

Classification Method for Railway Tunnel Secondary Lining Cold Joint Detection based on CNN-BiLSTM-SVM Model with Improved Hybrid Leader Algorithm

Honggu Zhu¹, Jiaye Wu^{2,*}

¹ School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin, China

² School of Mechanical Engineering, Southwest Petroleum University, Chengdu, China

* Corresponding author: Jiaye Wu

Abstract: Cold joints pose great safety risks to the safe operation of railways. In view of the existing cold joint detection methods, which have low detection efficiency and difficulty in data analysis, a tunnel secondary lining cold joint detection classification method based on the improved hybrid leader CNN-BiLSTM-SVM model is proposed. First, the Rayleigh wave method is used to extract the waveform information of the cold joints. Secondly, CNN-BiLSTM is used to perform feature extraction and fusion processing on the waveform information and then input into the support vector machine, and the improved hybrid leader algorithm is used to optimize the parameters in the SVM. Finally, the information is input into the optimized CNN-BiLSTM-SVM to obtain the cold joints detection classification results. In order to verify the effectiveness of this method, the waveform data collected using the Rayleigh wave method in the tunnel under construction and the verified coring detection results are used as the data set. The results show that the results of this method are higher than the unoptimized CNN-BiLSTM-SVM and the CNN-BiLSTM-SVM optimized by the seagull optimization algorithm and the sparrow search optimization algorithm.

Keywords: Cold Joints; Improved Hybrid Leader Algorithm; Convolutional Neural Network; Bidirectional Long Short-Term Memory Network; Support Vector Machine.

1. Introduction

Cold joints are common concrete defects [1], which appear as concrete weak interface zones with certain visible distribution characteristics. In the past two years, railway derailments caused by pieces of tunnel lining caused by cold joints have attracted great attention from the country. In the past, the detection of cold joints in tunnels mainly relied on manual detection. Now Wu [2] has proposed the Rayleigh wave in the basic shock elastic wave. As one of the methods of cold joints detection, the principle is that cold joints will change the energy and speed of Rayleigh waves [3], and cold joints are identified based on this characteristic. However, this method has higher requirements for later data analysis.

At present, deep learning is increasingly integrated with traditional disciplines. BP neural networks [4], convolutional neural networks[5], recurrent neural networks[6], etc. in deep learning can deeply mine the relevant information between data and reduce the difficulty of data analysis. A new solution that can also be used to achieve the classification task of cold joints detection. The convolutional neural network can identify the spatial characteristics of the detection signal and avoid manual extraction of physical signs. BiLSTM[7] is a type of recurrent neural network that can automatically learn the timing-related features in the cold joints detection signal. Support vector machine [8] is a common classification method in machine learning. It adapts to different data types and problems by introducing different kernel functions to further improve the accuracy and performance of the model. However, the selection and penalty of kernel function parameters Parameters have a great impact on the model.

Through searching domestic and foreign literature, no relevant articles on the combination of cold joint non-

destructive testing of concrete and deep learning or machine learning have been found. Based on the above problems, this paper proposes a model that proposes an improved hybrid leader CNN-BiLSTM-SVM. That is, the signal detected by Rayleigh waves is first processed by CNN-BiLSTM to obtain its spatial characteristics and timing information, and then the improved hybrid leader optimization algorithm is used to optimize the support vector machine parameters, and finally the optimal parameters are obtained by inputting The detection and classification of cold joints are implemented in SVM. Finally, through specific measured data analysis, it is shown that this method can achieve higher classification accuracy and can be applied to engineering practice.

2. Algorithm Theory

2.1. Hybrid Leader Algorithm

Optimization is the process of minimizing or maximizing a certain variable or multiple variables. Optimization is a random process, and existing modern intelligent optimization algorithms rely on random operators to avoid local optima through random methods. They all start the optimization process by creating a random solution or set of solutions for a given problem. Currently, there are many types of intelligent optimization algorithms: first, evolutionary optimization algorithms (genetic algorithms, differential evolution algorithms, etc.), second, population based optimization algorithms (sparrow search algorithms, seagull algorithms, etc.), and third, physics based optimization algorithms (simulated annealing algorithms, etc.).

The hybrid leader algorithm[9] is a method proposed by Mohammad Dehghani in 2022. Its main idea is that in population based algorithms, each member of the population

represents a potential solution in the problem solving space. Through the iterative process of the algorithm and information transmission, population members can continuously optimize their own positions, thereby providing better solutions. However, algorithms may overly rely on specific members (such as the best or worst members in the population) when updating the population, which may lead to the algorithm falling into local optima too early and unable to identify the global optimal solution region in the problem solving space. These limitations may lead to the algorithm converging too fast and preventing the discovery of better solutions. The hybrid leader algorithm, like population based optimization algorithms, also has two stages of exploration and development. It uses unique hybrid leaders to update and guide each member of the algorithm population in the search space. A mixed leader is generated based on three different members, including the best member, a random member, and corresponding members.

The initialization position of the hybrid leader algorithm:

$$x_{i,j} = lb_j + r(ub_j - lb_j), j = 1, 2, \dots, m \quad (1)$$

In the equation, $x_{i,j}$ represents the value of the j th variable in the i -th candidate solution. r is a random number between $[0,1]$, the upper and lower bounds of the j th problem.

The optimization problem should determine the objective function:

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix} \quad (2)$$

In the equation, F represents the objective function, X_i is the i -th candidate solution, and X is calculated as the number of population in the middle. F_i is the value of the objective function in the i -th candidate solution.

In the exploration phase: Hybrid Leader Algorithm (HLBO) uses mixed leaders to update group members. When constructing a random leader, there are three members in the group: the best member, one random member, and the corresponding member. The participation coefficients of these three members in the production of mixed leaders are determined by the quality of providing better value for the objective function. The quality of each member's presentation of candidate solutions is determined by equation (3):

$$q_i = \frac{F_i - F_{worst}}{\sum_{j=1}^{j=N} (F_j - F_{worst})}, i \in \{1, 2, \dots, N\} \quad (3)$$

Calculate the participation coefficient of each member using equation (3) as follows:

$$\begin{aligned} PC_i &= \frac{q_i}{q_i + q_{best} + q_k} \\ PC_{best} &= \frac{q_{best}}{q_i + q_{best} + q_k} \\ PC_k &= \frac{q_k}{q_i + q_{best} + q_k} \end{aligned} \quad (4)$$

Among them, q_i is the value of the i -th candidate solution, F_{worst} is the value of the objective function of the worst candidate solution, PC_i, PC_{best}, PC_k is the participation coefficient of the i -th member, the best member and the k -th member (k is randomly selected from the set $\{1, 2, \dots, N\}$ value).

After determining the participation coefficient, use

equation (5) members to generate hybrid leaders:

$$HL_i = PC_i.X_i + PC_{best}.X_{best} + PC_k.X_k \quad (5)$$

where: HL_i is the hybrid leader of the i -th member, X_k is a randomly selected group member, and index k is the number of rows of the member in the group matrix

The new position of each member is calculated using Equation (6) to calculate the population position in the search space under the guidance of the hybrid leader. If the value of the objective function is improved compared to the previous position, the new position is acceptable for the corresponding member, otherwise Stay at the previous position. Update using equation (7):

$$x_{i,j}^{new,p1} = \begin{cases} x_{i,j} + r.(HL_{i,j} + I.x_{i,j}), & F_{HL_i} < F_i \\ x_{i,j} + r.(x_{i,j} - HL_{i,j}), & else \end{cases} \quad (6)$$

$$X_i = \begin{cases} X_i^{new,p1}, & F_i^{new,p1} < F_i \\ X_i, & else \end{cases} \quad (7)$$

In the formula, $x_i^{new,p1}$ is the new position of the i -th member, r belongs to $[0,1]$, and I is a random integer in $[1,2]$.

In the development phase i.e. local search, in order to find better solutions close to the obtained solution. Consider a neighborhood around each member of the group that allows that member to change position by locally searching in the area and finding a position with a better value of the objective function, using equation (8) to improve and increase HLBO development Ability, if the newly calculated position improves the value of the objective function, the position is updated, as shown in Equation (9).

$$x_{i,j}^{new,p2} = x_{i,j} + (1 - 2r).R.(1 - \frac{t}{T}).x_{i,j} \quad (8)$$

$$X_i = \begin{cases} X_i^{new,p2}, & F_i^{new,p2} < F_i \\ X_i, & else \end{cases} \quad (9)$$

In the formula: $x_i^{new,p2}$ is the new position of the i -th member, R equals 0.2, t is the current iteration, and T is the maximum number of iterations.

2.2. Improved Hybrid Leader Algorithm

Improvement 1; The initialization position of the original algorithm is random, which may cause the initial position to not evenly cover the solution space. This paper proposes a sinusoidal mapping to replace the initialization position in the original algorithm. The introduction of cubic mapping is on the one hand because the population it generates is evenly distributed, which can avoid falling into the local optimal solution during the optimization process, thus improving the convergence speed of the algorithm. On the other hand, it ensures the diversity of understanding. The population generated by sinusoidal mapping has a high degree of randomness and uncertainty, which helps ensure the diversity of solutions. Its location distribution map is shown in Figure 1

$$x_{n+1} = \alpha \sqrt{x_n^2 \sin(\pi x_n)}, x_n \in \{lb, ub\} \quad (10)$$

where α is the random coefficient of the mapping.

Improvement 2: In the local exploration stage of the algorithm, it can be seen from Equation (8) that the new position keeps the specific direction unchanged based on the original position and increases the length of the search. However, this will lead to a lack of diversity in exploration and it is easy to fall into local Among the optimal solutions., This paper proposes an improvement on the search direction

based on it, that is, searching along the optimal direction of the current iteration, which increases the search range and makes it easier to jump out of the local optimal solution.

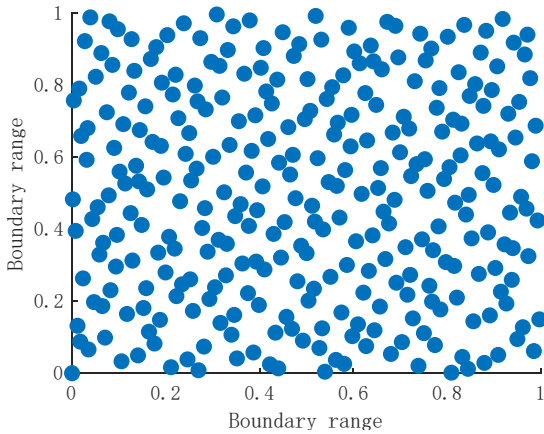


Figure 1. sinusoidal location distribution map

$$x_{i,j}^{new,p2} = x_{i,j} + (1 - 2r) \cdot R \cdot \left(1 - \frac{t}{T}\right) \cdot x_{i,j} + \beta \cdot (x_{best} - x_{i,j}), \quad \beta \in [0,1] \quad (11)$$

Two improved measures are introduced into the original algorithm to form an improved hybrid leader algorithm. By realizing the two stages of exploration and development, all HLBO members are updated, and the best candidate solution experienced during the generation process is introduced as the solution to the problem. and the iteration of the algorithm is completed, the algorithm enters the next iteration, and the iteration of the algorithm is completed.

2.3. Convolutional Neural Network

Convolutional neural network is a model of deep learning, which consists of input layer, convolution layer, pooling layer, activation function, fully connected layer output layer, etc. It has the characteristics of local connection and weight sharing. The convolutional neural network performs feature extraction through convolution operations, reducing the number of parameters of the model and greatly reducing the computational burden of the model. The convolutional neural network can capture through multi-layer convolution operations. The spatial hierarchy of information, the gradual abstraction from low-order features to high-order features, extracts rich features, which have excellent performance in image and video processing tasks, and are widely used in image classification, target detection, image segmentation and image generation. It has a wide range of applications in other fields.

2.4. BiLSTM

Bidirectional long short-term memory network is a variant of recurrent neural network, which is an improvement of LSTM, as shown in Figure 1. Compared with traditional one-way LSTM, BiLSTM has more obvious advantages. It consists of forward and reverse LSTM, which can share parameters and reduce the complexity of the model and the amount of calculation during training. The BiLSTM structure can capture richer contextual information by simultaneously propagating forward and reverse information in time series. Forward LSTM can capture information before the current moment, while reverse LSTM can capture information after the current moment. At the same time, it can effectively handle temporal dependencies in sequences. By integrating

forward and reverse information, this helps to improve the model's ability to model long-term dependencies in sequences, which makes BiLSTM better than other sequence models. training is more efficient. To sum up, BiLSTM has many advantages in sequence modeling and time series data processing. It can make full use of the contextual information of the sequence and improve the representation and prediction capabilities of the model. This makes BiLSTM an important tool for tasks such as natural language processing, audio processing, and time series prediction classification. The calculation process is as follows:

$$\begin{cases} i_t = \text{Sigmoid}(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}C_{t-1} + b_i) \\ f_t = \text{Sigmoid}(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}C_{t-1} + b_f) \\ C_t = f_t C_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ o_t = \text{Sigmoid}(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}C_t + b_o) \\ h_t = o_t \tanh(C_t) \end{cases} \quad (12)$$

In the formula: i_t, f_t, o_t are the input gate, C_t is forget gate and h_t is output gate respectively; and are the state and output of the memory cell respectively; W are the weight matrix; b are the offsets.

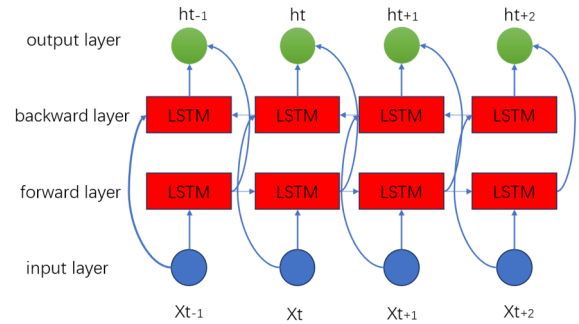


Figure 2. BiLSTM network

2.5. Support Vector Machine

Support vector machine is a commonly used machine learning algorithm. Support vector machine has the advantages of processing high-dimensional spatial data, strong generalization ability, avoiding over-fitting, good robustness to noise, strong interpretability and flexibility. It is widely used in both classification and regression problems. Kernel function SVM can achieve nonlinear separability by using the kernel function to map samples from low-dimensional space to high-dimensional space. By introducing the kernel function, SVM can handle linearly inseparable data sets and can learn more complex decision boundaries. However, the parameters in the kernel support vector machine (SVM) have a certain impact on its performance and results. The kernel function defines the way to map samples from low-dimensional space to high-dimensional space. Common kernel functions include linear kernel function, polynomial kernel function, Gaussian kernel function, etc. Different kernel functions are suitable for different types of data distributions. Kernel function parameter values will cause the decision boundary to be smoother, which may cause underfitting. Choosing appropriate parameter values can balance the fitting ability and generalization ability of the model. The regularization parameter C controls the degree of punishment for misclassified samples. A larger C value will lead to a reduced tolerance of the model to misclassified samples, which may result in a smaller decision boundary interval and make it easy to overfit. A smaller C value will produce a larger decision boundary interval, which is prone to underfitting. Choosing an appropriate C value can balance the

tolerance and generalization ability of the model.

3. Establishment of CNN Bilstm SVM Model based on Improved Hybrid Leader Optimization Algorithm

3.1. Data Source and Processing

In this article, there are a total of 1920 pieces of data in the dataset, all of which have been validated through core sampling and divided into training and testing sets in an 8:2 ratio. According to engineering practice, the detection results are divided into mild cold jointss, non cold jointss, and cold jointss, corresponding to labels 1, 2, and 3. The data is collected based on Rayleigh waves, and the sampling points are set to 512. The detection data is shown in Table 1. In order to accelerate the convergence of the network loss function and improve the model training speed, the data is normalized. The normalization formula is:

$$x_0 = \frac{x - x_{\min}}{x_{\max} - x} \quad (13)$$

In the formula: x refers to the sample data before normalization; x_0 is the normalized value; x_{\max} is the maximum and x_{\min} is minimum values in the sample data before normalization.

Table 1. Cold joints detection data

type	label	Sampling point 1 (cm)	Sampling point 512 (cm)
Slight cold joint	1	23.6	0.9
Non cold joint	2	0.7	3.5
cold joint	3	1.1	4.9

3.2. Acknowledgment

First, the detected data is normalized and input into the CNN. Then the model starts data training and learning. The parameters of the model are set as follows: First, the modulus format of the data is [512,1,1], and enters the sequenceInputLayer layer. Here, the sequence is not divided into blocks by default and the output form is [512,1,1,1]. The model here is divided into two parts. The first part enters the convolution operation, sets the number of convolution kernels to 16, and the convolution kernel size is [3,1], the output form of the model after entering the batch normalization layer is [512,1,16,1], the model performs secondary convolution, sets the number of convolution kernels to 32, and the convolution kernel size is [3,1], the output form of the model after entering the batch normalization layer is [512,1,32,1], the second part is directly connected to the sequenceUnfoldingLayer layer, and the final model form after splicing and flattening is [16384 ,1,1], and then after BiLSTM layer processing, the model form becomes [10,1]. After continuing with batch normalization processing and fully connected layer processing, the model form is [3,1], which we use SVM to replace softmax layer, and finally use the improved hybrid leader algorithm combined with cross-validation to optimize the SVM parameters, thus completing the entire model learning and training.

The model result diagram is shown in the figure 3. In the figure, seqfold is the sequenceInputLayer layer in matlab, which is used to receive sequence data as the input of the

model. The main function of the sequence input layer is to define the format and shape of the input data and pass the sequence data to the next layer of the network for processing. It is usually the first layer of the neural network, used to receive the original sequence data, and preprocess and format it to adapt to the input requirements of the network. Sequnfold is the sequenceUnfoldingLayer layer, SequenceUnfoldingLayer) is a kind of recurrent neural network. Commonly used layers. Its main function is to expand the input sequence data into a series of time step inputs. Ba is batchNormalizationLayer, flatten is flattening layer, and fc is fully connected layer.

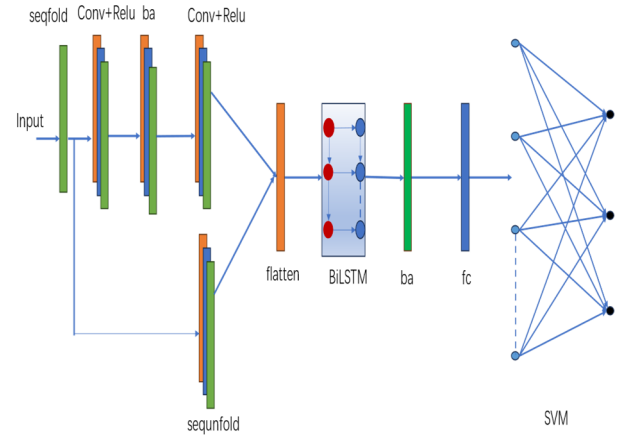


Figure 3. CNN-BiLSTM-SVM model diagram

3.3. Experimental Results

In order to verify the model accuracy of the method proposed in this article, comparative experiments were conducted on the parameters of SVM using the CNN-BILSETM-SVM optimized by the Seagull optimization algorithm and the Sparrow search optimization algorithm. The training set was cross-validated five times, and the three optimization algorithm models were the fitness function is the mean of five root mean square errors. The fitness value and SVM parameters are shown in Table 2. The fitness function trend is shown in Figure 4

$$E_{RMSE} = \frac{\sum_{k=1}^5 \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2}}{5} \quad (14)$$

In the formula: y_i represents the real label value, y_i^* represents the predicted label value, n represents the number of samples, and k represents the number of cross-validation times.

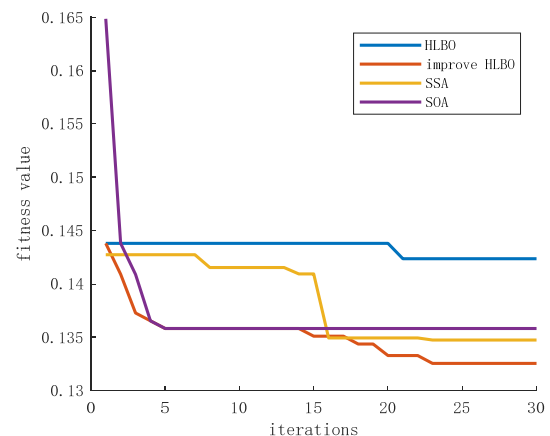


Figure 4. The fitness function trend

Table 2. SVM parameters and their fitness

index	Improved HLBO	HLBO	SOA	SSA
Regularization coefficient C	10.31	10.22	11.24	12.28
Kernel function parameters γ	2.41	1.25	3.21	2.73
adaptability	0.132	0.142	0.135	0.134

Results analysis, this article proposes CNN-BiLSTM-SVM based on the improved hybrid leader algorithm (IHLBO), CNN-BiLSTM-SVM based on the hybrid leader algorithm (HLBO), CNN-BiLSTM-SVM based on the Seagull Optimization Algorithm (SOA) [10], The accuracy rates of the four CNN-BiLSTM-SVM models of the Sparrow Search

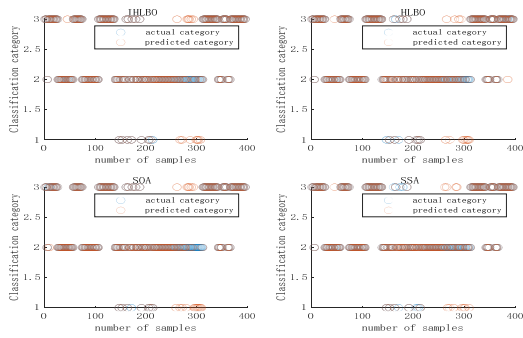
Optimization Algorithm (SSA) [11] are 92.4%, 89.9%, 91.4%, and 90.9% respectively. The model results are shown in Table 4.

Table 3. Model experimental parameter settings

Model parameters	content
MaxEpochs	500
MiniBatchSize	32
InitialLearnRate	1e-3
L2Regularization	1e-4
LearnRateSchedule	piecewise
LearnRateDropFactor	0.1
LearnRateDropPeriod	400

Table 4. Model results

	IHLBO	HLBO	SOA	SSA
type	label	Misjudgment number	Misjudgment number	Misjudgment number
Slight cold joint	1	28	26	24
Non cold joint	2	2	4	14
cold joint	3	0	4	6

**Figure 5.** Model testing results

The CNN-BiLSTM-SVM based on the improved hybrid leader algorithm proposed in this article is the highest among the models currently proposed. It is known from Figure 5 and Table 4: This model is more sensitive to the cold joints part and the non-cold joints part than the remaining three models. The reason is: when the various initialization parameters of the four models mentioned above are consistent, the improved hybrid leader algorithm has better parameters than the original leader algorithm, seagull optimization algorithm, and sparrow search algorithm (SSA). The search ability of the space is stronger, and it is less likely to fall into the local optimal solution. The robustness and generalization ability of the model have been strengthened, and the prediction effect of the model has been greatly improved.

In summary, the CNN-BiLSTM-SVM based on the improved hybrid leader algorithm proposed in this article can have high detection and diagnosis accuracy for suspected cold joint areas in the second lining of tunnels, and can be used in actual engineering.

4. Conclusion

In order to solve the current problems of low detection accuracy and high difficulty in data analysis in cold joints in railway tunnels, this paper proposes a CNN-BiLSTM-SVM cold joint detection classification model based on the improved hybrid leader algorithm. First, the detection data is normalized. The CNN network can be used for feature

extraction and BiLSTM can share parameters. This can reduce the complexity of the model and the amount of calculation during training, and capture richer contextual information. At the same time, the forward direction can be integrated with reverse information, this feature helps to improve the model's ability to model long-term dependencies in the sequence and extract the spatial and temporal characteristics of the cold joints detection signal. Finally, SVM is used to replace the softmax layer, which has the advantages of processing high-dimensional spatial data, strong generalization ability, avoiding over-fitting, good robustness to noise, strong interpretability and flexibility, and uses the improved hybrid leader algorithm to The SVM parameters are optimized to form a CNN-BiLSTM-SVM cold joints detection classification model with an improved hybrid leader algorithm. Experimental results show that the accuracy of this method is as high as 91.41%, which is higher than the other mentioned methods.

In the future, the data set can be gradually improved, a better fitness function can be designed, and through cross-validation, better SVM parameters can be obtained, the generalization ability of the model can be improved, and the detection accuracy can be further improved.

References

- [1] Deng, Huang-Shi, et al. "Study on deterioration, cracking mechanism and treatment measures of plain concrete lining with cold joints in high-speed railroad tunnel." *Case Studies in Construction Materials* 18 (2023): e01895.
- [2] Wu Jiaye, Jia Qisong, Chen Ting, etc. Detection and evaluation method of concrete cold joints based on impact elastic waves [J]. *Railway Construction*, 2023, 63(01): 88-91.
- [3] Xiao, Jing, and Fangsen Cui. "Machine learning enhanced characterization of surface defects using ultrasonic Rayleigh waves." *NDT & E International* (2023): 102969.
- [4] Ding, Shifei, Chunyang Su, and Junzhao Yu. "An optimizing BP neural network algorithm based on genetic algorithm." *Artificial intelligence review* 36 (2011): 153-162.
- [5] Alzubaidi, Laith, et al. "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions." *Journal of big Data* 8 (2021): 1-74.

- [6] Sherstinsky, Alex. "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network." *Physica D: Nonlinear Phenomena* 404 (2020): 132306.
- [7] Chen, Tao, et al. "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN." *Expert Systems with Applications* 72 (2017): 221-230.
- [8] Cherkassky, Vladimir, and Yunqian Ma. "Practical selection of SVM parameters and noise estimation for SVM regression." *Neural networks* 17.1 (2004): 113-126.
- [9] Dehghani, Mohammad, and Pavel Trojovský. "Hybrid leader based optimization: a new stochastic optimization algorithm for solving optimization applications." *Scientific Reports* 12.1 (2022): 5549.
- [10] Dhiman, Gaurav, and Vijay Kumar. "Seagull optimization algorithm: Theory and its applications for large-scale industrial engineering problems." *Knowledge-based systems* 165 (2019): 169-196.
- [11] Xue, Jiankai, and Bo Shen. "A novel swarm intelligence optimization approach: sparrow search algorithm." *Systems science & control engineering* 8.1 (2020): 22-34.