WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 2, Number 3, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v2n3.11
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Bearing Fault Diagnosis Based on SMA-VMD and CNN-LSTM

Nan Wang*

School of Electronic Engineering, Xi'an Shiyou University, Xi'an, Shaanxi 710065, China
*Corresponding Author: wangnan10789@163.com

ABSTRACT

To address the challenges of manual parameter setting dependency and low accuracy in bearing fault diagnosis using Variational Mode Decomposition (VMD), a method integrating Slime Mould Algorithm (SMA), VMD, and Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) network is proposed. Firstly, the SMA is employed to optimize the critical parameters of VMD using minimum information entropy as the fitness function, which resolves the issues related to manual parameter settings in VMD. Subsequently, the optimal parameter combination is used to process the bearing signals and extract the relevant Intrinsic Mode Functions (IMF). Finally, these IMFs are input into the CNN-LSTM network for fault diagnosis. The results demonstrate that this method significantly enhances the accuracy of motor bearing fault diagnosis, further improving the fault diagnosis accuracy.

KEYWORDS

Fault Diagnosis, Variational Mode Decomposition, CNN-LSTM

1. INTRODUCTION

Bearings are critical yet vulnerable components within rotating machinery, and their operational state is essential for ensuring the safe and reliable operation of mechanical equipment. Hence, monitoring and diagnosing the working condition of bearings holds significant importance.

In 1998, Norden E Huang[1] introduced the Empirical Mode Decomposition (EMD) method, which differs from traditional wavelet decomposition as it does not require predefined basis functions and exhibits adaptivity. However, EMD is prone to endpoint effects and mode mixing. These issues were later addressed through methods such as Ensemble Empirical Mode Decomposition (EEMD)[2] and Complete Ensemble Empirical Mode Decomposition (CEEMD)[3], which reduced these effects but at the expense of increased computational demand.

In 2014, Dragomiretskiy K [4]introduced the Variational Mode Decomposition (VMD) method, which improved the handling of endpoint effects and is suitable for non-stationary signal sequences. Nevertheless, the outcomes of VMD heavily depend on the initially set parameters by the user.

This paper utilizes the Slime Mould Algorithm to optimize the parameters of the VMD algorithm, selecting the intrinsic mode components with the smallest information entropy and integrating them with neural networks for diagnosing the state of bearings.

2. ALGORITHM INTRODUCTION

2.1. Slime Mould Algorithm (SMA)

In 2020, Li and others [5]were inspired by the foraging behavior of natural slime molds and proposed the Slime Mould Algorithm (SMA). The algorithm simulates the behavior and morphological changes of slime molds during foraging but does not model their complete life cycle. Additionally, the algorithm uses a weighting system to simulate the positive and negative feedback generated by the slime mold during its foraging process.

As it approaches food, the slime mold moves toward the scent of the food. When a vein of the slime mold contacts food, a diffusion wave quickly generates within the slime mold, accelerating the cytoplasmic flow, causing the vein to thicken. The mathematical model for this process is as follows:

$$X(t+1)=\begin{cases}X_b(t)+v_b\cdot(W\cdot X_A(t)-X_B(t)), r<p\\ \\ v_c\cdot X(t), r\geq p\end{cases} \tag{1}$$

Let t be the current iteration, $X(t)$ the position of the slime mold at iteration t,$X_b(t)$ the position with the highest concentration of food scent at time t, X the current position of the slime mold, $X_A(t)$ and$X_B(t)$ random individual positions of the slime mold, W the weight coefficient, and r a random value in the range [0,1]. The formula for p is:

$$p=\tanh|S(i)-D_{\mathrm{F}}| \tag{2}$$

where $S(i)$ is the fitness of X and D_{F} is the best fitness value among all iterations.

$$a=\operatorname{arctan}h\left(-\left(\frac{t}{\underset{t}{max}}+1\right)\right) \tag{3}$$

The weight coefficient W simulates the oscillation frequency of the biological oscillator of the slime mold when encountering food concentrations of varying densities:

$$W(\text{smellindex}(i))=\begin{cases}1+r\cdot\log\left(\frac{b_F-S(i)}{b_F-w_F}\right),\text{condition}\\ \\ 1-r\cdot\log\left(\frac{b_F-S(i)}{b_F-w_F}\right),\text{others}\end{cases} \tag{4}$$

$W(\text{smellindex}(i))$ for the condition $S(i)$ in the top half of the population, b_F andw_F are the best and worst fitness values in the current iteration, respectively.

During its search for food, the slime mold also divides some of its mass for random exploration. The position update formula for the slime mold, integrating these theories, is given by:

$$X^*=\begin{cases}\text{rand}\cdot(M_B-N_B)+N_B, \text{rand}<z\\ X_b(t)+v_b\cdot(W\cdot X_A(t)-X_B(t)), r<p\\ v_c\cdot X(t), r\geq p\end{cases} \tag{5}$$

where M_B and N_B are the upper and lower bounds of the search range. This study sets the value of z to 0.03 based on the reference [6].

2.2. Variational Mode Decomposition (VMD)

VMD is particularly effective for analyzing non-stationary signals as it captures transient and local frequency variations within the signal. Its mathematical formulation is as follows:

$$u_k(t) = A_k(t)\cos\varphi_k(t) \tag{6}$$

A constrained problem is constructed, resulting in a constrained variational model:

$$\begin{cases} min\{u_k\},[\omega_k]\left\{\sum_k \| \partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\|_2^2\right\} \\ \text{s.t.}\sum_k u_k(t) = f(t) \end{cases} \tag{7}$$

This constrained variational problem is then transformed into an unconstrained variational problem, yielding the augmented Lagrangian expression:

$$\begin{gathered} A = \alpha\sum_K ||\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}||_2^2 \\ B = ||f(t)-\sum_k u_k(t)||_2^2 \\ C = \left\langle\lambda(t), f(t)-\sum_k u_k(t)|\right\rangle \\ L(\{u_k\},\{\omega_k\},\lambda) = A + B + C \end{gathered} \tag{8}$$

2.3. SMA Optimization of VMD Parameters

Information entropy is a physical quantity that describes the level of uncertainty in a system. The greater the uncertainty of the probability distribution p, the higher the corresponding entropy value; conversely, the lower the uncertainty of p, the lower the entropy value.Therefore, if the decomposed Intrinsic Mode Functions (IMF) contain fault information, due to periodic impacts, they will appear more ordered and have lower entropy values[8].

In this context, the paper uses the SMA to optimize the parameters k and α of VMD. Using minimum information entropy as the fitness function and combining permutation entropy with mutual information entropy allows for a more comprehensive assessment of VMD's performance, making the optimization more adaptive.

The formula for information entropy is given as:

$$H(x) = -\sum_{i=1}^{N} p_i \lg p_i \tag{9}$$

The steps for SMA to optimize the VMD parameters are detailed in reference[5], and the optimization process is illustrated in Figure 1.

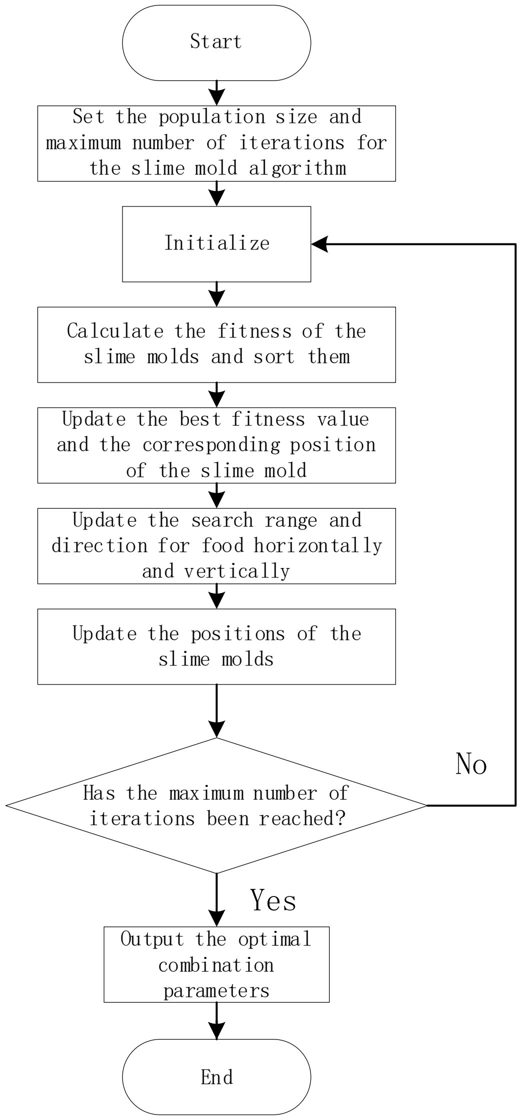

Figure 1: Flowchart of SMA Optimization for VMD

3. FAULT DIAGNOSIS MODEL BASED ON SMA-VMD AND CNN-LSTM

3.1. CNN-LSTM Model

Convolutional Neural Networks (CNN) are a widely used model in the field of deep learning. Typically, a CNN includes three fundamental layers: convolutional layers, pooling layers, and fully connected layers, along with optional layers such as normalization layers, activation layers, and dropout layers.

CNNs excel at extracting spatial features from data, but the learned features are confined to local areas, making it challenging to capture global information. On the other hand, LSTM is adept at extracting temporal features from data but requires significant computational resources and training time. Overall, this network structure combines the strengths of both CNN and LSTM, integrating features between the two to mitigate their individual weaknesses and enhance model performance and accuracy.

The structure of the CNN-LSTM model is shown in Figure 2.

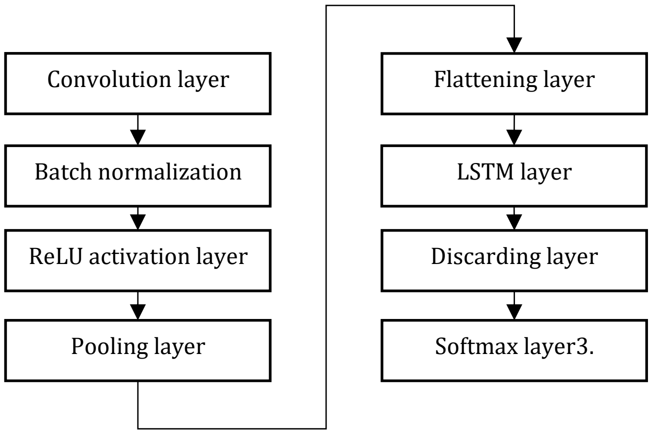

Figure 2: Structure of the CNN-LSTM Model

3.2. Diagnostic Method Process

The specific steps of the diagnostic method based on SMA-VMD and CNN-LSTM are as follows:

1) Collect fault signals from 10 different operating conditions of bearings;

2) Use information entropy as a metric to find the optimal parameter combination for VMD via the SMA algorithm and process the fault signals;

3) Following the principle of minimal information entropy, select the best Intrinsic Mode Function (IMF) component for each condition, resulting in 10 IMF components;

4) Use these 10 selected IMF components as feature vectors to train the CNN-LSTM model and perform fault diagnosis.

4. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental data in this paper is sourced from the rolling bearing vibration data collected by Case Western Reserve University, specifically focusing on the drive-end vibration signals. The vibration signals are classified into four types: inner race fault, rolling element fault, outer race fault, and normal signals. Faults of diameters 0.007 inches, 0.014 inches, and 0.021 inches for the inner race, rolling element, and outer race are respectively identified and sequenced from 1 to 10.

For each of these 10 different fault conditions, each sample data includes 2048 data points. There are 150 sample sets for each condition, comprising 120 training datasets and 30 testing datasets. Specific sample information is illustrated in Table 1.

Table 1: Specific Sample Information

Fault Diameter (inches)	Fault Type	Label	Training Data Samples	Test Data Samples
	Normal Condition	1	120*2048	30*2048
0.007	Inner Race Fault	2	120*2048	30*2048
	Rolling Element Fault	3	120*2048	30*2048
	Outer Race Fault	4	120*2048	30*2048
0.014	Inner Race Fault	5	120*2048	30*2048
	Rolling Element Fault	6	120*2048	30*2048
	Outer Race Fault	7	120*2048	30*2048
0.021	Inner Race Fault	8	120*2048	30*2048
	Rolling Element Fault	9	120*2048	30*2048
	Outer Race Fault	10	120*2048	30*2048

The primary parameters for VMD are K and α[9]. The completeness of the vibration signal decomposition relies on these two main input parameters. The initial population size of the SMA optimization algorithm is set to 30, with 100 iterations. An inner race fault sample with a fault diameter of 0.007 inches is selected for VMD parameter optimization using the SMA algorithm, resulting in the corresponding SMA iteration curve shown in the figure3. The optimal parameters K and α for VMD decomposition in this state are found to be 4 and 457, respectively.

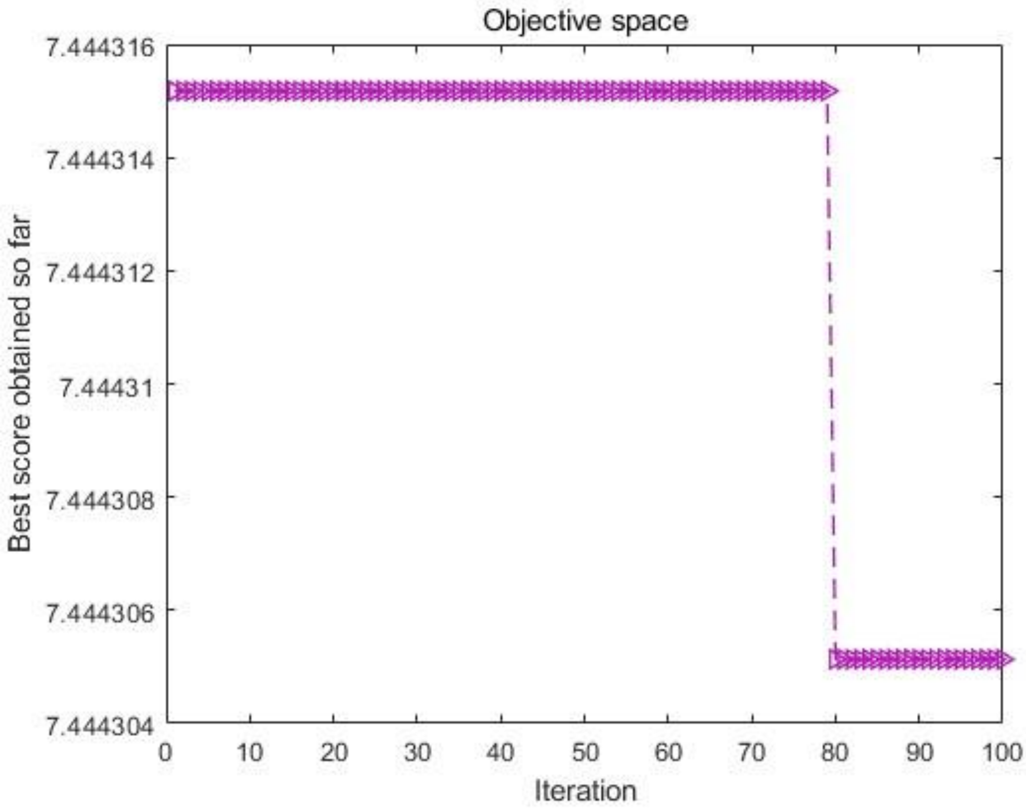

Figure 3: SMA Iteration Curve

Using the SMA-optimized VMD parameters to process this sample signal, the resultant time-domain graph is shown in Figure 4, and the central frequency variation curves for each IMF are displayed in Figure 5. The graph indicates that there is no overlap in the central frequencies of the IMFs.

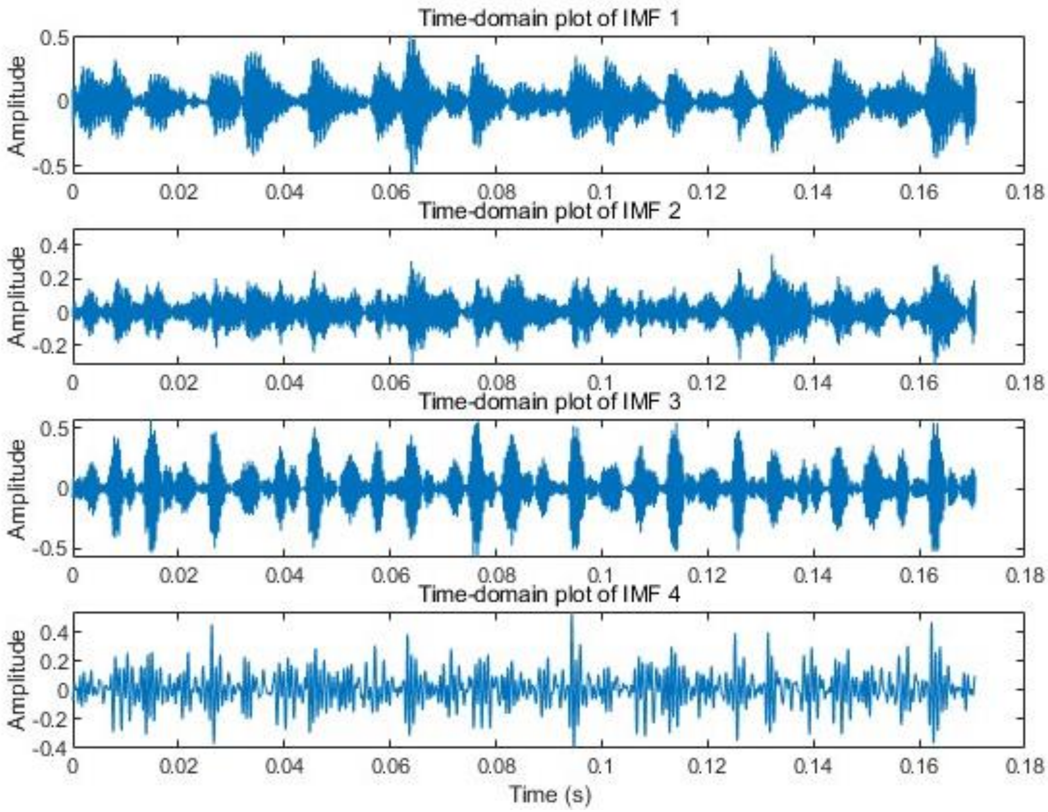

Figure 4: Time-Domain Graph

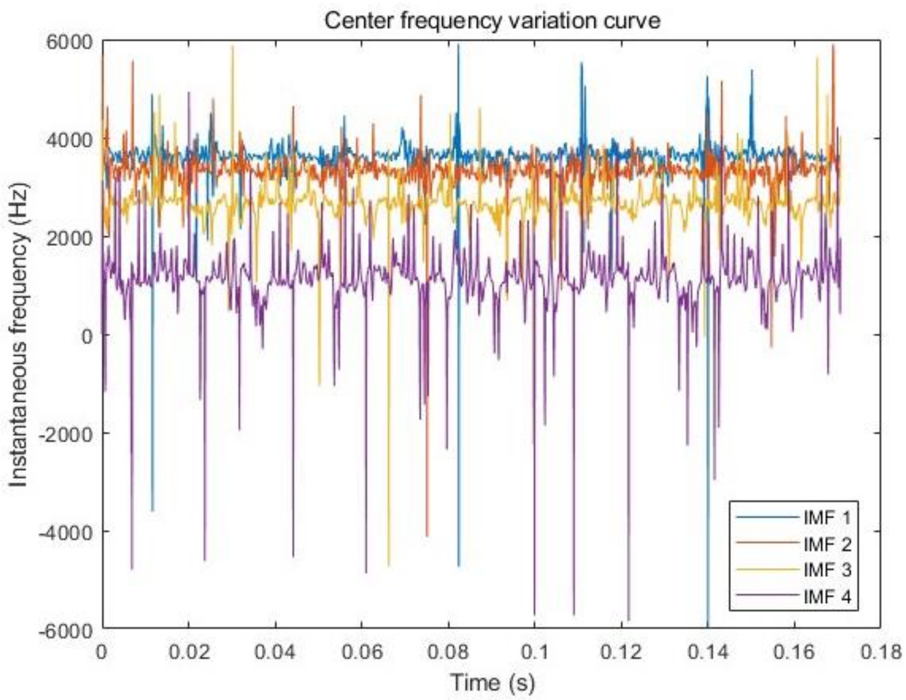

Figure 5: Central Frequency Change Graph

SMA is used for VMD parameter optimization for each fault condition, with the optimal K and α parameters for VMD decomposition of each fault state shown in Table 2.

Table 2: VMD Parameter Optimization Results

Bearing Fault Type	K Value	α Value
1	10	2413
2	3	457
3	6	2292
4	3	452
5	3	461
6	9	100
7	10	103
8	7	100
9	8	100
10	5	705

In MATLAB R2023b, using the same dataset and parameters, four different fault diagnosis models are trained and tested: SMA-VMD-CNN, SMA-VMD-LSTM, and SMA-VMD-CNN-LSTM. The population size is uniformly set to 30, with a maximum of 500 iterations.

The CNN model parameters are as follows: 10 filters of size 32×1 in the convolutional layer, with a stride of 1; a max pooling layer with a pooling window of 2×1 and a stride of 2. Training is conducted using the Adam optimizer, with 150 epochs and a learning rate of 0.01.

LSTM model parameters are set with 2 layers in the hidden layer, a dropout rate of 0.2, and 150 neurons, using the Adam optimization algorithm.

The classification results of the test sets for the three models are respectively shown in Figures 6, 7, and 8.

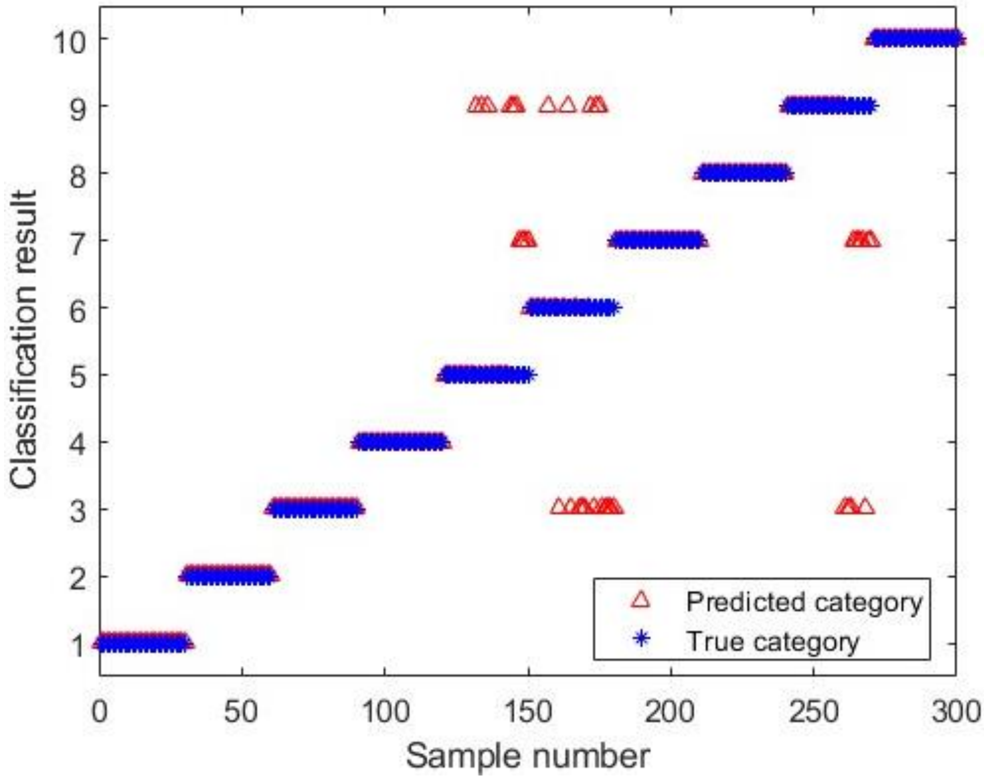

Figure 6: Classification Results of SMA-VMD-CNN

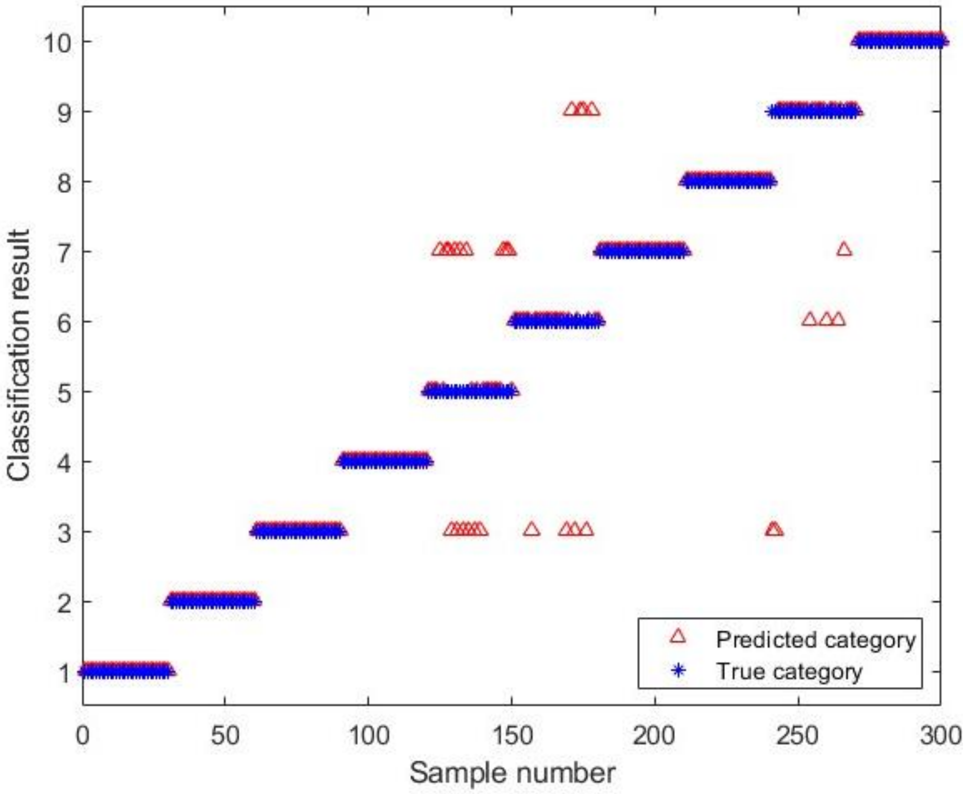

Figure 7: Classification Results of SMA-VMD-LSTM

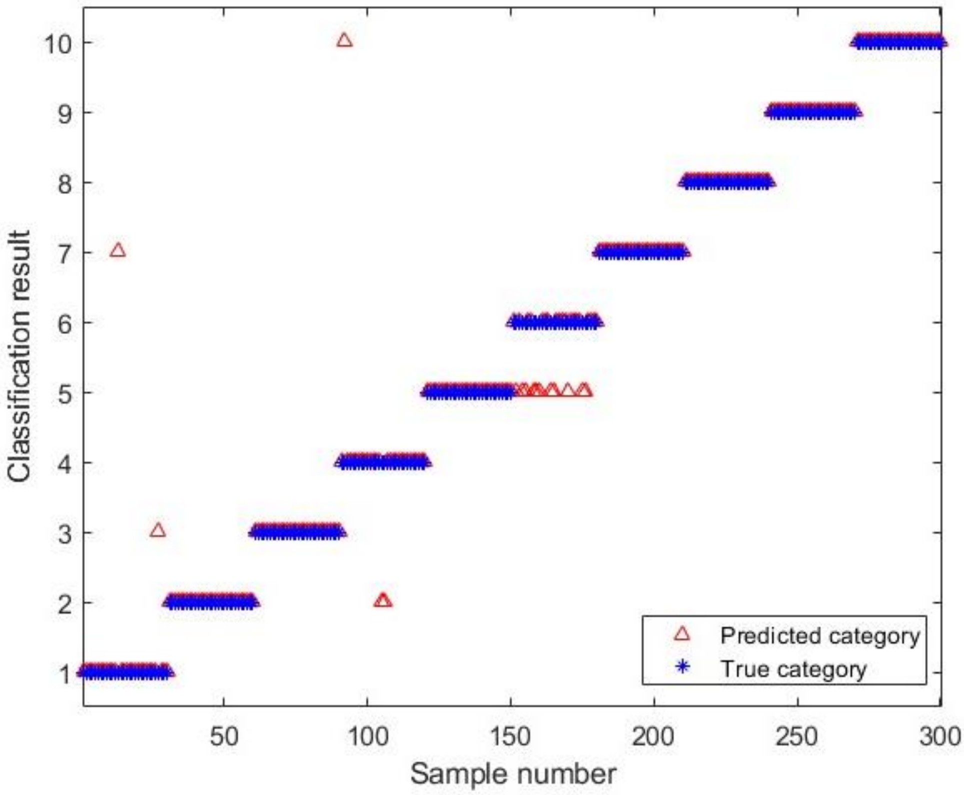

Figure 8: Classification Results of SMA-VMD-CNN-LSTM

Each model is run independently 10 times, and the average accuracy is recorded in Table 3.

Table 3: Accuracy of Different Models

Diagnostic Model	Accuracy (%)
SMA-VMD-CNN	88.7
SMA-VMD-LSTM	90.3
SMA-VMD-CNN-LSTM	94.6

As shown in Table 4, the average diagnostic accuracy of SMA-VMD-CNN is 88.7%, SMA-VMD-LSTM is 90.3%, and SMA-VMD-CNN-LSTM is 94.6%. Consequently, the SMA-VMD-CNN-LSTM fault diagnosis model achieves a higher recognition rate, further enhancing the fault diagnosis capability for rolling bearings.

5. CONCLUSION

To enhance the accuracy of bearing fault diagnosis, this paper proposes a fault diagnosis model utilizing SMA-VMD combined with CNN-LSTM. By using the SMA to optimize the main parameters K and α in the VMD algorithm, and by leveraging the characteristics of CNN and LSTM models, a hybrid CNN-LSTM model is constructed. This integration improves the predictive accuracy over that of using the individual models alone.

The combination of SMA-VMD data processing with the CNN-LSTM hybrid model culminates in the SMA-VMD-CNN-LSTM final model. Experimental results demonstrate that the proposed model significantly enhances the accuracy of bearing fault diagnosis, thereby confirming the effectiveness of integrating these methodologies for improved diagnostic performance.

REFERENCES

[1] Huang, N. E., Shen, Z., Long, S. R., et al. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 454, no. 1971, pp. 903-995, 1998.

[2] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," Advances in adaptive data analysis, vol. 1, no. 01, pp. 1-41, 2009.

[3] J. R. Yeh, J. S. Shieh, and N. E. Huang, "Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method," Advances in adaptive data analysis, vol. 2, no. 02, pp. 135-156, 2010.

[4] Dragomiretskiy, K., Zosso, D. J. "Variational mode decomposition." IEEE Transactions on Signal Processing, vol. 62, no. 3, pp. 531-544, 2013.

[5] S. Li, H. Chen, M. Wang, et al., "Slime mould algorithm: A new method for stochastic optimization," Future Generation Computer Systems, vol. 111, pp. 300-323, 2020.

[6] H. Chen, "Corrosion fault diagnosis of grounding grid based on multi-layer dimension reduction and FOA optimization," Insulators and Surge Arresters, no. 03, pp. 93-98, 2016.

[7] A. Dibaj, R. Hassannejad, M. M. Ettefagh, et al., "Incipient fault diagnosis of bearings based on parameter-optimized VMD and envelope spectrum weighted kurtosis index with a new sensitivity assessment threshold," ISA Transactions, vol. 114, pp. 413-433, 2021.

[8] H. Li, X. Wu, T. Liu, and Q. Chen, "Fault feature extraction of rolling bearings based on information entropy optimized variational mode decomposition," Journal of Vibration and Shock, vol. 37, no. 23, pp. 219-225, 2018.

[9] A. Dibaj, R. Hassannejad, M. M. Ettefagh, et al., "Incipient fault diagnosis of bearings based on parameter-optimized VMD and envelope spectrum weighted kurtosis index with a new sensitivity assessment threshold," ISA Transactions, 2020.