Fault Diagnosis of Roller Bearings Based on a Wavelet Neural Network and Manifold Learning

Lifeng Wu [1,2,3], Beibei Yao [1,2,3], Zhen Peng [4] and Yong Guan [1,2,3,*]

[1]   College of Information Engineering, Capital Normal University, Beijing 100048, China; wulifeng@cnu.edu.cn (L.W.); linhongyao897@163.com (B.Y.)
[2]   Beijing Engineering Research Center of Highly Reliable Embedded Systems, Capital Normal University, Beijing 100048, China
[3]   Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing 100048, China
[4]   Information Management Department, Beijing Institute of Petrochemical Technology, Beijing 102617, China; zhenpeng@bipt.edu.cn
*   Correspondence: guanyong@cnu.edu.cn; Tel.: +86-10-6890-6706

Abstract: In order to improve the accuracy of the fault diagnosis of roller bearings, this paper proposes a kind of fault diagnosis algorithm based on manifold learning combined with a wavelet neural network. First, a high-dimensional feature signal set is obtained using a conventional feature extraction algorithm; second, an improved Laplacian characteristic mapping algorithm is proposed to reduce the dimensions of the characteristics and obtain an effective characteristic signal. Finally, the processed characteristic signal is inputted into the constructed wavelet neural network whose output is the types of fault. In the actual experiment of recognizing data sets on roller bearing failures, the validity and accuracy of the method for diagnosing faults was verified.

Keywords: roller bearing; manifold learning; wavelet neural network; fault diagnosis

1. Introduction

Roller bearings are critical components of machinery and directly affect the performance of the whole system [1]. However, bearings inevitably degenerate and break down due to an overload of repetitive work. According to statistics, 30 percent of faults in rolling machinery originate from bearing faults; therefore, bearing fault prognostics is an important area of research [2–5]. Recently, studies on bearing faults prognostics have focused on fault signal feature extraction and fault classification. Commonly used methods for extracting the features of fault signals are Empirical Mode Decomposition (EMD), morphology, wavelet transform, signal value decomposition, and principal component analysis [6–10]. However, features extracted using the above methods are largely redundant, as bearing fault signals are unstable and nonlinear [11]. Manifold learning, such as Isomap [12], Locally Linear Embedding (LLE) [13], and Laplacian Eigenmap (LE) [14] can be used to extract effective features. Bearing vibration signals produce a great deal of noise and the first two methods referred to are very sensitive to noise. LE has a powerful local keeping ability and is not sensitive to noise [15]. Good results can be obtained with this method to extract features from roller bearings and reduce dimensionality. The local field selection of manifold learning is the key step and greatly affects performance [16]. Therefore, we propose an adaptive local field selection system based on Lalacian mapping to extract features of bearing fault signals. Classifications of fault signals and information fusion approaches such as firefly neural networks [17], recursive complex networks [18], and fuzzy C mean clustering [19] are commonly used. Neural networks have a great ability to fit nonlinear data

and recognize bearing faults. The proposed wavelet neural network with expansion and translation obtains better freedom and has a greater flexible effective function approximation ability than other conventional neural networks. It also learns more quickly and can learn from mutation function and discontinuous signal details.

From the above analysis, it is clear that LE is suitable for use in the nonlinear descending dimension method. The performance of LE depends greatly on the neighborhoods k. Generally, k is given according to experience. In this paper, we propose an improved LE that can be changed adaptively for any given k and obtain good results. We can also attain good characteristic signals using this method. Additionally, we propose a fault diagnosis method based on a wavelet neural network model where the merits of wavelet transform are integrated with that of an artificial neural network. The experimental results prove the effectiveness of the proposed methods.

2. Prognostics Method Based on Manifold Learning and a Wavelet Neural Network

A flow chart of the proposed prognostics method based on manifold learning and a wavelet neural network is shown in Figure 1. First, common feature extraction methods were used in the time domain signals to extract features to construct high-dimensional feature signal sets. Next, methods based on manifold learning were adopted to reduce the dimensionality and obtain effective feature parameters. Lastly, wavelet neural networks were set up to take feature parameters as inputs and fault types as outputs, and train networks until the required errors were reached and the prognostics tested.
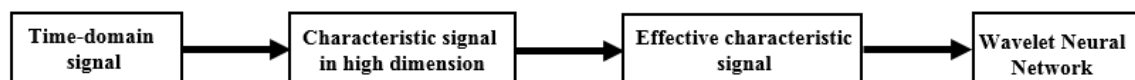

Figure 1. The fault diagnosis method based on a manifold learning and a neural network.

High-Dimensional Feature Signal Extraction

The fault prognostics of bearing faults can be considered as a map between two spaces where one is the fault feature set and the other is the state of faults. It is crucial to accurately find the corresponding fault features, i.e., feature extraction. Too many or too few features will produce difficulties in prognostics. In order to obtain fault features from nonlinear and unstable roller bearing vibration signals, traditional methods such as the Hilbert Huang transform, Fast Fourier Transformation (FFT), signal envelope, Empirical Mode Decomposition (EMD) and wavelet transform were taken to extract features from the monitoring time domain vibration signals. It is obvious that the feature signals extracted by these methods are redundant, nonlinear, and uncertain. Therefore, manifold learning is utilized to reduce dimensionality. Improved Laplacian Eigenmaps (ILE) are proposed when considering the local field selection of manifold learning.

3. Materials and Methods

3.1. ILE Algorithm

The idea of graph theory is introduced with the LE algorithm in this paper. By constructing weighted neighborhood graphs, the relation of the same positions of sample points in local neighborhoods is maintained in high and low dimensional space [14]. The neighborhoods k in LE is the key parameter in reducing dimensions. In view of the different characteristics of different manifold surfaces, it is easy to include the traditional fixed k value with the points in the same neighborhood graph. This results in the "short circuit phenomenon", as these points do not belong to the same manifold area. To solve this problem, an improved Laplace Eigen-mapping method dimensionality reduction algorithm is proposed and is shown in Figure 2. Detailed step-by-step instructions are as follows:

Step 1: Construct weighted neighborhood graphs.

From the *k*-nearest neighbors of the high-dimensional spatial data point X_i, the adjacency matrix of the data set is obtained.

1. Set the initial neighborhood value *k*, according to the principle of KNN for each of the sample points to determine the neighborhood, $\left\{ X_i^1,\ X_i^2,\ X_i^3 \ldots X_i^k \right\}$, $i = 1,\ 2,\ 3 \ldots N$;

2. For each of the resulting neighborhoods, we can calculate the tangent space coordinates that correspond to the neighborhood points [20].

$$\Phi_i^{(j)} = V^T N_i \left(I - \frac{1}{k} ee^T \right) \tag{1}$$

where Mapping function of neighborhood points to local tangent space coordinates can be expressed as a Equation (2) based on Taylor series expansion.

$$\| x_{ij} - \overline{x}_i - V_i V_i^T N_i (I - \frac{1}{k} ee^T) \|_2 \leq \lambda \| V_i V_i^T N_i (I - \frac{1}{k} ee^T) \|_2 \tag{2}$$

where $\overline{x}_i = \frac{1}{k} \sum_{i-1}^{k} x_{ij}$. λ is the scale factor. Θ is defined as the threshold. When $\lambda < \Theta$, the local neighborhood obtained from the *k*-value is considered to be linearly represented by the tangent plane, which is in accordance with the characteristics of the ideal neighborhood, $\Theta \in (0,\ 0.1)$.

3. Calculate the difference function of a high and low dimension distribution.

The difference function of the high and low dimension distribution of all points in the *i*-th neighborhood is defined as follows [21]:

$$C_i = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{3}$$

where p_{ij} is the distribution function of the *j*-th neighborhood point of the *i*-th high-dimensional sample point; and q_{ij} is the distribution function of the Euclidean distance of the low-dimensional neighborhood point. The *k*-value adjustment strategy is shown in Equation (4):

$$k_i^{new} = \begin{cases} (\frac{1}{n} \sum\limits_{i=1}^{n} k_i - min\{k\}) e^{-\frac{(C_i - \overline{C})^2}{2}} + k_i & \frac{1}{n} \sum\limits_{i=1}^{n} k_i > k_i \\ -(\frac{1}{n} \sum\limits_{i=1}^{n} k_i - min\{k\})(1 - e^{-\frac{(C_i - \overline{C})^2}{2}}) + k_i & \frac{1}{n} \sum\limits_{i=1}^{n} k_i < k_i \end{cases} \tag{4}$$

where $\overline{C} = \frac{1}{n} \sum_{i-1}^{n} C_i$

4. Assume that the neighborhood weights of each sample point is $W_i^k = \left\{ W_i^1,\ W_i^2,\ W_i^3 \ldots W_i^k \right\}$. Calculate the weight using the following equation:

$$W_i^c = max\left(0,\ min\left(1, W_i^c (1 + \gamma \Psi_{i,C})\right)\right) \tag{5}$$

The initial value of weight is one by default. γ is the adjustment factor, $\Psi_{i,C}$ is the update operator, and the small weight points will be removed as invalid neighborhood points while the remaining ones will be reserved as effective neighborhood points.

Step 2: Determine the weight of the neighborhood graph edge.

Determine the weight between the points and use the thermal kernel function to determine the weight size. If the two nodes are adjacent, the weight is one, otherwise the weight is zero.

Step 3: Calculate the feature map.

Compute the eigenvectors and eigenvalues of the Laplace matrix L:

$$Ly = \lambda Dy \tag{6}$$

where D is a diagonal matrix, the eigenvectors are outputted as the result of the reduced dimension that corresponds to the smallest m nonzero eigenvalue.
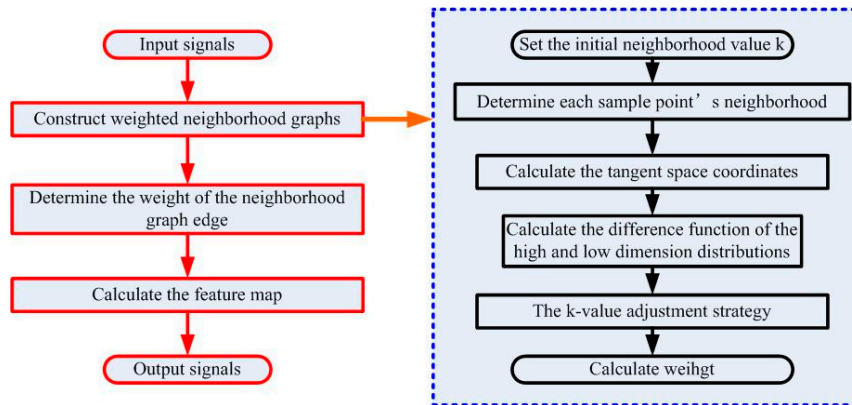

Figure 2. Diagram of the steps in the method.

3.2. Wavelet Neural Network

The wavelet neural network (WNN) is a forward network, which uses the wavelet basis function as the neuron activation function [22]. It adaptively adjusts the scaling and translation factors of the wavelet function, network connection weights, and the approximation function in batch mode. The fault diagnosis of the bearing based on the wavelet neural network is shown in Figure 3.
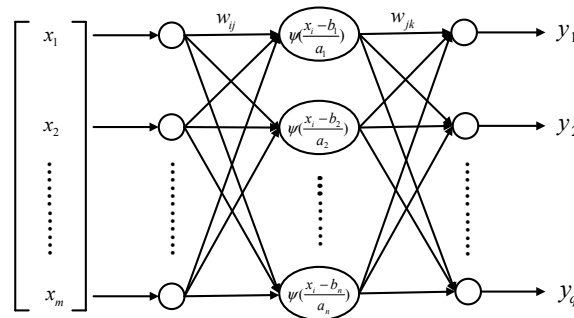

Figure 3. Fault diagnosis model based on the wavelet neural network.

where $\{x_1, x_2, \ldots, x_N\}$ is the bearing fault feature, which can botain by the ILE method proposed. $\{y_1 y_2 y_3 y_4\}$ is the bearing state ,where y_1 is the normal state; y_2 the ball failure; y_3 the inner ring failure; and y_4 is the outer ring failure four. "1" is in this state, while "0" is not in this state.

Ψ is the wavelet base; its Fourier transform $\hat{\Psi}(\omega)$ satisfies the admissible condition:

$$C_\Psi(\omega) = \int_{-\infty}^{+\infty} |\omega|^{-1} |\hat{\Psi}(\omega)|^2 d\omega < \infty \tag{7}$$

By scale scaling and translating for wavelet base functions $\Psi(x)$, the family of wavelet functions is obtained:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \Psi\left(\frac{t-b}{a}\right) \tag{8}$$

where $a, b \in R$, $a \neq 0$. a and b are scaled translation scale factors, respectively. The wavelet transform is:

$$(W_\Psi f)(a, b) = \int_R f(t) \Psi^*_{a,b}(t) dt \tag{9}$$

Discretizing a and b, when $a = a_0^m$, $b = nb_0 a_0^m$, $n, m \in Z$, then:

$$\Psi_{n,m}(t) = |a_0|^{-\frac{m}{2}} \Psi(a_0^{-m} t - nb_0), \ n, m \in Z \tag{10}$$

Equation (10) is the discrete wavelet transform. The discrete wavelet transform is:

$$(W_\Psi f)(a, b) = |a_0|^{-\frac{m}{2}} \int_{-\infty}^{+\infty} f(t) \Psi^*(a_0^{-m} t - nb_0) dt \tag{11}$$

Select the scaling parameters $a > 0$, $b > 0$ to ensure that the wavelet system $\{ \Psi_{n,m}(t) \mid (n, m) \in Z^2 \}$ satisfies the $L^2(R)$ frame condition:

$$A\|f\|^2 \leq \sum_{(m,n) \in Z^2} |\langle f, \Psi_{n,m} \rangle|^2 \leq B\|f\|^2 \tag{12}$$

In the formula, A, B are the frame boundaries. If $\widetilde{\Psi}_{n,m}$, $\Psi_{n,m}(t)$ are the dual function, then for any function $f(t) \in L^2(R)$, $f(t)$ can be expressed as a form of wavelet series:

$$f(t) = \sum_{(m,n) \in Z^2} \langle f, \widetilde{\Psi}_{n,m} \rangle \Psi_{n,m}(t) \tag{13}$$

The function can be expanded by a wavelet basis function and an approximate function:

$$f(t) = \sum_{j \in Z} w_j \frac{1}{a_j} \Psi\left(\frac{t - b_j}{a_j}\right) \tag{14}$$

The above equation can be realized by a neural network with a hidden layer, and by adjusting the weights a_j and b_j to an approximate function. In the formula, $\Psi_i(t)$ is the wavelet basis function $(i = 1, 2, \ldots, n)$; $y_k (k = 1, 2, \ldots, n)$ is the output of the network; and w_{ij}, w_{jk} is the connection weight between the input layer/the hidden layer and between the hidden layer/the output layer.

The wavelet basis function $\Psi_i(t)$ chose the Mexican Hat wavelet as follows:

$$\Psi(x) = \frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} (1 - x^2) e^{-\frac{x^2}{2}} \tag{15}$$

Take the energy function:

$$J = \frac{1}{2} \sum_{j-1}^{p} \sum_{i-1}^{q} \left(y_i^j - \hat{y}_i^j\right)^2 \tag{16}$$

In this formula, P is the number of samples, y is the actual value, and $\hat{y}$ is the output of WNN. The network output layer neuron function takes the Sigmoid function, that is $\varnothing(x) = 1/(1 + e^{-x})$. Then, the expression of the network output y is:

$$y_k = \varnothing\left(\sum_{k-1}^{q} w_{jk} \Psi_{a,b}\left(\left(\sum_{i-1}^{m} w_{ij} x_i - b_j\right)/a_j\right)\right) \tag{17}$$

$j = 1, 2, \ldots, n$, $k = 1, 2, \ldots, q$.

By using gradient descent method:

$$h(t) = h(t - 1) - \eta \frac{\partial J}{\partial h} + \alpha \Delta[h(t) - h(t - 1)] \tag{18}$$

In the formula, h shows that w_{ij}, w_{jk}, a_j, b_j, η is the learning rate, α is the momentum factor, and $\eta, \alpha \in (0, 1)$. Equation (16) is minimized by adjusting the parameters of the WNN.

4. Application

4.1. Experimental Subject

The bearing data was generated by the Industry/University Cooperative Research Centers Program (I/UCRC) for Intelligent Maintenance Systems. The test platform is shown in Figure 4. The rotation speed was kept constant at 2000 RPM by an alternating current motor and at a radial load of 6000 lbs. Data collection was facilitated by National Instruments Data Acquisition (NI DAQ) Card 6062E, with the sampling rate set at 20 kHz [23].
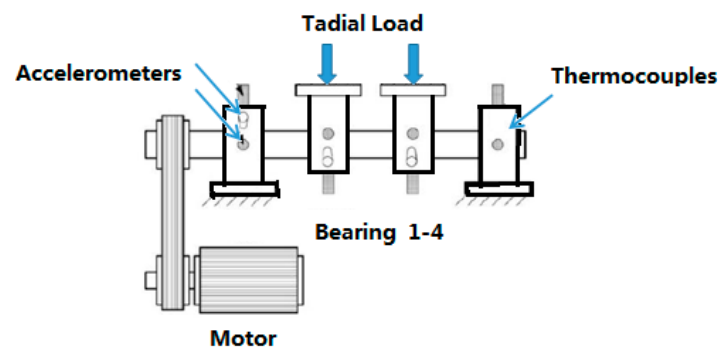

Figure 4. Illustration of the bearing test rig and sensor placement [23].

4.2. Roller Bearing Vibration Signals

The test was carried out for 35 days. Figure 5a–d shows the vibration signals for the normal state, a roller element failure, inner race failure, and outer race failure, respectively [23].
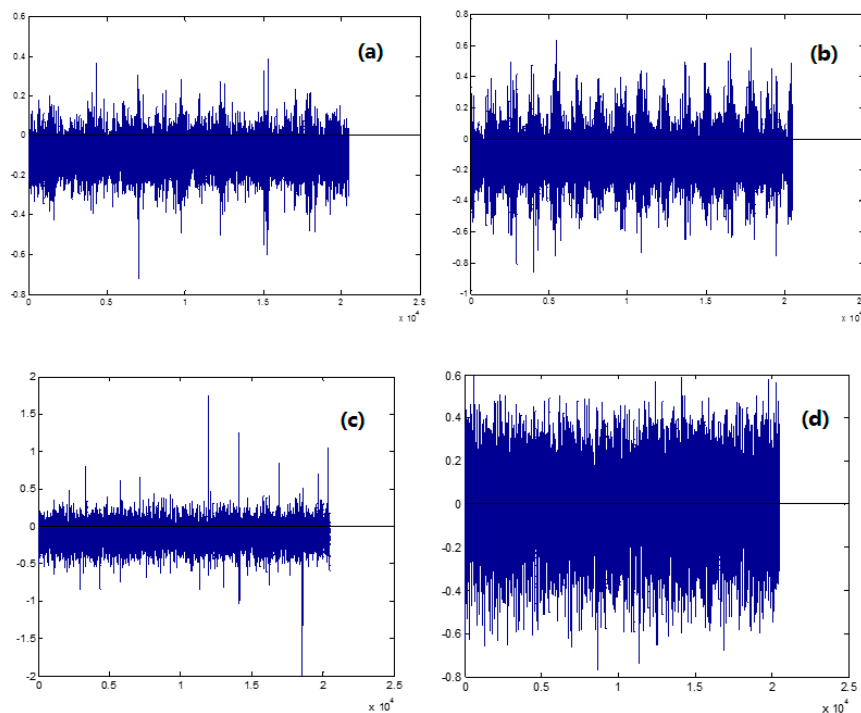

Figure 5. Vibration signal. (**a**) Normal state; (**b**) roller element failure; (**c**) inner race failure; (**d**) outer race failure.

At first, features were obtained using conventional methods as much as possible. Next, we reduce dimensions using the ILE method mentioned above. The result of the dimension reduction is shown in Table 1 where (1) is normal; (2) is a roller element failure; (3) is the inner race failure; and (4) is an outer race failure. All other characteristic signals found in all types of situations are also shown in Table 1.

Table 1. Fault data set.

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Feature 7	Feature 8	State
0.089374	0.047321	0.100898	1.958858	1.606291	1.813402	0.10817	6.301499	1
0.1036	0.04787	0.113916	1.665516	1.419555	1.560898	0.133881	11.49973	1
0.095195	0.046492	0.105729	1.817757	1.529219	1.698434	0.038316	8.864674	1
0.101547	0.047333	0.111828	1.703315	1.449659	1.596429	0.159635	10.19821	1
0.087449	0.047526	0.099293	1.996853	1.625361	1.84549	0.066327	6.246746	1
0.100355	0.046046	0.110214	1.708943	1.463397	1.607167	0.071301	10.474	1
0.097557	0.045718	0.107536	1.760794	1.500416	1.653888	0.157312	8.822157	1
0.088979	0.046411	0.100132	1.941436	1.603238	1.804188	0.037097	6.599557	1
0.103291	0.046228	0.112967	1.648921	1.420006	1.553027	0.116062	11.54525	1
0.091338	0.044016	0.101192	1.841057	1.559724	1.727983	0.124147	6.76533	1
0.114011	0.012182	0.114646	1.161855	1.152177	1.158599	0.102057	17.33023	2
0.114794	0.013028	0.115515	1.162091	1.151164	1.158401	0.132316	17.42503	2
0.11468	0.014126	0.115529	1.185616	1.172517	1.181195	0.097094	17.81285	2
0.114637	0.015178	0.115617	1.185164	1.170044	1.180042	0.125409	17.41817	2
0.114804	0.015979	0.115888	1.197769	1.180885	1.192032	0.116505	17.63886	2
0.113608	0.016515	0.114778	1.219157	1.200454	1.21281	0.099188	17.11847	2
0.11486	0.017319	0.116131	1.217056	1.196951	1.210202	0.137996	17.39344	2
0.113765	0.018268	0.115192	1.233035	1.210016	1.225195	0.091749	17.30835	2
0.113877	0.018827	0.11539	1.233015	1.208609	1.224676	0.126345	16.93885	2
0.114127	0.018397	0.115569	1.227595	1.204454	1.219682	0.1175	17.20449	2
0.002049	0.066937	0.066268	2.726322	1.851256	2.266006	0.001127	0.669745	3
0.004122	0.050691	0.05033	3.423602	2.139655	2.778715	0.107688	0.003864	3
0.00479	0.074184	0.073564	2.079078	1.654904	1.879734	0.048673	1.159074	3
0.008259	0.041219	0.041615	3.851942	2.399057	3.137867	0.099468	0.005723	3
0.002269	0.075672	0.074914	1.906661	1.55103	1.733169	0.060749	1.273057	3
0.015029	0.038022	0.040515	3.256373	2.091815	2.64823	0.031322	0.049579	3
0.00546	0.073968	0.073397	2.030281	1.623696	1.832346	0.001452	1.298614	3
0.009501	0.052639	0.052947	3.744457	2.227701	2.970326	0.014274	0.158361	3
0.007452	0.064623	0.064379	2.708752	1.917112	2.318829	0.084318	0.514988	3
0.002106	0.063102	0.062477	2.822905	1.894306	2.342631	0.001242	0.361359	3
0.114405	0.039467	0.120887	1.788677	1.637389	1.730164	0.15048	16.96202	4
0.116347	0.042593	0.123746	1.848408	1.67213	1.778463	0.149167	18.21358	4
0.109839	0.041349	0.117212	1.861561	1.67939	1.792125	0.120963	14.69025	4
0.116759	0.03221	0.121031	1.592266	1.506494	1.561613	0.182333	18.0051	4
0.115441	0.041217	0.122434	1.764115	1.604478	1.701672	0.152843	17.5869	4
0.112717	0.030918	0.116795	1.634773	1.546062	1.602	0.09182	16.66351	4
0.118077	0.033982	0.122771	1.740657	1.637494	1.702599	0.146755	19.34746	4
0.11363	0.032838	0.118185	1.703288	1.602074	1.666293	0.101843	17.10229	4
0.114089	0.032349	0.118495	1.530019	1.441842	1.497519	0.059789	17.91909	4
0.116349	0.032846	0.120804	1.725601	1.626705	1.688983	0.134462	18.36644	4

4.3. Diagnosis of a Wavelet Neural Network

According to the characteristics and state of the network, we set a wavelet neural network to include eight input nodes for the eight features and four output nodes to indicate the four types of state. We selected five groups from every state, with a total of 20 types of groups composed of training samples and 12 groups as test samples. We set ±0.01 as the training error of the network. Figure 6 highlights the design of the wavelet neural network.
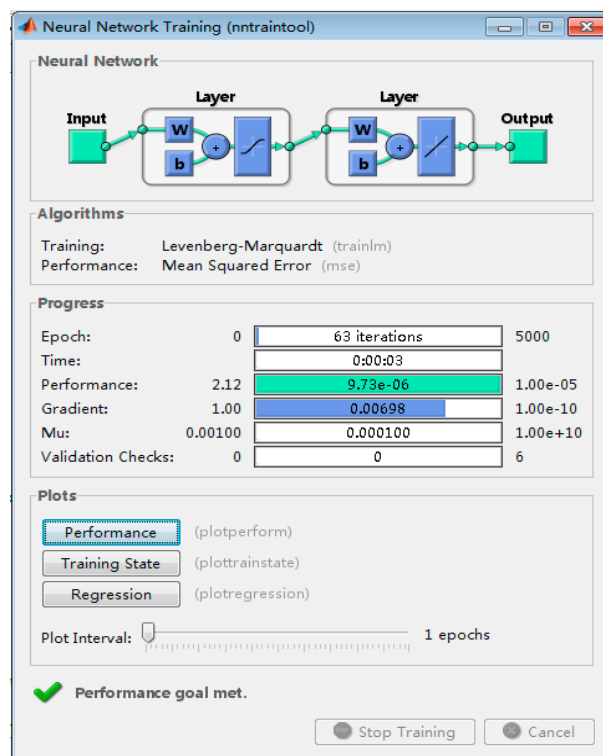
Figure 6. Diagnosis of the wavelet neural network.

When the network was trained, this meant that an error had met a demand. Test data was inputted into the wavelet neural network, and the test results are shown in Table 2.

Table 2. Fault diagnosis results.

Test Sample	Idea Output	Diagnosis Result			
No.	$y_1y_2y_3y_4$	y_1	y_2	y_3	y_4
1	1 0 0 0	1.0096	−0.0001	−0.0125	0.0030
2	1 0 0 0	1.0025	0.0001	0.0041	−0.0067
3	1 0 0 0	0.9799	0.0000	0.0225	−0.0024
4	0 1 0 0	0.0027	0.9990	−0.0118	0.0097
5	0 1 0 0	0.0004	1.0163	−0.0546	0.0382
6	0 1 0 0	0.0081	1.0044	−0.0011	−0.0114
7	0 0 1 0	0.0025	0.0000	0.9972	0.0003
8	0 0 1 0	0.0010	0.0000	0.9983	0.0007
9	0 0 1 0	0.0036	0.000	1.0044	−0.0009
10	0 0 0 1	−0.0379	−0.0099	−0.0267	1.0751
11	0 0 0 1	−0.0567	−0.0150	0.0639	1.0751
12	0 0 0 1	−0.0322	−0.0012	−0.0508	1.0844

As shown in Table 2, four states of the bearing can be well separated, with a 100% rate of recognition. This shows that: (1) the feature extraction is effective; (2) the wavelet neural network can be used for information fusion; and (3) good classification results can be achieved.

In order to further verify the effectiveness of the proposed method, we compared it with the methods described in Reference [24] based on the same bearing data. The authors of Reference [24] used principal component analysis (PCA), local and nonlocal preserving projection (LNPP), linear discriminant analysis (LDA), and supervised-learning-based local and nonlocal preserving projection (SLNPP) to extract features, respectively, before using the features as the inputs of *k*-nearest neighbor

algorithm (KNN). These results are shown in Table 3, where the proposed method has a higher recognition rate.

Table 3. Recognition rate of each method.

PCA [24]	LNNP [24]	LDA [24]	SLNPP [24]	ILE
93.7%	96.88%	98.44%	98.44%	100%

5. Conclusions

Roller bearing forms the core of rotating machinery, and its monitoring has always been of significant research interest. As roller bearing fault signals are nonlinear and non-stationary, we have done two things to resolve the issues of fault identification and fault classification.

(1) The performance of LE depends greatly on the neighborhoods *k*. Generally, *k* is valued according to experience. We proposed an improved LE algorithm that allows for adaptive change for any given *k*. The experimental results demonstrated that the proposed method can effectively obtain *k* and extract the feature signals of the roller bearings.

(2) Based on the feature signals and the integration of the merits of wavelet transform with that of an artificial neural network, we constructed a wavelet neural network for fault identification and classification. The experimental results indicate that this proposed method has excellent clinical practical value, with a classification accuracy rate of 100%.

Author Contributions: Yong Guan and Lifeng Wu performed the studies on fault diagnosis method based on a wavelet neural network and manifold learning. Zhen Peng and Beibei Yao analyzed the data. Lifeng Wu wrote the paper.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Xiang, D.; Cen, J. Method of roller bearing fault diagnosis based on information feature fusion of EMD entropy. *J. Aerosp. Power* **2015**, *30*, 1149–1154.
2. Li, Y.; Xu, M.; Wei, Y.; Huang, W. A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree. *Measurement* **2016**, *77*, 80–94. [CrossRef]
3. Urbanek, J.; Barszcz, T.; Uhl, T. Comparison of Advanced Signal-Processing Methods for Roller Bearing Faults Detection. *Metrol. Meas. Syst.* **2016**, *19*, 715–726.
4. Li, Y.; Xu, M.; Zhao, H.; Huang, W. Hierarchical fuzzy entropy and improved support vector machine based binary tree approach for rolling bearing fault diagnosis. *Mech. Mach. Theory* **2016**, *98*, 114–132. [CrossRef]
5. Cao, H.; Fan, F.; Zhou, K.; He, H. Wheel-bearing Fault Diagnosis of Trains using Empirical Wavelet Transform. *Measurement* **2016**, *82*, 439–449. [CrossRef]
6. Rauber, T.W.; De, A.B.F.; Varejao, F.M. Heterogeneous Feature Models and Feature Selection Applied to Bearing Fault Diagnosis. *IEEE Trans. Ind. Electron.* **2015**, *62*, 637–646. [CrossRef]
7. Sui, W.T.; Zhang, D.; Wang, W. Fault diagnosis of rolling element bearings based on EMD and MKD. *J. Vib. Shock* **2010**, *29*, 107–110.
8. Wu, Z.; Yang, S.; Ren, B. Rolling element bearing fault diagnosis method based on NAMEMD and multi-scale morphology. *J. Vib. Shock* **2016**, *35*, 127–133.
9. Wang, J.; Li, J.; Wan, X. Fault feature extraction method of rolling bearings based on singular value decomposition and local mean decomposition. *J. Mech. Eng.* **2015**, *3*, 104–110. [CrossRef]

10. Ma, L.; Kang, J.; Meng, Y. Research on feature extraction of rolling bearing incipient fault based on Morlet wavelet transform. *Chin. J. Sci. Inst.* **2013**, *34*, 920–926.

11. Huang, H.; Han, Z.; Zhang, Q. Method of fault diagnosis for rolling bearings based on Laplacian eigenmap. *J. Vib. Shock* **2015**, *5*, 128–134.

12. Rowei, S.T.; Saul, L.K. Nonlinear dimensionality reduction by locally linear embedding. *J. Sci.* **2000**, *290*, 2323–2326. [CrossRef] [PubMed]

13. Tenenbaum, J.B.; De Silva, B.; Lagford, J.C. A global geometric framework for nonlinear dimensionality reduction. *J. Sci.* **2000**, *290*, 2319–2323. [CrossRef] [PubMed]

14. Belkin, M.; Niyogi, P. Laplacian Eigen maps for dimensionality reduction and data representation. *J. Neural Comput.* **2003**, *15*, 1373–1396. [CrossRef]

15. Ge, C.; Liu, X.Y.; Ding, S. Analysis and compare in manifold learning. *J. Shandong Normal Univ. (Nat. Sci.)* **2010**, *25*, 15–17.

16. Huang, H.; Zhang, Q.; Han, Z.; Lin, X.; Liu, Q. Application of laplacian eigenmap in fault diagnosis of rolling bear. *J. China Meas. Test* **2015**, *41*, 94–98.

17. Xu, Q.; Liu, Y.; Tian, D. Fault diagnosis of rolling bearings using least square support vector regression based on glowworm swarm optimization algorithm. *J. Vib. Shock* **2014**, *10*, 8–12.

18. Sun, B.; Liang, C.; Shang, D. Rolling bearings fault diagnosis based on recurrence complex network. *J. Vib. Meas. Diagn.* **2015**, *3*, 578–584.

19. Wang, B.; Li, H.; Chen, Q. Rolling bearing performance degradative state recognition based on mathematical morphological fractal dimension and fuzzy center means. *Acta Armamentarh.* **2015**, *36*, 1982–1990.

20. Jing, W.; Zhen, Y.Z.; Hong, Y. Adaptive Manifold Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 253–265.

21. Martin, H.J.A.; De Lope, J.; Maravall, D. Robust high performance reinforcement learning through weighted *k*-nearest neighbors. *J. Neuro-Comput.* **2011**, *74*, 1251–1259. [CrossRef]

22. Zhang, Q.; Benveniste, A. Wavelet networks. *IEEE Trans. Neural Netw.* **1992**, *3*, 889–898. [CrossRef] [PubMed]

23. Hai, Q.; Lee, J.; Lin, J. Wavelet Filter-based Weak Signature Detection Method and its Application on Roller Bearing Prognostics. *J. Sound Vib.* **2006**, *289*, 1066–1090.

24. Yu, J. Local and nonlocal preserving projection for bearing defect classification and performance assessment. *IEEE Trans. Ind. Eelectron.* **2012**, *59*, 2363–2376. [CrossRef]