Research on the Fault Diagnosis Method of Rotating Machinery Based on Improved Variational Modal Decomposition and Probabilistic Neural Network Algorithm

Zhangjie Li [1,†], Chao Zou [1,†], Zhimin Chen [2], Hong Lu [1,*], Shiwen Xie [2], Wei Zhang [1] and Jiaqi He [1]

[1] School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan 430070, China; anhmcy@whut.edu.cn (Z.L.); zouchaomynmdrq@whut.edu.cn (C.Z.); 307487@whut.edu.cn (W.Z.); hejiaqi9863@163.com (J.H.)

[2] China Ship Development and Design Center, Wuhan 430070, China; minzc@163.com (Z.C.); xie_shiwen123@163.com (S.X.)

[*] Correspondence: landzh@whut.edu.cn

[†] These authors contributed equally to this study.

Abstract: The fault diagnosis of rotating machinery is vital in industry but traditionally depends on manual expertise, requiring substantial resources. To improve diagnostic accuracy, enable effective condition monitoring, and minimize the impact of faults on operations, advanced diagnostic techniques are essential. Hence, we propose an advanced fault diagnosis framework that leverages improved particle swarm optimization (IPSO), variational mode decomposition (VMD), and probabilistic neural networks (PNN) to accurately diagnose faults in rotating machinery using gear and rolling bearing vibration signals. Initially, the vibration signals are decomposed into intrinsic mode functions via VMD, enabling the capture of subtle but critical fault features. To address parameter selection challenges in VMD, we employed IPSO to optimize the VMD parameters, ensuring the optimal decomposition effect. Further, we refined the feature set by applying Laplace fraction optimization and feature dimensionality reduction, isolating sensitive features that serve as input to a PNN-based fault classification model. Experimental results demonstrated that this IPSO-VMD-PNN framework achieves high diagnostic accuracy for various fault types, establishing it as an effective tool for fault identification in rotating machinery.

Keywords: fault diagnosis; rotating machinery; variational mode decomposition (VMD); swarm optimization algorithm (IPSO); probabilistic neural network (PNN)

1. Introduction

New methods and techniques in technical diagnostics enable improved planning, reduced downtime, and a significant increase in the time between failures. Additionally, understanding the status of rotating machinery helps to boost profits [1]. Rotating machinery, as an important component of mechanical systems, plays a crucial role in power transmission and motion transformation. Due to their harsh working environments and the influence of factors such as vibration, noise, temperature, and humidity, rotating machinery frequently experiences failures during operation. These failures exhibit characteristics such as low magnitude, nonlinearity, complexity, and non-stationarity. Therefore, condition monitoring and fault diagnosis are essential for the normal operation of rotating machinery [2].

Early fault diagnosis techniques relied primarily on manual experience, making it difficult to identify accurately the causes of complex equipment and systems, and requiring significant human and material resources. Later, rule-based expert systems emerged, but they were limited by the formulation of rules and knowledge acquisition. In recent years, data mining-based machine learning has emerged, which can learn complex nonlinear

relationships from extracted features and establish diagnostic models for fault diagnosis, demonstrating good performance. However, the application of deep learning, which has gained popularity recently, is still subject to scrutiny in the field of fault diagnosis due to the greediness of data and the lack of model interpretability.

Scholars have developed various methods for analyzing and processing rotating machinery vibration signals, such as empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), and local mean decomposition (LMD) [3–7]. The EMD method can adaptively decompose signals without the need for any predefined basis functions. However, it lacks a rigorous mathematical foundation, leading to potential instability and inconsistency in the decomposition results. The EEMD method improves the stability of the decomposition results by averaging multiple decompositions. By adding white noise and performing multiple decompositions, it effectively reduces mode mixing. Nevertheless, this approach requires numerous decompositions, resulting in high computational costs and potentially introducing additional noise components that may affect the results. The LMD method yields clearer spectra for each component, offering excellent frequency resolution. However, it faces issues with endpoint effects, the decomposition process is relatively complex, and the computational load is significant.

VMD proposed in 2014 [8], effectively avoids mode mixing and is supported by a rigorous mathematical framework. The decomposition process is scientifically grounded and applicable to linear, nonlinear, and non-stationary signals, addressing the aforementioned limitations. Consequently, scholars have extensively employed the VMD method for fault diagnosis in rotating machinery [9,10]. Lin [11] optimized the decomposition layers and penalty parameters of VMD using improved envelope entropy, selecting components with high correlation to the original signal. By integrating sample entropy, Lin achieved fault diagnosis of bearings. Liu [12] determined the number of VMD decomposition layers by observing the center frequency of VMD components and utilizing fuzzy c-means clustering, achieving effective fault diagnosis. Li [13] applied the VMD algorithm for adaptive decomposition of vibration signals and used an improved extreme learning machine to diagnose faults in rolling bearings.

However, the two key parameters in VMD—the number of modes K and the penalty factor ω—directly affect the VMD decomposition effect and need to be optimized [14]. Wang et al. proposed the center frequency ratio method, which uses the center frequency ratio of adjacent modal components to select K, but ω cannot be determined [15]. Xiao [16] proposed using the average instantaneous frequency as the basis for selecting the number of decomposition modes K in the VMD algorithm. He utilized the unsupervised learning algorithm, self-organizing map (SOM), to classify gear faults, achieving effective gear fault diagnosis. Zhang [17] optimized the decomposition layers K of VMD by combining the artificial fish swarm algorithm with envelope entropy. Zhu [18] employed a swarm algorithm using kurtosis as a criterion for parameter optimization. Zhang [19] used the particle swarm optimization algorithm to optimize the decomposition layers of VMD. These studies indicate that most scholars focus on the impact of the number of VMD decomposition layers on the decomposition performance, while often neglecting the coupled effect of decomposition layers K and penalty parameters ω [20,21].

To address this issue, this paper proposes an improved VMD method based on PSO, which comprehensively analyzes the coupled effects of the decomposition layers and penalty factors on the decomposition performance. PSO is an intelligent search method that simulates the collaborative behavior of a group of organisms. It iteratively adjusts computing parameters by searching for local and global optimal solutions to approach the global optimal solution. However, the selection of various hyperparameters in particle swarm will greatly affect the optimization rate of the algorithm, and inappropriate hyperparameters will also cause the algorithm to fall into the local optimal solution prematurely. By incorporating the adaptive inertia weight strategy to accelerate the convergence speed and using the compression factor method to solve the problem of getting stuck in local optima, the optimized particle swarm algorithm is employed to search adaptively for the

optimal *K* and penalty parameters *α*. The optimized results are combined with feature extraction on the original signal, and finally, a probabilistic neural network is used for fault mode recognition, achieving high-precision fault diagnosis.

The remainder of this paper is structured as follows. To fully leverage the performance advantages of VMD and its suitability for local damage fault feature extraction, as well as to address the difficulty in selecting key parameters in VMD, Section 2 combines VMD with an improved particle swarm optimization algorithm to achieve the optimal combination of VMD key parameters. With this optimal parameter combination, VMD is utilized to decompose vibration signals. In Section 3, effective modal components are extracted based on the correlation coefficient threshold. To mitigate the loss of signal features due to the VMD strong penalty on signal discontinuities, multi-domain feature parameters are selected using the Laplacian score method, collectively constructing a fault-sensitive feature vector. Based on this, a probabilistic neural network model is built to achieve fault diagnosis of rotating machinery. In Section 4, the signal processing and fault diagnosis algorithm proposed in the previous sections are validated using public data of gears and rolling bearings, demonstrating that IPSO-VMD-PNN possesses strong feature extraction capabilities and high fault diagnosis accuracy.

In summary, the main contributions of this paper are as follows:

(1) A vibration analysis model for key components of rotating machinery is established by analyzing failure mechanisms and signal characteristics. VMD is proposed for fault feature extraction and its effectiveness is verified through simulated fault signals.

(2) The PSO is improved using adaptive inertia weights and compression factor methods. Optimal parameters *K* and penalty parameters *α* for VMD are determined. VMD decomposes the vibration signal, and effective components are identified using correlation coefficient thresholding. Multi-domain features are selected using the Laplace score method, constructing a fault-sensitive feature vector.

(3) The PNN is used for fault classification based on the extracted feature vector. The model's accuracy is verified using public data.

2. Vibration Signal Processing Method Based on VMD

2.1. IPSO-VMD Algorithm Design

VMD algorithm analyzes the original signal according to the number of modal decompositions, obtains *K* modal functions $u_k(t)$ with center frequency $\omega(t)$, and performs Hilbert transform on $u_k(t)$ to obtain the bandwidth and spectrum information of each IMF component. By adding an exponential term to $u_k(t)$, and determining the bandwidth of the IMF component according to the L^2 norm, the solution formula is as follows [22,23]:

$$
\begin{cases}
\displaystyle\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\
s.t. \displaystyle\sum_{k=1}^{K} u_k = f(t), u_k = u_1, u_2, \cdots, u_K
\end{cases}
\tag{1}
$$

where, u_k is the IMF component obtained after VMD decomposition, $\{\omega_k\} = \{\omega_1, \omega_2, \cdots, \omega_K\}$ is the center frequency of each IMF component, and $\delta(t)$ is the pulse function.

PSO [24] is widely used in parameter selection as a global search algorithm. Suppose that the domain of definition to be searched is a D dimensional space, where there is a population of m particles, and each particle itself contains information such as position and velocity. The position of the *i*-th particle in this space is represented by vector $X_i = (x_{i1}, x_{i2}, \cdots, x_{iD})^T$, and its velocity is represented by vector $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})^T$. The best position of particle *i* is $P_{best} = (p_{i1}, p_{i2}, \cdots, p_{iD})^T$, and the best position of all particles in the whole population can be expressed as $Pgbest = (p_{g1}, p_{g2}, \cdots, p_{gD})^T$. In the PSO algorithm, each particle can update its speed and position information iteratively

according to the information such as the best position of the individual and the best position of the population:

$$\begin{cases} v_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 \left(p_{id} - x_{id}^k \right) + c_2 r_2 \left(p_{gd} - x_{id}^k \right) \\ x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} t = x_{id}^k + v_{id}^{k+1} \end{cases} \tag{2}$$

where, $i = 1, 2, \cdots, m; d = 1, 2, \cdots, D; k$ is the current number of evolutions; c_1 and c_2 are learning factors; ω is the inertia weight; r_1 and r_2 are random numbers in $[0, 1]$, t is the time of each step of movement, and $t = 1$ is taken here.

In the particle swarm optimization algorithm, the parameters to be designed mainly include the following: population size m, learning factors c_1 and c_2, maximum speed V_{max}, inertia weight ω [25–28], among which the learning factor and inertia weight are the key objects of optimization.

1. Adaptive inertia weight strategy

First, the most important inertia weight ω in the PSO algorithm is optimized by using the adaptive inertia weight strategy, which can realize the automatic optimization of the inertia weight, and then make the algorithm converge to the optimal solution quickly. In the adaptive inertia weight strategy, if the minimum value of the objective function is solved, the more is the fitness and the closer is the distance to the optimal solution. At this time, a smaller weight is needed to facilitate local search. On the contrary, a larger weight should be adopted to improve the global search ability of particles [29,30]. The adjustment method is as follows:

$$\omega_i^k = \begin{cases} \omega_{\min} + (\omega_{\max} - \omega_{\min}) \frac{f\left(x_i^k\right) - f_{\min}^k}{f_{\text{average}}^k - f_{\min}^k} & , f\left(x_i^k\right) \le f_{\text{average}}^k \\ \omega_{\max} & , f\left(x_i^k\right) > f_{\text{average}}^k \end{cases} \tag{3}$$

where,

- $\omega_{\min}$ and $\omega_{\max}$ are the preset minimum and maximum values of inertial weight ω, taking $\omega_{\min} = 0.4$, $\omega_{\max} = 0.9$;
- $f_{\text{average}}^k = \sum\limits_{i=1}^{n} f\left(x_i^k\right) / n$ is the average value calculated by the fitness of the whole population at k-th iteration;
- $f_{\min}^k = \min\left\{f\left(x_1^k\right), f\left(x_2^k\right), \cdots, f\left(x_n^k\right)\right\}$ is the minimum fitness of the whole population at iteration.

2. Compressibility factor method

Different from inertia weight ω, the individual learning factor c_1 and social learning factor c_2 represent the influence of the historical information of particles themselves and other particles on the later behavior of particles, while the learning factor describes the information interaction between populations. Therefore, individual learning factor c_1 and social learning factor c_2 have great influence on the PSO method. In order to balance the global search ability and local search ability of the algorithm, the PSO method is optimized with compression factor, which can ensure that the PSO algorithm has strong convergence in solving, and can reduce its speed limit [31].

If $c_1 = c_2 = 2.05$, then $C = c_1 + c_2 = 4.1$, compression factor $\Psi = 2 / \left| \left(2 - C - \sqrt{C^2 - 4C} \right) \right|$. For the speed in Equation (3), it can be updated as follows:

$$v_{id}^{k+1} = \Psi \left[\omega v_{id}^k + c_1 r_1 \left(p_{id} - x_{id}^k \right) + c_2 r_2 \left(p_{gd} - x_{id}^k \right) \right] \tag{4}$$

3. Elite Learning Strategies

Elite learnt strategy is an improved algorithm that enhances search efficiency and accuracy by introducing an elite guidance mechanism. In each iteration, the best-performing

particles or a small number of good particles are selected as 'elites', and the other particles in the population are allowed to take into account the position information of these elites in addition to the p_{Best} and g_{Best} when updating themselves. Non-elite particles can update their velocity and position by introducing the influence of the positions of the elite particles:

$$v_i(t+1) = \omega \cdot v_i(t) + c_1 \cdot r_1 \cdot \left(p_{Best_i} - x_i(t)\right) + c_2 \cdot r_2 \cdot \left(g_{Best} - x_i(t)\right) + c_3 \cdot r_3 \cdot \left(Elite_i - x_i(t)\right) \tag{5}$$

where $Elite_i$ is the position of the *i*-th elite particle, c_3 is the elite learning factor, and r_3 is a random number between [0,1].

4. Fitness Distance Ratio Optimization Strategy

The Fitness Distance Ratio (FDR) is a measure of the relative strengths and weaknesses of particles in relation to other particles. It combines the fitness value of a particle with information about the distance between particles and is used to guide the search behaviors of the particles. The fitness distance ratio is defined as follows:

$$FDR_i = \frac{f(x_i)}{\left\| x_i - P_{Best_i} \right\|} \tag{6}$$

where $f\left(P_{Best_i}\right)$ is the fitness value of particle *i* and $\left\| x_i - P_{Best_i} \right\|$ is the Euclidean distance between the current position of the particle and its individual optimal position. FDR_i is a factor that adjusts the speed of the particle as it moves towards b to enhance the algorithm to pay more attention to particles with high adaptation whose distance is far away, and the updated formula for the speed is as follows:

$$v_i(t+1) = \omega \cdot v_i(t) + c_1 \cdot r_1 \cdot \left(p_{Best_i} - x_i(t)\right) + c_2 \cdot r_2 \cdot \left(g_{Best} - x_i(t)\right) \cdot FDR_i \tag{7}$$

This approach improves the ability of the particle swarm to concentrate towards more optimal solution regions, as the elite particles represent the currently searched optimal solution regions, and guiding the other particles towards these regions helps to avoid premature convergence and accelerates the finding of the global optimal solution. With this elite guidance mechanism, the ELPSO algorithm is able to balance the need for global and local searches more efficiently, in line with the need for efficient and reliable optimization methods in complex optimization problems.

For the selection of fitness function, a new fitness function *fitness* can be constructed based on envelope entropy Ep [32] and envelope spectrum sparsity S [33]:

$$fitness = Ep/S \tag{8}$$

It can be observed that the sparser the signal, the smaller the Ep, and the larger the S, the smaller is the value calculated by the fitness function *fitness*. To determine parameter combination $[K, \alpha]$ in VMD, after VMD decomposition of different parameter combinations for the original signal, the *fitness* value of each component under different parameter combinations is calculated, respectively, and the global minimum *fitness* value is the optimization object. The minimum *fitness* value corresponds to the optimal parameter combination $[K, \alpha]$, which is expressed as follows:

$$\left[\hat{K}, \hat{\alpha}\right] = \underset{(K,\alpha)}{\arg\min}\{fitness(i)\} \tag{9}$$

Based on the adaptive inertia weight and compression factor method, the original particle swarm optimization algorithm is improved in the early stage, and IPSO is obtained, while $fitness = Ep/S$ is determined as the fitness function.

Comparison experiments are designed to verify the improvement strategies for the several particle swarm algorithms above, and the best improvement strategy is selected as the improvement strategy for the particle swarm algorithm in the final diagnostic architecture, and the final iterative graph results are as follows:

Where IPSO stands for PSO based on adaptive inertia weight and compression factor method, PSO stands for unimproved PSO, EL-PSO stands for improved PSO based on elite learning strategy, and FDR-PSO stands for PSO based on fitness distance ratio optimization. According to Figure 1 it can be observed that the IPSO gives the best results and has the fastest convergence rate.
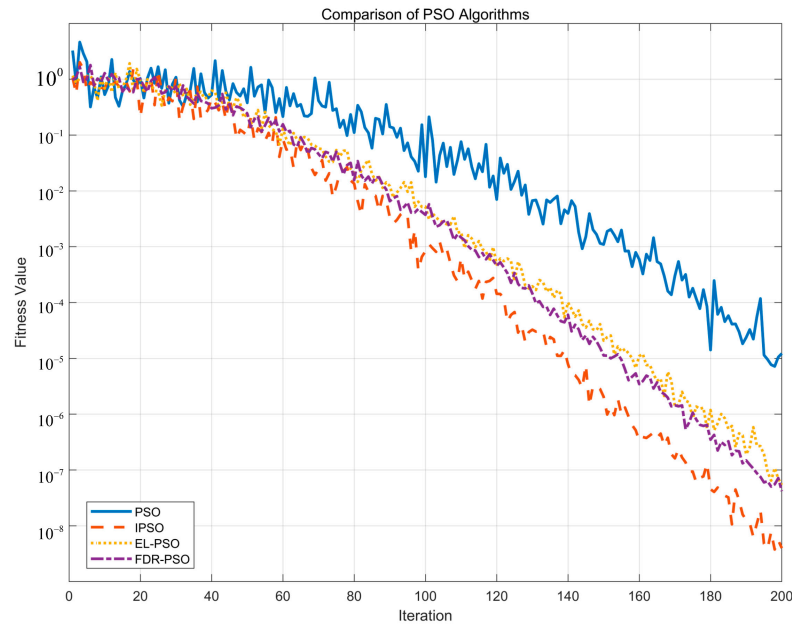

Figure 1. Comparison of PSO algorithms.

When VMD decomposition is carried out on the original vibration signal, the IPSO algorithm is used to carry out adaptive optimization on the two parameters K and α in VMD to obtain the best combination of parameters, while the corresponding VMD decomposition is carried out based on the best combination of parameters. The design idea of the IPSO-VMD algorithm shown in Figure 2 can be constructed.
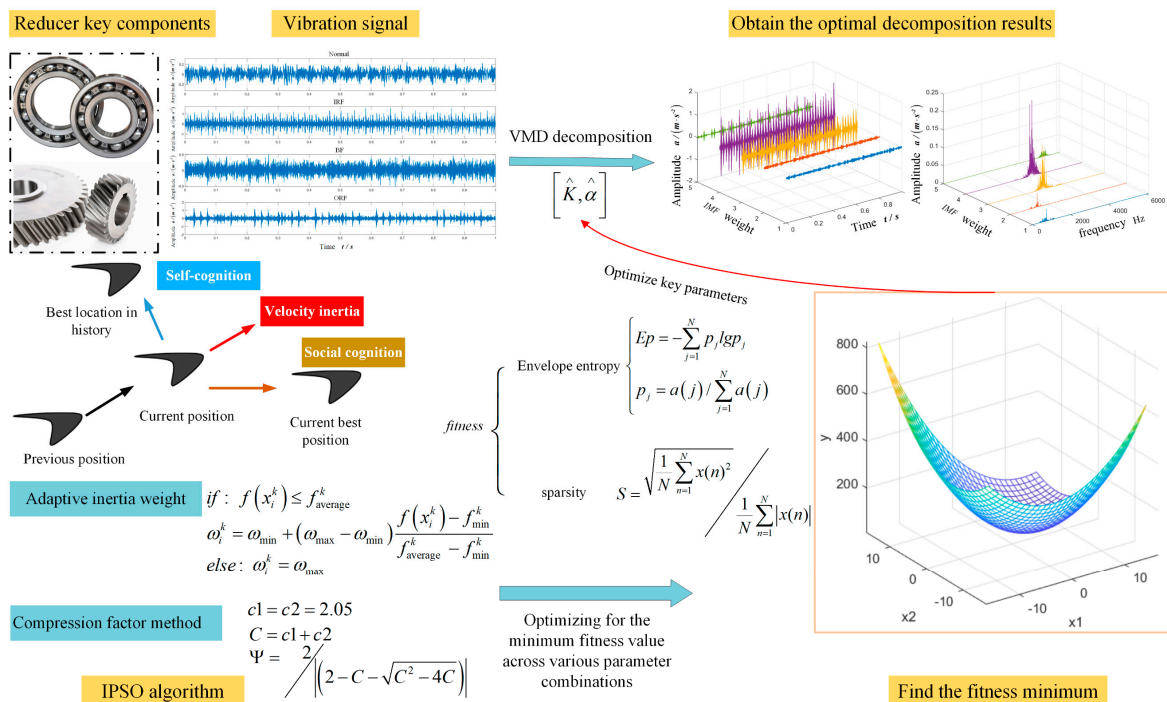

Figure 2. IPSO-VMD algorithm design diagram.

2.2. Fault Simulation Experiment

The real vibration signal of gear bearing usually contains three signal components—impact signal, noise signal, and harmonic interference signal—in order to verify the effectiveness of VMD decomposition for fault signals. The simulation signal $f(t)$ of rolling bearing outer ring local damage fault is constructed as follows:

$$\begin{cases} f(t) = x_1(t) + x_2(t) + x_3(t) \\ x_1(t) = 2exp(-at_0)cos(2\pi f_1 t) \\ x_2(t) = sin(2\pi f_2 t) + sin(2\pi f_3 t) \end{cases} \tag{10}$$

It can be observed from the formula that the simulation signal $f(t)$ is composed of $x_1(t)$, $x_2(t)$, and $x_3(t)$, where $x_1(t)$ is the simulation signal of impact attenuation caused by fault, which is represented by a pulse sequence with periodicity that obeys the exponential attenuation law. The amplitude of $x_1(t)$ is 2, $t_0 = mod(k/f_s, 1/f_m), k = 0, 1, \cdots, 2047$, the sampling frequency $f_s = 2048$ Hz, the number of sampling points $N = 2048$, the attenuation coefficient $a = 100$, the carrier frequency $f_1 = 200$ Hz, and the fault feature frequency $f_m = 16$ Hz. $x_2(t)$ is the sine superposition signal of frequency $f_2 = 20$ Hz and $f_3 = 30$ Hz, which is used to simulate harmonic interference. $x_3(t)$ is the Gaussian white noise of mean value $\overline{x} = 0$ and standard deviation $\sigma = 0.1$, which is used to simulate the background noise. As shown in Figure 3, the time-domain waveforms of $f(t)$, $x_1(t)$, $x_2(t)$, and $x_3(t)$ are given, respectively.
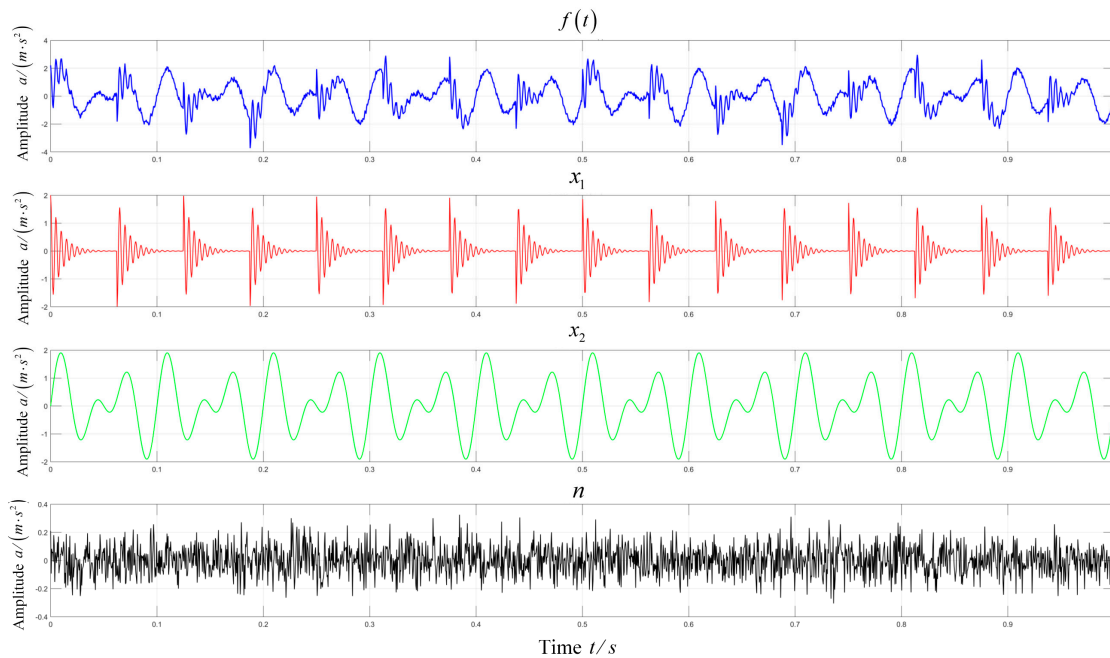

Figure 3. Time-domain waveform of rolling bearing fault simulation signal $f(t)$ and its components.

Figure 4 shows the time-domain waveform and spectrum of simulation signal $f(t)$, and the obvious fault feature frequencies cannot be obtained from both the time-domain and frequency-domain perspectives. In the spectrum diagram, obvious harmonic interference components of 20 Hz and 30 Hz can be observed while the fault feature frequency $f_m = 16$ Hz is submerged in the interference signal and noise. The envelope waveform and envelope spectrum of simulation signal $f(t)$ obtained based on Hilbert transform are shown in Figure 5. It is impossible to identify accurately whether there is a fault from the envelope waveform. In the envelope spectrum, there are both 16 Hz feature frequency with small amplitude and 10 Hz interference component with large amplitude in the envelope spectrum. The low amplitude spike around 200 Hz in the figure is a resonance frequency band.
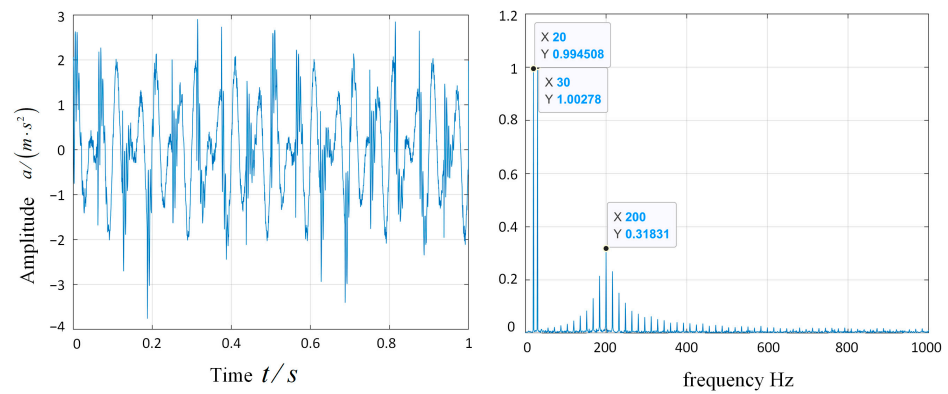
Figure 4. Simulation signal $f(t)$ time-domain waveform and its spectrum.
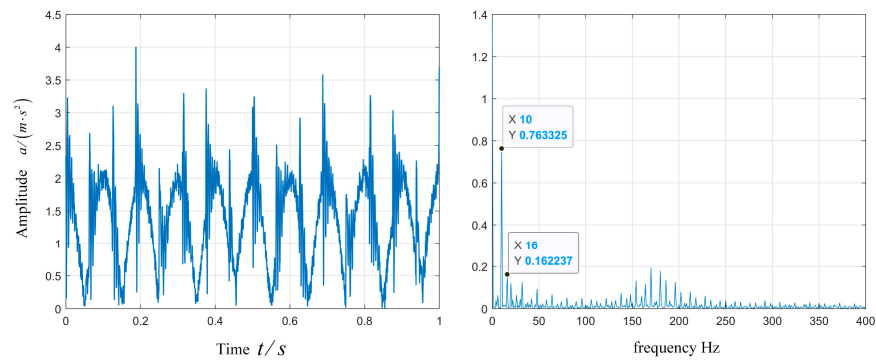

Figure 5. Envelope waveform and envelope spectrum of simulation signal $f(t)$.

VMD decomposition result of simulation signal $f(t)$ (the best parameter combination obtained after optimization by IPSO-VMD is $[\hat{K}, \hat{\alpha}] = [3, 5500]$), as shown in Figure 6, IMF1 perfectly corresponds to the harmonic interference $x_2(t)$ in the original signal, and the periodic shock attenuation can be observed in IMF2. The corresponding fault characteristic frequency 16 Hz and its frequency multiplication components (32 Hz, 48 Hz, etc.) can be clearly identified from the IMF2 and IMF3 components. Therefore, VMD can effectively decompose the $x_1(t)$ and $x_2(t)$ components of the rolling bearing fault simulation signal.
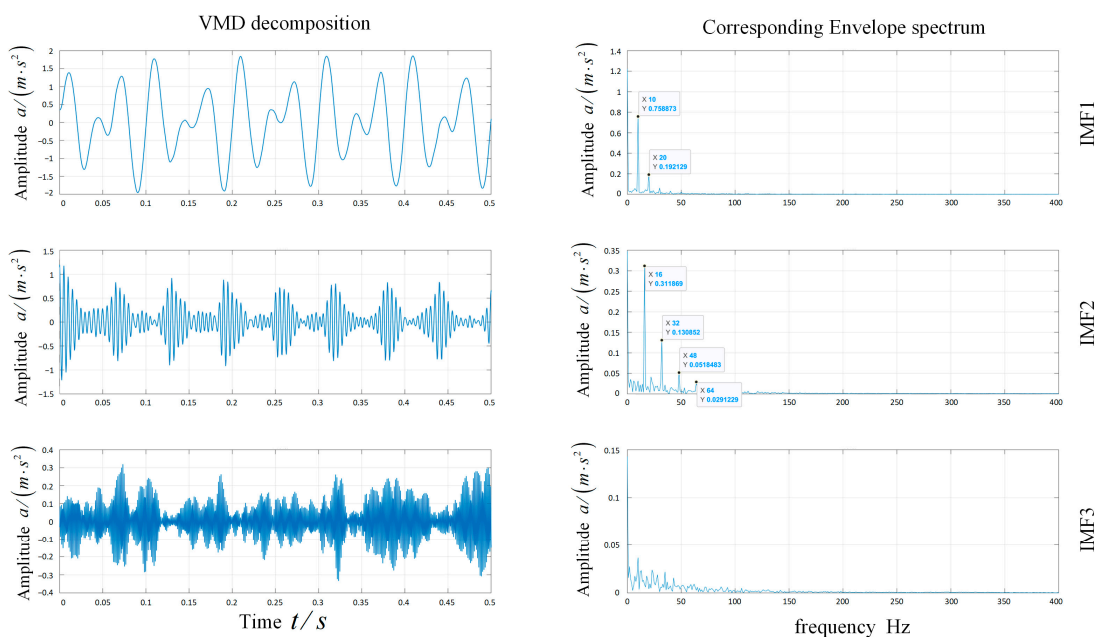

Figure 6. VMD decomposition results of rolling bearing fault simulation signal $f(t)$.

3. Feature Extraction and Recognition Based on IPSO-VMD

3.1. Feature Extraction Based on VMD

The failure of different components in the system will change the frequency distribution of the signal, further leading to the adjustment of the energy distribution of the signal. In order to avoid the dependence of traditional fault diagnosis methods on abnormal samples, VMD decomposition results under the optimal parameter combination are used to determine the effective components in combination with the correlation coefficient method. Finally, the energy ratio of the first five effective components is used as the feature parameter.

1. Correlation Coefficient

Correlation coefficient is a statistical indicator to measure the closeness of correlation between different variables. The correlation coefficient of sample X and sample Y is defined as follows:

$$\rho_{(X,Y)} = \frac{\text{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \tag{11}$$

where, $\text{Cov}(X,Y)$ is the covariance of sample X and sample Y, $D(X)$ and $D(Y)$ are the variances of sample X and Y, respectively. The correlation coefficient $\rho \in [-1,1]$; the larger the absolute value of the correlation coefficient, the closer is the correlation between the two samples.

For discrete digital signals, the correlation coefficients of signals $x(i)$ and $y(i)$ can be described as follows:

$$\rho_{(x(i),y(i))} = \sum_{i=0}^{\infty} x(i)y(i) / \left[\sum_{i=0}^{\infty} x^2(i) \sum_{i=0}^{\infty} y^2(i) \right]^{1/2} \tag{12}$$

For signals, the correlation coefficient ρ describes the similarity of two signals if $\rho = 1$ represents that the two signals are completely correlated. On setting the threshold value $\sigma = 0.1$, if the correlation coefficient between the i-th IMF component and the original signal is $\rho > \sigma$, the IMF component is considered as an effective component.

2. IMF Energy Ratio

First, the original signal is processed by IPSO-VMD to obtain K IMF components. The energy $imf\ E_i$ of each IMF component is calculated according to Equation (10):

$$imf\ E_i = \int_0^T imf_i^2(t)dt,\ i = 1, 2, \cdots, K \tag{13}$$

The obtained $imf\ E_i$ is normalized, and the normalized $imf\ E_i$ of the effective components is constructed as the eigenvector T:

$$T = \left[\frac{imf\ E_1}{E}, \frac{imf\ E_2}{E}, \cdots, \frac{imf\ E_n}{E} \right] \tag{14}$$

where, E is the energy of the original signal, and n is the number of effective components.

3.2. Multi-Dimension Sensitive Feature Optimization

For the time-domain parameters, they are closely related to the mechanical system parameters, operating conditions, and signal distribution. For the frequency-domain parameters, they can well reflect the spectral changes of the vibration signals of the mechanical components and determine the operating status of the mechanical components. Because the VMD algorithm will have a greater degree of punishment for abrupt signals, some abrupt information in the original signal will be lost in the VMD decomposition process. Considering the time domain or the frequency domain alone has its own limitations, be-

cause they can capture different information and characteristics of the signals, and only by considering them together can the sensitive feature vectors be better constructed.

1. Time-Domain Parameters

Time-domain parameters are closely related to mechanical system parameters, working conditions, and signal distribution. For time series signal x_i, the time-domain parameters used in this paper include 12 indicators: maximum $x_{\max}$, absolute mean $\overline{x}$, effective value x_{rms}, peak to peak value x_{p-p}, average amplitude MA, variance σ_x^2, deviation S_c, waveform index S_f, peak index C_f, pulse index I_f, margin index CL_f, kurtosis β.

2. Frequency-Domain Parameters

Based on the frequency-domain parameters, the frequency spectrum changes of the vibration signals of mechanical components can be well reflected, and the running state of the mechanical components can be judged. The frequency-domain parameter indexes adopted in this paper include: mean frequency Mf, center frequency Fc, root mean square frequency $Rmsf$, standard deviation frequency Sdf, frequency-domain amplitude skewness index Scf, and frequency kurtosis βf.

3. Other Parameters

Energy operator [34] is as follows:

$$Teo = \frac{N \sum\limits_{n=1}^{N} (r(n) - \overline{r}(n))^4}{\left(\sum\limits_{n=1}^{N} (r(n) - \overline{r}(n))^2 \right)^2} \tag{15}$$

where, $\overline{r}(n)$ represents the mean value of $r(n)$, and $r(n)$ represents the Teager energy differential signal, $r(n) = x^2(n) - x(n-1)x(n+1)$.

A total of 19 fault feature parameters are extracted based on time domain and frequency domain. Although they can reflect different information of vibration signals and system faults, the sensitivity of different fault feature parameters is different, and there is partial irrelevance or redundancy. The Laplacian score (LS) method is used to extract the inherent information architecture of the feature set, map the complex and high-dimensional feature space to the low dimensional space, and achieve the optimization of multi-dimensional feature [35]. The specific steps of the LS method are as follows:

Step 1: Input the training sample feature matrix $F \in \Re^{m \times n}$, where m is the number of samples, n is the feature dimension, and f_{ri} is the r-th dimension feature of the i-th sample.

Step 2: Build a neighbor graph G, where G has m sample points; x_i corresponds to the i-th node. If x_i and x_j are "neighbors", they are connected at two points. x_i and x_j are defined as neighbors. On the contrary, the two points are not connected.

Step 3: Define the weight matrix S_{ij} as below:

$$S_{ij} = \begin{cases} e^{\frac{\|x_i - x_j\|^2}{t}} & if \ x_i \ close \ to \ x_j \\ 0 & otherwise \end{cases} \tag{16}$$

where t is a suitable constant.

Step 4: L_r is defined as the Laplace score of the r-th feature f_r. The calculation method of L_r is as follows:

$$L_r = \frac{\sum_{ij} (f_{ri} - f_{rj})^2 S_{ij}}{var(f_r)} \tag{17}$$

where, $f_r = [f_{r1}, f_{r2}, \cdots f_{rm}]^T$, $D = diag(SI)$, $I = [1, 1, \cdots 1]^T$, $L = D - S$, matrix L is the Laplace matrix of neighborhood graph G:

Where

$$
\begin{aligned}
\sum_{ij}(f_{ri} - f_{rj})^2 S_{ij} \quad &= \sum_{ij}\left(f_{ri}^2 S_{ij} - 2f_{ri}f_{rj}S_{ij} + f_{rj}^2 S_{ij}\right) = \sum_{ij}\left(2f_{ri}^2 S_{ij} - 2f_{ri}f_{rj}S_{ij}\right) \\
&= 2\sum_{ij}(f_{ri}^2 S_{ij} - f_{ri}f_{rj}S_{ij}) = 2f_r^T D f_r - 2f_r^T S f_r = 2f_r^T L f_r
\end{aligned}
\tag{18}
$$

$$
var(f_r) = \sum_i (f_{ri} - \mu_r)^2 D_{ii}
\tag{19}
$$

where, $var(f_r)$ is the variance of the r-th feature f_r, and standardization f_r can obtain the following:

$$
\widetilde{f} = f_r - \left(\frac{f_r^T D I}{I^T D I}\right) I
\tag{20}
$$

Further, we can find the Laplacian score of the r-th feature:

$$
L_r = \frac{\widetilde{f}_r^T L \widetilde{f}_r}{\widetilde{f}_r^T D \widetilde{f}_r}
\tag{21}
$$

Step 5: Sort L_r in ascending order and output L_r in turn.

According to Equation (14), the smaller the molecular $(f_{ri} - f_{rj})$, the smaller is the feature difference within the representative sample. The greater the variance of the r-th feature f_r, the greater is the feature difference between samples and the higher is the separability. Therefore, the smaller the Laplace score L_r corresponding to the feature, the more important it is. The final multi-dimensional sensitive feature vector T_{multi} can be obtained by combining the first five order IMF component ratios of feature vector T constructed earlier and the five feature parameters extracted in this section:

$$
T_{multi} = \left[\frac{imf\,E_1}{E}, \frac{imf\,E_2}{E}, \frac{imf\,E_3}{E}, \frac{imf\,E_4}{E}, \frac{imf\,E_5}{E}, S_c, Teo, Sdf, \overline{x}, Scf\right]
\tag{22}
$$

3.3. Probabilistic Neural Networks

PNN is an artificial neural network model utilizing Bayes law and Palzen window. The PNN model has the advantages of simple structure, fast calculation speed, and high classification accuracy. Furthermore, the PNN learning algorithm is based on sample distribution, it does not require extensive parameter tuning like other machine learning algorithms, and it can automatically learn and classify based on the provided data without complex parameter tuning. At the model architecture level, the PNN model consists of four parts: input, mode, sum, and output, and its network model structure is shown in Figure 7. For pattern recognition, the input layer is used to receive input samples, and the number of nodes depends on the dimension of the input feature matrix. The function of the mode layer is to calculate the matching relationship between input samples and each node of the mode layer, and the number of nodes depends on the number of training samples. The function of the summation layer is to calculate the probability belonging to the same category in the model layer, whose number of nodes equals the number of categories of input samples, while the output layer outputs the probability density estimates of different categories through a competitive method and determines that the corresponding type of the maximum value is the judgement result [36].

The spread parameter is one of the most important parameters in PNN, which controls the smoothness of the radial basis function. The smaller the spread parameter, the more complex is the decision boundary, which may lead to overfitting. The larger the spread parameter, the smoother is the model, which may lead to underfitting.
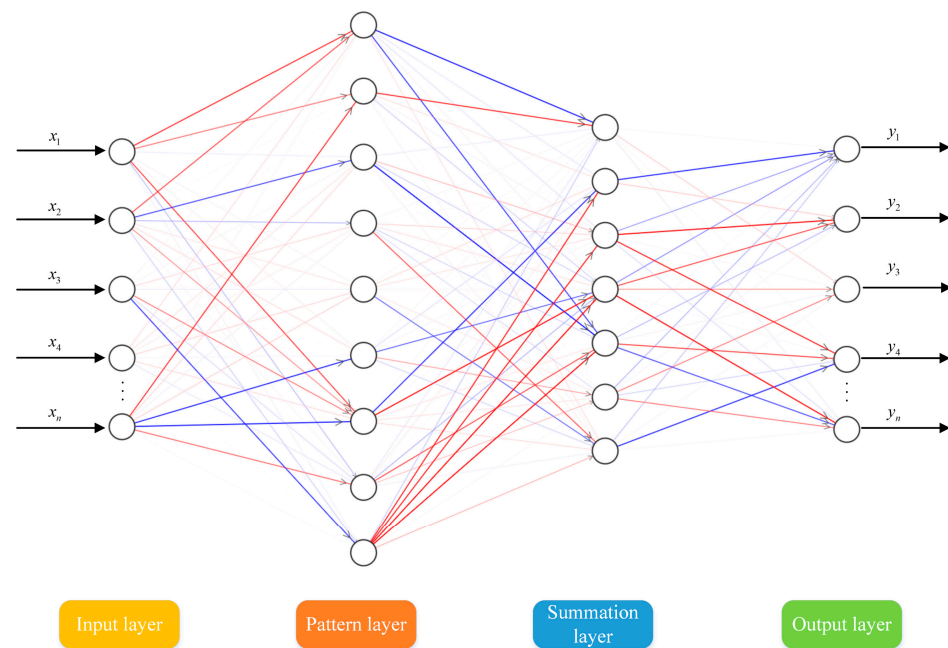
Figure 7. Basic structure of PNN.

At the same time the size of the data volume and the data dimension have a significant impact on the PNN network, which is directly related to the training effect and generalization ability of the model. Larger data volume can provide richer information to help the PNN model better learn and understand the distribution of data, thus improving the training effect and prediction accuracy of the model. A larger data volume also helps to improve the generalization ability of the model and reduce the risk of overfitting. When the data volume is small, the model may overlearn the noise in the training data, leading to performance degradation on new data. However, with too large data dimensions, it is difficult for the network to maintain sparse invariance, and it is difficult to train the PNN network under high-dimensional data, and some irrelevant or redundant features may interfere with the learning process of the PNN model, leading to performance degradation.

4. Experimental Verification

To verify the effectiveness of the fault diagnosis algorithm proposed in this paper, two different cases of open data from rotating machinery parts were used for fault diagnosis.

4.1. Case 1

Gears are the core components of rotating machinery. The quality of gears can affect the performance of rotating machinery. In this case, the gear failure is analyzed by spur gear failures from the University of Connecticut gear public data set as an example [37]. This data set contains a full set of gearbox vibration data. The experimental bench used is composed of drive motor, reducer, brake, and speed controller.

As shown in Figure 8, the transmission diagram of the gearbox in the University of Connecticut data test bench is a two-stage reduction structure. The gear box contains input shaft, intermediate shaft, and output shaft. ER-10K rolling bearings are used on both sides of each stage shaft to connect with the reducer box. The number of teeth of gear 1 is 32, gear 2 is 80, gear 3 is 48, and gear 4 is 64. Therefore, the first stage reduction ratio $i_1 = 2.5$, the second stage reduction ratio $i_2 = 4/3$, and the detailed parameters of data acquisition are shown in Table 1.
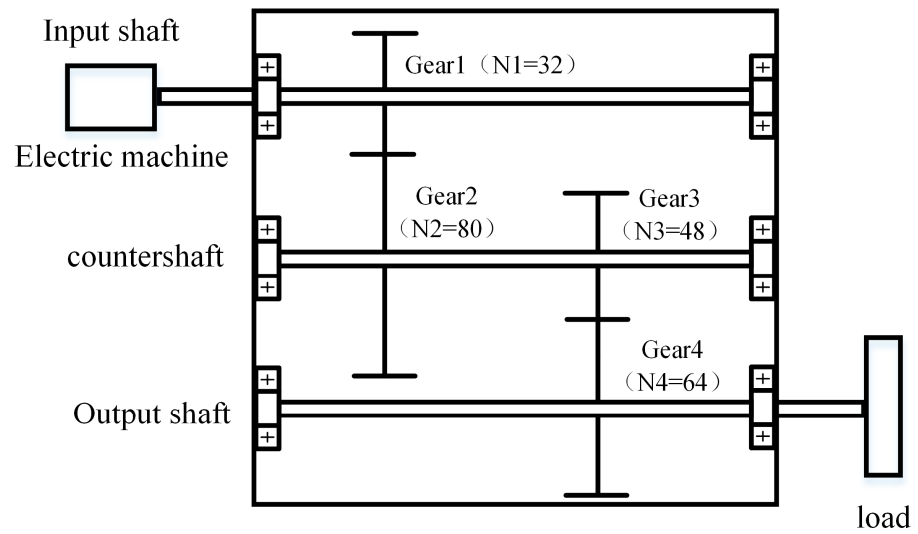
Figure 8. Gearbox transmission diagram.

Table 1. Detailed parameters of experiment.

Acquisition System	Processing Board	Sampling Frequency	Sampling Time	Speed
dSPACE	DS1006	20 kHz	10 s	2820 rpm

In this experiment, the information of the experiment is that only gear fault data collected from this data set are used, bearing #1~bearing #6 are normal, input and output shaft are normal, gear fault includes flaking and pitting, and fault occurs on gear 1 and 3, respectively. Fault label and status information are shown in Table 2.

Table 2. PHM gear fault label and status information.

Label	Gears 1	Gears 2	Gears 3	Gears 4
1	Normal	Normal	Normal	Normal
2	Flake	Normal	Normal	Normal
3	Pit	Normal	Normal	Normal

In the experiment, 100 samples are collected for each status, with each sample consisting of 3600 data points covering three periods of experimental data. As shown in Figure 9, the time-domain diagram of three different state data after down-sampling is shown.

IPSO adaptive search is performed on three types of fault data and the best parameter combination $\left[\hat{K}, \hat{\alpha}\right]$ is shown in Table 3. The best combination of parameters obtained by optimizing label 1 data is $[10, 2000]$, label 2 is $[10, 3150]$, and label 3 is $[8, 4950]$.

Table 3. Optimum parameter combination $\left[\hat{K}, \hat{\alpha}\right]$ for different fault data.

Label 1	Label 2	Label 3
$[10, 2000]$	$[10, 3150]$	$[8, 4950]$

The data of the three labels are processed by IPSO-VMD and each IMF component is calculated. The correlation coefficient between each IMF component and the original signal is shown in Figure 10. It can be observed from the diagram that the correlation coefficient between all IMF components and the original signal under the three labels is greater than the set threshold $\sigma = 0.1$. It is concluded that all IMF components from the three different fault data of the gear wheel are effective components.
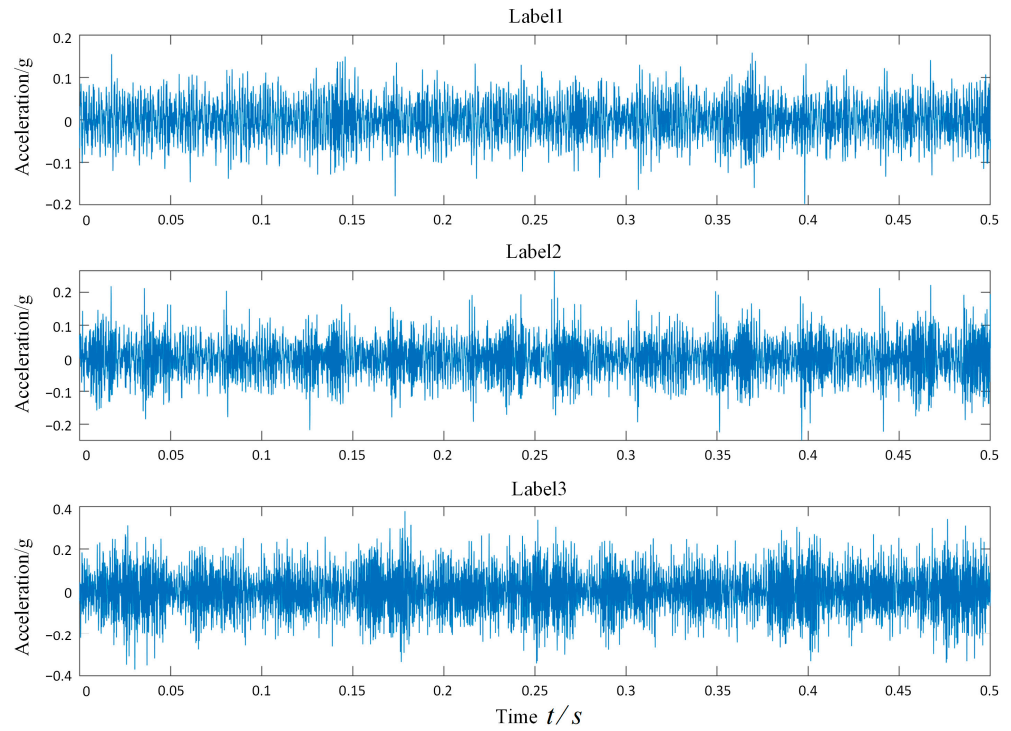
Figure 9. Time-domain diagram of gear data in three different conditions.
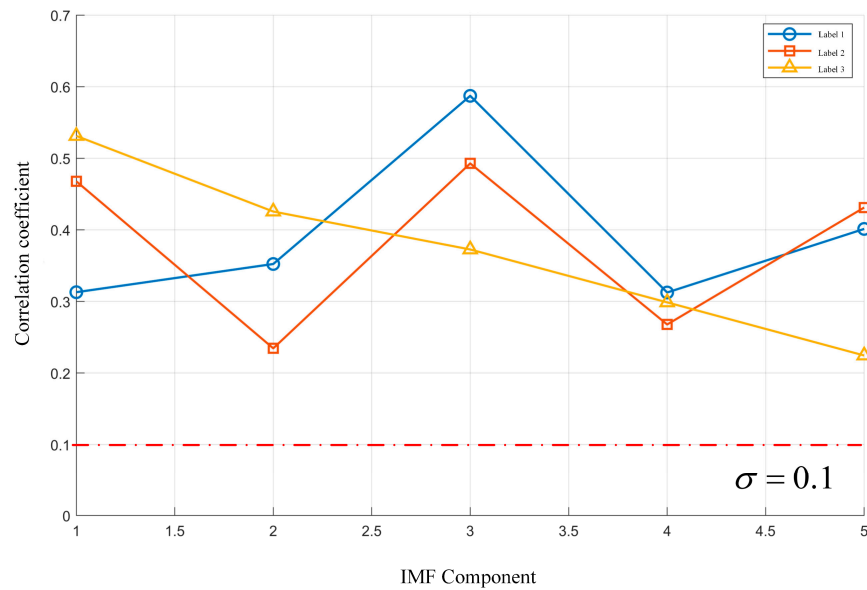

Figure 10. IMF component correlation coefficient of gear data.

Further, the energy ratio feature of the effective components is shown in Figure 11. It can be observed that the energy ratio of the IMF3 component of label 1 data is the largest, accounting for 0.325, while the IMF5 component is the smallest, accounting for 0.06. IMF1 and IMF3 are the first two components in the proportion of label 2 data, 0.376 and 0.4039, respectively, and the other three components are close to each other. The proportion of label 3 data in IMF1 and IMF2 is the first two, 0.3259 and 0.2035, respectively. The proportion of other components decreases as they move forward.
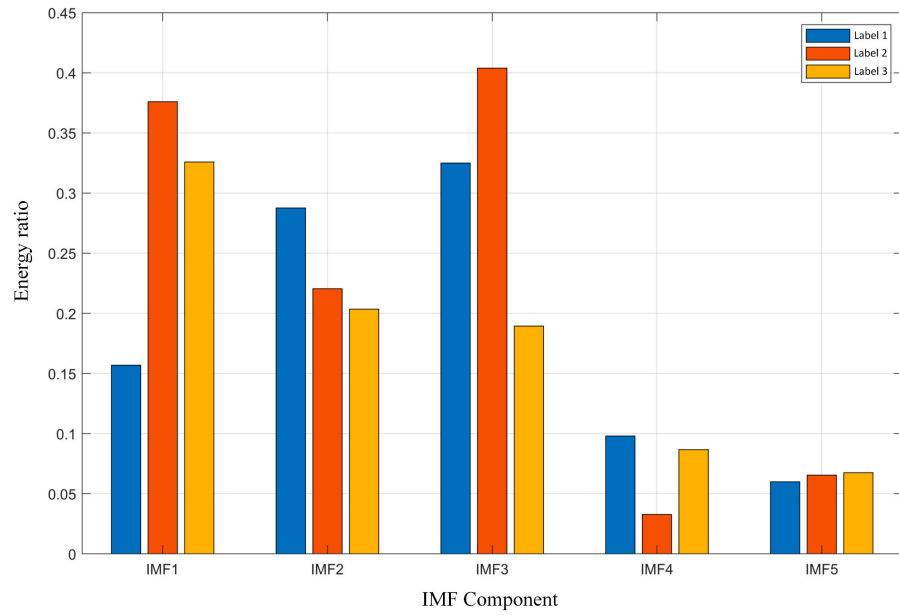
Figure 11. Energy ratio of IMF components for gear data in different conditions.

The extracted IMF component energy ratio feature and the time–frequency domain feature are collected as feature vectors and input into PNN model. Then, 70% of the 300 sets of gear fault data are randomly selected as training set data and the remaining 30% as test set data. Using the rule of thumb to determine the size of the spread parameter the following is carried out: first select the initial value of spread = 0.1, and then gradually increase the size of the parameter, and finally in spread = 1.5 when the final diagnostic effect is the best, then select spread = 1.5.

After the fault diagnosis, the confusion matrices for the training and test sets of the gear fault data are shown in Figure 12. At the model training level, label 1 in the training set has a total of 70 data and there are two misclassifications (both sample 7 and sample 16 classify label 1 as label 2) with an accuracy of 97.14%. Label 2 has a total of 70 data with one misclassification (sample 19 classifies label 2 as label 1) with an accuracy of 98.57%. Label 3 has 70 data and has an accuracy rate of 100%. Label 1 with 30 data in the test set is misclassified by one with an accuracy of 96.67%. Label 2 has 30 data, one misclassification and 96.67% accuracy. Label 3 has 30 data with 100% accuracy. The misclassification that occurred is that Sample 27 predicted label 1 as label 2 and sample 7 predicted label 2 as label 1, with some confusion between label 1 and label 2.
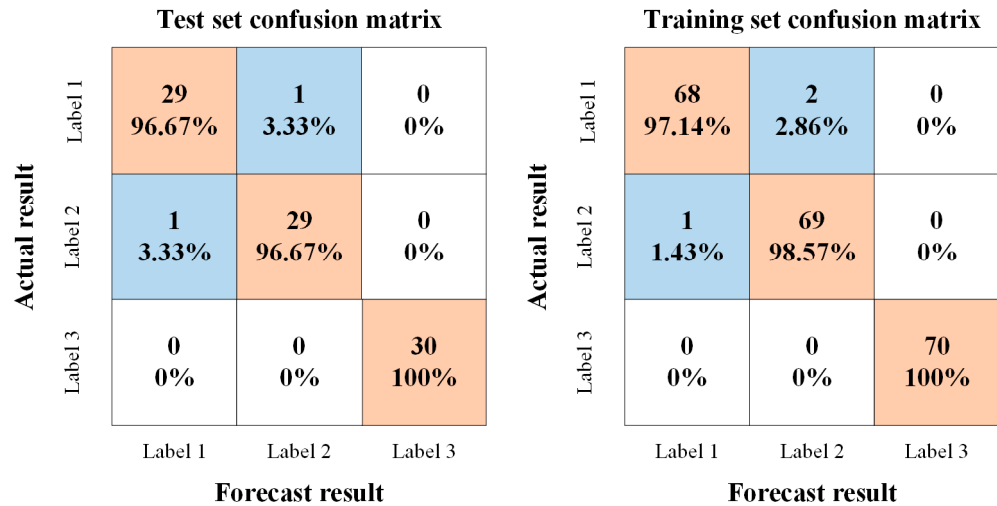

Figure 12. Gear fault diagnosis confusion matrix.

The PSO-VMD feature extraction algorithm is compared with the optimized IPSO-VMD, EMD and compared with PNN, support vector machines using extreme learning machine (ELM) [38]. The diagnostic results for the gear fault samples are shown in Table 4. It can be seen that for the gear data, the IPSO-VMD-PNN algorithm is the most effective with an overall accuracy of 97.78%.

Table 4. Diagnostic results of gear failure samples.

Gears Status	Feature Extraction Algorithm	Classification Algorithm	Number of Training Sets	Number of Test Sets	Right	Misjudgment	Accuracy Rates
Label 1	IPSO-VMD	PNN	70	30	29	1	96.67%
	PSO-VMD	PNN	70	30	28	2	93.33%
	PSO-VMD	ELM	70	30	26	4	86.67%
	PSO-VMD	SVM	70	30	27	3	90.00%
	EMD	PNN	70	30	24	6	80.00%
Label 2	IPSO-VMD	PNN	70	30	29	1	96.67%
	PSO-VMD	PNN	70	30	27	3	90.00%
	PSO-VMD	ELM	70	30	24	6	80.00%
	PSO-VMD	SVM	70	30	26	4	86.67%
	EMD	PNN	70	30	24	6	80.00%
Label 3	IPSO-VMD	PNN	70	30	30	0	100%
	PSO-VMD	PNN	70	30	28	2	93.33%
	PSO-VMD	ELM	70	30	26	4	86.67%
	PSO-VMD	SVM	70	30	28	2	93.33%
	EMD	PNN	70	30	25	5	83.33%

4.2. Case 2

As a critical component of rotating machinery, rolling bearings can significantly impact its performance. This case conducts signal processing research on the open bearing fault data set [39] of Case Western Reserve University (CWRU) bearing data center and the rolling bearing data measured in the laboratory. The test bench used by the mechanism is shown in Figure 13, including drive motor, load motor, coupling, torque sensor, and other components, providing measured fault data of different fault positions and sizes. Among them, the drive end bearing is an SKF 6205, and the fan end bearing is an SKF 6203. The acceleration sensor is used for data acquisition and installed above the bearing pedestal at the drive end and fan end. The sampling frequency of the data acquisition system is 12 kHz.
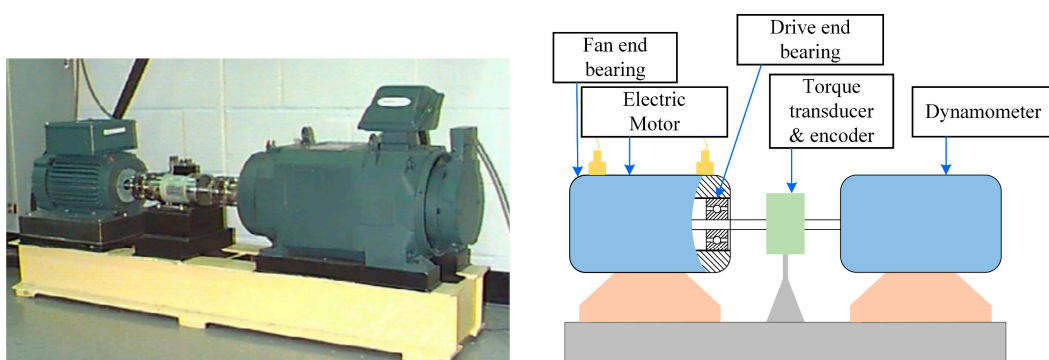

Figure 13. CWRU bearing data center test bench.

Based on the PNN model, this section carries out diagnostic experiments on four states of the rolling bearing in the CWRU open data set. The technical information is shown in Table 5.

Table 5. Technical information of failure samples of rolling bearing.

Bearing Status	Fault Diameter (mm)	Rotational Speed (rpm)	Data Labels	Number of Samples	Failure Label
Normal	0	1750	*Normal_2*	200	1
Rolling element Failure	0.3556	1750	*B*014_2	200	2
Inner ring Failure	0.3556	1750	*IR*014_2	200	3
Outer ring Failure	0.3556	1750	*OR*014@6_2	200	4

The data of four states are searched by IPSO adaptive search, and the best parameter combination $[\hat{K}, \hat{\alpha}]$ is shown in Table 6.

Table 6. Optimum parameter combination for different status data.

Normal	Ball Failure	Inner Failure	Outer Failure
[10, 2950]	[8, 3000]	[9, 1850]	[6, 3400]

Each IMF component of four status data processed by IPSO-VMD is obtained, and the correlation coefficient between each IMF component and the original signal is calculated as shown in Figure 14. The figure shows that all other IMF components are valid components.
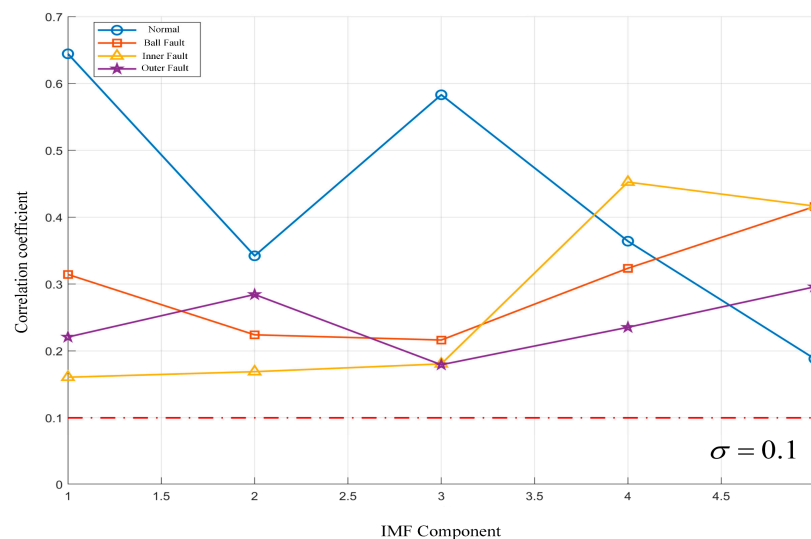

Figure 14. IMF component correlation coefficient of rolling bearing data.

Further, the energy ratio feature of effective components is shown in Figure 15. It can be observed that the distribution trend of the five IMF components of the rolling element fault and inner ring fault is similar. The proportion of IMF2 and IMF3 is small while IMF5 is the largest, but the proportional energy ratio of the inner ring fault is larger than that of the rolling element fault. For outer ring faults, IMF1 and IMF5 are 0 in the first five order components, and the ratios of IMF1 and IMF2 are close, while there is little difference in energy; IMF3 accounts for the largest proportion, about 4.5 times of IMF3. For normal conditions, IMF1 accounts for the largest proportion, exceeding 0.5, while IMF5 accounts for the smallest. In summary, the energy ratios of IMF components of each order for different faults can effectively distinguish different types of faults.
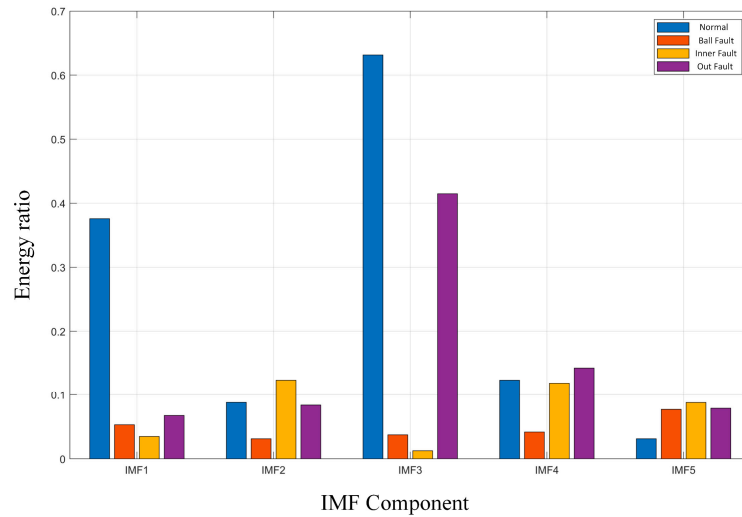
Figure 15. Energy ratio of IMF components for rolling bearing data of different states.

Consistent with the previous method of constructing the training set and test set, the confusion matrix of the rolling bearing fault data training set and test set is obtained as shown in Figure 16. Using the rule of thumb to determine the size of the spread parameter, the following is carried out: first select the initial value of spread = 0.1, and then gradually increase the size of the parameter, and finally in spread = 2.3 when the final diagnostic effect is the best, then select spread = 2.3.
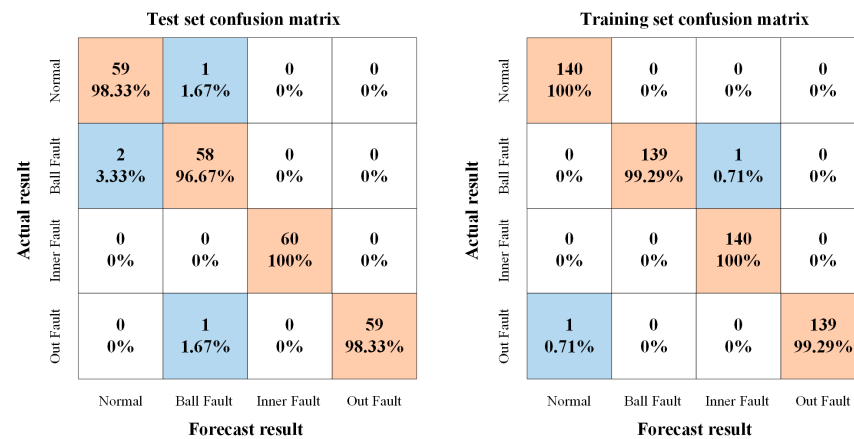

Figure 16. Confusion matrix of rolling bearing fault diagnosis.

At the model training level, sample number 57 in the normal class is mis-predicted as an outer ring failure and sample number 35 in the inner ring failure class is mis-predicted as a rolling element failure. In the test set, the model is able to diagnose the inner ring of the rolling bearing with 100% accuracy, while there is one misclassification in the normal condition, which predicted sample No. 90 as a rolling element fault. There are two misclassifications for rolling element faults; samples #4 and #25 are predicted as normal, and in the case of the outer ring fault, sample #56 is predicted as a rolling element fault.

PSO-VMD, optimized IPSO-VMD and EMD, PNN and SVM, and ELM methods are compared and analyzed. The diagnostic results for the rolling bearing fault samples are shown in Table 7. In general, for the rolling bearing fault data, IPSO-VMD-PNN has the best diagnostic performance in the test set, with an overall accuracy of 98.3%. In summary, the above proposed IPSO-VMD fault feature extraction method can be combined with the PNN model to achieve higher fault diagnosis accuracy.

Table 7. Diagnostic results of rolling bearing fault samples.

Gears Status	Feature Extraction Algorithm	Classification Algorithm	Number of Training Sets	Number of Test Sets	Right	Misjudgment	Accuracy Rate
Normal	IPSO-VMD	PNN	140	60	59	1	98.33%
	PSO-VMD	PNN	140	60	57	3	95.00%
	PSO-VMD	ELM	140	60	55	5	91.67%
	PSO-VMD	SVM	140	60	58	2	96.67%
	EMD	PNN	140	60	54	6	90.00%
Rolling element failure	IPSO-VMD	PNN	140	60	58	2	96.67%
	PSO-VMD	PNN	140	60	56	4	93.33%
	PSO-VMD	ELM	140	60	54	6	90.00%
	PSO-VMD	SVM	140	60	57	3	95.00%
	EMD	PNN	140	60	53	7	88.33%
Inner ring Failure	IPSO-VMD	PNN	140	60	60	0	100%
	PSO-VMD	PNN	140	60	59	1	98.33%
	PSO-VMD	ELM	140	60	58	2	96.67%
	PSO-VMD	SVM	140	60	58	2	96.67%
	EMD	PNN	140	60	56	4	93.33%
Outer ring Failure	IPSO-VMD	PNN	140	60	59	1	98.33%
	PSO-VMD	PNN	140	60	58	2	96.67%
	PSO-VMD	ELM	140	60	57	3	95.00%
	PSO-VMD	SVM	140	60	58	2	96.67%
	EMD	PNN	140	60	55	5	91.67%

5. Conclusions and Future Work

In this study, a fault diagnosis framework for rotating machinery based on IPSO-VMD-PNN was established by the combination of numerical simulation and experimental verification, and the effectiveness of the framework was verified with the open data of gears and rolling bearings.

- In view of the difficulty in selecting the two key parameters K and α in VMD, the adaptive inertia weight strategy and compression factor method were used to improve the PSO algorithm, and the adaptive optimization of parameters was carried out by combining the fitness function, so as to avoid the instability of manual value taking and cumbersome operation.
- In the extracted multi-dimensional feature vectors, the sensitivity of different fault feature parameters varies, and some parameters may be irrelevant or redundant. Laplace score was used to select sensitive feature parameters from multi-domain feature parameters, and the fault sensitive feature vector was constructed and input into the PNN model to realize fault diagnosis.
- Through the validation of simulated and real signals, the proposed IPSO-VMD-PNN framework shows excellent robustness and adaptability, demonstrating its potential for application in practical engineering.

Future research will focus on further optimizing the proposed method to enhance its fault diagnosis accuracy under more complex conditions. Additionally, the potential application of this method in intelligent diagnosis under varying rotational speeds will be explored.

Author Contributions: Conceptualization, Z.L.; methodology, C.Z.; software, C.Z.; validation, Z.L. and C.Z.; investigation, J.H.; data curation, S.X.; writing—original draft preparation, Z.L. and C.Z.; writing—review and editing, Z.L. and Z.C.; visualization, W.Z.; supervision, H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author/s.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Desnica, E.; Ašonja, A.; Radovanović, L.; Palinkaš, I.; Kiss, I. Selection, Dimensioning and Maintenance of Roller Bearings. In *31st International Conference on Organization and Technology of Maintenance (OTO 2022)*; Lecture Notes in Networks and Systems; Blažević, D., Ademović, N., Barić, T., Cumin, J., Desnica, E., Eds.; Springer: Cham, Switzerland, 2023; Volume 592. [CrossRef]
2. Xu, X.; Yan, X.; Yang, K.; Zhao, J.; Sheng, C.; Yuan, C. Review of condition monitoring and fault diagnosis for marine power systems. *Transp. Saf. Environ.* **2021**, *3*, 85–102.
3. Chen, Q.; Dai, S.; Bi, X. Fault diagnosis of rolling bearing based on EEMD. *Comput. Simul.* **2021**, *38*, 361–364.
4. Wang, J.; Du, G.; Zhu, Z.; Shen, C.; He, Q. Fault diagnosis of rotating machines based on the EMD manifold. *Mech. Syst. Signal Process.* **2020**, *135*, 106443.
5. Xu, M.; Yao, H. Fault diagnosis method of wheelset based on EEMD-MPE and support vector machine optimized by quantum-behaved particle swarm algorithm. *Measurement* **2023**, *216*, 112923.
6. Zhao, H.; Li, X.; Liu, Z.; Wen, H.; He, J. A double interpolation and mutation interval reconstruction LMD and its application in fault diagnosis of reciprocating compressor. *Appl. Sci.* **2023**, *13*, 7543. [CrossRef]
7. Goyal, D.; Choudhary, A.; Sandhu, J.K.; Srivastava, P.; Saxena, K.K. An intelligent self-adaptive bearing fault diagnosis approach based on improved local mean decomposition. *Int. J. Interact. Des. Manuf.* **2022**, *16*, 1–11. [CrossRef]
8. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544.
9. Li, F.; Li, R.; Tian, L.; Chen, L.; Liu, J. Data-driven time-frequency analysis method based on variational mode decomposition and its application to gear fault diagnosis in variable working conditions. *Mech. Syst. Signal Process.* **2019**, *116*, 462–479.
10. Lin, S.L. Intelligent fault diagnosis and forecast of time-varying bearing based on deep learning VMD-Densenet. *Sensors* **2021**, *21*, 7467. [CrossRef] [PubMed]
11. Li, L.; Meng, W.; Liu, X.; Fei, J. Research on rolling bearing fault diagnosis based on variational modal decomposition parameter optimization and an improved support vector machine. *Electronics* **2023**, *12*, 1290. [CrossRef]
12. Liu, C.; Wu, Y.; Zhen, C. Rolling bearings fault diagnosis based on variational mode decomposition and fuzzy c means clustering. *Proc. CSEE* **2015**, *35*, 3358–3365.
13. Li, K.; Su, L.; Wu, J.; Wang, H.; Chen, P. A rolling bearing fault diagnosis method based on variational mode decomposition and an improved kernel extreme learning machine. *Appl. Sci.* **2017**, *7*, 1004. [CrossRef]
14. Wang, X.; Yang, Z.; Yan, X. Novel particle swarm optimization-based variational mode decomposition method for the fault diagnosis of complex rotating machinery. *IEEE/ASME Trans. Mechatron.* **2017**, *23*, 68–79.
15. Wang, C.; Li, H.; Huang, G.; Ou, J. Early fault diagnosis for planetary gearbox based on adaptive parameter optimized VMD and singular kurtosis difference spectrum. *IEEE Access* **2019**, *7*, 31501–31516.
16. Xiao, D.; Ding, J.; Li, X.; Huang, L. Gear fault diagnosis based on kurtosis criterion VMD and SOM neural network. *Appl. Sci.* **2019**, *9*, 5424. [CrossRef]
17. Zhang, J.; Ji, J.; Xu, T. A bearing fault diagnosis method based on variational mode decomposition parameter optimization. *Sci. Technol. Eng.* **2021**, *21*, 3601–3605.
18. Zhu, S.; Xia, H.; Peng, B.; Zio, E.; Wang, Z.; Jiang, Y. Feature extraction for early fault detection in rotating machinery of nuclear power plants based on adaptive VMD and Teager energy operator. *Ann. Nucl. Energy* **2021**, *15*, 108392.
19. Zhang, Q.; Chen, S.; Fan, Z.P. Bearing Fault Diagnosis based on Improved Particle Swarm Optimized VMD and SVM models. *Adv. Mech. Eng.* **2021**, *13*, 16878140211028451.
20. Song, X.; Wang, H.; Chen, P. Weighted kurtosis-based VMD and improved frequency-weighted energy operator low-speed bearing-fault diagnosis. *Meas. Sci. Technol.* **2020**, *32*, 035016.
21. Chang, Y.; Bao, G.; Cheng, S.; He, T.; Yang, Q. Improved VMD-KFCM algorithm for the fault diagnosis of rolling bearing vibration signals. *IET Signal Process.* **2021**, *15*, 238–250.
22. Jin, Z.; He, D.; Wei, Z. Intelligent fault diagnosis of train axle box bearing based on parameter optimization VMD and improved DBN. *Eng. Appl. Artif. Intell.* **2022**, *110*, 104713.
23. Sharma, V.; Parey, A. Extraction of weak fault transients using variational mode decomposition for fault diagnosis of gearbox under varying speed. *Eng. Fail. Anal.* **2020**, *107*, 104204.
24. Lv, X.; Zhou, D.; Ma, L.; Zhang, Y.; Tang, Y. An Improved Lagrange Particle Swarm Optimization Algorithm and Its Application in Multiple Fault Diagnosis. *Shock Vib.* **2020**, *2020*, 1091548.
25. Zhu, J.; Wang, C.; Hu, Z.; Kong, F.; Liu, X. Adaptive variational mode decomposition based on artificial fish swarm algorithm for fault diagnosis of rolling bearings. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2017**, *231*, 635–654.

26.   Liu, Y.; Mu, C.; Kou, W.; Liu, J. A simple multi-population evolutionary algorithm using PSO strategy for constrained engineering design optimization. *Int. J. Digit. Content Technol. Appl.* **2012**, *6*, 532–541.

27.   Li, K.; Zhou, G.; Zhai, J.; Li, F.; Shao, M. Improved PSO_AdaBoost ensemble algorithm for imbalanced data. *Sensors* **2019**, *19*, 1476. [CrossRef] [PubMed]

28.   Mjahed, S.; El Hadaj, S.; Bouzaachane, K.; Raghay, S. Improved PSO based K-means clustering applied to fault signals diagnosis. In Proceedings of the 2018 International Conference on Control, Automation and Diagnosis (ICCAD), Marrakech, Morocco, 19–21 March 2018; pp. 1–6.

29.   Zhou, S.; Li, G.; Wang, Q.; Zheng, C.; Wen, Y. Improved particle swarm optimization algorithm based driving strategy research for permanent magnet spherical motor. *Trans. China Electrotech. Soc.* **2023**, *38*, 166–176+189.

30.   Feng, Z.; Zhang, C.; Li, B. Optimization of MFAC parameters based on improved particle swarm optimization algorithm. *Control Eng. China* **2021**, *28*, 766–773.

31.   Yang, J.; He, L.; Fu, S. An improved PSO-based charging strategy of electric vehicles in electrical distribution grid. *Appl. Energy* **2014**, *128*, 82–92.

32.   Yan, X.; Xu, Y.; She, D.; Zhang, W. A Bearing Fault Diagnosis Method Based on PAVME and MEDE. *Entropy* **2021**, *23*, 1402. [CrossRef] [PubMed]

33.   Peeters, C.; Antoni, J.; Helsen, J. Blind filters based on envelope spectrum sparsity indicators for bearing and gear vibration-based condition monitoring. *Mech. Syst. Signal Process.* **2020**, *138*, 106556.

34.   Xu, Y.; Wang, Y.; Lingzhi, W.; Qu, J. Bearing fault detection using an alternative analytic energy operator: A fast and non-filtering method. *Meas. Sci. Technol.* **2021**, *32*, 105101.

35.   Yan, X.; Liu, Y.; Huang, D.; Jia, M. A new approach to health condition identification of rolling bearing using hierarchical dispersion entropy and improved Laplacian score. *Struct. Health Monit.* **2021**, *20*, 1169–1195.

36.   Ma, J.; Li, Z.; Li, C.; Zhan, L.; Zhang, G.-Z. Rolling bearing fault diagnosis based on refined composite multi-scale approximate entropy and optimized probabilistic neural network. *Entropy* **2021**, *23*, 259. [CrossRef] [PubMed]

37.   Jing, L.; Zhao, M.; Li, P.; Xu, X. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. *Measurement* **2017**, *111*, 1–10.

38.   He, C.; Wu, T.; Gu, R.; Jin, Z.; Ma, R.; Qu, H. Rolling bearing fault diagnosis based on composite multiscale permutation entropy and reverse cognitive fruit fly optimization algorithm–extreme learning machine. *Measurement* **2021**, *173*, 108636.

39.   Case Western Reserve University (CWRU), Cleveland, Ohio, USA [EB/OL]. Available online: https://engineering.case.edu/bearingdatacenter/download-data-file (accessed on 22 January 2022).