machines

MDPI

Article

Feature Extraction of a Planetary Gearbox Based on the KPCA Dual-Kernel Function Optimized by the Swarm Intelligent Fusion Algorithm

Yan He [1], Linzheng Ye [1,2,*] and Yao Liu [1,2]

[1] School of Mechanical Engineering, North University of China, Taiyuan 030051, China; zbhy91@163.com (Y.H.); liuyao@nuc.edu.cn (Y.L.)
[2] Shanxi Key Laboratory of Advanced Manufacturing Technology, North University of China, Taiyuan 030051, China
* Correspondence: yelinzheng@nuc.edu.cn

Abstract: The feature extraction problem of coupled vibration signals with multiple fault modes of planetary gears has not been solved effectively. At present, kernel principal component analysis (KPCA) is usually used to solve nonlinear feature extraction problems, but the kernel function selection and its blind parameter setting greatly affect the performance of the algorithm. For the optimization of the kernel parameters, it is very urgent to study the theoretical modeling to improve the performance of kernel principal component analysis. Aiming at the deficiency of kernel principal component analysis using the single-kernel function for the nonlinear mapping of feature extraction, a dual-kernel function based on the flexible linear combination of a radial basis kernel function and polynomial kernel function is proposed. In order to increase the scientificity of setting the kernel parameters and the flexible weight coefficient, a mathematical model for dual-kernel parameter optimization was constructed based on a Fisher criterion discriminant analysis. In addition, this paper puts forward a swarm intelligent fusion algorithm to increase this method's advantages for optimization problems, involving the shuffled frog leaping algorithm combined with particle swarm optimization (SFLA-PSO). The new fusion algorithm was applied to optimize the kernel parameters to improve the performance of KPCA nonlinear mapping. The optimized dual-kernel function KPCA (DKKPCA) was applied to the feature extraction of planetary gear wear damage, and had a good identification effect on the fuzzy damage boundary of the planetary gearbox. The conclusion is that the DKKPCA optimized by the SFLA-PSO swarm intelligent fusion algorithm not only effectively improves the performance of feature extraction, but also enables the adaptive selection of parameters for the dual-kernel function and the adjustment of weights for the basic kernel function through a certain degree of optimization; so, this method has great potential for practical use.

Keywords: swarm intelligence optimization; kernel principal component analysis; dual-kernel function; planetary gearbox; feature extraction

check for updates

Citation: He, Y.; Ye, L.; Liu, Y. Feature Extraction of a Planetary Gearbox Based on the KPCA Dual-Kernel Function Optimized by the Swarm Intelligent Fusion Algorithm. *Machines* **2024**, *12*, 82. https://doi.org/10.3390/machines12010082

Academic Editors: Weidong Zhu, Yawen Wang, Gang Li and Jing Wei

Received: 13 December 2023
Revised: 15 January 2024
Accepted: 18 January 2024
Published: 21 January 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Feature extraction is an important prerequisite for pattern recognition. It can be divided into the methods based on signal processing and those based on learning [1]. The former was developed earlier, and its typical algorithms include the Fourier transform, Gabor filter, and wavelet transform [2]. The latter has emerged in the last ten years; this method uses the dimension reduction method to map the data from the original space onto a certain low-dimensional space. The data in the low-dimensional space can greatly reflect the essential characteristics of the original space. The typical algorithms include principal component analysis (PCA), kernel principal component analysis (KPCA) [3], linear discriminant analysis (LDA) [4], etc. The feature extraction method based on signal processing involves extracting features from the transform domain and emphasizing the

individual information from the samples, while the method based on learning mainly considers the statistical characteristics of all the samples and extracts the typical features that can express the original information.

In recent years, some achievements have been made using the proposed kernel learning method in the research fields of feature extraction, pattern recognition, data mining, and image and signal processing. To a certain extent, it solves the nonlinear problem in an actual system and improves the accuracy of pattern recognition and prediction. For example, a kernel principal component analysis (KPCA) and the observer method were applied to detect faults in a hydraulic system by analyzing the variation in water volume, and achieved results for the detection of faults in complex and nonlinear systems [5]. A rolling bearing fault diagnosis method based on the Volterra series and a kernel principal component analysis (KPCA) has been proposed [6]. However, the kernel learning method has encountered a bottleneck problem that affects the performance of the algorithm, that is, the selection of the kernel function and its parameters. The kernel method is based on the kernel function, which can be divided into the local kernel and global kernel at present. For the so-called local kernel, only when the distance between the data points is very close or the two features are similar will the value of the kernel function be greatly affected; for the global kernel, on the contrary, a large distance between the data points or a significant feature difference can also have a great impact on the value of the kernel function. In general, the radial basis function (RBF) kernel is the typical local kernel, and the polynomial kernel is the global kernel, and their mapping performances are different when their parameters are changed. The different kernel functions and their parameters reflect the differences in their mapping performance for the overall or local data.

In the practical application of the kernel analysis method, the type of kernel function is generally selected according to the data characteristics of practical problems [7]. Each kernel function has its own limitations. In reality, one kernel function mapping can only reflect a certain feature, and often cannot effectively describe the whole feature. Moreover, with the complexity and diversification of engineering problems, there is an increasing number of problems, such as large sample data sets or uneven high-dimensional data, and the defects of the single-kernel function are more obvious. Therefore, researchers began to try to combine the multi-kernel functions. Lei et al. proposed a relevance vector machine prediction method based on an adaptive multi-kernel combination, which was applied to the remaining useful life prediction of machinery [8]. Fu et al. proposed a combined kernel correlation vector machine and a quantum particle swarm optimization algorithm for transformer fault diagnosis [9]. Deng et al. modified kernel principal component analysis using a dual-weighted local outlier factor and applied this method to linear process monitoring [10]. Pan et al. put forward a mean-class kernel principal component fault diagnosis method with a combined kernel function for monitoring the distillation process [11].

In recent years, it has been proven that the performance of nonlinear mapping using multiple kernels is better than that using the single-kernel model or the single-kernel machine combination model. The construction methods for the combinatorial kernel can be divided as follows: the first method is a multi-kernel linear combination synthesis, in which the shape is like $K = \sum_{i=1}^{q} \beta_j K_i, \beta_j \geq 0, \sum_{j}^{q} \beta_j = 1$, where q is the number of basic kernel functions, K_i is the basic kernel function, and β is the weight coefficient; the second is a multi-kernel extended synthesis method, which is an attempt to achieve the fusion of the different kernel matrices through the concept of summation "averaging" [12]. For example, Lv et al. derived a hybrid kernel function combining the RBF kernel with the polynomial kernel, and introduced it into an extreme learning machine to improve the accuracy of an intrusion detection system [13]. Nithya et al. proposed a method for kidney disease detection and segmentation using an artificial neural network and a multi-kernel K-means clustering algorithm for ultrasound images [14]. Ouyang et al. proposed a hybrid improved kernel LDA and PNN algorithm for efficient face recognition [15]. Afzal et al. achieved deep multilayer kernel learning in kernel vector machines for classification [16].

For the combinatorial kernel function, the scientific setting of the kernel parameters is the key to improving the performance of the kernel learning algorithm. In past studies, the kernel parameters were set according to the experimental or empirical value with considerable blindness. However, the current theoretical research is mainly based on the experimental correction method, which is inefficient, as well as data-related methods and intelligence based on optimization methods [17]. Applying the swarm intelligence optimization method to the study of the optimization of the kernel function and its parameters represents a new research hotspot for improving the status quo and enhancing the performance of the kernel learning method [18,19].

Particle swarm optimization (PSO) and the shuffled frog leaping algorithm (SFLA) are categorized as swarm intelligence optimization algorithms [20,21]. PSO is a method derived from the behavior characteristics of biological populations, and is used to solve optimization problems. In order to optimize the search, the behavioral characteristics of birds are simulated, and competition among the individuals is achieved. The SFLA simulates the optimization of the group foraging of frogs by introducing a meme evolution mechanism into the group search. Organizational social behavior is used to replace the natural selection mechanisms of evolutionary algorithms, and this method has been widely used in optimization problems, path selection, and fault diagnosis [22].

As complex compound gear transmission systems, planetary gearboxes with a compact structure, large transmission ratio, high transmission efficiency, smooth operation, and other characteristics, have been widely used in automobile gearboxes, wind turbines [23], and aviation engine settings. They play an important role in industries such as coal, energy, advanced manufacturing, and wind power. Due to the harsh working environment, gear and bearing damage and failures often occur and, due to different degrees of damage, diverse fault modes, and complex vibration signal transmission paths, the problems of feature extraction, state recognition, and the fault diagnosis of planetary gear multi-fault-mode coupled vibration signals have not been effectively solved. At present, KPCA is commonly used to solve nonlinear feature extraction problems. This paper focuses on the transmission boxes of wind turbines and establishes a planetary gearbox simulation fault experimental platform for feature extraction and fault diagnosis research.

In this paper, aiming at the deficiency of the nonlinear mapping of the single-kernel function for KPCA feature extraction, a dual-kernel function with a flexible linear combination of the RBF and polynomial function is proposed. In order to increase the scientificity of the setting of the kernel parameters and the flexible weight coefficient, a mathematical model for the optimization of the kernel parameters was constructed, based on Fisher's discriminant thought. Making full use of the advantages of the swarm intelligence fusion algorithm for solving optimization problems, the SFLA-PSO fusion algorithm is presented, and is used to find the optimal solution for kernel parameters and the flexible weight coefficient. Lastly, a KPCA of the dual-kernel function optimized by the SFLA-PSO fusion algorithm was applied to the feature extraction of planetary gear wear.

The paper is organized as follows: First, it introduces the kernel learning method and summarizes the combination kernel function and its kernel parameter's important role in feature extraction with KPCA. The necessity of kernel parameter optimization with the swarm intelligence optimization algorithm is put forward. In Section 2, the SFLA and PSO fusion mechanism is described, and the flow chart for the swarm intelligence algorithm is given. The constitution of the dual-kernel function of the KPCA and its optimization model, as well as the simulation and comparative analysis of the Iris data, are described in Section 3. In addition, the experimental scheme for planetary gearbox failure is described in Section 4. The feature extraction for the planetary gearbox based on the KPCA dual-kernel parameter optimized by the SFLA-PSO is presented, and the simulation results are discussed in detail in Section 5. Finally, we present the conclusion in Section 6.

2. Swarm Intelligence Fusion Algorithm

2.1. The SFLA and PSO Fusion Mechanism

As a new branch of evolutionary algorithms, the swarm intelligence optimization algorithm has many advantages, such as its simplicity, easy implementation, low requirements for the mathematical behavior of the target problem, and high efficiency, so it has attracted the attention of scholars [24]. One form of the swarm intelligence optimization algorithm, PSO, has the advantage of fast convergence speed, but it easily falls into local convergence, while the SFLA has a strong global search and optimization ability. Similarly, as in evolution, it has to go through a series of processes, such as frog population generation, evolution, crossover, and information exchange. The SFLA evolution mechanism is more flexible, but its computational complexity is high, and its convergence speed is slow. Therefore, the fusion strategy of different algorithms aims to utilize the unique advantages of the original algorithms, achieving mutual coordination and complementary advantages, and forming a robust and more efficient algorithm [25]. The fusion of the swarm intelligence algorithm involves the construction of a new fusion algorithm, replacing the original algorithm, which involves the strategies of population size, the population structure relationship, individual learning or interaction, and the population evolution mode (selecting, eliminating, and updating individuals). According to the concept of elite strategy, the SFLA algorithm is combined with the PSO algorithm in this study, the strategy of "two-level optimization and internal and external circulation" is implemented, and the SFLA-PSO swarm intelligent fusion algorithm is formed. Compared to other improved PSO and SFLA methods, the concept of the fusion algorithm is simple, no additional parameters are added, and it is easy to implement.

The SFLA-PSO fusion strategy is as follows: Firstly, the randomly generated particles are divided into n_{pso} subgroups of equal size. The particles in each subgroup evolve independently according to the PSO mechanism to achieve the optimization of the first level. Then, the optimal particles in each subgroup are taken out to form a new population, i.e., the initial frog group. According to the SFLA mechanism, the optimal frog position is that which realizes the optimization of the second level. Finally, the best position in the SFLA subgroup is reflected in the speed of the particle swarm optimization update, which can effectively guide PSO evolution and avoid falling into the local optimum. Using this strategy not only reduces the blindness of the initial SFLA group, but also enables a fine search within the SFLA.

The particle velocity and position update for the PSO optimization of the first level are shown in Equations (1) and (2) [20]:

$$v_{jid}(t+1) = \omega \times v_{jid}(t) + c_1 r_1 \left(p_{jid} - x_{jid}(t) \right) + c_2 r_2 \left(p_{jgd} - x_{jid}(t) \right) \tag{1}$$

$$x_{jid}(t+1) = x_{jid}(t) + v_{jid}(t+1) \tag{2}$$

where $j = 1, 2, \ldots, n_{\mathrm{pso}}$ is the group number of the particle swarm; $i = 1, 2, \ldots, n_{\mathrm{p}}$ denotes the ith particle; p_{jid} and p_{jgd} represent the current optimal position and the global optimal position of the particles in the jth group, respectively; and x_{jid} is the current position of the ith particle in the jth group. v_{jid} is the current velocity of the ith particle in the jth group, $v_{jid} \in [-v_{\mathrm{limit}}, v_{\mathrm{limit}}]$, v_{limit} is the maximum speed limit. C_1 and c_2 are the learning factors, r_1 and r_2 are uniform random numbers within the range [0, 1], and ω is the inertia weight. t is the evolution generation.

In the SFLA optimization of the second stage, the frog's velocity and position update are shown in Equations (3)–(5) [21].

The step length update formula is

$$S_i = r \times (P_{\mathrm{b}} - P_{\mathrm{w}}) \tag{3}$$

$$S_i = r \times (P_g - P_{\mathrm{w}}) \tag{4}$$

where r is a uniform random number within [0, 1], P_b is the optimal frog position in the meme group, P_w is the worst frog position in the meme group, P_g is the best frog position in the whole group, and S_i is the frog step length, $S_{min} \leq S_i \leq S_{max}$.

The position update formula for the worst-positioned frog is

$$P'_w = P_w + S_i \quad (5)$$

where P'_w is the position of the updated frog, which is within the range $[P_{min}, P_{max}]$. When the position of the worst-positioned individual is not improved through this evolution, P_g, the best of the entire frog group, is used to replace P_b, as shown in Equation (4), and then Formula (5) is used for further calculations.

2.2. Fusion Algorithm Flow

The flow chart for the SFLA-PSO fusion algorithm is shown in Figure 1. The specific optimization process is as follows:

(1) Relevant parameters of the SFLA-PSO fusion algorithm are set.

The parameters for PSO include the total number of particles in the initial particle swarm N, the number of particles per particle swarm n_p, the number of particle groups n_{pso}, the particle dimension d, the total number of iterations T_{max}, the inertia weight ω, learning factors c_1, c_2, and the velocity limit v_{limit}. The parameters of the SFLA include the number of meme groups m, the number of frogs in each meme group n, the total number of iterations of the SFLA G_{max}, the number of local cycles L_{max}, and the maximum step S_{max} of the SFLA, where $N = n_{pso} \times n_p$ and $n_{pso} = n \times m$.

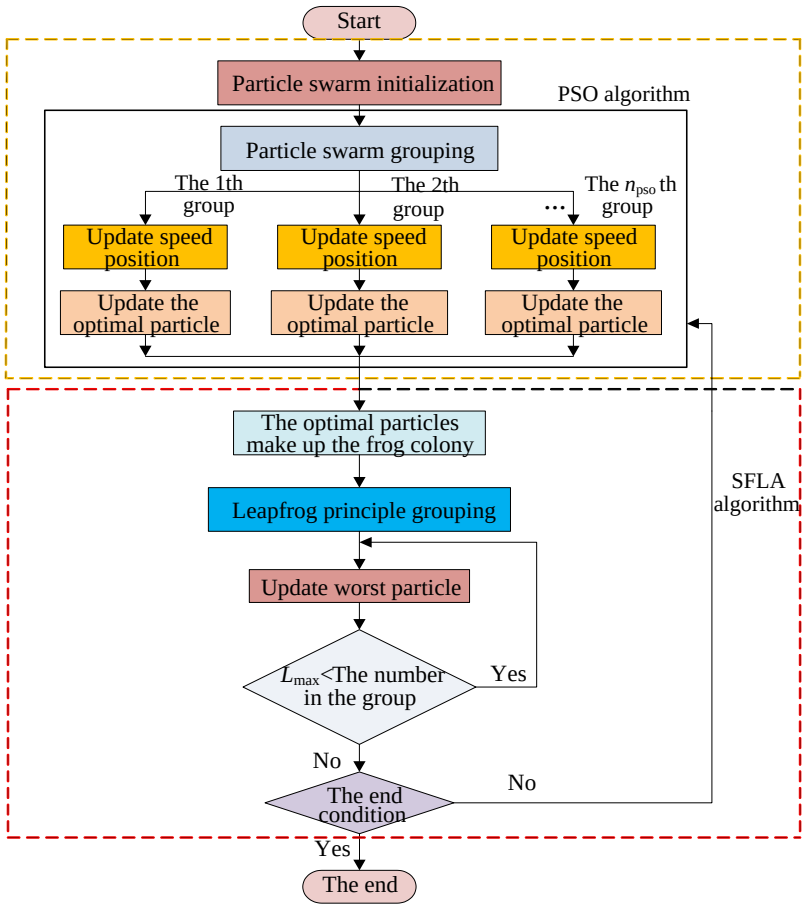


Figure 1. SFLA-PSO fusion algorithm flow chart.

(2) The particle swarm is initialized, and each particle is initialized. After the fitness is sorted, each particle is divided into n_{pso} groups with n_{p} particles in each group.
(3) The PSO mode iterative optimization is carried out for each group, the fitness of each particle is evaluated, and the optimal particle P_{jid} of each group is recorded.
(4) According to Formulas (1) and (2) for PSO, the velocity and position of each group particle are updated to complete the optimization of the first level.
(5) Each group of optimal particles (called elites) optimized by PSO are regarded as all the frogs of the SFLA, and their fitness is evaluated and ranked.
(6) According to the SFLA grouping mechanism, the elite particles are divided into m groups with n particles in each group.
(7) According to the updating Formulas (3)–(5) of the SFLA, the local optimal value and the local worst value for the population are recorded. After all the groups are updated, the global optimal value is updated to find the best X_{gd}, and the second-level optimization is completed.
(8) If the end condition is satisfied, that is, the predetermined accuracy requirement is met or the set maximum iteration number is reached, X_{gd} is the output, and the algorithm is stopped; otherwise, the procedure returns to (4).

3. The Composition and Optimization of the Dual-Kernel Function of KPCA

3.1. The Principle of KPCA Algorithm Feature Extraction

The detection data can be abstracted as vector $\boldsymbol{y_i} = \left(y_i^1, y_i^2, \cdots, y_i^L\right)$, where y_i^L is the Lth-dimension data of $\boldsymbol{y_i}$. The KPCA feature extraction steps are as follows [26]:

(1) The data are mapped to the high-dimensional feature space via the nonlinear transformation $\Phi(\boldsymbol{y_i})$, and the kernel matrix can be calculated. The covariance matrix $\boldsymbol{C}^{\mathbf{F}}$ in the feature space of the training sample after mapping is

$$\boldsymbol{C}^{\mathrm{F}} = \frac{1}{M}\sum_{i=1}^{M}\Phi(\boldsymbol{y}_i)\Phi(\boldsymbol{y}_i)^{\mathrm{T}} \tag{6}$$

where M is the number of training samples. The kernel matrix $\boldsymbol{K}$ is defined as

$$[\boldsymbol{K}]_{ij} = K_{ij} = \left\langle \Phi(\boldsymbol{y}_i), \Phi\left(\boldsymbol{y}_j\right)\right\rangle = K(\boldsymbol{y}_i\boldsymbol{y}_j), i, j = 1, 2, \cdots, M \tag{7}$$

where K is the kernel function, which is defined as a function that represents the inner product $\left\langle \Phi(\boldsymbol{y}_i), \Phi\left(\boldsymbol{y}_j\right)\right\rangle$ of vector $\boldsymbol{y_i}$ and $\boldsymbol{y_j}$ through nonlinear mapping to a feature space, using two vectors in the original space as a function in order to achieve nonlinear mapping.

(2) Data centralization: If

$$\sum_{i=1}^{M}\Phi(\boldsymbol{y}_i) \neq 0$$

then the centralized processing is carried out. $\boldsymbol{K}$ is replaced by $\boldsymbol{K^*} = \boldsymbol{K} - \boldsymbol{AK} - \boldsymbol{KA} + \boldsymbol{AKA}$. In the formula, $A_{ij} = \frac{1}{\boldsymbol{M}}$.

(3) The eigenvalues λ and eigenvectors $\boldsymbol{v}$ of the covariance matrix $\boldsymbol{C}^{\mathbf{F}}$ are calculated, satisfying the following conditions:

$$\lambda\boldsymbol{v} = \boldsymbol{C^F v} \tag{8}$$

By substituting Equation (7) into Equation (8), the following equation can be obtained:

$$\boldsymbol{M}\lambda\boldsymbol{\alpha} = \boldsymbol{K\alpha} \tag{9}$$

The eigenvalues greater than zero, $\lambda_1, \lambda_2, \ldots, \lambda_p$, and the corresponding eigenvectors, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_p$, are obtained by solving Equation (9), giving

$$\boldsymbol{V}^k = \sum_{i=1}^{M} \alpha_i^k \Phi(\boldsymbol{y}_i) \tag{10}$$

(4) Extraction of the principal component: the projection $\Phi(\boldsymbol{y})$ on the feature vector space $\boldsymbol{V}^k$ is calculated:

$$\boldsymbol{V}^k \cdot \Phi(\boldsymbol{y}) = \sum_{i=1}^{M} \alpha_i^k \Phi(y_i)\Phi(\boldsymbol{y}) \tag{11}$$

where $\Phi(\boldsymbol{y_i})\cdot\Phi(\boldsymbol{y})$ can be calculated using the kernel technique, which is a type of function that makes $\boldsymbol{K}_{ij} = \Phi(\boldsymbol{y}_i)\cdot\Phi(\boldsymbol{y}_j)$ hold. Let

$$g_k(\boldsymbol{y}) = V^k \cdot \Phi(\boldsymbol{y}) = \sum_{i}^{M} \alpha_i^k K(\boldsymbol{y}_i, \boldsymbol{y}) \tag{12}$$

where K is the kernel function and $g_k(\boldsymbol{y})$ becomes the kth nonlinear kernel principal component corresponding to $\Phi(\boldsymbol{y})$, and all projected values are used as the feature vectors of the samples.

$$g_k(\boldsymbol{y}) = \left[g_1(\boldsymbol{y}_i), g_2(\boldsymbol{y}_i), \cdots, g_p(\boldsymbol{y}_i) \right] \tag{13}$$

As long as the first r maximum eigenvalues of $\boldsymbol{K}$ satisfy $\sum_{j=1}^{r} \lambda_j / \sum_{j=1}^{p} \lambda_j \geq 85\%$, then r feature vectors can be extracted from the above p feature components; thus, the feature vectors of the sample can be extracted as

$$\boldsymbol{y}_i' = \left[g_1(\boldsymbol{y}_i), g_2(\boldsymbol{y}_i), \cdots, g_r(\boldsymbol{y}_i) \right]$$

The common kernel functions are as follows:
Polynomial kernel function (Poly):

$$K(y_i, y_j) = [(y_i \cdot y_j) + 1]^{d_1} \tag{14}$$

Radial basis kernel function (RBF):

$$K(y_i, y_j) = \exp(-\|y_i - y_j\|^2 / 2\sigma^2) \tag{15}$$

Sigmoid kernel function (Sigmoid):

$$K(y_i, y_j) = \tanh[\nu(y_i, y_j) + c] \tag{16}$$

3.2. The Composition of the Dual-Kernel Function of Flexible Weight Linear Combination

At present, the combination kernel method typically employs a new kernel function generated by a linear combination of the simple kernel functions, which satisfies the Mercer condition and has had some successful applications [26,27]. However, there is no theoretical basis for the parameter selection or the combination of kernel functions, and the uneven distribution of samples cannot be solved satisfactorily, which greatly limits the expression ability of the decision function. Therefore, there is an urgent need to establish a mathematical model for combinatorial kernel optimization using theory and to theoretically optimize the weight coefficient and parameters of the combined kernel function using an intelligent algorithm [28].

Due to the different characteristics of different kernel functions, their performance at solving specific problems varies significantly. Therefore, the combined kernel is the best

method with which to select the kernel function, and it can avoid the problem of kernel function selection to some extent. The dual kernel can adopt the direct sum kernel, weighted sum kernel, and weighted polynomial extension kernel. In many practical applications, the Gaussian kernel function shows excellent properties. In this study, the Gaussian kernel function was used as the main function, and other kernel functions were selected for linear combination. The Gaussian kernel function is a typical local kernel. In order to balance the local kernel with the global kernel, the typical global kernel–polynomial kernel is used for the combination. In the Gaussian kernel function, the RBF is a typical local kernel. So, the RBF and the polynomial kernel were used in this study as a flexible linear combination. The dual-kernel function expression is shown in Equation (17):

$$K(y_i \cdot y_j) = \gamma \cdot [(y_i \cdot y_j) + 1]^{d_1} + (1 - \gamma) \cdot \exp(-\left\|y_i - y_j\right\|^2 / 2\sigma^2) \tag{17}$$

where γ is the flexible weight of the two kernel functions, with $1 \geq \gamma \geq 0$, and σ and d_1 are the parameters of the combined kernel function.

With the help of Fisher's criteria (FC) of minimum within-class distance and maximum class distance [29], a model with which to measure the class discrimination of the feature space data was constructed. For two kinds of problems, the classification can be determined by constructing the discriminant function and criterion [30].

In the feature space, the data samples are $\boldsymbol{y}_1(y_{11}, y_{12}, \cdots, y_{1i}), \boldsymbol{y}_2(y_{21}, y_{22}, \cdots, y_{2j})$ ($i = 1, 2, \ldots, n_1, j = 1, 2, \ldots, n_2$), and the mean vectors of the two groups of samples are

$$\boldsymbol{\mu}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} \Phi(y_{1i}) \tag{18}$$

$$\boldsymbol{\mu}_2 = \frac{1}{n_2} \sum_{j=1}^{n_2} \Phi(y_{2j}) \tag{19}$$

The square of the distance between the two groups is as follows:

$$\begin{aligned} D_b &= \|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|^2 = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \\ &= \tfrac{1}{n_1^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} k(y_{1i}, y_{1j}) - \tfrac{2}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} k(y_{1i}, y_{2j}) + \tfrac{1}{n_2^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} k(y_{2i}, y_{2j}) \end{aligned} \tag{20}$$

The square of the dispersion in $\boldsymbol{y}_\mathrm{k}$ is

$$D_{w1} = \sum_{i=1}^{n_1} \left\|\Phi(y_{1i}) - \boldsymbol{\mu}_1\right\|^2 = \sum_{i=1}^{n_1} \Phi(y_{1i})^T \Phi(y_{1i}) - n_1 \boldsymbol{\mu}_1^T \boldsymbol{\mu}_1 = \sum_{i=1}^{n_1} k(y_{1i}, y_{1i}) - \frac{1}{n_1} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} k(y_{1i}, y_{1j}) \tag{21}$$

$$D_{w2} = \sum_{j=1}^{n_2} \left\|\Phi(y_{2j}) - \boldsymbol{\mu}_2\right\|^2 = \sum_{j=1}^{n_2} \Phi(y_{2j})^T \Phi(y_{2j}) - n_2 \mu_2^T \mu_2 = \sum_{j=1}^{n_2} k(y_{2j}, y_{2j}) - \frac{1}{n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} k(y_{2i}, y_{2j}) \tag{22}$$

According to Fisher's criterion, the objective function is established; that is, the fitness function of the swarm intelligence optimization is

$$J_{Fisher}(\gamma, d_1, \sigma) = \frac{D_{w1} + D_{w2}}{D_b} \tag{23}$$

Because the mapping process is realized by means of the kernel function, the final transformation is to find γ, d_1, and σ to obtain the minimum value of the J_{Fisher} function. The optimal kernel parameters γ^*, d_1^*, and σ^* can be found via the SFLA-PSO to achieve the minimum value of J_{Fisher}.

3.3. The Specific Steps of Dual-Kernel Parameter Optimization

The FC discriminant function J_{Fisher} is used as the fitness of the SFLA-PSO algorithm to optimize the parameters γ, d_1, and σ of the dual-kernel function. The specific process is as follows:

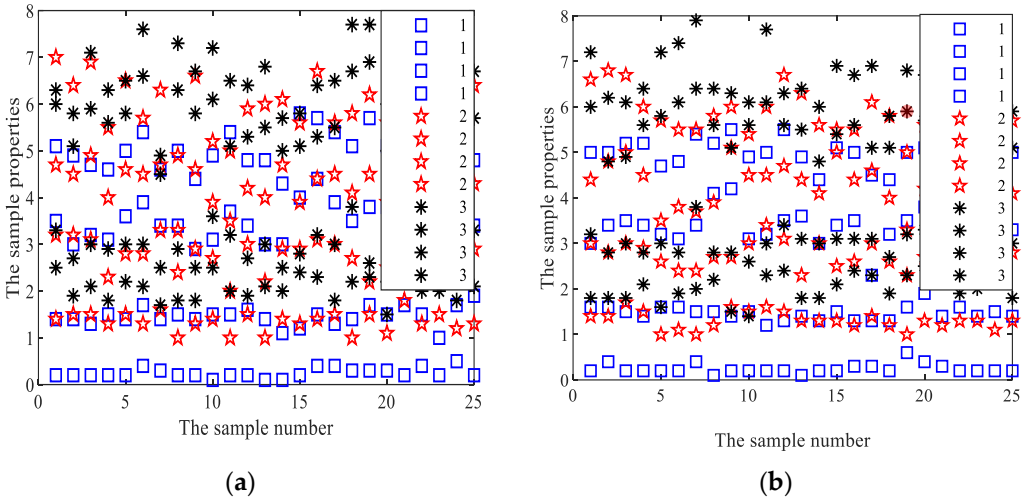
(1) According to Fisher's criterion, the samples are input, and the sum of the square of the distance between class D_{b} and within class D_{w1} and D_{w2} are calculated using Formulas (20)–(22).
(2) The optimized objective function J_{Fisher} is constructed using Equation (23) and is used as the fitness of the swarm intelligence optimization.
(3) The parameters of the SFLA-PSO swarm intelligent fusion algorithm are set, and the particle swarm optimization is initialized.
(4) The initial population is generated randomly, the fitness of the individuals are calculated, and the velocity and position are updated according to Formulas (1) and (2) of the PSO strategy.
(5) The optimal particles optimized in the first layer of PSO are taken as all the initial frogs of the SFLA and are grouped once more.
(6) According to the fitness size, the frog's velocity and position are updated based on the frog leaping update Formulas (3)–(5) to determine the best advantage.
(7) If the objective function J_{Fisher} satisfies the termination condition, the optimal value J_{Fisher} (γ^*, $d_1{}^*$, σ^*) is output, and the algorithm is stopped. Otherwise, it returns to (4).

3.4. Simulation and Comparative Analysis

In order to compare the application effect of the KPCA feature extraction method based on the dual-kernel function proposed in this paper (DKKPCA for short), a comprehensive comparison test was carried out.

3.4.1. Iris Plant Database

In the simulation experiment, the Iris data set was selected, and its characteristics were used as the data source. It is divided into three types of patterns, and there are 50 data sets in each category. Each data point includes four data sets, namely, the length and width of the calyx and the length and width of the petals. They are commonly taken as test sets and training sets in data mining and data classification. The first class is linearly separable from the second and the third classes, while the second and the third classes are nonlinearly separable. The distribution diagram is shown in Figure 2.



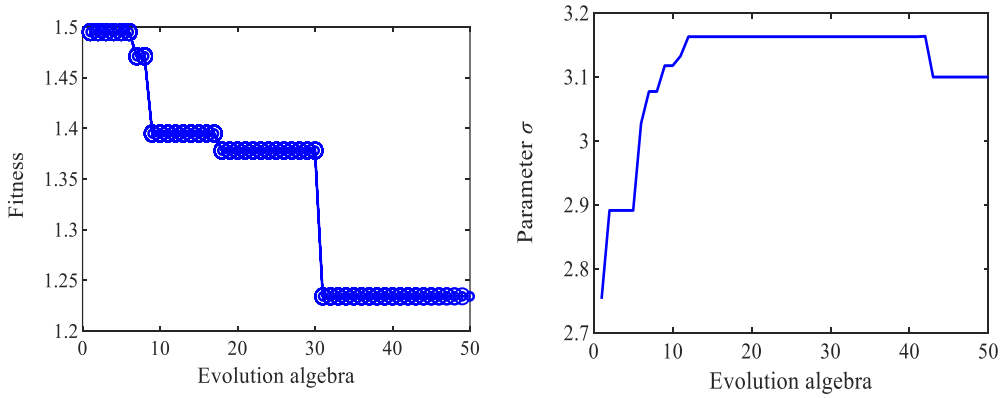
(**a**) (**b**)

Figure 2. Iris sample distribution: (**a**) the training sample distribution; (**b**) the test sample distribution.

3.4.2. Parameter Optimization of the Iris Data Set with the Dual-Kernel Function Based on the SFLA-PSO

In the simulation analysis, the first 25 samples of the Iris data set were taken as the training samples, and the SFLA-PSO was used to optimize the kernel parameters. The parameters of the SFLA-PSO were as follows: N = 200, n_{PSO} = 20, n_p = 10, c_1 = c_2 = 2, the PSO iteration number was 10, F = 20, m = 5, n = 4, L_{max} = 10, T_{max} = 50, d = 10, and S_{max} = 20.

The optimization iteration process is shown in Figure 3. It can be seen that the minimum value can be obtained in the 50 evolution algebras, and the three optimized parameters are γ^* = 0.057, d_1^* = 1, and σ^* = 3.1.



(**a**) (**b**)

Figure 3. Evolution process of kernel parameters: (**a**) SFLA-PSO evolution course; (**b**) evolution course of kernel parameters σ.

3.4.3. DKKPCA Feature Extraction of the Iris Data

The optimized parameters γ^*, d_1^*, and σ^* were used with the Iris data for the DKKPCA analysis, and the results are shown in Figure 4a,b. Figure 4a shows the DKKPCA scatter diagram for the principal component of the Iris data, and Figure 4b shows the histogram for the kernel principal component contribution rate. It can be seen that, due to the combination of the RBF kernel and the polynomial kernel function, the microscopic effect of the local kernel function and the amplification effect of the global kernel function are comprehensively used, the weight γ and kernel parameters d_1 and σ are optimized, and the clustering effect of the Iris samples is very obvious. Three types of Iris patterns are clustered in their respective centers, and the data projection points for categories 1 and 2 are very compact; the distance between classes is large, and the regional boundaries between classes are obvious.

For the comparative analysis, the analysis results of the single-kernel KPCA are shown in Figure 5a,b, which are the polynomial kernel function and the RBF kernel function, respectively. From Figure 5a, it can be found that the three types of data are mixed together in the feature space using KPCA and cannot be distinguished. It can be observed from Figure 5b that class 1 can be obviously separated from class 2 and class 3. However, some of the projection points for the data from classes 2 and 3 are still mixed due to their nonlinear inseparability, and the boundary between class 2 and class 3 is not obvious. The KPCA results are listed in Table 1 for the dual-kernel function (DKKPCA), the polynomial kernel function (KPCA_Poly), and the radial basis kernel (KPCA_RBF). The contribution rate of the first two principal components using KPCA is more than 85%; that is, Iris can completely replace the original four attributes by using two attributes, and the features are reduced by half.

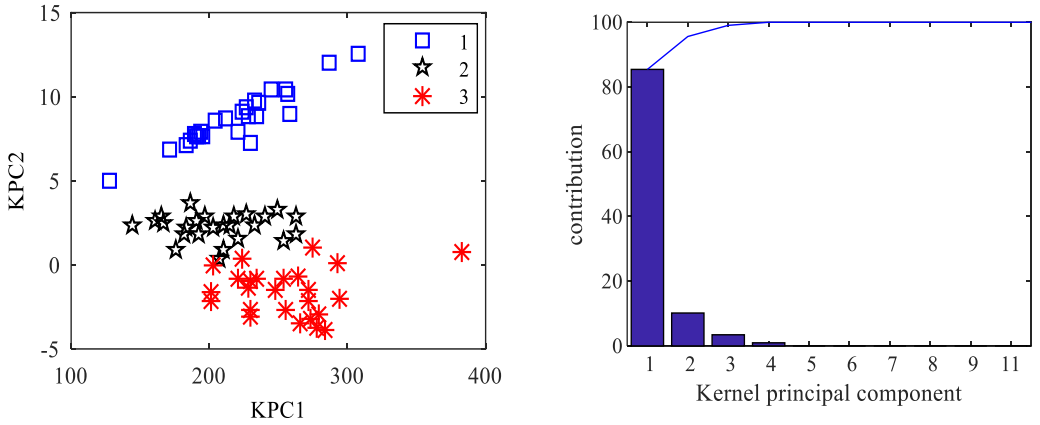


Figure 4. KPCA results for Iris data using the optimized dual-kernel parameters (**a**) $\gamma = 0.057$, $d_1 = 1$, and $\sigma = 3.1$; (**b**) histogram of the DKKPCA contribution rate.

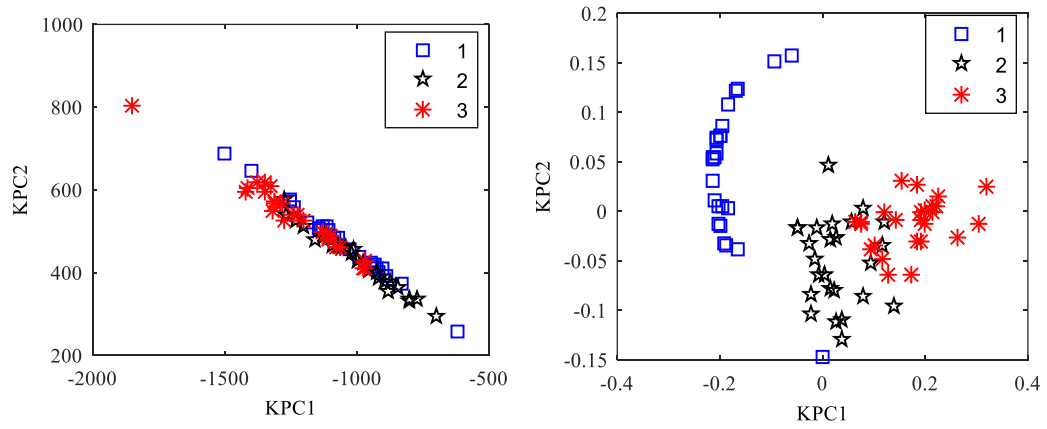


Figure 5. Iris data single-kernel principal component analysis results: (**a**) polynomial kernel function ($d_1 = 1$); (**b**) RBF kernel function ($\sigma = 3.1$).

Table 1. KPCA results for Iris data.

Algorithm	Serial Number	1	2	3	4
DKKPCA	Kernel principal component eigenvalue	2.518	0.2996	0.1014	0.0284
	Contribution %	85.431	10.165	3.441	0.963
	Cumulative contribution rate %	85.431	95.595	99.036	100.000
KPCA_Poly	Kernel principal component eigenvalue	49.566	4.0559	0.3441	0.0217
	Contribution %	91.8099	7.5126	0.6373	0.0402
	Cumulative contribution rate %	91.8099	99.3225	99.9558	100.00
KPCA_RBF	Kernel principal component eigenvalue	3.0601	0.5177	0.0946	0.0127
	Contribution %	83.0404	14.0486	2.5666	0.3444
	Cumulative contribution rate %	83.0404	97.089	99.6556	100.000

In order to compare the effect of KPCA before and after the parameter optimization of the dual-kernel function, the KPCA results before the optimization of the dual kernel are given, as shown in Figure 6. The results for parameters $\gamma = 0.5$, $d_1 = 1$, and $\sigma = 3.389$ are shown in Figure 6a, and those for $\gamma = 0.06$, $\sigma = 3.333$, and $d_1 = 1$ are shown in Figure 6b. It can be seen from these that due to the influence of the three parameters, the Iris data cannot be well separated linearly after nonlinear mapping onto the feature space, especially for

classes 2 and 3. The optimized results for the dual-kernel parameters shown in Figure 4a are obviously better than those of the KPCA before optimization in Figure 6.

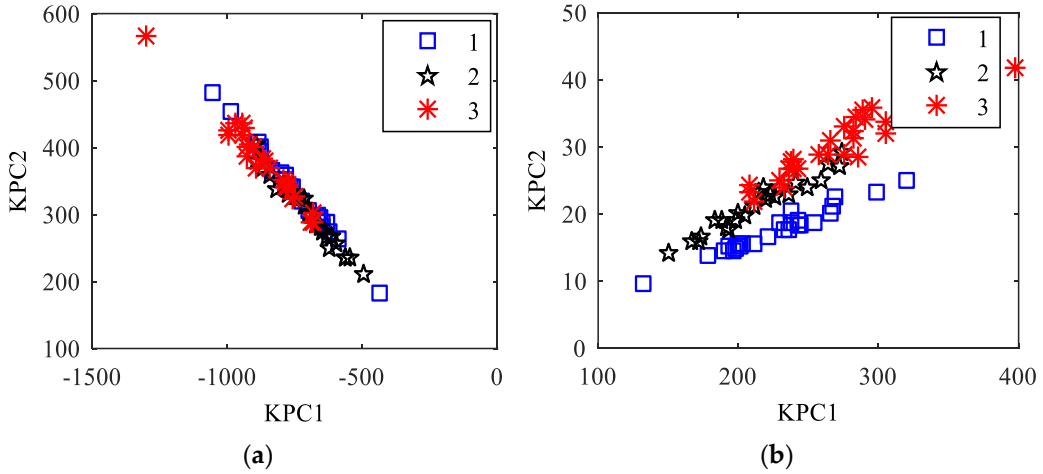


Figure 6. KPCA results for Iris data before dual-kernel function optimization: (**a**) $\gamma = 0.5$, $d_1 = 1$, and $\sigma = 3.389$; (**b**) $\gamma = 0.06$, $d_1 = 1$, and $\sigma = 3.333$.

Therefore, the DKKPCA optimized by the intelligent fusion algorithm is not only suitable for solving the nonlinear feature extraction problem, but can also provide better feature quality than the linear dimension reduction method and can greatly enhance the ability of nonlinear data processing.

4. The Simulation Failure Experiment for the Planetary Gearbox

4.1. Experimental Scheme of the Planetary Gearbox

As shown in Figure 7, the planetary gearbox fault diagnosis testbed consists of a control cabinet, a motor, a two-stage transmission box, a magnetic powder brake, etc. The structural diagram for the transmission system of the experimental platform is shown in Figure 8. The first transmission is a helical gear transmission, and the second stage is a 2K-H planetary gear transmission, which includes an inner gear ring, a solar wheel, and three planetary wheels. In this study, the fault modes for the planetary gear transmission system of a wind power system were simulated, and the speed range was 75–3000 r/min. Table 2 illustrates the equipment's technical parameters.

The motor provides power, the helical gearbox plays a deceleration role, and the planetary gearbox is a key research area. The magnetic powder brake acts as the load, and they are connected by an elastic coupling in the middle. We controlled the motor speed to adjust the shaft speed, and simulated different loads by setting the parameters of the magnetic powder brake. According to the testing plan, six measuring points, shown in Figure 9, were arranged. Four measuring points (1 to 4) were distributed on the helical gearbox, and the others (5 to 6) were on the planetary gearbox. The signal lines for the six sensors were connected to the DASP signal acquisition instrument and computer, and the entire collection system was fully built. The vibration signal was measured using six unidirectional piezoelectric acceleration sensors (CA-YD-186G) in this experiment.

Table 2. Technical parameters of main components.

Component	Name	Parameter
Helical gear box	Gear	Large gear: modules = 2, number of teeth = 77 Pinion: modules = 2, number of teeth = 55
	Bear	Deep groove ball rolling bearing 6206

Table 2. *Cont.*

Component	Name	Parameter
Planetary gearbox	Gear	Inner gear ring: modules = 2, number of teeth = 72
		Planetary gear: modules = 2, number of teeth = 27, quantity = 3
		Sun gear: modules = 2, number of teeth = 18
	Bear	Deep groove ball rolling bearing: for the planet wheel 6202, for the planet shelf = 6206, for the sun wheel 6205
Brake	The loading form is magnetic, and the loading torque is 0–100 N·m	
Motor	Converter motor	2.2 kW, the rotational speed is 1500 RPM, and the rated speed is 1410 RPM

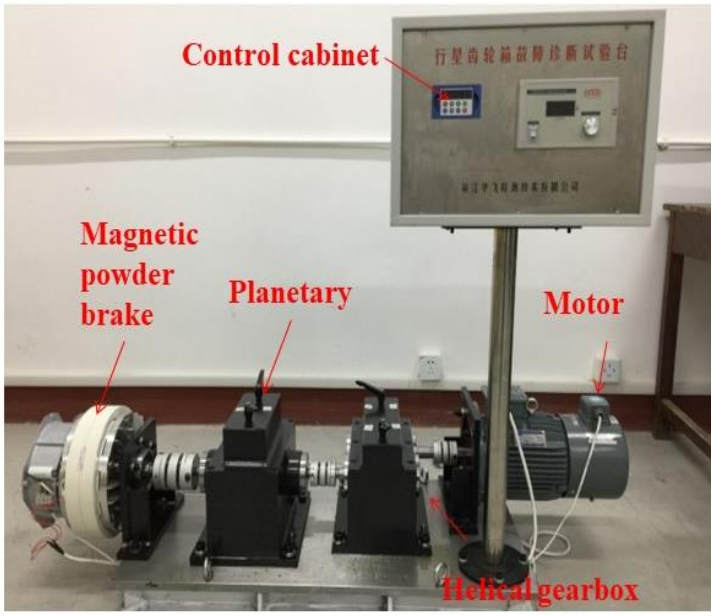
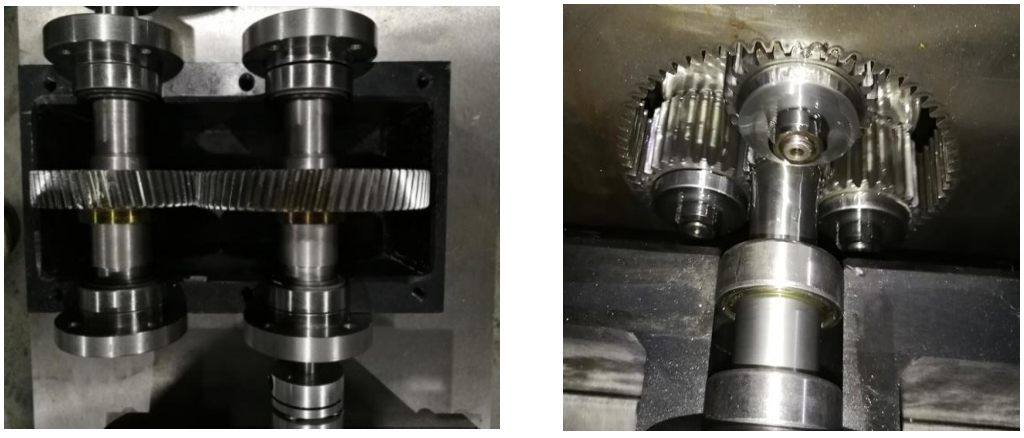


Figure 7. The planetary gear test bed.



(**a**) (**b**)

Figure 8. Structural diagrams of the transmission system of the experimental platform: (**a**) helical gear transmission; (**b**) planetary gear transmission.

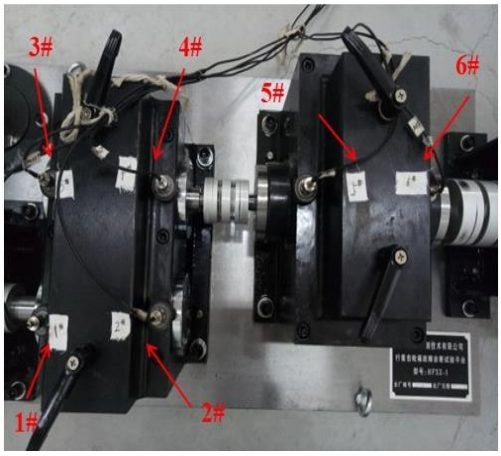


Figure 9. Arrangement of measuring points.

4.2. Analysis of Vibration Signal

The vibration signals for each measuring point on the box body were measured. According to the experimental scheme, the vibration signals for the normal state and three kinds of planetary gear tooth surface wear states were mainly measured and analyzed, and the characteristic parameters were extracted. They include 21 time-domain features, such as the mean value, mean square value, maximum value, minimum value, variance, root-mean-square value, root amplitude, absolute average amplitude, skewness, kurtosis, peak, and six-order moments, and six frequency-domain features, namely, the frequency domain variance, correlation factor, power spectrum barycenter index, mean square spectrum, harmonic factor, and origin moment of spectrum. Sixty groups of training samples and sixty groups of test samples were extracted. After their standardization, the KPCA dual-kernel parameter optimization and feature extraction were carried out.

5. Feature Extraction of the Planetary Gearbox Based on the KPCA Dual-Kernel Parameter Optimized by the SFLA-PSO

5.1. Optimization of the Dual-Kernel Parameters of KPCA Based on the SFLA-PSO

In order to improve the mapping performance of KPCA, the SFLA-PSO fusion algorithm was used to optimize the parameters of the dual-kernel function for the data sets for the normal state, one-tooth wear, two-tooth wear, and three-tooth wear of the planetary gear (referred to as models A, B, C, and D, respectively). The parameters of the SFLA-PSO fusion algorithm are shown in Table 3, and the optimization was carried out after substituting the sample data. The optimization results for the kernel parameters of the different wear modes are shown in Table 4.

Table 3. The parameters of the SFLA-PSO fusion algorithm.

Algorithm	Parameters
SFLA-PSO	$N = 200$, $n_{PSO} = 20$, $n_P = 10$, $c_1 = c_2 = 2$, $F = 20$, $m = 5$, $n = 4$, $L_{max} = 10$, $T_{max} = 100$, $d = 20$, $S_{max} = 20$

Table 4. The optimization results of dual-kernel parameters of planetary gear wear state.

Type	Model	KPCA Kernel Parameter			J_{Fnsher}
		γ^*	d_1^*	σ^*	
3-type	ABC	0.005	0.8	23.37	0.3497
	BCD	0.038	0.896	5.4438	0.8970
4-type	ABCD	0.055	1.03	13.1	0.9949

The evolutionary optimization process is shown in Figure 10a,b, from which it can be seen that the fitness function J_{Fisher} can reach the minimum value and obtain the optimal parameter in 100 iteration steps. It can be seen that different flexible weight coefficients γ^* and kernel parameters d_1^* and σ^* should be adopted when DKKPCA mapping is applied to identify the different wear models.

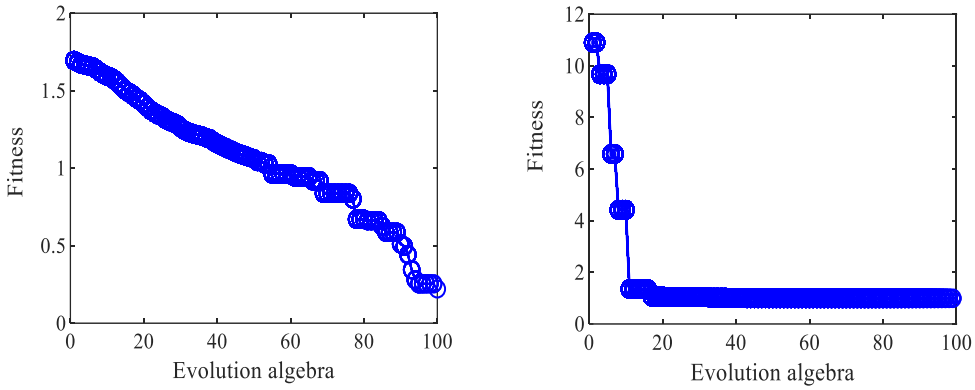


Figure 10. Evolution curves of the kernel parameter optimization process: (**a**) ABC model; (**b**) ABCD model.

5.2. KPCA Feature Extraction of the Planetary Gear Wear Based on the Dual-Kernel Optimization

Based on the test results of the normal state and the planetary gear wear in the experiment, the kernel parameters optimized by the SFLA-PSO were substituted into the DKKPCA to analyze the characteristics of the different wear models of the planetary gear. The analysis flow is shown in Figure 11.

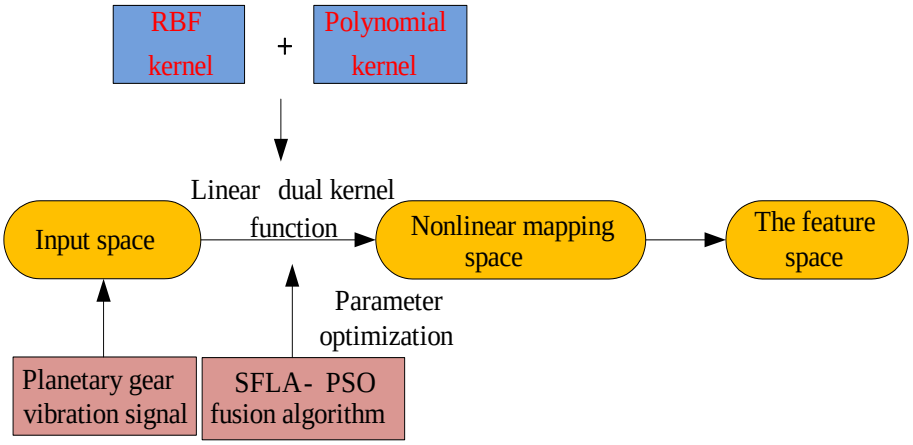


Figure 11. A feature extraction flow diagram for dual-kernel optimization.

The KPCA analysis results are shown in Figure 12; Figure 12a_1–c_1 show the scatter diagrams for the KPCA principal component before the optimization of the RBF kernel parameters. The kernel parameter σ was empirically set as follows: ABC: σ = 10.25; BCD: σ = 5.4438; ABCD: σ = 23.1. Figure 13a_2–c_2 show those for the DKKPCA principal component before the dual-kernel parameter optimization, for which the three parameters were set according to experience: ABC: γ = 0.07, σ = 23.37, and d_1 = 1.2; BCD: γ = 0.009, d_1 = 1, and σ = 5.4438; ABCD: γ = 0.2, d_1 = 2, and σ = 10. Figure 14a_3–c_3 show those for the DKKPCA principal component after the dual-kernel parameter optimization, and the three parameters are the optimization results γ^*, d_1^*, and σ^*.

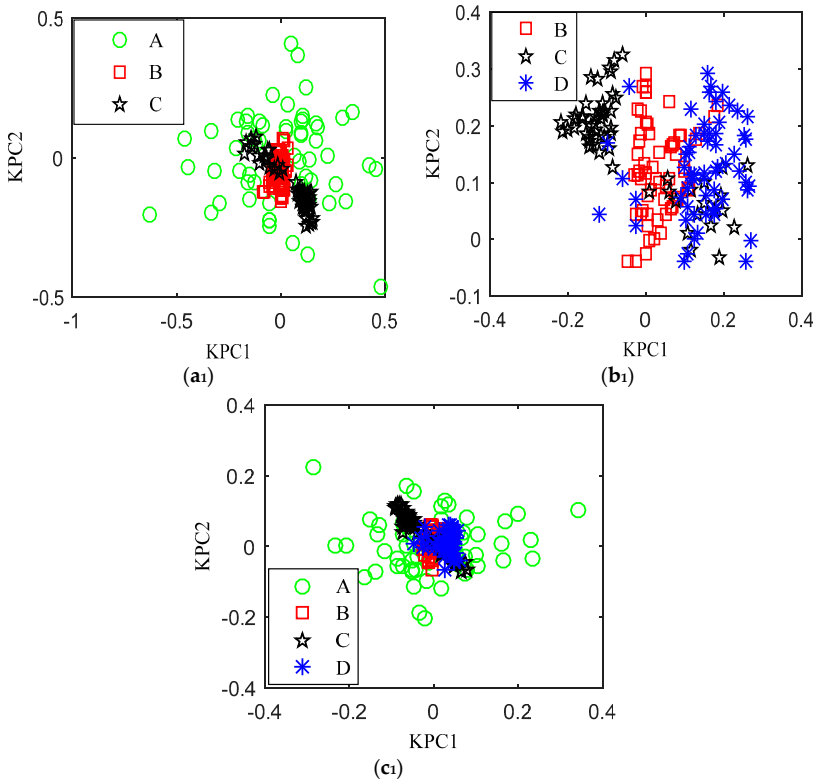


Figure 12. KPCA scatter diagram of planetary faults before single-kernel parameter optimization: A: normal state, B: one-tooth wear, C: two-tooth wear, D: three-tooth wear. (**a1**) ABC: σ = 10.25; (**b1**) BCD: σ = 5.4438; (**c1**) ABCD: σ = 23.1.

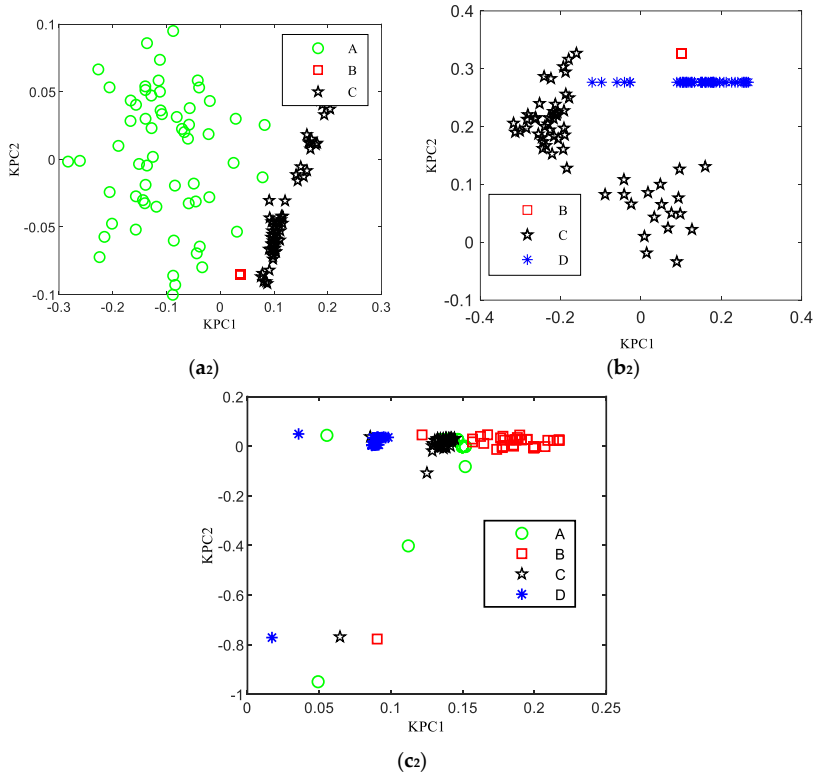


Figure 13. DKKPCA scatter diagram of planetary faults before dual-kernel parameter optimization: (**a2**) γ = 0.07, d = 1.2, and σ = 23.37; (**b2**) γ = 0.009, d_1 = 1, and σ = 5.443; (**c2**) γ = 0.2, d = 2, and σ = 10.

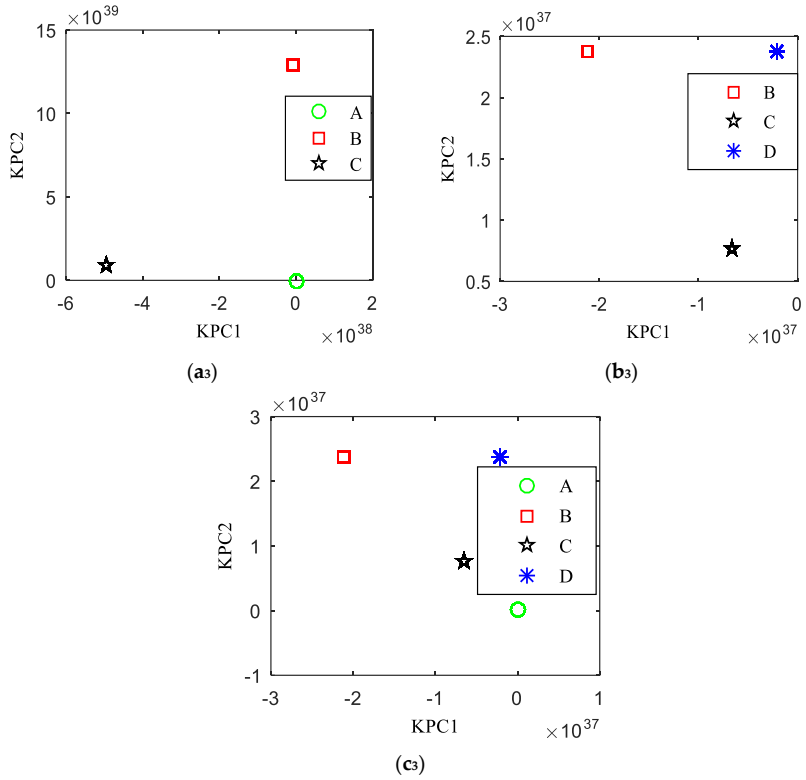


Figure 14. DKKPCA scatter diagram of planetary gear faults after dual-kernel parameter optimization: (**a3**) γ = 0.005, d_1 = 0.8, and σ = 23.37; (**b3**) γ = 0.038, d1 = 0.896, and σ = 5.4438; (**c3**) γ = 0.055, d_1 = 1.03, and σ = 13.1.

5.3. The Analysis of the Obtained Results

As can be seen from Figure 12a_1–c_1, the data points for the normal model are relatively scattered and interspersed with the three models of teeth wear. Although B, C, and D have their own clustering centers, some of the samples with the three models of wear are interwoven, and there is no obvious boundary line to accurately distinguish the wear models. Thus, it can be seen from Figure 13a_2–c_2 that the DKKPCA greatly improved the wear recognition of several planetary gears via the dual-kernel function, but some features of ABCD still intersected and were not completely distinguished, especially for the normal state (model A) and the slight wear (model B) of one tooth. Therefore, the KPCA scatter diagrams' comparison of the planetary faults using the single-kernel KPCA and DKKPCA indicate that the DKKPCA has a better nonlinear mapping performance than the single-kernel KPCA. The DKKPCA is suitable for intricate processes and nonlinear dimensionality reduction.

It can be seen from Figure 14a_3–c_3 that regardless of whether three or four models are employed, they can be highly clustered to a point after DKKPCA mapping. There is no gap in any category of B, C, and D for the three different wear models, the class spacing is large, and the categories are clear and divisible. This is because the interference from the small sample features of the contribution rate is eliminated after high-dimensional nonlinear mapping. It can be seen that when the dual-kernel function after a flexible weight linear combination is adopted for the DKKPCA, the kernel parameters are optimized to avoid the blindness of parameter selection, and the accuracy of the fault feature extraction and identification are significantly improved.

The KPCA method proposed in this paper was applied to a gearbox (JZQ-250 type) [30], for which the simulation failures included normal working conditions, a bearing outer-ring crack in the intermediate shaft, a broken tooth in the bearing cage fracture gear, and a broken tooth in the gear combined with a fracture of the bearing outer ring. The analysis

results of the gearbox fault diagnosis demonstrate its effectiveness. So, the method has been verified for other conditions of other gearboxes.

6. Conclusions

KPCA is a nonlinear dimensionality reduction technique that maps data onto a high-dimensional feature space using kernel functions, enabling the detection of faults in complex and nonlinear systems. Its advantages lie in its ability to handle nonlinearity, making it suitable for intricate processes. However, KPCA's performance heavily relies on the appropriate choice of kernel function and its associated parameters, which can be challenging to determine in practice.

Aiming at the disadvantages of the KPCA single-kernel function and the advantages of the multi-kernel method, this paper proposes a KPCA with a flexible weight linear combination of the dual-kernel function (DKKPCA). In order to improve KPCA's performance and solve the optimization problem of dual-kernel parameters, an optimization model of dual-kernel parameters was constructed by referring to Fisher's criterion and defining the optimization variables and objectives. Then, taking full advantage of the SFLA-PSO fusion algorithm, which is simple, and a global search, the kernel parameters were optimized. A comprehensive comparison test was carried out using the Iris data to verify the application effect of the DKKPCA feature extraction method, which could be separated well after nonlinear mapping onto the feature space, especially for classes 2 and 3. The simulation results indicate that the DKKPCA optimized by the intelligent fusion algorithm is not only suitable for solving the nonlinear feature extraction problem, but can also provide better feature quality than the linear dimension reduction method, and greatly enhances the ability of nonlinear data processing.

A planetary gear simulation fault diagnosis experiment was conducted, and the vibration signals of the normal state and three kinds of planetary gear tooth surface wear states were mainly measured and analyzed. Then, the proposed DKKPCA method was applied to the feature extraction of multi-fault-mode coupled vibration signals of a planetary gear. The KPCA scatter diagrams for the planetary faults using the single-kernel KPCA and dual-kernel KPCA (DKKPCA) were compared and analyzed, and the results show that the DKKPCA has a better nonlinear mapping performance than the single-kernel KPCA. Similarly, a comparative analysis of the DKKPCA before and after the dual-kernel parameter optimization was completed shows that the scattering can be highly clustered to a point, and the wear damage state of a planetary gear can be accurately identified using the DKKPCA after parameter optimization; so, it has better nonlinear mapping performance than before parameter optimization. Therefore, the DKKPCA method has been adapted to feature extraction and state recognition for the nonlinear behavior of other mechanical equipment.

Author Contributions: Writing—original draft preparation and writing—review and editing, Y.H.; visualization and software, L.Y.; writing—review and editing, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Young Science Foundation of Shanxi province, China, grant number 201901D211201; the National Natural Science Foundation of China, grant numbers 52375470 and 52005455; and the Central Guidance on Local Science and Technology Development Fund of Shanxi Province, grant number YDZJSX2022C005.

Data Availability Statement: The data used to support the findings of this study are included within the article.

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

References

1. Pan, H.; Zheng, J.; Yang, Y.; Cheng, J. Nonlinear sparse mode decomposition and its application in planetary gearbox fault diagnosis. *Mech. Mach. Theory* **2021**, *155*, 104082. [CrossRef]
2. Wang, C.; Li, H.; Ou, J.; Hu, R.; Hu, S.; Liu, A. Identification of planetary gearbox weak compound fault based on parallel dual-parameter optimized resonance sparse decomposition and improved momeda. *Measurement* **2020**, *165*, 108079. [CrossRef]
3. Schölkof, B.S.; Muller, K.R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.* **1998**, *10*, 1299–1319. [CrossRef]
4. Xiao, Y.; Feng, L. A novel neural-network approach of analog fault diagnosis based on kernel discriminant analysis and particle swarm optimization. *Appl. Soft Comput.* **2012**, *12*, 904–920. [CrossRef]
5. Fatma, L.; Lotfi, M.; Wiem, A.; Hedi, D.; Khaled, Y. Investigating Machine Learning and Control Theory Approaches for Process Fault Detection: A Comparative Study of KPCA and the Observer-Based Method. *Sensors* **2023**, *23*, 6899.
6. Wang, Y.; Dong, R.; Wang, X.; Zhang, X. Research on Rolling Bearing Fault Diagnosis Based on Volterra Kernel Identification and KPCA. *Shock. Vib.* **2023**, *2023*, 1–9. [CrossRef]
7. Shao, R.; Hu, W.; Wang, Y.; Qi, X. The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform. *Measurement* **2014**, *54*, 118–132. [CrossRef]
8. Lei, Y.; Chen, W.; Li, N.; Lin, J. A relevance vector machine prediction method based on adaptive multi-kernel dual and its application to remaining useful life prediction of machinery. *J. Mech. Eng.* **2016**, *52*, 87–93. [CrossRef]
9. Fu, H.; Ren, R.; Yan, Z.; Ma, Y. Fault diagnosis method of power transformers using multi-kernel RVM and QPSO. *High Volt. Appar.* **2017**, *53*, 131–135.
10. Deng, X.; Lei, W. Modified kernel principal component analysis using dual-weighted local outlier factor and its application ton on linear process monitoring. *ISA Trans.* **2018**, *72*, 218–228. [CrossRef]
11. Pan, C.; Li, H.; Chen, B.; Zhou, M. Fault Diagnosis method with class mean kernel principal component analysis based on combined kernel function. *Comput. Simul.* **2019**, *36*, 414–419.
12. Wang, H.; Cai, Y.; Sun, F.; Zhao, Z. Adaptive sequence learning and application of multi-scale kernel method. *Pattern Recognit. Artif. Intell.* **2011**, *24*, 72–81.
13. Lv, L.; Wang, W.; Zhang, Z.; Liu, X. A novel intrusion detection system based on an optimal hybrid kernel extreme learning machine. *Knowl.-Based Syst.* **2020**, *195*, 105648. [CrossRef]
14. Nithya, A.; Appathurai, A.; Venkatadri, N.; Ramji, D.R.; Anna, P.C. Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images. *Measurement* **2020**, *149*, 106952. [CrossRef]
15. Ouyang, A.; Liu, Y.; Pei, S.; Peng, X.; He, M.; Wang, Q. A hybrid improved kernel LDA and PNN algorithm for efficient face recognition. *Neurocomputing* **2020**, *393*, 214–222. [CrossRef]
16. Afzal, A.L.; Asharaf, S. Deep multiple multilayer kernel learning in core vector machines. *Expert Syst. Appl.* **2018**, *96*, 149–156.
17. Li, X.; Gao, X.; Li, K.; Hou, Y. Prediction for dynamic fluid Level of oil well based on GPR with AFSA optimized combined kernel function. *J. Northeast. Univ. (Nat. Sci.)* **2017**, *38*, 11–15.
18. Xie, F.; Chen, H.; Xie, S.; Jiang, W.; Liu, B.; Li, X. Bearing state recognition based on kernel principal component analysis of particle swarm optimization. *Meas. Control Technol.* **2018**, *37*, 28–35.
19. Bernal, D.; Lázaro, J.M.; Prieto, M.A.; Llanes, S.O.; da Silva, N.A.J. Optimizing kernel methods to reduce dimensionality in fault diagnosis of industrial systems. *Comput. Ind. Eng.* **2015**, *87*, 140–149. [CrossRef]
20. Kennedy, J.; Eberhart, R. Particle Swarm Optimization. In Proceedings of the IEEE International Conference on Neural Networks; IEEE Computer Society: Washington, DC, USA, 1995; pp. 1942–1948.
21. Eusuff, M.M.; Lansey, K.E. Optimization of water distribution network design using shuffled frog leaping algorithm. *J. Water Resour. Plan. Manag.* **2003**, *129*, 210–225. [CrossRef]
22. Jaafari, A.; Zenner, E.K.; Panahi, M.; Shahabi, H. Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wild fire probability. *Agric. For. Meteorol.* **2019**, *266–267*, 198–207. [CrossRef]
23. Wang, Y.; Sun, W.; Liu, L.; Wang, B.; Bao, S.; Jiang, R. Fault Diagnosis of Wind Turbine Planetary Gear Based on a Digital Twin. *Appl. Sci.* **2023**, *13*, 4776. [CrossRef]
24. Kennedy, J.; Eberhart, R.C. *Swarm Intelligence*; Morgan Kaufmann Division of Academic Press: San Francisco, CA, USA, 2001; pp. 11–15.
25. Rajamohana, S.P.; Umamaheswari, K. Hybrid approach of improved binary particle swarm optimization and shuffled frog leaping for feature selection. *Comput. Electr. Eng.* **2018**, *67*, 497–508. [CrossRef]
26. Wang, H.; Sun, F.; Cai, Y.; Chen, N.; Ding, L. Multi-kernel learning method. *Acta Autom. Sin.* **2010**, *36*, 1037–1050. [CrossRef]
27. Li, J.; Qiao, J.; Yin, H.; Liu, D. *Kernel Adaptive Learning and Application in Pattern Recognition*; Publishing House of Electronics Industry: Beijing, China, 2013; pp. 23–25.
28. Chen, W.; Panahi, M.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Panahi, S.; Li, S.; Jaafari, A.; Bin Ahmad, B. Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility. *Catena* **2019**, *172*, 212–231. [CrossRef]

29. Zhao, Y. *Pattern Recognition*; Shanghai Jiao Tong University: Shanghai, China, 2013; pp. 12–16.
30. He, Y.; Wang, Z. Regularized kernel function parameter of kpca using WPSO-FDA for feature extraction and fault recognition of gearbox. *J. Vibroengineering* **2018**, *20*, 225–239. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.