.91	0.93

Analyze the results in TABLE 4: single LVQ has 5% missed alarms, but its FAR is the lowest at 1.43%, and F1- score is second only to ensemble LVQ. From the perspective of algorithm principle, the classification of LVQ in the competitive layer depends only on the distance between input vectors. If two input vectors are very close, they are likely to be classified into one category, and then supervised learning is used in the output layer to achieve accurate classification. This makes LVQ have a certain degree of noise suppression, which can effectively reduce FAR, but also affects the PSR to some extent. After ensemble learning based on bagging, the problem of unbalanced small-scale sample learning environments can be alleviated, and the classification variance of LVQ model is improved, so that the PSR of ensemble LVQ increased to 100%, the FAR remained at the lowest 1.43%, and F1-score reached the best 0.98.In contrast, single SVM has a 100% PSR, but its FAR also reaches at 4.29%, indicating that SVM is more sensitive to positive samples in this learning environment, resulting in higher FAR, so the F1-score is lower than single LVQ. It is worth noting that the SVM after ensemble learning does not improve the classification effect of single SVM, and even slightly increases FAR. This is because the final decision of SVM is determined by a small number of support vectors, and the change of negative samples has limited impact on the classification results. Therefore, it is difficult to optimize the classification variance of SVM model to improve the overall prediction performance through ensemble learning based on bagging. Therefore, the LVQ after ensemble learning is more suitable for the prewarning discrimination environment in this article.

D. PREWARNING CORRECTNESS TEST AND ANALYSIS

For all failure cases with successful prewarning in the previous section, first infer the action sequence range of future failures occurrence, and then calculate the ratio of action sequence when the actual failure occurrence within the inferred range as α. Obviously, when α is higher and ε is smaller, the failure prewarning is more accurate, but the fact is that α will rise with the increase of ε. therefore, an equation (22) that can balance the value of α and ε is defined here as the prewarning accuracy metric *Mac*:

$$Mac = \frac{1}{2}\left(\alpha - \log_l \varepsilon + 1\right) \tag{22}$$

Equation (22) limits *Mac* to [0,1]. The closer *Mac* is to 1, the more accurate the prewarning is.

Next, calculate and analyze the change in *Mac* when ε takes different values according to equations (20) and (22), where $\bar{O}$ is equal to 0.6. The test result is shown in FIGURE 9.

From FIGURE 9, when ε is 2, *Mac* reaches the highest **0.774**. At this time, **85%** of the cases have the action sequence within the inferred range when the actual failures occurrence, and the inference error of the remaining cases is one action outside the range. In addition, combined with the train headway at the time of each failure, the inferred time range of the failure occurrence spans can be accurate to within **10-24minutes**.

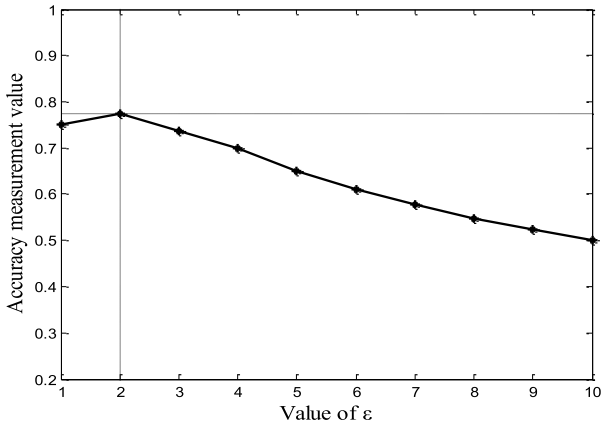

FIGURE 9. Prewarning accuracy test result.

E. PREWARNING TIMELINESS TEST AND ANALYSIS

According to Chapter II, whether the prewarning is timely or not is mainly determined by the single prewarning calculation time $\bar{T}$ and the prewarning lead time $\widehat{T}$. The following provides a detailed analysis of $\bar{T}$ and $\widehat{T}$ obtained from the actual testing:

1) SINGLE PREWARNING CALCULATION TIME $\bar{T}$

In the entire method, the predictor and classifier need to be trained in advance, while in practical application, the time of performing a single prewarning calculation is mainly consumed in two aspects: load index calculation and prediction, prewarning classification and discrimination. A large number

of measured statistics show that the average time for load index calculation and prediction is **3.83 seconds**, the average time for prewarning classification and discrimination is **0.08 seconds**, and the average total time is **3.91 seconds**, which clearly meets the requirements of real-time analysis.

2) PREWARNING LEAD TIME $\widehat{T}$

Firstly, it can be determined that the lower limit $\widehat{T}$ min of $\widehat{T}$ is the time of one train headway (that is, the prewarning will be given at the last turnout action before the failure at least). Since train headways on different subway lines may be inconsistent, the minimum design headway of 120 seconds for train turnback in the CBTC (communication based train control) signal system is taken as the theoretical value for $\widehat{T}$ min here. According to the prewarning discrimination test for all failure cases, the prewarning lead time varies from **18 to 60 minutes.** Then, calculate $\widehat{T}/\bar{T}$ as the metric value of prewarning timeliness for each failure case. The test results for partial cases are shown in FIGURE 10.

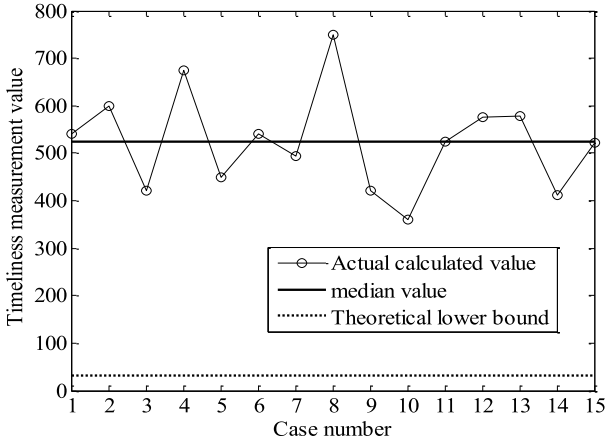

FIGURE 10. **Prewarning timeliness test result.**

It can be seen that the prewarning timeliness metric value of each test case in FIGURE.10 ranges from 360 to 750, with a median of 525. Therefore, it can meet the requirement that the prewarning calculation time is far less than the prewarning lead time.

In summary, the method proposed in this article has good prewarning timeliness, and can infer the time range of turnout jamming failure occurrence within a small span in advance, facilitating timely and targeted emergency measures for failure prevention on site. For example, in the case of tight time constraints, it is possible to directly replace the turnback line to avoid failure, and when time is abundant,, temporary measures such as oiling can be used to reduce the resistance of turnout conversion.

V. CONCLUSION

This article focuses on the research of real-time prewarning of subway turnout jamming during operation, which has important practical significance. The research focuses on the three core issues: the correctness, accuracy and timeliness of prewarning, so that the method can issue prewarning successfully within the appropriate time and effectively control false alarms. At the same time, it tries to infer the time range of failure occurrence, breaking through the limitations of existing methods in terms of the prewarning accuracy. In order to achieve these purposes, the following work has been done:1. established PSO-WGPM model to realize the short-term accurate prediction of action load index; 2. optimized the structure of load index series, combined time-domain characteristics with overload statistical features to obtain a multi-dimensional hybrid prewarning feature set, providing a basis for effective prewarning analysis; 3. designed an ensemble LVQ learning model based on hybrid voting strategy for final prewarning discrimination; 4. employed the overload rate of the predicted index series to infer the time range of failure occurrence. Comprehensive experimental testing using real failure/non-failure case data from subway systems has verified that the method is simple and effective, with good practical value.

However, due to the uncertainty of index prediction and the lack of learning samples, the method still has a certain proportion of false alarms, and the accuracy of prewarning also needs to be improved. The potential solution for the former is to directly use uncertainty prediction instead of deterministic prediction to improve the fault tolerance of the method. The potential solutions for the latter are mainly to expand the data source set or use the cutting-edge small-scale sample learning technology for prewarning discrimination. How to apply the method migration to other types of failure prewarning tasks is also the next research goal.

REFERENCES

[1] C. Chen, "Condition monitoring and fault prediction of railway turnout," M.S. thesis, Beijing Jiaotong Univ., Beijing, China, 2020.

[2] D. Ou, R. Xue, and K. Cui, "A data-driven fault diagnosis method for railway turnouts," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 4, pp. 448–457, Apr. 2019, doi: 10.1177/0361198119837222.

[3] X. Hu, Y. Cao, T. Tang, and Y. Sun, "Data-driven technology of fault diagnosis in railway point machines: Review and challenges," *Transp. Saf. Environ.*, vol. 4, no. 4, Dec. 2022, Art. no. tdac036, doi: 10.1093/tse/tdac036.

[4] O. F. Eker, F. Camci, A. Guclu, H. Yilboga, M. Sevkli, and S. Baskan, "A simple state-based prognostic model for railway turnout systems," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1718–1726, May 2011, doi: 10.1109/TIE.2010.2051399.

[5] O. F. Eker and F. Camci, "State-based prognostics with state duration information," *Qual. Rel. Eng. Int.*, vol. 29, no. 4, pp. 465–476, Jun. 2013, doi: 10.1002/qre.1393.

[6] Y. Chen, Q. Dai, and J. Li, "Research on the fault prognostics model of the switch machine based on HSMM optimized by SA-CPSO," *J. Railway Sci. Eng.*, vol. 16, no. 4, pp. 1050–1057, 2019.

[7] V. Atamuradov, K. Medjaher, F. Camci, P. Dersin, and N. Zerhouni, "Railway point machine prognostics based on feature fusion and health state assessment," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 8, pp. 2691–2704, Aug. 2019, doi: 10.1109/TIM.2018.2869193.

[8] Q. Meng, "Research on fault diagnosis and degradation state prediction for electro-hydraulic switch equipment," Beijing Jiaotong Univ., Tech. Rep., 2023.

[9] C. Chen, T. Xu, G. Wang, and B. Li, "Railway turnout system RUL prediction based on feature fusion and genetic programming," *Measurement*, vol. 151, Feb. 2020, Art. no. 107162, doi: 10.1016/j.measurement.2019.107162.
[10] H. Wen and B. Yang, "Intelligent decision system for urban rail transit switch repairing based on abnormal granularity and three-way decision theory," *Urban Mass Transit*, vol. 24, no. 12, pp. 136–140, 2021.
[11] J. Gao, "Fault diagnosis and early warning system of turnout equipment based on PHM," Dalian Jiaotong Univ., Tech. Rep., 2023.
[12] C. Li and L. Zhao, "CACC-RF-based risk prediction of railway switch gap jam fault," *J. China Railway Soc.*, vol. 44, no. 6, pp. 46–55, 2022.
[13] H. Kim, J. Sa, Y. Chung, D. Park, and S. Yoon, "Fault diagnosis of railway point machines using dynamic time warping," *Electron. Lett.*, vol. 52, no. 10, pp. 818–819, May 2016, doi: 10.1049/el.2016.0206.
[14] Z. Zhong, J. Chen, T. Tang, T. Xu, and F. Wang, "SVDD-based research on railway-turnout fault detection and health assessment," *J. Southwest Jiaotong Univ.*, vol. 53, no. 4, pp. 842–849, 2018.
[15] Y. Zhang, T. Xu, C. Chen, G. Wang, Z. Zhang, and T. Xiao, "A hierarchical method based on improved deep forest and case-based reasoning for railway turnout fault diagnosis," *Eng. Failure Anal.*, vol. 127, Sep. 2021, Art. no. 105446, doi: 10.1016/j.engfailanal.2021.105446.
[16] X. C. Wu and X. Chu, "Research on division of degradation stage of turnout equipment based on wavelet packet decomposition and GG fuzzy clustering," *J. China Railway Soc.*, vol. 44, no. 1, pp. 79–85, 2022.
[17] Y. Sun, Y. Cao, and P. Li, "Contactless fault diagnosis for railway point machines based on multi-scale fractional wavelet packet energy entropy and synchronous optimization strategy," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 5906–5914, Jun. 2022, doi: 10.1109/TVT.2022.3158436.
[18] Y. Sun, Y. Cao, P. Li, and S. Su, "Entropy feature fusion-based diagnosis for railway point machines using vibration signals based on kernel principal component analysis and support vector machine," *IEEE Intell. Transp. Syst. Mag.*, vol. 15, no. 6, pp. 96–108, Nov./Dec. 2023, doi: 10.1109/MITS.2023.3295376.
[19] X. Han and N. Zhao, "Generative adversarial network-based fault diagnosis model for railway point machine in sustainable railway transportation," *Int. J. Sensor Netw.*, vol. 43, no. 1, pp. 50–62, Jan. 2023, doi: 10.1504/IJSNET.2023.133816.
[20] H. Wen, J. Mao, and Y. Xu, "Adaptive discrimination of metro turn back switch locking jamming critical state based on switch machine power," *Urban Mass Transit*, vol. 24, no. 10, pp. 205–209, 2021.
[21] F. P. García, D. J. Pedregal, and C. Roberts, "Time series methods applied to failure prediction and detection," *Rel. Eng. Syst. Saf.*, vol. 95, no. 6, pp. 698–703, Jun. 2010, doi: 10.1016/j.ress.2009.10.009.
[22] Y. Zhang, X. Jiang, and B. Zhao, "Turnout fault prediction based on rough set and grey model," *J. Railway Sci. Eng.*, vol. 16, no. 9, pp. 2331–2338, 2019.
[23] G. Zhang, Y. Si, G. Chen, and G. Wei, "Turnout fault prediction method based on gated recurrent units model," *J. Meas. Sci. Instrum.*, vol. 12, no. 3, pp. 304–313, 2021.
[24] S. Huang, Z. Zhang, F. Zhang, and L. Yan, "Prediction method of ARMA-based turnout operation current fault curve," *Urban Mass Transit*, vol. 25, no. 12, pp. 52–55, 2022.
[25] H. Zhang, X. Gan, S. Li, and Z. Chen, "UAV safe route planning based on PSO-BAS algorithm," *J. Syst. Eng. Electron.*, vol. 33, no. 5, pp. 1151–1160, Oct. 2022, doi: 10.23919/JSEE.2022.000111.
[26] L. H. Zhao and Q. Lu, "Method of turnout fault diagnosis based on grey correlation analysis," *J. China Railway Soc.*, vol. 36, no. 2, pp. 69–74, 2014.
[27] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods an results for class a and class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999, doi: 10.1109/18.761256.
[28] Q. Sun, L. Jiang, and H. Xu, "Expectation-maximization algorithm of Gaussian mixture model for vehicle-commodity matching in logistics supply chain," *Complexity*, vol. 2021, Jan. 2021, Art. no. 9305890, doi: 10.1155/2021/9305890.
[29] H. Jiang, Y. Chen, L. Kong, G. Cai, and H. Jiang, "An LVQ clustering algorithm based on neighborhood granules," *J. Intell. Fuzzy Syst.*, vol. 43, no. 5, pp. 6109–6122, Sep. 2022, doi: 10.3233/JIFS-220092.
[30] G. Li, C. Xu, X. Gong, P. Lu, and H. Dong, "Fewshot deep learning on turnout fault diagnosis model based on masked autoencoder," *China Railway Sci.*, vol. 43, no. 6, pp. 175–185, 2022.
[31] Y. Cao, Y. Sun, G. Xie, and P. Li, "A sound-based fault diagnosis method for railway point machines based on two-stage feature selection strategy and ensemble classifier," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 12074–12083, Aug. 2022.
[32] Z. -H. Zhou, *Machine Learning*. Beijing, China: Tsinghua University Press, 2016.
[33] A. R. Al-Shabeeb, A. Al-Fugara, K. M. Khedher, A. N. Mabdeh, and R. Al-Adamat, "Spatial mapping of landslide susceptibility in jerash governorate of Jordan using genetic algorithm-based wrapper feature selection and bagging-based ensemble model," *Geomatics, Natural Hazards Risk*, vol. 13, no. 1, pp. 2252–2282, Dec. 2022, doi: 10.1080/19475705.2022.2112096.
[34] H. Chen, J. Tao, Y. Jiang, and D. Chen, "Amplitude probability density functions for non-Gaussian random vibrations based on a Gaussian mixture model," *J. Vib. Shock*, vol. 33, no. 5, pp. 115–119, 2014.
[35] K. Vastola, "Threshold detection in narrow-band non-Gaussian noise," *IEEE Trans. Commun.*, vol. COM-32, no. 2, pp. 134–139, Feb. 1984, doi: 10.1109/TCOM.1984.1096037.
[36] D. Liu, "The effectiveness of three-way classification with interpretable perspective," *Inf. Sci.*, vol. 567, pp. 237–255, Aug. 2021, doi: 10.1016/j.ins.2021.03.030.

HAO WEN received the M.S. degree in signal and information processing from Southwest Jiaotong University, China, in 2016. He is currently a Lecturer with Wuhan Railway Vocational College of Technology. His research interests include rail transit equipment fault prediction and machine learning.

JIE XIAO received the Ph.D. degree in control science and engineering from Huazhong University of Science and Technology, Hubei, China, in 2010. She is currently a Professor with Wuhan Railway Vocational College of Technology. Her current research interest includes computer vision based on bionics.

GUANGXIANG XIE received the M.S. degree in signal and information processing from Southwest Jiaotong University, China, in 2017. He is currently engaged in the urban rail transit industry and mainly responsible for the maintenance and management of urban rail transit signal equipment.