

## Vibration Signal Features Prediction of GIS Equipment Based on Improved Slime Mold Optimization Algorithm Optimizing CNN-BiLSTM



Xiping Jiang<sup>1</sup>, Qian Wang<sup>1</sup>, Yingkai Long<sup>1</sup>, Shiling Zhang<sup>1</sup>, Yun Fang<sup>2</sup>, Dong Hu<sup>2\*</sup>

<sup>1</sup> Electric Power Research Institute of State Grid Chongqing Electric Power Company, Chongqing 401123, China

<sup>2</sup> School of Engineering and Technology, Southwest University, Chongqing 400715, China

Corresponding Author Email: [hhyyh@swu.edu.cn](mailto:hhyyh@swu.edu.cn)

<https://doi.org/10.18280/ts.390403>

### ABSTRACT

**Received:** 10 May 2022

**Accepted:** 28 July 2022

#### Keywords:

*GIS equipment, feature prediction, slime mold optimization algorithm, differential evolution strategy*

The change of vibration signal of GIS equipment can reflect the internal mechanical state. In order to improve the prediction accuracy of vibration signal characteristics of GIS equipment, this paper proposes an improved slime mold algorithm to optimize CNN-BiLSTM GIS vibration characteristics prediction method. First, the vibration characteristic parameters are extracted in the frequency domain based on the GIS historical vibration signal through Fourier transformation. Secondly, in order to enhance the feature utilization ability of BiLSTM model, 1D CNN is used to extract feature parameters; the differential evolution strategy is integrated into the slime mold optimization algorithm and the parameters such as the number of hidden layer neurons and learning rate of CNN-BiLSTM are optimized. Finally, an improved slime mold algorithm is established to optimize the GIS feature prediction model of CNN-BiLSTM. The experimental results show that the root mean square error and mean absolute percentage error of the DESMA-CNN-BiLSTM model are 1.7915 and 0.1317%, respectively, which are better than other methods in prediction accuracy. The improved algorithm proposed in this paper has strong global search ability and fast convergence speed.

## 1. INTRODUCTION

GIS equipment is the core component of the power system. Its main role is to strengthen the reactive power compensation, improve the power quality of high voltage lines, reduce line loss, etc. [1] In the long-term operation process, the GIS vibration may cause the looseness of the internal fasteners core, winding and other key components of the GIS. In extreme cases, it may even lead to serious equipment failure [2-5]. The reactor acoustic vibration features are closely related to its internal mechanical state [6]. When the internal mechanical state of GIS equipment changes, it will cause local amplitude fluctuations [7]. Therefore, the operation status of GIS equipment can be evaluated based on the GIS vibration signal, which is of great significance to maintain the stability of the power system and ensure the efficient transmission of energy [8].

In recent years, researchers have carried out extensive researches on equipment condition assessment for the vibration features of transformers and reactors [9-11]. Wei Xu et al. proposed a reactor fault detection method based on vibration signal, which could reflect its internal operating state. Chen et al. [12] put forward a reactor vibration signal processing method based on recursive quantitative analysis. However, the research was carried out only in the time domain, not in the frequency domain. Li et al. [13]. Proposed a fault diagnosis model based on vibration signal power spectrum and Markov chain, and used phase space coefficients to identify the state features of vibration signals. However, other sensitive features of the reconstructed signal in high-dimensional space were not considered, resulting in limited identification ability.

Huang and Gan [14], based on the percentage of 90 degree and non-90 degree angles of phase trajectory points as the feature extraction standard, extracted the features of vibration signals and realized fault classification in a satisfactory way. However, time delay and embedding dimension poses a significant impact on the determination of phase trajectory point angle, and the reliability of feature extraction is low [15].

With the development of sensor technology and artificial intelligence, deep learning methods are being increasingly used for reactor fault diagnosis and trend prediction. The deep learning model requires a large number of samples for training. However, it is not easy to obtain enough fault samples in actual scenarios. Zhu et al. [16] used DCGAN network to generate comprehensive fault samples to train convolutional neural network (CNN), which improved the performance of reactor fault diagnosis model. Zhu et al. [17] proposed a deep learning method (EP-CNN) based on coding position and feature fusion by collecting vibration signals under different working conditions, which effectively reduced the workload and achieved high recognition accuracy. Gao et al. [6] carried out prediction research on multi-channel acoustic and vibration features of reactors by using Gated Cycle Unit (GRU) and Long Short Term Memory Neural Network (LSTM), and achieved high prediction accuracy. However, the hyper-parameters of the model are set manually, and there are certain contingencies. Liu [18] used particle swarm optimization (PSO) to optimize the input weight of LSTM network in order to avoid blindness in the selection of LSTM hyper-parameters. However, the influence of the weight of particles themselves was not taken into account, which might easily lead to local optimization. Similar to other neural networks, the selection of

hyper-parameters of BiLSTM neural networks depends on empirical values and has poor generalization performance [19]. However, the existing deep learning neural network models can learn the long-term historical dependence of data, but cannot characterize the potential relationship of local features in time sequence data in a satisfactory way [20].

In summary, in order to improve the accuracy of GIS equipment vibration signal feature prediction, this paper proposes an improved slime mold algorithm to optimize CNN-BiLSTM GIS vibration feature prediction method. First, the vibration characteristic parameters are extracted in the frequency domain based on the GIS historical vibration signal through Fourier transformation. Secondly, in order to enhance the feature utilization ability of the BiLSTM model, 1D CNN is used to extract feature parameters; at the same time, in order to enhance the population diversity of the swarm intelligence optimization algorithm and avoid falling into local optimum, the differential evolution strategy is integrated into the slime mold optimization algorithm (SMA). The SMA algorithm optimizes the number of hidden layer neurons, learning rate and other hyperparameters of CNN-BiLSTM, avoiding the blindness of hyperparameter selection. Finally, a GIS feature prediction model based on the improved slime mold algorithm was constructed to optimize CNN-BiLSTM. The effectiveness of the method in this paper is verified by analyzing the multi-channel vibration signal of GIS equipment in a substation by an example.

## 2. ALGORITHMIC THEORY

### 2.1 Slime Mould Algorithm

The slime mold algorithm (SMA) is a meta-heuristic optimization algorithm developed by Li et al. in 2020. It is inspired by the oscillation mode of slime mold in nature [21]. When the organism in the slime mold comes into contact with the food, it will wrap it first, and then release the enzymes needed for digestion. The author uses weight adjustment to simulate the feeding mechanism and optimization process of SMA method.

The SMA algorithm is generally divided into the following three steps:

#### (1) Approaching food

The approaching behavior of slime molds is modeled as a mathematical equation through the formula. Formula (1) is used to simulate the contraction mode:

$$\vec{X}(t+1) = \begin{cases} \vec{X}_b(t) + \vec{vb} \cdot (\vec{W} \cdot \vec{X}_A(t) - \vec{X}_B(t)), r < p \\ \vec{vc} \cdot \vec{X}(t), r \geq p \end{cases} \quad (1)$$

where: the range of  $\vec{vb}$  is  $[-a, a]$ ;  $\vec{vc}$  linearly decreases from 1 to 0;  $\vec{X}_b$  is the location of the individual with the highest fitness value currently found;  $\vec{X}$  is the location of slime mold;  $\vec{X}_A$  and  $\vec{X}_B$  are two individuals randomly selected from slime molds;  $\vec{W}$  is the weight of slime mold, and the Formula  $p$  is defined as follows:

$$p = \tanh|S(i) - DF| \quad (2)$$

wherein,  $S(i)$  is the fitness value of  $\vec{X}$ ;  $DF$  represents the best

fitness value in the iteration process. The calculation of  $\vec{vb}$  is shown in Formula (3), (4), and  $T$  represents the maximum number of iterations:

$$\vec{vb} = [-a, a] \quad (3)$$

$$a = \arctan h\left(-\frac{t}{T} + 1\right) \quad (4)$$

The calculating formula of  $\vec{W}$  is as follows:

$$\vec{W}(SmellIndex(i)) = \begin{cases} 1 + r \cdot \log\left(\frac{F_b - S_i}{F_b - F_w} + 1\right), \text{condition} \\ 1 - r \cdot \log\left(\frac{F_b - S_i}{F_b - F_w} + 1\right), \text{others} \end{cases} \quad (5)$$

$$SmellIndex = \text{sort}(S) \quad (6)$$

*Condition* is the top half of  $S(i)$  in the group;  $r$  is a random number at  $[0, 1]$ ;  $F_b$  is the best fitness value obtained in the current iteration process;  $F_w$  is the worst fitness value obtained in the current iteration process; *SmellIndex* is the sorted fitness value sequence (in ascending order in the minimization problem).

#### (2) Wrapping food

In the process of approaching the food, the slime mold mainly realizes the position update by simulating the contraction mode of venous tissue structure, as shown in Formula (7):

$$\vec{X}^* = \begin{cases} \text{rand} \cdot (B_u - B_l) + B_l, \text{rand} < z \\ \vec{X}_b(t) + \vec{vb} \cdot (\vec{W} \cdot \vec{X}_A(t) - \vec{X}_B(t)), r < p \\ \vec{vc} \cdot \vec{X}(t), r \geq p \end{cases} \quad (7)$$

where:  $\vec{X}^*$  is the updated position.  $B_l$  and  $B_u$  are respectively the upper and lower boundaries of the search space, and  $\text{rand}$  is the random value between  $[0, 1]$ .

#### (3) Oscillating food

As the number of iterations increases, the value of  $\vec{vb}$  oscillates randomly between  $[-a, a]$  and finally approaches 0. The value of  $\vec{vc}$  will oscillate between  $[-1, 1]$  and will eventually also approach zero. The synergistic effect of the two simulates the selection behavior of slime molds.

### 2.2 Differential evolution strategy

In the ordinary SMA algorithm, the position update of slime molds is determined by the fitness of individuals in the slime mold population. Whether the individual is updated is determined by comparing the fitness value and the size of the individual slime mold. However, this method will lower the diversity of slime mold population in the process of algorithm iteration. Differential Evolution Algorithm (DE) based on genetic algorithm and other evolutionary ideas has the ability to increase population diversity [22]. The basic process of DE includes mutation, crossover and selection.

#### (1) Mutation operator

The variation vector in previous iterations is generated according to the following formula:

$$v_i^G = x_{r_1}^G + F \times (x_{r_2}^G - x_{r_3}^G) \quad (8)$$

where,  $r_1, r_2$ , and  $r_3$  are randomly selected from the population  $[1, 2, \dots, N]$ , and are not equal to  $i$ .  $F$  is the control parameter of the proportional vector. The value range proposed by Storn et al. is within  $[0, 2]$ .

### (2) Crossover operator

The crossover operator is to mix the target vector and mutation vector, and obtain the following through Formula (9):

$$u_{j,i}^G = \begin{cases} v_{j,i}^G & \text{if } (\text{rand}_{j,i} \leq CR \text{ or } j = j_{rand}) \\ x_{j,i}^G & \text{otherwise} \end{cases} \quad (9)$$

where,  $CR$  is the number of control crossovers, which controls the probability of generating parameters from mutation vectors.

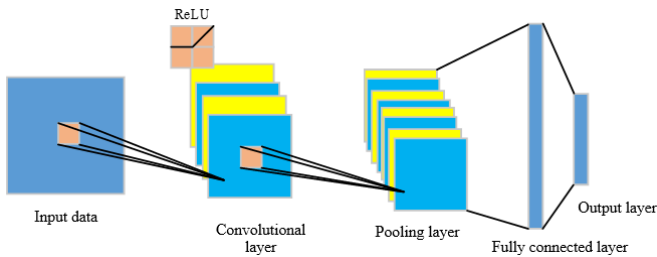
### (3) Selection operator

In the selection operator, the greedy strategy is used to discuss whether to choose  $u_i^G$  or  $x_i^G$ . If the fitness value of  $u_i^G$  is better, it will be selected as the next generation individual. Otherwise,  $x_i^G$  is selected as the next generation individual.

$$x_i^{G+1} = \begin{cases} u_i^G & \text{Fit}(u_i^G) \leq \text{Fit}(x_i^G) \\ x_i^G & \text{otherwise} \end{cases} \quad (10)$$

## 2.3 CNN and BiLSTM

### 2.3.1 Convolutional neural network



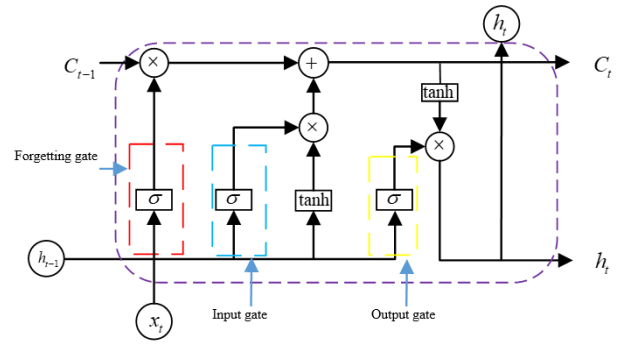
**Figure 1.** Model structure of Convolutional Neural Network

Convolution neural network (CNN) is a kind of feedforward neural network, which is mainly designed to process two-dimensional data. CNN can be effectively applied to time sequence prediction and can extract features in an adaptive manner [23]. CNN's local awareness and weight sharing can greatly reduce the number of parameters, thus improving the efficiency of learning models. CNN is mainly composed of three parts: convolution layer, pooling layer and fully connected layer, as shown in Figure 1. The pooling layer after the convolution layer can reduce the feature dimension and retain strong features, so as to reduce the training cost of the network.

### 2.3.2 Bidirectional long short-term memory neural network

Long short-term memory neural network (LSTM) is a network model proposed by Schmidhuber et al. in 1997, which aims to solve the problem of gradient disappearance and gradient explosion that exist for a long time in recurrent neural network (RNN) [24]. LSTM modules are similar to the standard RNN model, and they operate in a special interactive way. As shown in Figure 2, LSTM memory unit is composed of three parts: forgetting gate, input gate and output gate.

BiLSTM neural network has an additional reverse LSTM layer on the basis of LSTM [25]. As shown in Figure 3, The bidirectional network enables recursive response training for the states of the hidden layer of the starting point and the end point of the sequence, which can further excavate some relationship between the current data and the data at the previous moment, as well as the data at the future moment, so as to further process the reverse information, optimize the long-term dependency, and improve the prediction accuracy of the model.



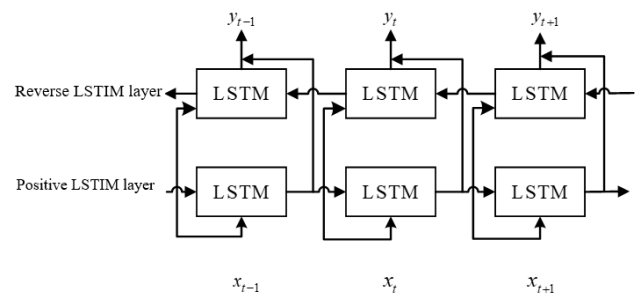
**Figure 2.** LSTM model structure

The output value of the BiLSTM model is jointly determined by the forward and reverse hidden layers. Formula (11) is the output of the forward hidden layer, and Formula (12) is the output of the reverse hidden layer:

$$\vec{h}_t = \sigma(W_{x\vec{h}} x_t + W_{\vec{h}\vec{h}} \vec{h}_{t-1} + b_{\vec{h}}) \quad (11)$$

$$\vec{h}_t = \sigma(W_{x\vec{h}} x_t + W_{\vec{h}\vec{h}} \vec{h}_{t-1} + b_{\vec{h}}) \quad (12)$$

In Formula (11),  $W_{x\vec{h}}$  is the output weight of the hidden layer of the unit;  $W_{\vec{h}\vec{h}}$  is the weight of the state quantity from the previous moment to the current moment;  $b_{\vec{h}}$  is offset;  $\vec{h}_{t-1}$  is the output value of the hidden layer state at the previous moment.



**Figure 3.** BiLSTM model structure

## 3. SLIME MOLD OPTIMIZATION ALGORITHM BASED ON DIFFERENTIAL EVOLUTION STRATEGY

In SMA algorithm, DE strategy is added to improve the diversity of the population and enhance the local search ability of slime molds [26]. The calculation formula for each individual slime mold is as follows:

$$u_{j,i}^G = \begin{cases} v_{j,i}^G & \text{if } (rand_{j,i} \leq CR \text{ or } j = j_{rand}) \\ x_{j,i}^G + \bar{v}b \times (\bar{W} \times x_{j,A}^G - x_{j,B}^G) & \text{otherwise} \end{cases} \quad (13)$$

$$m = \frac{t}{T} \quad (15)$$

wherein,  $r$  and  $j,i$  is the random number between  $[0,1]$ , CR is the cross control number, and the calculation of  $\bar{W}$  is shown in formula (14):

$$\bar{W} = \begin{cases} 1 + r \cdot \log\left(\frac{F_b - S_i}{F_b - F_w} + 1\right), & \text{if } rand < m \\ 1 - r \cdot \log\left(\frac{F_b - S_i}{F_b - F_w} + 1\right), & \text{otherwise} \end{cases} \quad (14)$$

Parameter  $v_i^G$  is defined as follows:

$$v_i^G = x_{r_1}^G + \bar{v}b \times (x_{r_2}^G - x_{r_3}^G) \quad (16)$$

where,  $r_1, r_2$ , and  $r_3$  are randomly selected from the population  $[1,2,\dots,N]$ , and are not equal to  $i$ . The value of  $\bar{v}b$  is between  $[-a, a]$ , and the calculation of  $\bar{v}b$  and  $a$  is shown in Formula (3) - (4). The flow chart of slime mold optimization algorithm based on differential evolution strategy (DESMA) is shown in Figure 4.

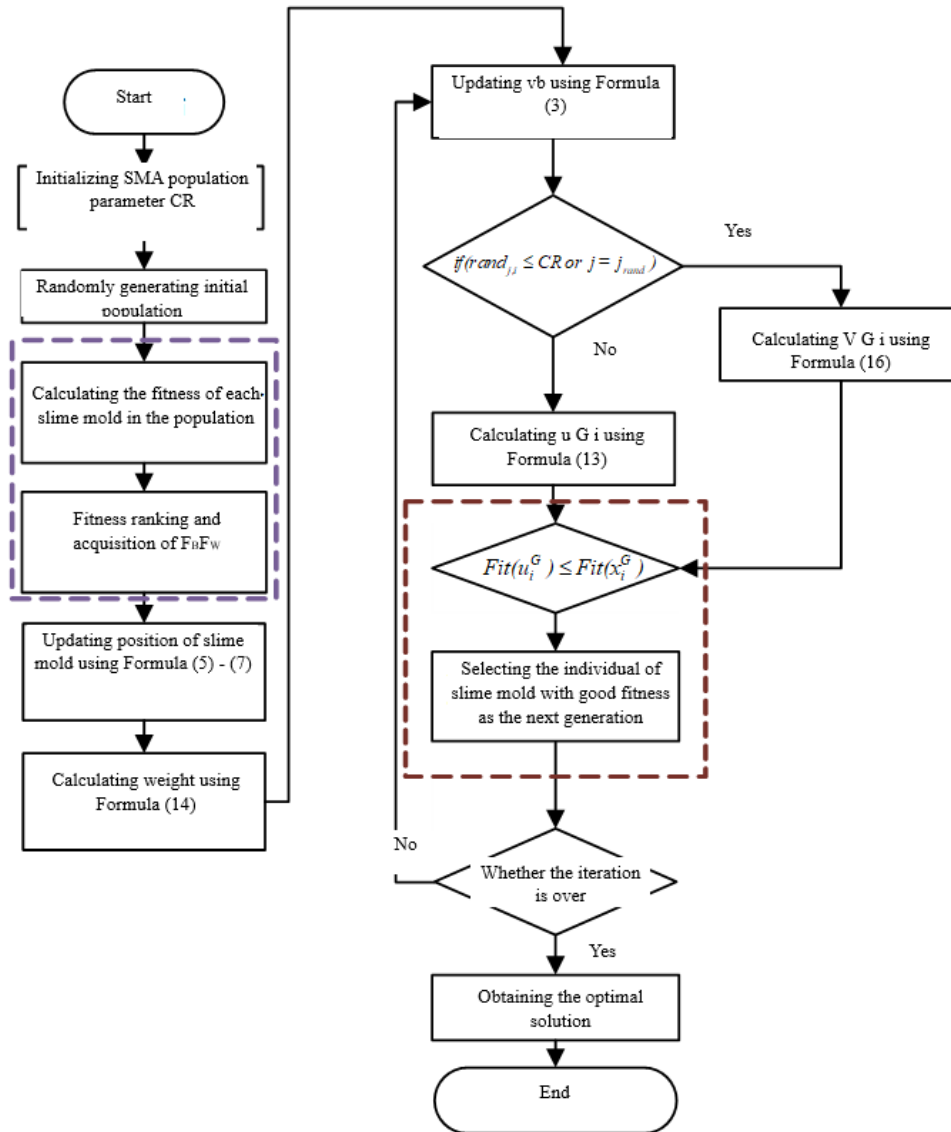


Figure 4. Flow chart of DESMA optimization algorithm

#### 4. DESMA-CNN-BILSTM PREDICTION MODEL

##### 4.1 Experimental data

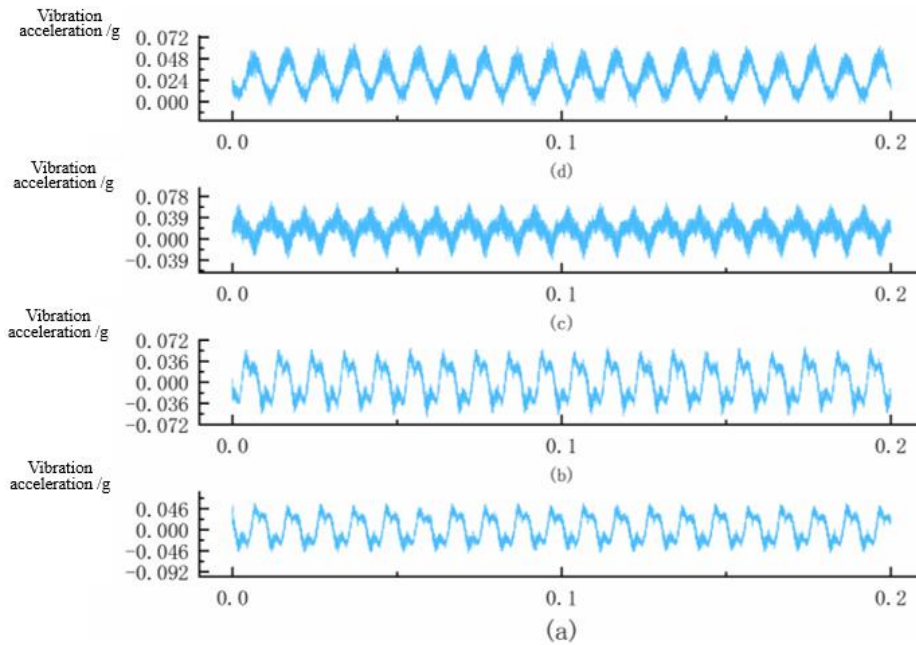
The experimental data in this paper are obtained from the field measured vibration signals of a 500kV GIS equipment in an electrical research institute in China. The frequency is

125KHz, with four channels. Twenty-four-hour sampling was carried out on the GIS equipment, at a sampling interval of 1 hour. Each channel has 24 time domain waveforms in a day. Figure 5 shows the time domain data of the first 0.2S of the GIS equipment. (a), (b), (c) and (d) are respectively the vibration data of channel 1, channel 2, channel 3 and channel 4. As the field test data, the data in this paper includes not only

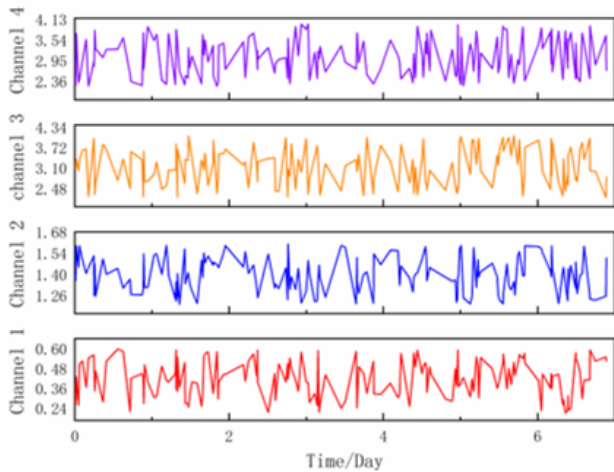
the noise signal with fundamental frequency of 100Hz, but also harmonic signals with frequency of 200 and 300Hz.

As the time domain waveform of field data is complex, this paper mainly looks at the prediction of frequency domain waveform. In order to comprehensively measure its operation

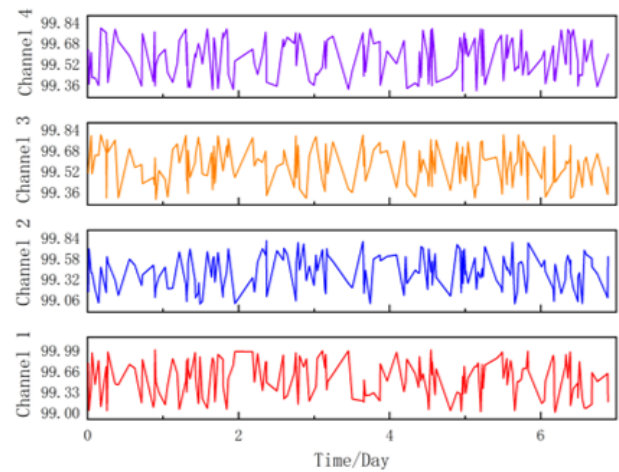
state, in this paper, three feature parameters, including the odd-even harmonic ratio, fundamental frequency proportion and vibration entropy, are selected as the composition of time sequence prediction data.



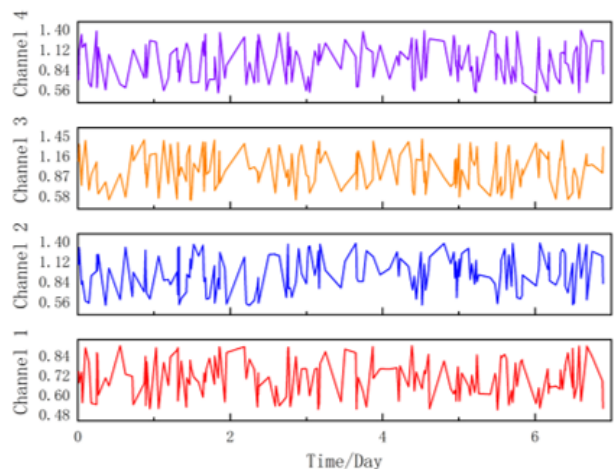
**Figure 5.** Time domain diagram of signals of each channel: (a), (b), (c) and (d) correspond to channel 1, channel 2, channel 3 and channel 4 respectively



(a) Odd-even harmonic ratio feature historical data



(c) Fundamental frequency ratio historical data



(b) Vibration entropy feature historical data

**Figure 6.** Three characteristics of historical data of each channel

Figure 6 shows the historical data of the three feature parameters including the odd-even harmonic ratio, fundamental frequency proportion and vibration entropy, of each channel. The historical data of each feature parameter has a certain degree of fluctuation and a certain degree of time sequence. The eigenvalue of each channel is related to the ready-made operation condition of reactor, environmental factors and other factors.

#### 4.2 Model in this paper

For the four-channel vibration signals, 12 time sequence data are constructed for three feature parameters, namely odd-

even harmonic ratio, fundamental frequency proportion and vibration entropy. In this paper, the training set and test set are divided into 6 days of data as the feature parameters of the model training set to predict the feature parameter of the next day. The training set time sequence length is 144, and the test set time sequence length is 24. Based on the DESMA-CNN BiLSTM GIS vibration feature prediction model constructed in this paper, 12 feature data are grouped into training set and test set, and CNN is used to extract features from the training set data, and then the data are sent to the BiLSTM model. The hyper-parameter range of the BiLSTM model is set as follows: the batch size range  $batch\_size \in [1,50]$ , the number range of the three hidden layer nodes  $h1, h2, h3 \in [1,100]$ , and the learning rate  $learning\_rate \in [0.001,0.2]$ . The number of iterations of the DESMA-CNN BiLSTM model is set to 100. In the model training process, the mean square error is used as the fitness function of the DESMA algorithm, and finally each weight value approaches the optimal solution. Based on the optimized CNN-BiLSTM model, the test set data is used for testing to finally obtain the prediction results. The prediction process of DESMA-CNN BiLSTM model is shown in Figure 7.

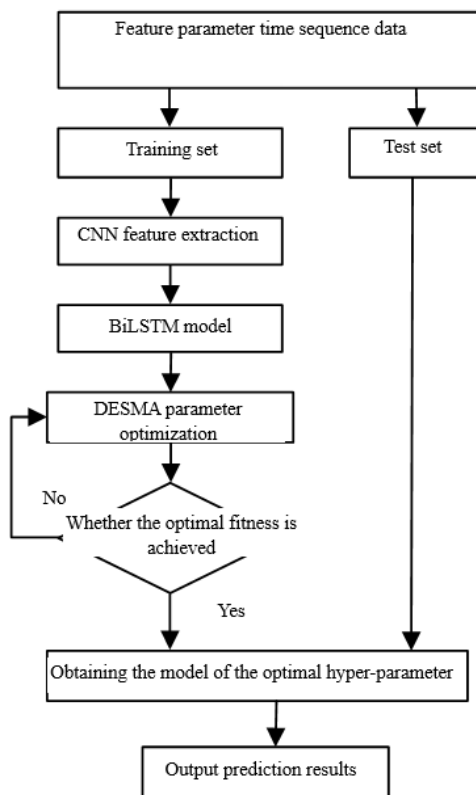


Figure 7. Forecast process of DESMA-CNN BiLSTM model

## 5. EXPERIMENTAL RESULTS AND COMPARATIVE ANALYSIS

### 5.1 Experiment result

Based on the feature prediction results of the GIS equipment vibration signal mentioned in this paper, as shown in Figure 8, the feature prediction results of each channel show that the DESMA-CNN-BiLSTM model can accurately capture the historical trend of the three feature parameters, namely, odd-even harmonic ratio, fundamental frequency proportion, and

vibration entropy. In addition, the prediction results are roughly the same as the actual results. In order to verify the performance of the DESMA-CNN-BiLSTM prediction model in this paper, three evaluation indicators, namely, mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE), are selected for verification. Table 1 through Table 3 show the prediction effect evaluation of each feature.

It can be seen from Table 1 through Table 3 that the MAE values of vibration entropy in each channel are 0.057, 0.069, 0.0833 and 0.0786 respectively. The lower the vibration entropy is, the more concentrated the spectral energy is. When DC magnetic biasing occurs to the GIS equipment, odd harmonic components in the vibration signal will increase. This feature can be used to determine the degree of DC magnetic biasing of the GIS. Table 3 shows that the MAPE of odd-even harmonic ratio of channel 2 is 0.03%, which is the lowest among the values of each channel. However, the overall error is low. The result of using DESMA algorithm to optimize the hyper-parameter is that the learning rate is 0.01, the number of three hidden layer nodes is 29, 35, 35, and the batch size is 24.

Table 1. Prediction performance evaluation of fundamental frequency proportion

Evaluation indicator	Channel 1	Channel 2	Channel 3	Channel 4
MAE	0.0836	0.0664	0.0533	0.0784
RMSE	2.58	1.67	1.83	1.9
MAPE(%)	0.1046	0.0367	0.0595	0.0457

Table 2. Vibration entropy prediction performance evaluation

Evaluation indicator	Channel 1	Channel 2	Channel 3	Channel 4
MAE	0.057	0.069	0.0833	0.0786
RMSE	1.41	1.63	1.72	1.93
MAPE(%)	0.0346	0.0417	0.095	0.0657

Table 3. Prediction performance evaluation of odd-even harmonic ratio

Evaluation indicator	Channel 1	Channel 2	Channel 3	Channel 4
MAE	0.0481	0.0643	0.0511	0.0483
RMSE	1.05	2.43	1.89	1.53
MAPE(%)	0.0413	0.0324	0.0695	0.0387

### 5.2 Comparative analysis

In order to verify the superiority of the method in this paper, five models, SMA-CNN BiLSTM, CNN-BiLSTM, BiLSTM, LSTM and PSO-BiLSTM, are selected for comparative experiments. The prediction results are shown in Figure 9, which is only displayed in terms of the feature parameters of odd-even harmonic ratio. It can be seen from Figure 9 that the DESMA-CNN BiLSTM method is more suitable for real data than other methods, and the prediction accuracy of this method is higher than other prediction models. Figure 10 shows the loss value changes with the number of iterations when using different methods to predict odd-even harmonic ratio features. The loss value of this method tends to converge when the number iterations reaches 20, and the loss value is 0.0691 when iteration is completed. Compared with SMA-CNN

BiLSTM method, this method has a higher convergence speed, and the loss value is reduced by 0.021. It shows that the improved strategy proposed in this paper increases the diversity of the SMA algorithm population and improves its ability to search the optimal solution globally.

Table 4 shows the evaluation indicators for the prediction results of odd-even harmonic wave ratio, fundamental

frequency proportion and vibration entropy of each channel using different methods. It can be seen from Table 4 that the MAPE values of SMA-CNN-BiLSTM, PSO-BiLSTM, CNN-BiLSTM, BiLSTM and LSTM are 0.1359%, 0.1449%, 0.1548%, 0.1642% and 0.1691% respectively, which are 3.1%, 9.1%, 14%, 17% and 22.1% higher than the those obtained using the above methods.

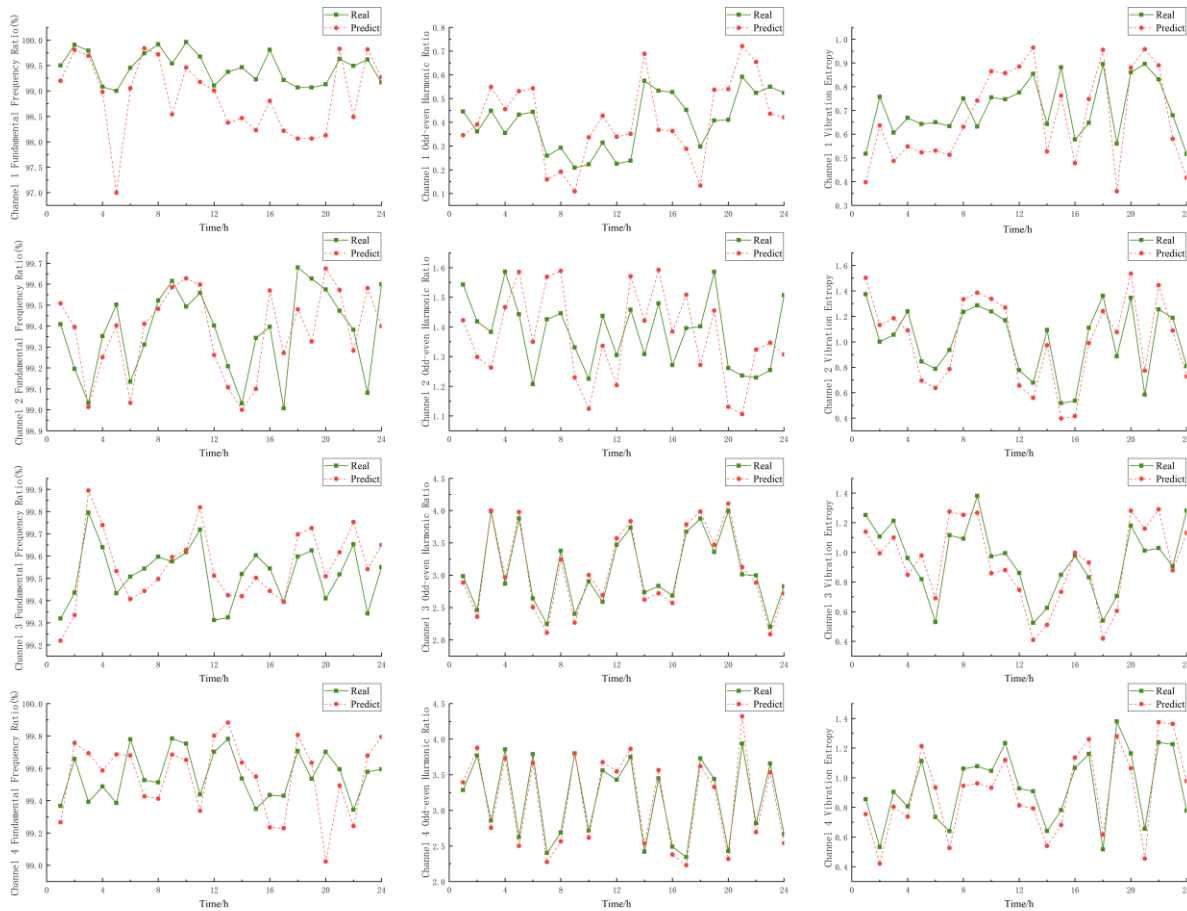


Figure 8. Prediction results of feature parameters of each channel

Table 4. Mean value and standard deviation of evaluation indicator results of different methods

Model	Evaluation Indicators		
	MAE	RMSE	MAPE(%)
LSTM	0.2419±0.71	4.3261±0.67	0.1691±0.48
BiLSTM	0.1823±0.6	4.0216±0.51	0.1642±0.32
CNN-BiLSTM	0.1573±0.39	3.1264±0.42	0.1548±0.31
PSO-CNN-BiLSTM	0.1547±0.47	2.0364±0.61	0.1449±0.24
SMA-CNN-BiLSTM	0.1461±0.43	1.793±0.52	0.1359±0.14
DESMA-CNN-BiLSTM	0.1412±0.01	1.7915±0.39	0.1317±0.02

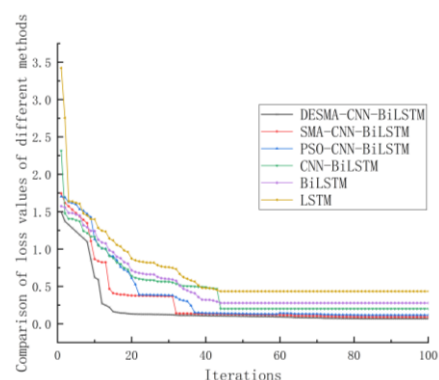
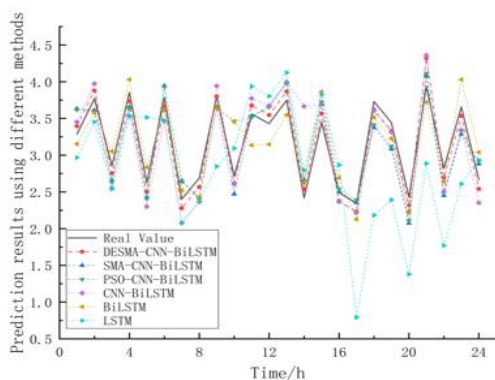


Figure 9. Comparison of prediction results by different methods

Figure 10. Comparison of loss values by different methods

## 6. CONCLUSION

In order to improve the accuracy of GIS equipment vibration signal feature prediction, this paper proposes a GIS vibration feature prediction method based on improved slime mold algorithm to optimize CNN-BiLSTM. In the first place, based on the historical vibration signal of the GIS equipment, the vibration feature parameters are extracted in the frequency domain through Fourier transformation. Furthermore, in order to enhance the feature utilization ability of the BiLSTM model, 1D CNN is used to extract feature parameters; At the same time, in order to enhance the population diversity of swarm intelligence optimization algorithm and avoid being trapped in local optimum, differential evolution mechanism is integrated into slime mold optimization algorithm (SMA). The improved SMA algorithm is used to optimize the number of neurons on the hidden layer, learning rate and other hyper-parameters of CNN-BiLSTM, avoiding the blindness of hyper-parameter selection. Finally, a feature prediction model of GIS equipment based on improved slime mold algorithm optimized CNN-BiLSTM is built. The multi-channel vibration signal of GIS equipment in a substation is analyzed through an example. The experimental results show that:

1) The prediction method of vibration features of GIS equipment based on DESMA-CNN-BiLSTM can accurately capture the change trend of different features of vibration signals in each channel, and the average prediction error is only 0.1317%.

2) DESMA algorithm boasts a high convergence speed and satisfactory population diversity. It can effectively avoid being trapped in local optimum in the optimization process. The improved strategy of swarm intelligence algorithm proposed in this paper improves the search performance of SMA algorithm.

## ACKNOWLEDGMENT

This work was supported by the Electric Power Research Institute of State Grid Chongqing Electric Power Company GIS/GIL Equipment Mechanical Vibration State Online Data Deep Learning and Software Visualization Project (Project No.: 4412100668).

## REFERENCES

- [1] Hou, P., Ma, H., Ju, P., Chen, X., Zhu, C. (2021). A new vibration analysis approach for monitoring the working condition of a high-voltage shunt reactor. *IEEE Access*, 9: 46487-46504. <https://doi.org/10.1109/ACCESS.2021.3068264>
- [2] Bojić, S., Babić, B., Uglešić, I. (2018). Comparative research into transients by switching of high voltage shunt reactor. *Electric Power Systems Research*, 162: 74-82. <https://doi.org/10.1016/j.epsr.2018.04.018>
- [3] Khan, M.A., Zheng, T. (2019). Modelling and design of a low-level turn-to-turn fault protection scheme for extra-high voltage magnetically controlled shunt reactor. *Energies*, 12(24): 4628. <https://doi.org/10.3390/en12244628>
- [4] Minh, T.P., Duc, H.B., Hoai, N.P., Cong, T.T., Cong, M.B., Thanh, B.D., Quoc, V.D. (2021). Finite element modeling of shunt reactors used in high voltage power systems. *Engineering, Technology & Applied Science Research*, 11(4): 7411-7416. <https://doi.org/10.48084/etasr.4271>
- [5] Minh, T.P., Duc, H.B., Quoc, V.D. (2022). Analysis of leakage inductances in shunt reactors: Application to high voltage transmission lines. *Engineering, Technology & Applied Science Research*, 12(3): 8488-8491. <https://doi.org/10.48084/etasr.4826>
- [6] Gao, S.G., Ji, S.C., Meng, L.M., Tian, Y., Zhang, Y.K. (2022). Evaluation method of operating state of high-voltage shunt reactor based on online monitoring system and acoustic-vibration characteristic prediction model. *Chinese Journal of Electrotechnical Technology*, 37(9): 2179-2189. <https://doi.org/10.19595/j.cnki.1000-6753.tces.210834>
- [7] Das, S., Sidhu, T.S., Zadeh, M.R.D., Zhang, Z. (2017). A novel hybrid differential algorithm for turn to turn fault detection in shunt reactors. *IEEE Transactions on Power Delivery*, 32(6): 2537-2545. <https://doi.org/10.1109/TPWRD.2017.2680456>
- [8] Zhu, M., Huang, Q.Q., Qi, Y.K., Zhou, K.J., Mei, J., Zhou, W., Zhang, J. (2020). Health assessment method of high voltage shunt reactor based on total discrete spectrum of vibration signal. *Electrical Measurement and Instrumentation*, 1-9. <https://kns.cnki.net/kcms/detail/23.1202.TH.20210616.1723.008.html>
- [9] Yao, C., Zhao, Z., Li, C., Chen, X., Zhao, Y., Zhao, X., Li, W. (2015). Noninvasive method for online detection of internal winding faults of 750 kV EHV shunt reactors. *IEEE Transactions on Dielectrics and Electrical Insulation*, 22(5): 2833-2840. <https://doi.org/10.1109/TDEI.2015.004969>
- [10] Jiang, N., Hao, B.X., Zhao, R.Y., Ma, H.Z., Xu, L., Li, L. (2019). Application of empirical wavelet transform in vibration signal analysis of UHV shunt reactor. In 2019 IEEE Milan Power Tech, 1-5. <https://doi.org/10.1109/PTC.2019.8810623>
- [11] Deng, W., Liu, H., Xu, J., Zhao, H., Song, Y. (2020). An improved quantum-inspired differential evolution algorithm for deep belief network. *IEEE Transactions on Instrumentation and Measurement*, 69(10): 7319-7327. <https://doi.org/10.1109/TIM.2020.2983233>
- [12] Chen, X., Pan, X., Liu, Z., Ma, H. (2019). Analysis of vibration signals of HUV shunt reactor based on CRP and RQA. In 2019 IEEE Milan Power Tech, 1-5. <https://doi.org/10.1109/PTC.2019.8810648>
- [13] Li, Q., Zhao, T., Zhang, L., Lou, J. (2011). Mechanical fault diagnostics of on-load tap changer within power transformers based on hidden Markov model. *IEEE Transactions on Power Delivery*, 27(2): 596-601. <https://doi.org/10.1109/TPWRD.2011.2175454>
- [14] Huang, W., Gan, C. (2018). A vector angle method of rolling bearing fault classification by phase-space reconstruction technique. *Journal of Testing and Evaluation*, 48(4): 2624-2638. <https://doi.org/10.1520/JTE20180280>
- [15] So, E., Verhoeven, R., Dorpmanns, L., Angelo, D. (2015). Traceability of loss measurements of extra high voltage three-phase shunt reactors. *IEEE Transactions on Instrumentation and Measurement*, 64(6): 1344-1349. <https://doi.org/10.1109/TIM.2015.2398956>
- [16] Zhu, M., Zhang, Z., Mei, J., Zhou, K., Chen, P., Qi, Y., Huang, Q. (2021). Data augmentation using DCGAN for



- improved fault detection of high voltage shunt reactor. In *Journal of Physics: Conference Series*, 1944(1): 012012. <https://doi.org/10.1088/1742-6596/1944/1/012012>
- [17] Zhu, M., Zhang, Z., Qi, Y., Fu, M., Mei, J., Huang, Q. (2020). Fault diagnosis of GIS equipment based on encoding position machine learning. 2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI), pp. 752-755. <https://doi.org/10.1109/CISAI54367.2021.00152>
- [18] Liu, B. (2020). Application research of short-term power load forecasting based on PSO-LSTM algorithm, Jilin University, 2020.
- [19] Shao, K.J., Zhou, Y., Song, H., Shan, H.Q. (2022). Output power prediction of combined cycle power plant based on PSO optimization BiLSTM. *Electric Power Science and Engineering*, 38(2): 9-17. <https://doi.org/10.3969/j.ISSN.1672-0792.2022.02.002>
- [20] Huang, D.M., Wang, C., Hu, A.D., Sun, J.Z., Sun, Y., Li, J.F. (2021). Tidal level prediction of tidal power stations based on CNN-BiLSTM. *Hydroelectric Power Generation. Water Power*, 47(10): 80-84. <https://doi.org/10.3969/j.issn.0559-9342.2021.10.016>
- [21] Li, S., Chen, H., Wang, M., Heidari, A.A., Mirjalili, S. (2020). Slime mould algorithm: A new method for stochastic optimization. *Future Generation Computer Systems*, 111: 300-323. <https://doi.org/10.1016/j.future.2020.03.055>
- [22] Ding, Q.F., Yin, X.Y. (2017). Review of differential evolution algorithms. *Journal of Intelligent Systems*, 12: 431-442. <https://kns.cnki.net/kcms/detail/23.1538.tp.20170606.1114.006.html>
- [23] Li, X.S., Ma, H.W., Lin, Y.Z. (2019). Structural damage recognition based on convolutional neural network. *Vibration and Shock*, 38(1): 159-167. <https://doi.org/10.13465/j.cnki.jvs.2019.01.023>
- [24] Han, X., Ning, S.C., Li, J.F., Fu, F., Wu, D.X. (2020). Trend prediction of telemetry parameters based on long-term and short-term neural network. *Measurement and Control Technology*, 39(12): 105-110, 137. <https://doi.org/10.19708/j.ckjs.2020.12.018>
- [25] Dai, J.G., Jiang, N. (2021). Cotton yield prediction method based on CNN-BiLSTM. *Chinese Journal of Agricultural Engineering*, 37(17): 152-159. <https://doi.org/10.11975/j.issn.1002-6819.2021.17.017>
- [26] Hu, P., Wu, Z.J., Zhou, X.Y., Deng, C.S. (2014). Dynamic differential evolution algorithm based on elite region learning. *Acta Electronica Sinica*, 42(8): 1522-1530. <https://doi.org/10.3969/j.issn.0372-2112.2014.08.010>