BEARING FAULT DIAGNOSIS BASED ON CNN-MGO-LSSVM

Yiming Xue[1], Chuanbo Wen[2]*

[1]*School of Electrical Engineering,Shanghai DianJi University, Shanghai,China*
[*]wencb@sdju.edu.cn

Keywords:FAULT DIAGNOSIS,MGO MOSS GROWTH OPTIMIZATION ALGORITHM,PSO PARTICLE SWARM ALGORITHM,LEAST SQUARES SUPPORT VECTOR MACHINE,TWO-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORK

Abstract

Bearings play a critical role in rotating machinery, directly determining the safety and operational health of equipment. However, with the advancement of industrial society, bearing failures due to malfunctions have become increasingly prominent.Therefore, this study proposes an integrated model for bearing fault diagnosis by optimizing two key parameters of the Least Squares Support Vector Machine (LSSVM) using a combination of the MGO moss growth optimization algorithm and the PSO particle swarm algorithm: the regularization parameter and the kernel parameter. The high-dimensional feature vectors from the output layer of a two-dimensional convolutional neural network (2D-CNN) are used as input for the LSSVM. This model demonstrates higher accuracy and automation.

1 Introduction

With the development of industrial technology, research on bearing fault diagnosis has been continuously deepening, and one of the most commonly used methods is convolutional neural networks (CNN)[1].Convolutional neural networks in deep learning can process large amounts of data and effectively extract data features.Hoang et al. proposed a feature-free bearing fault diagnosis method based on convolutional neural networks (CNNs).1-D vibration signals are converted into 2-D data, referred to as vibration images. The vibration images are then fed into the CNN for bearing fault classification[2]. This method effectively identifies the fault damage categories of bearings, but the model may perform poorly due to overfitting when data is scarce.Wang et al. proposed a method for bearing fault diagnosis using a multi-attention one-dimensional convolutional neural network (MA1DCNN)[3]. MA1DCNN can adaptively recalibrate the features of each layer and enhance the feature learning of fault pulses. The network incorporates multiple attention modules, with each layer adding additional parameters. The JAM module significantly increases the weight calculation of channels and time through cascaded EAM and CAM.However, when the number of JAM layers exceeds five, the diagnostic accuracy decreases, and the high complexity may lead to overfitting.Zhang et al. proposed a fault diagnosis method based on deep learning[4]. The model adopts a hybrid architecture of one-dimensional convolutional neural networks (1DCNN) and support vector machines (SVM) to achieve dynamic adaptive feature extraction through 1DCNN and input the optimized feature vectors into the classifier.In this process, the particle swarm optimization (PSO) algorithm is introduced to adaptively adjust the hyperparameters of SVM, constructing a composite classification system that integrates deep feature learning and intelligent parameter optimization. However, SVM requires solving convex quadratic programming (QP) problems, with a time complexity of $O\left(n^2\right)$ or higher, resulting in slower training speeds.

1D-CNN can process raw time series data but struggle to capture frequency-domain information or time-frequency features. 2D-CNN can simultaneously extract features from both the time domain and the frequency domain (or other transform domains). Bearing fault signals are typically transformed into two-dimensional images using short-time Fourier transform (STFT), wavelet transform (CWT), or time-frequency analysis. 2D-CNNs can better capture local patterns and spatial relationships within these images.Furthermore, 2DCNNs perform more effectively in complex fault patterns, extracting richer fault features from time-frequency maps. Combined with the powerful expressive capabilities of deep learning models, they typically achieve higher diagnostic accuracy[5][6][7][8].Zhang et al. proposed a method for diagnosing rotary machine imbalance faults by combining time-frequency feature oversampling (TFFO) with convolutional neural networks (CNN)[9]. They used synthetic minority oversampling technique (SMOTE) to expand the minority samples to achieve TFFO, obtaining a balanced dataset that was then input into an improved 2DCNN based on LeNet-5 to achieve fault diagnosis.This paper proposes a bearing fault diagnosis model combining 2DCNN with MGO_PSO-LSSVM. An optimization algorithm combining MGO and PSO is used to optimize the two parameters of LSSVM. LSSVM transforms the inequality constraints in SVM into equality constraints, converting the optimization problem into solving a system of linear equations, thereby avoiding the complex quadratic programming issues in traditional SVM. The optimization objective of LSSVM is more explicit,minimizing the mean squared error, making the parameter optimization objective more intuitive. Optimal values can be quickly determined through cross-validation or grid search. In bearing fault diagnosis, LSSVM combined with CWT time-frequency. features can rapidly classify different fault types, achieving higher computational efficiency than SVM.

2 Methodology

2.1Continuous wavelet transform (CWT)

Continuous Wavelet Transform (CWT)[10] uses wavelet basis functions and adjusts the scale factor and time shift factor of the wavelet basis functions to analyze the characteristics of signals at different frequency scales and time positions. Compared with the short-time Fourier transform (STFT), CWT is more flexible. The mathematical expression of CWT is:

$$CWT(a,\tau) = \left[f(t),\psi_{a,\tau}(t)\right] = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)\psi^*\left(\frac{t-\tau}{a}\right)dt \tag{1}$$

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-\tau}{a}\right), \quad a,\tau \in R, a>0 \tag{2}$$

In this equation, represents the scaling factor, represents the translation factor, and equation (2) represents the selected wavelet basis function.

2.2Convolutional Neural Network (CNN)

Convolutional Neural Network[11] is a deep learning model mainly composed of convolution layers, pooling layers, and fully connected layers, which performs well in fields such as image recognition.

*2.2.1 Convolutional Layer:*The convolution layer is the core component of CNN, consisting of multiple convolution kernels that extract features from the input data through convolution operations.The mathematical expression for convolution operations is:

$$x_j^{\text{out}} = f_c\left(\sum_{i \in M_j} x_i^{\text{input}} \cdot k_{ij} + b_j\right)x \tag{3}$$

Among them, x_i^{input} represents the local features of the $i-th$ convolutional layer input; k_{ij} and b_j are the weight matrix and bias term corresponding to the $j-th$ convolution layer, respectively; M_j is the set of input feature maps, and x_j^{out} is the output feature map.After convolution, a nonlinear activation function is typically used to introduce nonlinearity.

*2.2.2 Pooling Layer:*Pooling layers reduce the dimensionality of feature maps, which reduces computation while preventing overfitting and improving network robustness. Pooling is typically divided into average pooling and max pooling.

*2.2.3Fully Connected Layer:*The fully connected layer maps the feature maps extracted by the convolution layer and pooling layer to the target space. The $soft\max$ function is typically used for classification tasks during output. The mathematical expression for the $soft\max$ function is:

$$f(z_i) = \frac{\exp(z_i)}{\sum_{i=1}^{c}\exp(z_i)} \tag{4}$$

Among them, $z_i = \sum_{i \in M_j}\left(a_i^{\text{im}} \cdot w_{ij}\right) + b_j$ is the logarithm of the output values of neurons in the fully connected layer.

2.3MGO_PSO Optimization Algorithm

Moss Growth Optimization (MGO)[12] is a novel meta-heuristic algorithm inspired by the growth process of moss. MGO develops a creative mechanism called "determining wind direction," which utilizes the positional relationship between most individuals and the optimal individual to determine the evolutionary direction of all individuals in the population, effectively avoiding the problem of getting stuck in local optima.This study combines the MGO moss growth optimization algorithm with the PSO particle swarm optimization algorithm[13] to develop an optimized algorithm for bearing fault diagnosis. This algorithm is used to optimize two key parameters of LSSVM and minimize classification error. The optimized algorithm formula is as follows:

Particle Velocity: $Vi(0) = 0, \quad \forall \mathrm{i} \in \{1,2,....N\}$ (5)

Global Optimal Position:

$$g(0) = \arg\min_{X_i(0)} f(X_i(0)) \tag{6}$$

Dynamic Weighting Update: $w(t) = 0.9 - 0.5 \cdot \dfrac{t}{T}$ (7)

Hybrid Update Strategy: For each iteration $t \in \{1,2,....T\}$, the update rule for each particle i is:

$$\begin{cases} if\ t \equiv 0^{(\bmod 3)}: \begin{cases} V_i(t) = w(t)V_i(t-1) + r_1 \otimes \left(g(t-1) - X_i(t-1)\right) \\ X_i'(t) = X_i(t-1) + V_i(t) \end{cases} \\ otherwise: X_i'(t) = X_i(t-1) + w(t)\left(g(t-1) - X_i(t-1)\right) \end{cases} \tag{8}$$

where $r_1 \in [0,1]^{\dim}$ is a uniform random variable and $\otimes$ denotes the ***Hadamard*** area.

Boundary Constraints: $X_i'(t) = \text{clip}\left(X_i'(t), lb, ub\right)$ (9)

Adaptability Assessment and Updating:

$$\begin{cases} if\ f\left(X_i'(t)\right) < f\left(X_i(t-1)\right):\ X_i(t) = X_i'(t) \\ \begin{cases} if\ f\left(X_i'(t)\right) < f\left(g(t-1)\right): g(t) = X_i'(t) \\ otherwise: g(t) = g(t-1) \end{cases} \\ otherwise: X_i(t) = X_i(t-1) \end{cases} \tag{10}$$

N is the number of particles; T is the maximum number of iterations; dim is the dimension of the search space; lb and ub are the lower and upper bounds of the variables, and $g(t)$ is the global optimal position at the t iteration.This algorithm combines a dynamic weight adjustment mechanism with the speed update strategy of particle swarm optimization (used every 3 iterations) and a direct displacement strategy guided by the global optimum, thereby enhancing local exploration capabilities while maintaining population diversity.

2.4Least Squares Support Vector Machine (LSSVM)

Least Squares Support Vector Machine[14] transforms the optimization problem of SVM into a constrained quadratic programming problem and uses the least squares method for optimization and solution, thereby handling nonlinear classification and regression problems.

The optimization objective formula is as follows:

$$\min_{w,b,e} \mathcal{J}(w,e)=\frac{1}{2}w^{\top}w+\gamma\cdot\frac{1}{2}\sum_{i=1}^{N}e_i^2 \text{s.t.} \quad y_i\left[w^{\top}\varphi(x_i)+b\right]=1-e_i,\quad \forall i \tag{11}$$

The Lagrange multiplier method is given by the following formula:

$$\mathcal{L}(w,b,e;\alpha)=\mathcal{J}(w,e)-\sum_{i=1}^{N}\alpha_i\left\{y_i\left[w^{\top}\varphi(x_i)+b\right]-1+e_i\right\} \tag{12}$$

In this equation, a_i represents the Lagrange multiplier, which is also the support value.

Solve for the optimal conditions using the following formula:

$$\begin{cases}\dfrac{\partial\mathcal{L}}{\partial w}=0 & \rightarrow w=\sum_{i=1}^{N}\alpha_i y_i\varphi(x_i)\\ \dfrac{\partial\mathcal{L}}{\partial b}=0 & \rightarrow \sum_{i=1}^{N}\alpha_i y_i=0\\ \dfrac{\partial\mathcal{L}}{\partial e_i}=0 & \rightarrow \alpha_i=\gamma e_i,\quad i=1,\ldots,N\\ \dfrac{\partial\mathcal{L}}{\partial \alpha_i}=0 & \rightarrow y_i\left[w^{\top}\varphi(x_i)+b\right]-1+e_i=0,\quad i=1,\ldots,N\end{cases} \tag{13}$$

Solve the dual problem using the following formula:

$$\begin{bmatrix}0 & y^{\top}\\ y & \Omega+\dfrac{I}{\gamma}\end{bmatrix}\begin{bmatrix}b\\ \alpha\end{bmatrix}=\begin{bmatrix}0\\ I_v\end{bmatrix} \tag{14}$$

$$where\Omega_{ij}=y_i y_j\varphi(x_i)^{\top}\varphi(x_j)=y_i y_j K(x_i,x_j), \quad for(i,j=1,\ldots,N) \tag{15}$$

$$and y=[y_1;\ldots;y_N], I_v=[1;\ldots;1] \tag{16}$$

3 Model Establishment

This paper establishes a hybrid model combining CNN and optimized LSSVM, which is efficient and robust, capable of achieving high-precision diagnosis under different operating conditions and complex fault scenarios. The CNN part is responsible for feature extraction, while LSSVM uses an optimization algorithm combining MGO and PSO to adjust parameters and improve classification performance.By extracting multi-level features through a convolutional neural network (CNN), the model introduces an optimization algorithm combining MGO and PSO to dynamically adjust the kernel parameters of LSSVM, overcoming the efficiency bottleneck of traditional parameter tuning and improving classification accuracy.By combining the deep features of CNN with the statistical learning advantages of LSSVM, the model achieves high accuracy in fault classification tasks and converges quickly on the test set. When the training data scale is limited, the model enhances its generalization ability through batch normalization and Dropout technology, effectively alleviating the overfitting problem in small sample scenarios. Its structural model is shown in Fig. 1.

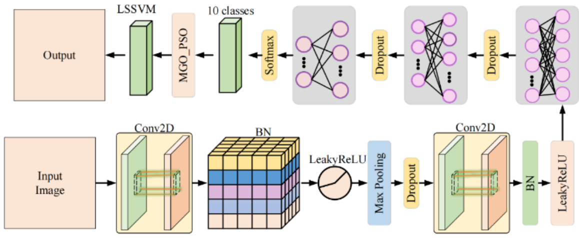


Fig. 1 Model structure diagram

This paper uses continuous wavelet transform (CWT) to convert the original vibration signal into a time-frequency map, inputs the time-frequency image into a two-dimensional convolutional neural network (2D-CNN) for adaptive fault feature extraction, takes the results of the fully connected layer as input for the least squares support vector machine, and uses a combined MGO and PSO algorithm to optimize the two key parameters of LSSVM to complete the multi-level classification task.

3.1Continuous wavelet transform (CWT)

This paper employs color three-channel time-frequency images generated using continuous wavelet transform. Wavelets are selected as wavelet basis functions due to their excellent time-frequency localization properties, making them suitable for analyzing non-stationary signals. A total of 256 wavelets of different scales are generated, and the wavelet center frequencies are obtained along with the scale parameters, ultimately yielding the scale sequence and its corresponding frequencies.By loading the original vibration signal data, the signal is extracted from each sample, the wavelet coefficient matrix is calculated, and the wavelet coefficients are processed using absolute values. A time-frequency plot is plotted, and all images are uniformly scaled to 64×64×3 (RGB three channels) to meet the subsequent CNN input requirements.

3.2 2DCNN-LSSVM

The input layer receives the time-frequency maps generated by continuous wavelet transform. The first convolutional layer uses a convolutional kernel with 16 channels to extract local features. The second convolutional layer uses a convolutional kernel with 32 channels to capture more complex spatial patterns.Each convolutional layer uses the LeakyReLU activation function. The LeakyReLU activation function has a small positive slope when $x<0$. Compared to the ReLU activation function, LeakyReLU avoids the issue of neuron death, allowing gradients to propagate. Its formula is as follows:

$$\text{LeakyReLU}(x)=\begin{cases}x, & x>0\\ ax, & x\le 0\end{cases} \tag{17}$$

Batch normalization is performed to improve stability. It also includes a Dropout layer with a dropout rate of 0.1, which means that 10% of the neurons are discarded to prevent overfitting. The pooling layer is max pooling, which gradually reduces the feature map size and enhances translation invariance.The fully connected layer flattens the multi-dimensional feature maps into vectors, which are then fed into the LSSVM module. The kernel function of the LSSVM module uses radial basis functions (RBF) to map features to a high-dimensional space, with the formula as follows:

$$\boldsymbol{K}(\boldsymbol{x}_i,\boldsymbol{x}_j)=\exp\left(-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{2\sigma^2}\right) \tag{18}$$

By adjusting the kernel parameters (γ: regularization coefficient, σ^2: kernel width) using the MGO_PSO algorithm, the classification error is minimized. The model structure diagram is shown in Fig. 2.

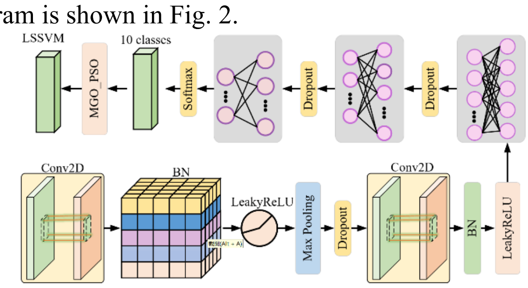


Fig. 2 2DCNN-LSSVM model structure diagram

4 Results

To validate the performance of the fault diagnosis model adopted in this study for common bearing faults, this paper uses the bearing fault dataset from Case Western Reserve University (CWRU) [15]. The CWRU dataset includes:

Ball faults: As a critical internal component of bearings, ball faults can directly interfere with the smooth operation of the bearing. Their vibration signals typically exhibit high-frequency characteristics in spectral analysis.

Inner race (Inner Race) faults: Damage to the inner race alters the dynamic characteristics of the bearing system. The corresponding vibration signals are primarily low-frequency components, but as the damage worsens, the signal amplitude and frequency distribution exhibit significant nonlinear changes.

Outer race (Outer Race) failure: This type of failure often causes instability in equipment operation. Through frequency domain analysis, characteristic frequency components related to the failure location can be observed, and the presence of such components can serve as criteria for outer race damage.

Normal operating condition (Normal): Data from fault-free operation serves as a baseline reference. By comparing and analyzing the data, the distinctive features of signal differences under different fault modes can be effectively identified, providing a theoretical basis for fault diagnosis.

4.1Preprocessing of experimental data

The drive end bearings of the CWRU motors are SKF 6205-2RS, which were damaged by EDM, with artificial damage on the inner, ball and outer races of the bearings. The damage points on the outer ring were located at 3, 6 and 12 o'clock respectively. The vibration signals were collected by accelerometers at a sampling frequency of 12 kHz under loads of 0, 1, 2, and 3 HP and at speeds of 1797 r/min, 1772 r/min, 1750 r/min, and 1730 r/min, respectively. This experiment was conducted at 12 kHz sampling frequency and the bearing data were collected at a rotational speed of 1797 r/min and at 0, 1, 2, and 3 HP load conditions. The operating conditions cover four categories: Normal, Inner Race Fault, Outer Race Fault, and Ball Fault. Among them, each type of fault is set to 0.007, 0.014 and 0.021 inches wide three damage levels, a total of 10 states. The size of the fault is 0.007 inches ball fault named B007, inner ring fault named IR007, outer ring fault named OR007, normal bearing named Normal, and so on the damage degree of 0.014 inches and 0.021 inches.The experimental setup of the dataset is shown in Fig. 3. The generated time-frequency plots are shown in Fig. 4.

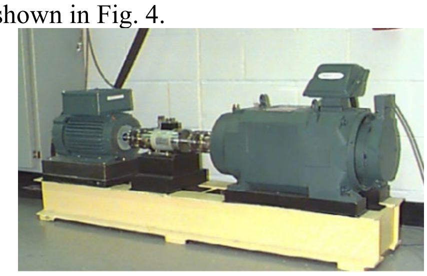


Fig. 3 Data set experiment settings

Its components are: a 1.5 kW (2 horsepower) electric motor; a torque sensor/encoder; a power tester; and an electronic controller.

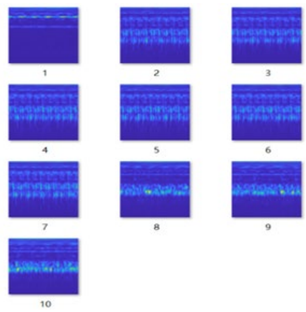
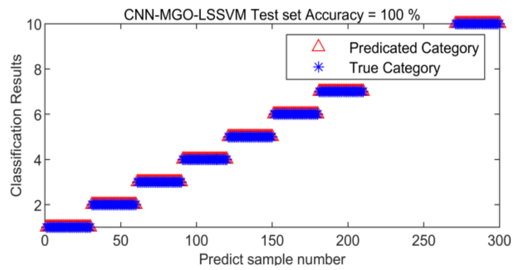


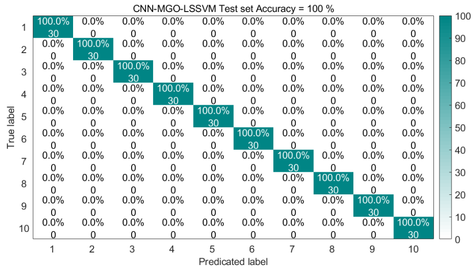
Fig. 4 Time-frequency diagram

4.2Experimental Validation

The model used in this study achieved an average small-batch accuracy of 99.92% after several rounds of training, with a final small-batch accuracy of 100%, a small-batch distortion of 2.2713×10^{-5}, and a base learning rate of 0.0010. The accuracy plot is shown in Fig. 5(a), and the confusion matrix is shown in Fig.5(b).



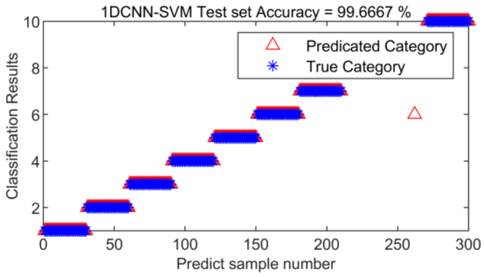
a. Accuracy chart



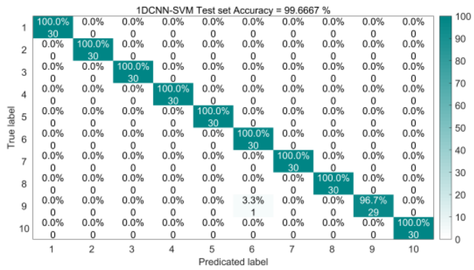
b. Matrix confusion diagram

Fig. 5 test results of CNN-MGO-LSSVM

Compared to the 1DCNN-SVM model, the mini-batch loss value of the 1DCNN-SVM model on this test set is 0.0064, with a base learning rate of 0.0010 and an accuracy rate of 99.6667%. The accuracy rate plot and confusion matrix of the 1DCNN-SVM test set are shown in Fig. 6(a) and Fig. 6(b) below.As can be seen, the model adopted in this study performs well in sample generation, with a significantly higher accuracy rate than the 1DCNN-SVM model, and its loss value is significantly lower than that of the 1DCNN-SVM model, indicating that the model's performance and generalization ability have been improved.Upon examining the confusion matrix, the accuracy rate for the 10 fault detection cases in the confusion matrix of the model used in this study is 100%, which is clearly superior to the confusion matrix of the 1DCNN-SVM model.



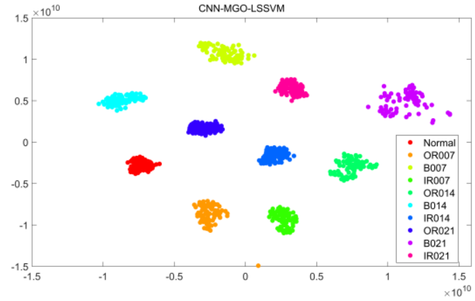
a. 1DCNN-SVM Accuracy chart



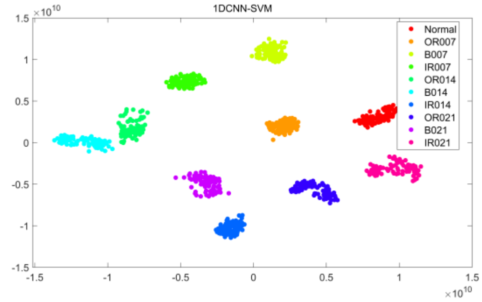
b. 1DCNN-SVM Matrix confusion diagram

Fig. 6 test results of 1DCNN-SVM

According to the two network models shown in Fig. 7(a) and Fig. 7(b), where 10 colors represent 1 normal bearing state and 9 fault states, comparing the model used in this study with the 1DCNN-SVM model, it can be seen that the model used in this study is ideal in terms of class separation and class compactness, with all categories clearly separated and the distances between classes more obvious.



a. This study uses the t-SNE plot of the model



b. 1DCNN-SVM model t-SNE plot

Fig. 7 Comparison of T-SNE for different data feature processing methods

5 Conclusion

This paper addresses the issue of bearing fault diagnosis by proposing a hybrid diagnostic model based on a two-dimensional convolutional neural network (2D-CNN) and an optimized least squares support vector machine (LSSVM). The original one-dimensional vibration signal is converted into a three-channel time-frequency image using continuous wavelet transform (CWT). The 2D-CNN is then employed to adaptively extract deep spatio-temporal features.and combines the MGO lichen growth optimization algorithm and PSO particle swarm algorithm to dynamically optimize the regularization parameter (γ) and kernel parameter (σ^2) of LSSVM, effectively addressing the bottlenecks of low efficiency and susceptibility to local optima in traditional parameter tuning. The 2D-CNN employs a two-layer convolutional structure (3×3 and 5×5 convolutional kernels) combined with LeakyReLU activation functions and batch normalization techniques,gradually compressing feature dimensions through maximum pooling layers, with the high-dimensional features output by the fully connected layer serving as input for LSSVM. The LSSVM module employs a radial basis kernel function (RBF) to map features to a high-dimensional space, and uses the MGO and PSO hybrid optimization algorithm to search for the optimal parameter combination globally, significantly improving classification performance.

The experiment was conducted using the bearing dataset from Case Western Reserve University (CWRU), covering 10 operating conditions (including normal conditions and different damage diameters of the inner ring, outer ring, and rolling elements).The results validated the model's excellent generalization ability and robustness. t-SNE visualization analysis further showed that the features extracted by the model exhibit significant inter-class separability and intra-class compactness in low-dimensional space, indicating its superior discriminative ability in high-dimensional feature mapping.Additionally, the model effectively mitigates overfitting risks in small-sample scenarios by incorporating a Dropout layer (dropout rate of 0.1) and batch normalization techniques.

6 References

[1]Duy-Tang Hoang, Hee-Jun Kang,A survey on Deep Learning based bearing fault diagnosis,Neurocomputing, Volume 335,2019,Pages 327-335,ISSN 0925-2312.

[2]Convolutional Neural Network Based Bearing Fault Diagnosis Hoang, DT and Kang, HJ 13th International Conference on Intelligent Computing (ICIC)2017INTELLIGENT COMPUTING THEORIES AND APPLICATION, ICIC 2017, PT IIarrow_drop_down 10362 , pp.105-111.

[3]H. Wang, Z. Liu, D. Peng and Y. Qin, "Understanding and Learning Discriminant Features based on Multiattention 1DCNN for Wheelset Bearing Fault Diagnosis," in IEEE Transactions on Industrial Informatics, vol. 16, no. 9, pp. 5735-5745, Sept. 2020.

[4]X. Zhang, P. Han, L. Xu, F. Zhang, Y. Wang and L. Gao, "Research on Bearing Fault Diagnosis of Wind Turbine Gearbox Based on 1DCNN-PSO-SVM," in IEEE Access, vol. 8, pp. 192248-192258, 2020.

[5]X. Peng, B. Zhang and D. Gao, "Research on fault diagnosis method of rolling bearing based on 2DCNN, " 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 2020, pp. 693-697.

[6]Yang, S.; Yang, P.; Yu, H.; Bai, J.; Feng, W.; Su, Y.; Si, Y. A 2DCNN-RF Model for Offshore Wind Turbine High-Speed Bearing-Fault Diagnosis under Noisy Environment. Energies 2022, 15, 3340.

[7]X. Zhao, L. Wang, M. Yang, Y. Chen and J. Xiang, "A Novel Small-Sample Fault Diagnosis Method for Rolling Bearings via Continuous Wavelet Transform and Siamese Neural Network," in IEEE Sensors Journal, vol. 24, no. 15, pp. 24988-24996, 1 Aug.1, 2024.

[8]Chen, H.-Y.; Lee, C.-H. Deep Learning Approach for Vibration Signals Applications. Sensors 2021 , 21, 3929.

[9]Zhang, L.; Liu, Y.; Zhou, J.; Luo, M.; Pu, S.; Yang, X. An Imbalanced Fault Diagnosis Method Based on TFFO and CNN for Rotating Machinery. Sensors 2022, 22, 8749.

[10]H. Zhao et al.,“Intelligent Diagnosis Using Continuous Wavelet Transform and Gauss Convolutional Deep Belief Network,” in IEEE Transactions on Reliability, vol. 72, no. 2, pp. 692-702,2023,6.

[11]Wang Chenglin , Liu Suchun , Wang Yawei , Xiong Juntao , Zhang Zhaoguo , Zhao Bo , Luo Lufeng , Lin Guichao , He Peng, Application of Convolutional Neural Network-Based Detection Methods in Fresh Fruit Production: A Comprehensive Review, Frontiers in Plant Science, Volume 13 – 2022,ISSN1664-462X.

[12]Boli Zheng, Yi Chen, Chaofan Wang, Ali Asghar Heidari, Lei Liu, Huiling Chen, The moss growth optimization (MGO): concepts and performance, Journal of Computational Design and Engineering, Volume 11, Issue 5, October 2024, Pages 184–221.

[13]Nor Azlina Ab. Aziz, Zuwairie Ibrahim, Marizan Mubin, Sophan Wahyudi Nawawi, Mohd Saberi Mohamad, Improving particle swarm optimization via adaptive switching asynchronous – synchronous update, Applied Soft Computing, Volume 72,2018,Pages 298-311,ISSN 1568-4946.

[14]Chuanfa Chen, Changqing Yan, Yanyan Li,A robust weighted least squares support vector regression based on least trimmed squares, Neurocomputing, Volume 168, 2015,Pages 941-946,ISSN 0925-2312.

[15]X. Zhang, B. Zhao and Y. Lin, "Machine Learning Based Bearing Fault Diagnosis Using the Case Western Reserve University Data: A Review," in IEEE Access, vol. 9, pp. 155598-155608, 2021.