

Article

# A Deep Learning Model for Fault Diagnosis with a Deep Neural Network and Feature Fusion on Multi-Channel Sensory Signals

Qing Ye <sup>1,\*</sup>, Shaohu Liu <sup>2</sup> and Changhua Liu <sup>3</sup>

<sup>1</sup> School of Computer Science, Yangtze University, Jingzhou 430023, China

<sup>2</sup> School of Mechanical Engineering, Yangtze University, Jingzhou 430023, China; liushaoh@126.com

<sup>3</sup> General Office, Yangtze University, Jingzhou 430023, China; yangtzelch@163.com

\* Correspondence: yeqing@yangtzeu.edu.cn; Tel.: +86-0716-18986647012

Received: 22 June 2020; Accepted: 30 July 2020; Published: 1 August 2020



**Abstract:** Collecting multi-channel sensory signals is a feasible way to enhance performance in the diagnosis of mechanical equipment. In this article, a deep learning method combined with feature fusion on multi-channel sensory signals is proposed. First, a deep neural network (DNN) made up of auto-encoders is adopted to adaptively learn representative features from sensory signal and approximate non-linear relation between symptoms and fault modes. Then, Locality Preserving Projection (LPP) is utilized in the fusion of features extracted from multi-channel sensory signals. Finally, a novel diagnostic model based on multiple DNNs (MDNNs) and softmax is constructed with the input of fused deep features. The proposed method is verified in intelligent failure recognition for automobile final drive to evaluate its performance. A set of contrastive analyses of several intelligent models based on the Back-Propagation Neural Network (BPNN), Support Vector Machine (SVM) and the proposed deep architecture with single sensory signal and multi-channel sensory signals is implemented. The proposed deep architecture of feature extraction and feature fusion on multi-channel sensory signals can effectively recognize the fault patterns of final drive with the best diagnostic accuracy of 95.84%. The results confirm that the proposed method is more robust and effective than other comparative methods in the contrastive experiments.

**Keywords:** array signal processing; feature fusion; deep neural network; multi-channel sensory signals; intelligent fault diagnosis

## 1. Introduction

Final drive is the core component in the rear axle of automobile and is always running under complex operating conditions and inevitably faults. Any fault of final drive may cause severe human injury, production stoppage and economic loss [1,2]. The change of status is hidden in the vibrational characteristics produced during the operating period. Therefore, fault diagnosis with vibration signal is extremely efficient in status detection and failure recognition of machinery [3–8]. P.K. Wong used Ensemble Empirical Mode Decomposition (EEMD) in the failure recognition of gears [9]. Stephen McLaughlin acquired gear features from modulating signal [10]. D.P. Jena et al. used variant technology of wavelet transform to implement fault diagnosis [11]. M Amarnatha utilized vibration signals to detect the severity of fault of a helical gear tooth [12].

Diagnosing fault using a single-channel vibration signal may cause the decrease of diagnostic accuracy due to the direction of sensor and installation position [13]. Array signal processing is extensively applied in sonar, wireless communications, medical diagnosis and engineering applications [14–17]. Array signal processing could enhance useful signal and suppress noise and

interference by placing a set of sensors in different positions to form sensor array [18–21]. With the purpose of improving the diagnostic accuracy, it is essential to obtain multi-channel sensory signals by setting some sensors along different directions. The diagnostic accuracy could be more superior and reliable than using single signals.

Nowadays, many researches contributed to information mining and the utility of massive data [22]. Moreover, vibrational signals collected from multiple sensors are widely used in the field of fault diagnosis [23–26]. Yaguo Lei used multi-sensor data in fault detection of gearbox [27]. M.S. Safizadeh studied multi-sensor data fusion to improve the performance of fault recognition for rolling element bearings [28]. Luyang Jing combined deep neural networks and multi-sensor data fusion in the fault detection of planetary gearbox [29]. Zhixiong Li researched in fault recognition of gear cracks with multi-channel sensory signals [30]. Joao A. Duro adopted multi-sensor signals in machinery condition monitoring [31]. Khazae M. adopted a feature-level fusion in fault diagnosis of planetary gearbox [32]. Yunusa-Kaltungo developed a novel framework of fault characterization for rotating machines [33]. Zhiwen Liu used support vector machine and ant colony into fault diagnosis of gearbox [34]. Previous researches usually adopted traditional machine learning technologies in machinery fault diagnosis. Few papers have studied the deep feature learning and fusion model with multi-channel data.

The vibration signals are unavoidably polluted by disturbance originated from other vibration components and paroxysmal noise [35,36]. Traditional fault diagnosis employs complex signal process techniques to extract the shallow features manually. However, the representativeness of features obviously affects the capability of traditional fault recognition. Moreover, the manual selection of most sensitive features for a specific issue is time consuming [37]. These shortcomings encourage researchers to find out a new method to adaptively and automatically learn deep features from original data.

After a series of non-linear transformations, representative and essential features can be effectively mined by deep learning. Deep learning includes three basic types: Convolutional Neural Network (CNN), Deep Neural Network (DNN) and Deep Belief Network (DBN) [38–41]. Recently, deep learning using deep architectures is widely applied for intelligent fault diagnosis [42–44]. Compared with DBN and CNN, DNN consisted of a stack of auto-encoders that is purely unsupervised and easy to implement [45,46]. Due to the deep architectures, DNN can mine fault-sensitive features from raw signals so as to effectively ferret out the non-linear relationship between symptoms and faults. Jia Feng applied deep learning technology into the field of rotating machinery [37]. J. Zabalza et al. employed stacked auto-encoder to extract feature in hyper spectral imaging [47]. P. Xiong et al. employed de-noising auto-encoder for enhancement of signal [48].

Despite the fact that deep features extracted from multi-channel sensory data are fault sensitive and representative, features of different channels are usually heterogeneous and redundant. In order to avoid performance reduction caused by information redundancy, Locality Preserving Projection (LPP) is used in the fusion of extracted features obtained from multi-channel sensory signals. LPP is a typical dimensionality reduction algorithm based on local structure of data which can effectively preserve the neighborhood information and local characteristics during the fusion procedure [49–52]. The fused deep features of multi-channel sensory data will be fed into a diagnostic model based on softmax [53–55].

In this paper, aiming at solving the limitation of single sensor and automatically extracting fault-sensitive features without manual pre-processing, a deep learning model with multi-channel sensory signals is proposed and verified in application of intelligent fault diagnosis of automobile final drive. The contrast analysis and experimental results show that the proposed method can adaptively extract deep features from original multi-channel signals and effectively improve the accuracy and capability of failure recognition. The contributions and key techniques of our work are summarized in three aspects:

- (1) To solve the limitation of single sensor, we collect multi-channel sensory signals by installing several sensors in different positions along horizontal and vertical direction so as to implement reliable monitoring.
- (2) To solve the limitation of using traditional signal process techniques to manually extract features, we employ deep learning technique in learning representative features from original multi-channel signal adaptively.
- (3) In order to avoid the heterogeneity and redundancy of deep features from multi-channel data, we fuse these deep features by using locality preserving projection.

The following is organized in four sections. In Section 2, the fundamental theory used in the research is given. Application of the diagnostic method based on multiple DNNs (MDNNs) is presented in the next section. A set of experiments is implemented to evaluate the accuracy and efficiency in Section 4. In Section 5, a set of conclusions is provided.

## 2. The Fundamental Theory of Auto-Encoder

An auto-encoder is a special neural network based on unsupervised learning with three necessary layers of artificial neural network. The training process of an auto-encoder always contains two stages, encoding and decoding [56].

Given a sample set  $X = \{x^1, \dots, x^i, \dots, x^M\}$ ,  $x^i \in \mathbb{R}^N$ , where  $M$  denotes the sample size. In encoding stage, input data  $x^i$  is transformed into a low-dimensional feature space to learn an approximation of input data as shown in below:

$$h(i) = f_{\theta}(x^i) \quad (1)$$

where  $f_{\theta}$  denotes encoding function and  $\theta = [W, b]$  denotes the parameters between the first layer and the second layer.

In decoding stage, input data  $x^i$  can be reconstructed in output layer as follows:

$$x^{i'} = g_{\theta'}(h(i)) \quad (2)$$

in which  $g_{\theta'}$  denotes decoding function,  $x^{i'}$  denotes the reconstruction of original data  $x^i$ , and  $\theta' = [W', b']$  is the parameters between the second layer and the third layer.

By minimizing the average reconstruction error of  $M$  samples, the parameter set  $\{\theta, \theta'\}$  of the encoding stage and decoding stage can be optimized:

$$\mathcal{O}_{AE}(\theta, \theta') = \frac{1}{M} \sum_{i=1}^M L(x^i, x^{i'}) \quad (3)$$

$$L(x, x') = \|x - x'\|^2 \quad (4)$$

where  $L(x, x')$  represents reconstruction error function to measure the discrepancy between original data and its reconstruction.

## 3. Proposed Diagnostic Model

A deep learning method with deep architecture and feature fusion on multi-channel sensory signals is developed in this paper. It includes three parts: the construction of deep architecture for feature learning, the fusion of the deep features extracted from multi-channel sensory data and the construction of intelligent diagnostic model using softmax.

### 3.1. Construction of Deep Neural Network for Deep Feature Learning

By piling up  $N$  auto-encoders, construct a DNN with  $N$  hidden layers so as to hierarchically learn essential characteristics from sensory data. The first auto-encoder consists of input layer and the first hidden layer, and so on. The training of DNN contains two sections: pre-training and

fine-tuning [57–59]. During the process of the first section based on unsupervised learning, the encoder vector of  $x(i)$  obtained from the first auto-encoder is:

$$h(i)^1 = f_{\theta^1}(x(i)) \quad (5)$$

where  $\theta^1$  represents the parameter of the first auto-encoder.

Then, the first encoder vector  $h(i)^1$  is inputted into the second auto-encoder, which consists of the first hidden layer and the second hidden layer, and so on. In this way, the Nth encoder vector of  $x(i)$  is obtained as follows:

$$h(i)^N = f_{\theta^N}(h(i)^{N-1}) \quad (6)$$

where  $\theta^N$  is the parameter of the Nth auto-encoder.

Compared with random parameters, by minimizing the reconstruction error, the unsupervised pre-training of DNN can improve the generalization of parameters. In the process of fine-tuning parameters backwards based on supervised learning through the BP (Back Propagation) algorithm, the ability of feature learning is further enhanced. BP algorithm compares the output of output layer with the corresponding label so as to calculate the loss value. The loss function is shown below:

$$\varnothing_{\text{DNN}}(\theta) = \frac{1}{M} \sum_{i=1}^M L(x(i), h(i)^N) \quad (7)$$

where  $\theta = \{\theta^1, \theta^2, \dots, \theta^N\}$  corresponds to N hidden layers. The partial derivative of the parameter is solved, and then the parameter is updated by the gradient descent algorithm. The parameter set can be optimized with learning rate of  $\mu$  as follows:

$$\theta = \theta - \mu \frac{\varnothing_{\text{DNN}}(\theta)}{\partial \theta} \quad (8)$$

During the training of DNN, a series of non-linear transformations of pre-training captures the local variation of input data, and fine-tuning mines the discriminated information from input data.

### 3.2. Fusion of Deep Features Extracted from Multi-Channel Signal

As shown in Figure 1, with the purpose of obtaining more complete and reliable data,  $S$  sensors are installed so as to collect multi-channel signals and construct multiple DNNs (MDNNs). In MDNNs, raw timing signals collected from different sensors are inputted into different DNNs. The deep features extracted from MDNNs are usually redundant and high dimensioned. LPP is a dimension-deduction method that can be used to fuse deep features and effectively preserve local structure in lower dimension [60,61].

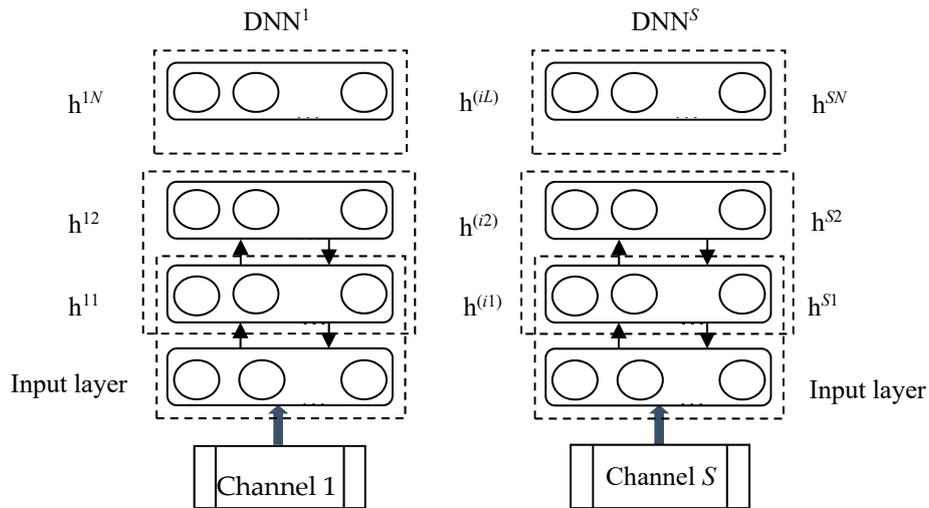


Figure 1. Architecture of multiple deep neural networks (MDNNs).

Let  $F = \{f_1, \dots, f_i, \dots, f_n\}$  to represent the deep feature learned from MDNNs, in which  $f_i = [f_i^1, \dots, f_i^s, \dots, f_i^S] \in \mathbb{R}^D$ ,  $n$  represents the sample size and  $S$  represents the number of sensors. Using a projection matrix  $A \in \mathbb{R}^{D \times d}$ , deep features with a dimension of  $D$  are transformed into dimension of  $d$  as below:

$$f_i^* = A^T f_i, i = 1, \dots, n \tag{9}$$

The objective function of LPP is the following formula:

$$\min \sum_{ij} (f_i^* - f_j^*)^2 W_{ij} \tag{10}$$

$$W_{ij} = \begin{cases} \exp(-\|f_i - f_j\|^2 / t) \\ 0 \end{cases} \tag{11}$$

where  $W_{ij}$  represents the matrix to measure relations of different components in deep feature. With algebraic formulas, transform the objective function into the following formula:

$$\min A^T F L F^T A \tag{12}$$

where  $L = \varnothing - W$  is the Laplacian matrix in which  $\varnothing$  is the diagonal matrix of  $W$ .

Then, the above problem with constraint is converted into a generalized eigenvalue problem. The solution of above objective function can be achieved from eigenvectors corresponding to minimum eigenvalues.

### 3.3. The Procedure of Intelligent Diagnostic Model

The fused deep features  $F^* = A^T F$ , which can reflect the variation of conditions will be inputted into the following fault classifier based on softmax. As shown in Figure 2, the proposed method with multi-channel sensory data contains two parts: a training diagnostic model based on fused deep features of multi-channel data and fault recognition. The detailed procedure of the proposed method is organized in the following six steps:

- Step 1: Collect multi-channel sensory data from multiple sensors installed in different directions and positions.
- Step 2: Without manually extracting features by using traditional signal process techniques, raw data is split up into training subset and testing subset.

- Step 3: Construct deep architecture with multiple DNNs to learn fault-sensitive and representative features adaptively from multi-channel sensory signals.
- Step 4: Fuse the deep features learned from MDNNs constructed in Step 3 by using LPP, and acquire the representative low-dimensional features.
- Step 5: Feed the fused features obtained in Step 4 into the fault classifier based on softmax.
- Step 6: Implement fault recognition on the testing set to verify the classification ability and generalization of the method.

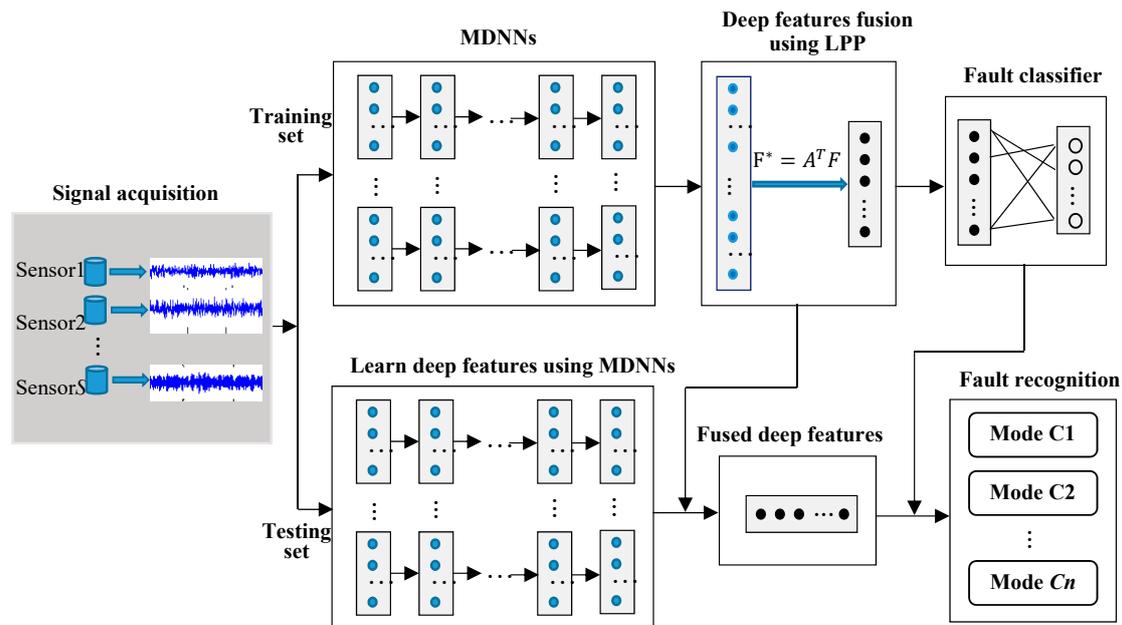


Figure 2. Procedure of the proposed model.

## 4. Experiments and Discussion

### 4.1. Experimental Arrangement

As shown in Figure 3, a set of contrastive experiments are implemented on the test rig of automobile final drive. The test rig contains a control cabinet section, a drive section to activate the driving motor and a fixture section. The rotating speed can be controlled by the cabinet section. The fixture section is to simulate the running situation under a specific rotating speed. To carry out reliable monitoring, two vibration acceleration sensors are located in orthogonal positions so as to collect more stable multi-channel sensory data along a different transmission path. This kind of sensor is frequently used to acquire high-frequency signals in mechanical vibration engineering. The characteristics of vibration acceleration sensors are a wide frequency response, a wide dynamic range, and high reliability. Figure 4 shows the locations of multiple sensors.

The most common failure modes of final drive include gear crack, gear error, gear tooth broken, gear burr, misalignment and gear hard point [12]. These faults usually occur in the gears pair as shown in Figures 5 and 6. Signals from the above 6 fault modes and normal mode are collected. To collect more stable and fault-sensitive vibrational signals, based on previous research and experiences, the frequency of sampling is 12,000 Hz and the rotational speed of the motor is 1200 r/min. The frequency of sampling must be higher than the pre-estimated gear meshing frequency or the detailed information of failure will be lost.

In order to collect enough and reliable data to represent each fault modes, 10 simulations of data acquisition are implemented for each mode. In the processing of data acquisition, the time of duration is 2 s and 25 signals are collected. Each simulation takes 10 s. In this way, 1750 vibration signals

corresponding to each different mode are obtained from each sensor. The collected signals are split into training subset with 1400 signals and testing subset with 350 signals. Each vibration signal contains 1024 data points. Descriptions about the final drive fault modes are listed in Table 1.

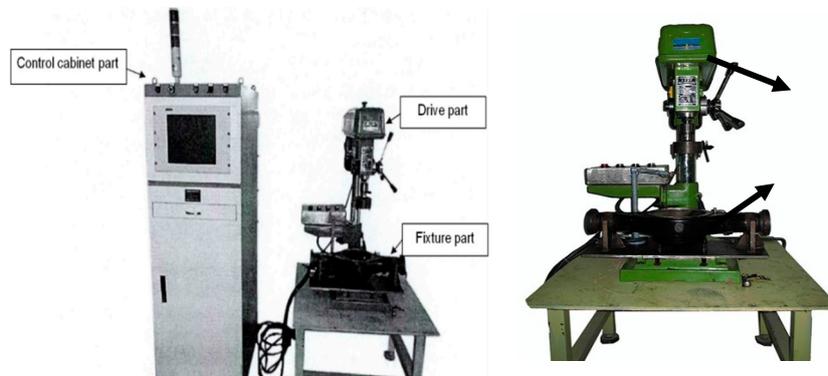


Figure 3. Test rig of automobile final drive.

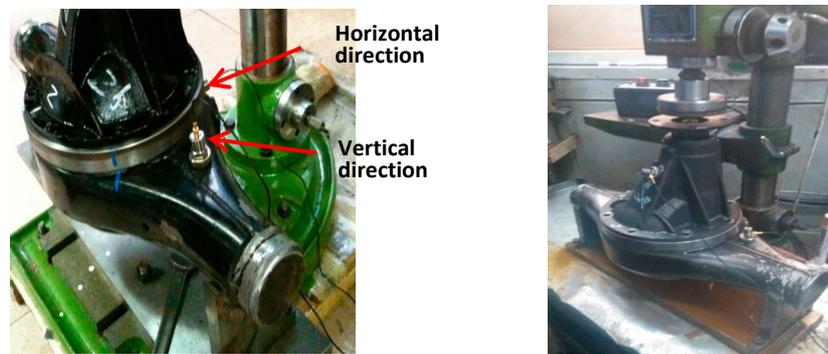
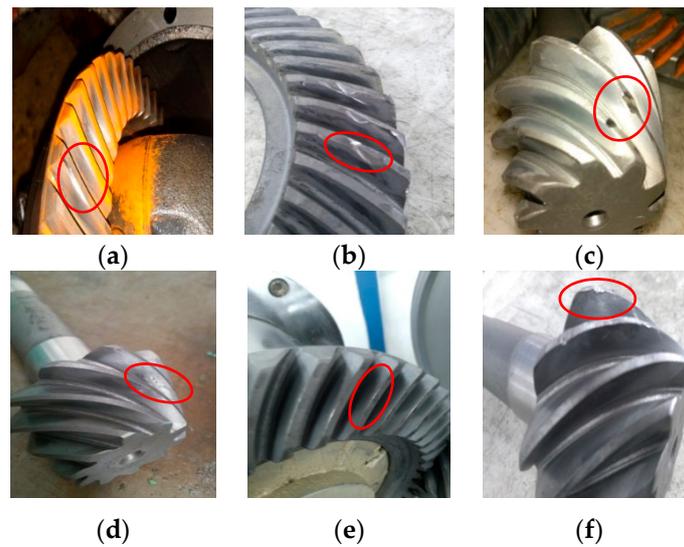


Figure 4. Installation of multiple sensors.



Figure 5. Pivotal part in final drive: gear pair.



**Figure 6.** Faulty gears pairs: (a) gear crack, (b) gear error, (c) gear tooth broken, (d) gear burr, (e) misalignment, (f) gear hard point.

**Table 1.** Fault modes' description.

Label	Fault Patterns	Size of Training Set	Size of Testing Set
C1	Normal status	1400	350
C2	Gear crack	1400	350
C3	Gear error	1400	350
C4	Gear tooth broken	1400	350
C5	Gear burr	1400	350
C6	Misalignment	1400	350
C7	Gear hard point	1400	350

A series of experiments is performed under Matlab 7.0 on the PC with the configuration of 3.4GHz CPU and 4GB RAM.

To evaluate the superiority of the proposed model, we also utilize BPNN (Back-Propagation Neural Networks) and shallow architecture SVM (Support Vector Machine) to construct diagnostic models for automobile final drive.

To verify the effectiveness of fusing sensory data from multiple sensors, the first set of experiments are implemented between diagnostic models based on multi-channel data and single sensory data. To verify the superiority of deep architecture in feature learning, the second set of experiments is implemented between the proposed method and diagnostic model based on BP neural networks with a shallower structure. To analyze the representativeness of the features learned from hidden layers in deep architecture, the third set of experiments is implemented to reveal most typical features using LPP and Principle Component Analysis (PCA). Ten trials are carried out for each experiment.

#### 4.2. Models Design

With regard to a diagnostic model based on MDNNs, the layout of deep architecture has an important impact on accuracy. In order to design the optimal model, we investigate the accuracy variation of different deep architectures. As shown in Figure 7, the diagnostic accuracy improves obviously as the deep architectures changes from containing only one hidden layer to four hidden layers. However, as the hidden layer changes from including 100 neurons to including 1000 neurons with an interval of 100, the diagnostic accuracy is fluctuating up and down. According to the results, the optimal model based on MDNNs consists of three hidden layers and there are 400 units in each hidden layer.

In order to train each DNN, the epoch times in pre-training are set to 100, and the epoch times of fine-tuning are setting to 200 by experience. In the fusion procedure based on LPP, the dimension of fused deep features  $d$  and hyper-parameter  $k$  of nearest neighbors are set to 20 and 12 using a cross-validation algorithm.

For the diagnostic model based on BPNN, the shallow architecture contains two hidden layers with 100 units per hidden layer. For the diagnostic model based on SVM, the optimum parameters are chosen by using 10-fold cross validation. The hyper-parameter  $C$  is set as  $10^\alpha$ , in which  $\alpha$  is in the range of 0 to 2. To achieve non-linear classification, kernel function RBF (Radial Basis Function) with  $C = 10$  and  $r = 2$  is adopted.

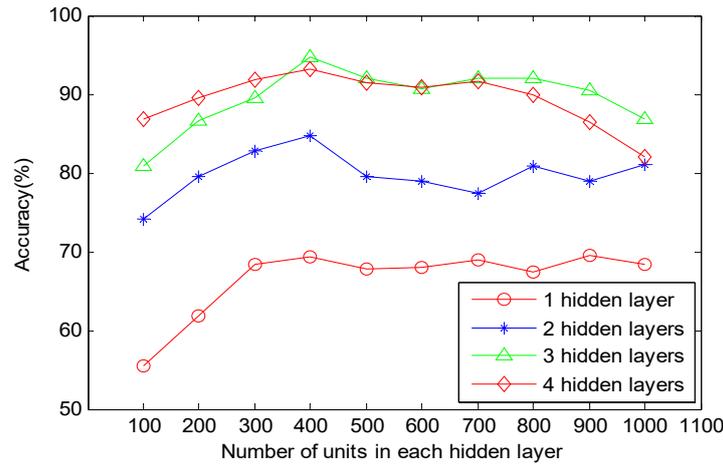


Figure 7. Performances of different deep architectures.

### 4.3. Contrast Analysis and Discussion

#### 4.3.1. Validity of the Fusion Strategy

To verify the validity of the fusion strategy on multi-channel sensory data, we implement the first set of experiments to compare several diagnostic models based on multi-channel data and single sensory data, respectively. Figures 8–10 show the detailed diagnostic results of these contrastive models based on the proposed method, BPNN and SVM in ten trials.

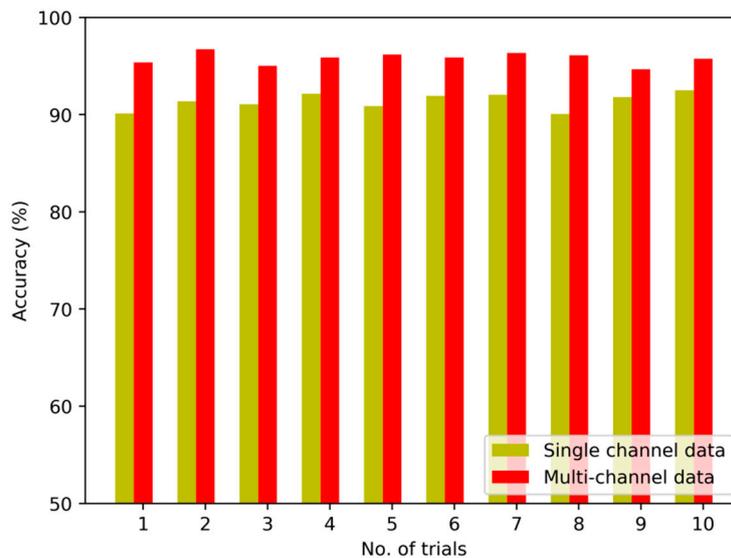
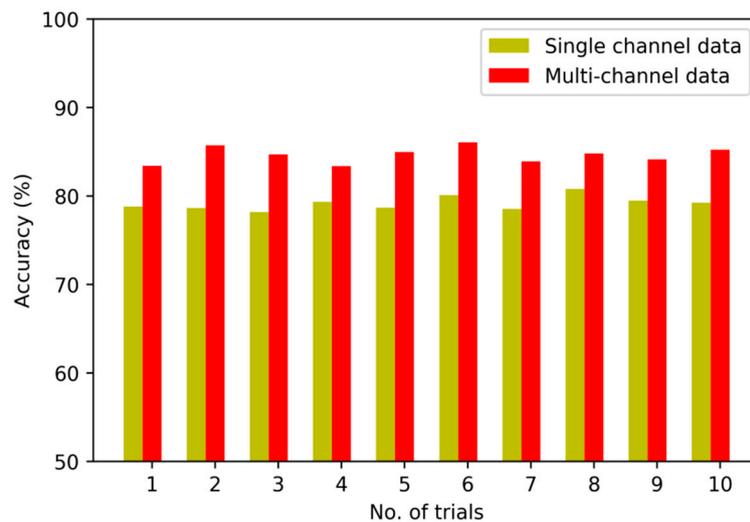
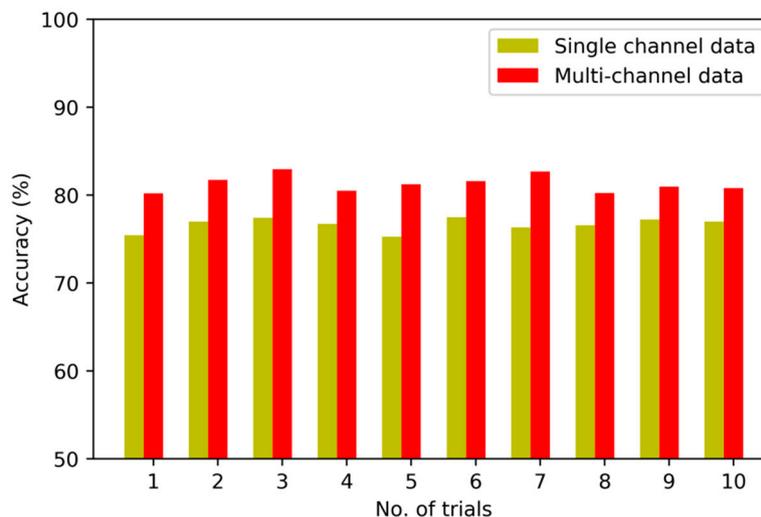


Figure 8. Diagnostic results of 10 trials for the proposed method.



**Figure 9.** Diagnostic results of 10 trials for BPNN-based model.



**Figure 10.** Diagnostic results of 10 trials for Support Vector Machine (SVM)-based model.

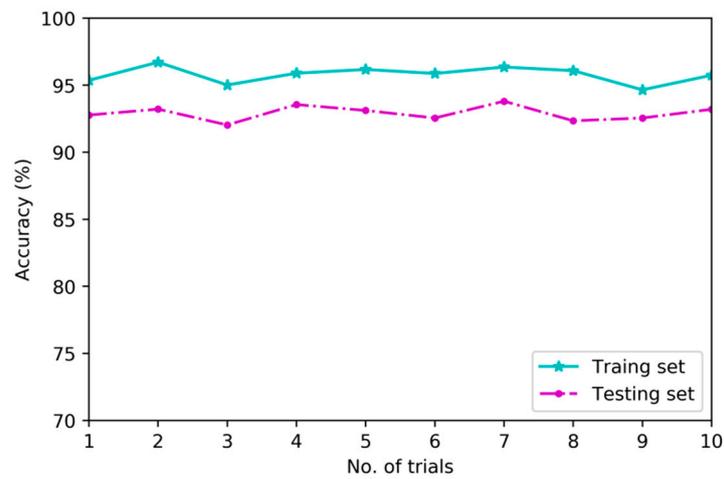
The diagnostic results of the proposed method with multi-channel data and single-channel data are 95.8% and 91.4%. It means that the accuracy is improved after the multi-channel data fusion process. The diagnostic results of the model based on BPNN are 79.1%, and the diagnostic results of model based on SVM are 76.6% without fusion of multi-channel data. On closer inspection, the accuracy is increased to 84.6% and 81.3% after fusing the sensory data of multiple sensors.

The comparison result indicates that the multi-channel data collected from multiple sensors are more complete and abundant than single-channel data. Moreover, the fusion strategy can reflect the interdependency between multi-channel data and the fault modes effectively so as to improve the performance of diagnostic model by nearly 8%. In particular, by constructing multiple DNN to learn representative features from multi-channel sensory data and fusing the learned features using LPP algorithm effectively, the diagnostic model using the proposed method exhibits its superiority in fault recognition.

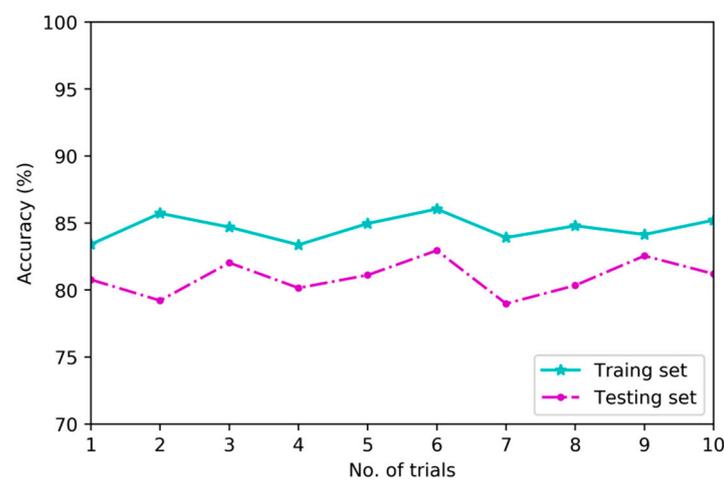
#### 4.3.2. Validity of the Deep Architecture

To verify the superiority of deep architecture in feature learning, this research implements a second set of experiments to compare diagnostic models based on deep architecture and shallow architecture using training set and testing set which are consisted of multi-channel sensory data.

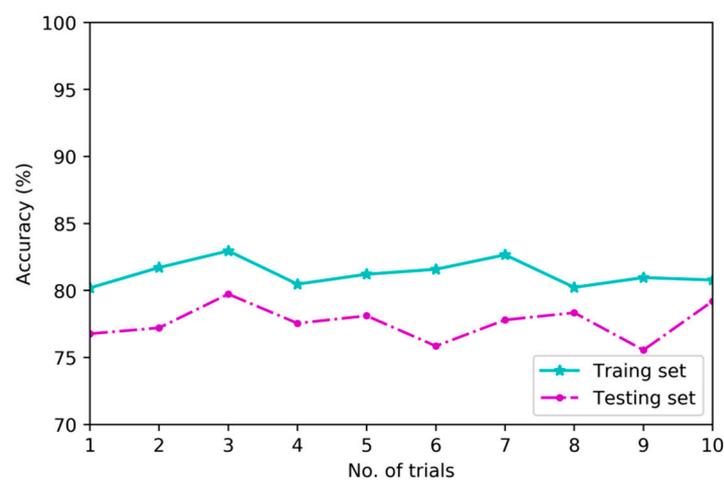
Figures 11–13 show the differences in classification results of training subset and testing subset among three contrastive diagnostic models based on the proposed method, BPNN and SVM. Each experiment is carried out for ten trials.



**Figure 11.** Classification results of MDNNs.



**Figure 12.** Classification results of BPNN.



**Figure 13.** Classification results of SVM.

As shown in Figure 11, the average results of training subset and testing subset in ten trials for diagnostic model based on MDNNs are 95.8% and 92.8%. As shown in Figures 12 and 13, the average results of training subset for another two models based on BPNN and SVM are 84.6% and 81.3%, and the testing accuracies are 81.1% and 77.6%, respectively.

From Figures 11–13, the deep architecture of MDNNs can effectively automatically learn intrinsic features which are sensitive to fault patterns and the testing accuracy of ten trials are very stable in the range of 92.03% to 93.8%. However, the testing accuracy of diagnostic model based on BPNN is unstable and fluctuating from 76.2% to 82.6%. The comparison results indicate that deep feature learning architecture outperforms shallow architecture in feature learning performance on average without using any signal processing and expert experience [24].

### 4.3.3. Representative of Deep Features

With the purpose of analyzing the representativeness of features learned from the proposed deep architecture with three hidden layers, we implement the third set of experiments to reveal the first two typical features learned from different hidden layers using LPP. Figures 14–16 visualize the typical 2-dimensional deep features learned from three hidden layers in the deep architecture.

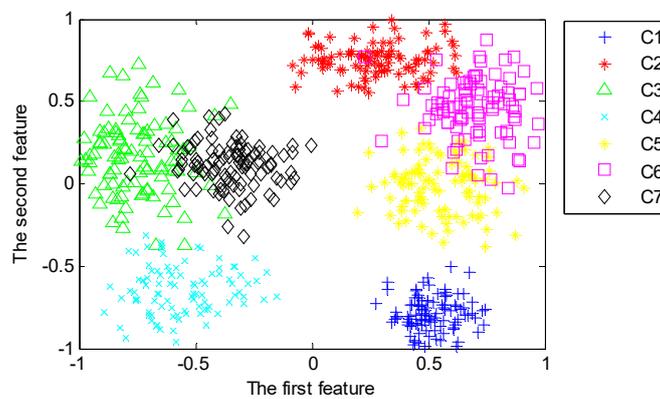


Figure 14. Typical features fused from the first hidden layer using Locality Preserving Projection (LPP).

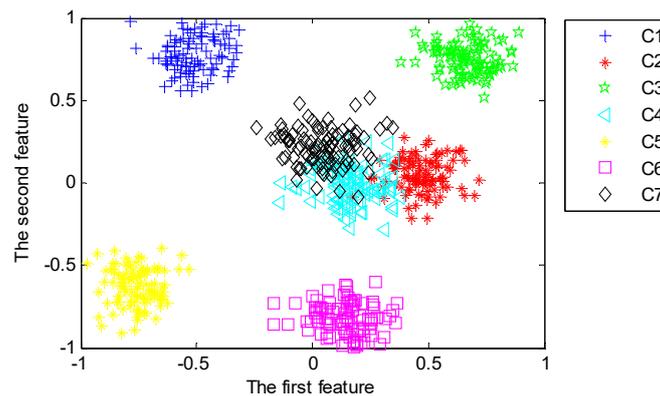
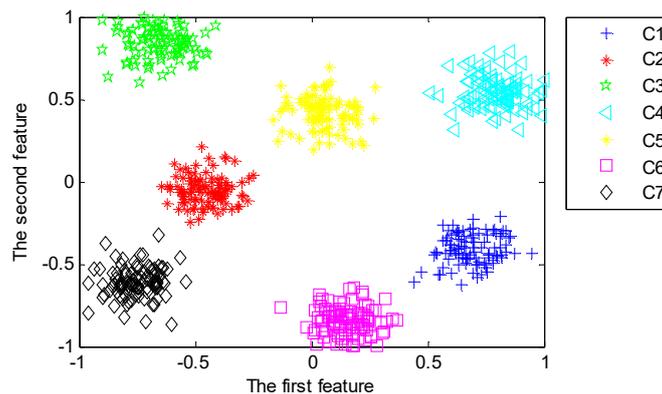


Figure 15. Typical features fused from the second hidden layer using LPP.

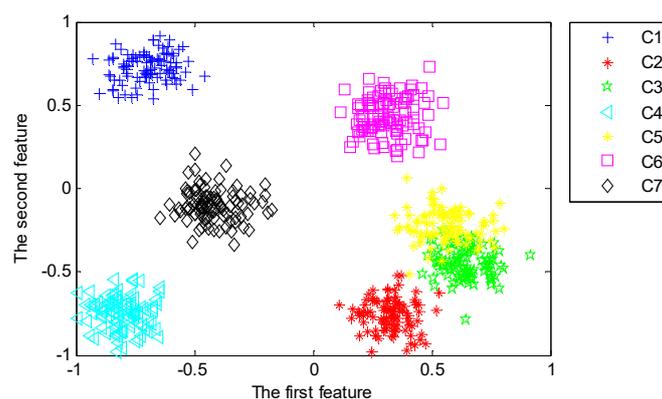


**Figure 16.** Typical features fused from the third hidden layer using LPP.

From Figures 14–16, we can find that the first two dimensional features fused from the deeper hidden layer by using LPP are more discrete between different patterns than that of the shallow hidden layer. This demonstrates that the deep architecture can learn more differentiable features from the input data through multiple non-linear transformation layer-by-layer [47]. In particular, the first two dimensional features of the last hidden layer can almost separate samples of each pattern precisely.

For comparison, we preserve the most typical two-dimensional features that are fused by using LPP and Principle Component Analysis (PCA). With the purpose of dimension reduction, we employ PCA in the fusion of deep features learned from the last hidden layer. With the features learned from the last hidden layer with the dimensionality of 400 for each sensor, the output of MDNNs is in the dimension of  $2 \times 400$ . By using PCA, the dimensionality of features extracted from MDNNs with multi-channel data is reduced from  $2 \times 400$  into 20, which is the same dimension of fused features using LPP.

In order to intuitively judge the fault-sensitivity of the learned features, the first two principal components of the fused features are presented for visualization. The typical components fused by PCA of the last hidden layer are shown in Figure 17. Some features of pattern C3 and pattern C5 are mixed and failed to distinguish. By comparing Figure 16 with Figure 17, we can conclude that LPP is more superior in retaining the local structure of input data that is valuable for fault pattern recognition than PCA. LPP achieve fusion and extraction using manifold learning [42,53].



**Figure 17.** Principal components fused by Principle Component Analysis (PCA).

#### 4.3.4. Contrastive Analysis and Discussion

In this study, we mainly focus on the intelligent fault diagnosis with raw vibrational data without previous traditional signal pre-processing. The representative features are automatically extracted using deep structure of networks. However, in the traditional fault diagnosis of machinery, features in time domain and frequency domain can be extracted manually from vibrational signals and be inputted

into the fault recognition model [38]. To prove the sensitivity to above two forms of features between different diagnostic models, we compare the proposed model based on MDNNs and diagnostic models based on BPNN and SVM, separately with input of features manually extracted with signal pre-processing techniques and features adaptively extracted from raw data without pre-processing.

Manual features are extracted from vibrational signals using wavelet package transform with a maximum layer of 5 and db4. For each frequency-band, 19 features are obtained. In this way, the dimension of manual features for each vibration signal is  $19 \times 32$ . The average diagnostic accuracies of these diagnostic models are shown in Table 2.

**Table 2.** Average accuracy of different diagnostic models.

Models	Average Testing Accuracy (%)	
	Features Manually Extracted with Signal Pre-Processing	Features Adaptively Extracted without Pre-Processing
The proposed model	94.23 ( $\pm 1.54$ )	93.84 ( $\pm 0.73$ )
Model based on BPNN	84.27 ( $\pm 3.81$ )	81.46 ( $\pm 4.32$ )
Model based on SVM	79.62 ( $\pm 2.88$ )	76.49 ( $\pm 2.92$ )

From Table 2, we can find that diagnostic models based on BPNN and SVM are more sensitive to the forms of features inputted into the diagnostic model [41]. The average testing accuracies of these two shallow structured models with the input of manual features and features without pre-processing are 84.27% and 81.46%, 79.62% and 76.49%, respectively. However, the accuracies of the proposed model are very close, which are 94.23% and 93.84%. It indicates that without complex previous signal pre-processing, the proposed model can still recognize the vast majority of testing samples. It indicates the advantage of the proposed deep learning model on processing raw data adaptively and efficiently.

With the purpose of proving superiority of the proposed model combining MDNNs with multi-channel signal fusion, a series of experiments with the same sample set is implemented to compare some frequently used techniques in intelligent fault diagnosis. For each experiment, ten trials are implemented. The average training accuracy and testing accuracy of these different diagnostic models are given in Table 3. The average training time of all the diagnostic models are given in Table 4.

**Table 3.** Average accuracy of different diagnostic models.

Models	Average Training Accuracy (%)		Average Testing Accuracy (%)	
	Single-Channel Data	Multi-Channel Data	Single-Channel Data	Multi-Channel Data
The proposed method	91.42 ( $\pm 1.11$ )	95.84 ( $\pm 0.94$ )	90.17 ( $\pm 1.29$ )	92.76 ( $\pm 1.13$ )
Model based on BPNN	79.09 ( $\pm 4.59$ )	84.56 ( $\pm 3.72$ )	76.29 ( $\pm 4.80$ )	81.14 ( $\pm 4.05$ )
Model based on SVM	76.62 ( $\pm 2.97$ )	81.28 ( $\pm 2.63$ )	74.32 ( $\pm 3.47$ )	77.63 ( $\pm 3.06$ )

**Table 4.** Average training time of different diagnostic models.

Models	Average Training Time (s)	
	Without Fusion Multi-Channel Data	Fusion Multi-Channel Data
The proposed model	58.56	39.49
Model based on BPNN	30.39	18.61
Model based on SVM	10.22	5.93

As shown in Table 3, the average training accuracy and testing accuracy of the proposed model with single-channel data is 91.42% and 90.17%. The average training accuracy and testing accuracy of the proposed model based on multiple DNNs with deep architecture for multi-channel feature learning are 95.84% and 92.76%, which are higher than other diagnostic models with different kinds of data. From the contrastive results in Table 3, it reveals that the capability significantly exceeds the diagnostic

model based on SVM with shallow architecture for multi-channel feature learning, which are 81.28% and 77.63%. By contrast analysis, we can summarize that:

- (1) In general, deep architecture of neural networks can effectively extract essential and useful features from raw data. However, it is hardly to obtain favorable results by using BPNN with deep architecture. Furthermore, as shown in Table 3, the variance of testing accuracy and training accuracy is obviously higher than other models. It indicates that the performance of BPNN is unstable due to the local minimum problem. The reason for this disadvantage is that the stability of BPNN will be affected by the initial value of network parameters [10]. It may also lead to obvious deviation in the procedure of error back propagation.
- (2) From Table 3, it shows that the deep learning method shows its obvious superiority in feature learning. However, even with multiple hidden layers and multi-channel data, the diagnostic performance of the model based on BPNN is still far from satisfactory, which are 84.56% and 81.14%. Feature learning with deep architecture contains a process of pre-training and a process of fine-tuning. The local minimum problem of traditional BPNN can be obviously solved through the optimization of initial weights layer-by-layer during the procedure of pre-training in unsupervised way and the adjustment of weights during the procedure of fine-tuning in supervised way, which are the typical characteristics of deep neural networks [49].
- (3) From the separation of typical features fused from the hidden layers shown in Figures 14–16, we can find that a diagnostic model based on MDNNs can automatically extract fault-sensitive features that directly affect the final diagnostic results. In addition, it can adaptively mine the complex non-linear relevance between raw data and several fault modes, which are crucial for condition monitoring [29]. The capability of the model is independent on engineering experiences and prior knowledge in application area.

As shown in Table 4, the average training time of the proposed model with multi-channel data is 39.49 s. From the contrastive results in Table 4, we can find that the training time of the diagnostic model based on SVM with shallow architecture which is 5.93 s is obviously less than diagnostic models based on MDNNs and BPNN with deep architecture. The training time of the proposed model is the largest, which is 58.56 s.

The computational complexity of the proposed model is a typical shortcoming when the number of hidden layers increased. The reason is that the number of weights of shallow architecture that should be adjusted is much less than deep architecture containing multiple hidden layers [47]. Though the training time of the proposed model is the longest, the training process is accomplished within one minute.

Moreover, compared with the training time of models without fusion strategy, which is 58.56 s, the training time of models with the fusion of multi-channel data using the LPP algorithm is obviously decreased to improve the training efficiency. The reason behind the improvement is that the fusion process using LPP can decrease the dimension of high-dimensional features so as to decrease the computational complexity [43,44].

## 5. Conclusions

In this article, a deep learning method for multi-channel sensory data combining a deep neural network for feature mining and feature fusion is proposed. First, a deep structure MDNNs made up of multiple auto-encoders is constructed to adaptively extract significant features from sensory signals and to mine complex relation between symptoms and fault patterns. Secondly, LPP is used to fuse representative deep features learned from multi-channel sensory data using MDNNs. Finally, an intelligent diagnostic model is constructed by inputting the fused deep features into softmax.

The diagnostic model is employed into the application of intelligent failure recognition for automobile final drive. The average training accuracy and testing accuracy of the proposed model based on multi-channel data are 95.84% and 92.76%, which are higher than other contrastive models.

Compared with fault recognition with single-channel signals, the proposed method can effectively enhance the capability of failure classification from 90.17% to 92.76%. Moreover, the variance of the training accuracy and testing accuracy is obviously the smallest among these contrastive models. Through the contrastive analysis of experimental results, it proves that the proposed diagnostic model is more superior and stable than other models in fault recognition and condition monitoring. The deep structure of feature learning from multi-channel data can effectively solve the limitation of a single sensor and the adaptive fusion of feature can avoid the heterogeneity and redundancy of deep features extracted from multi-channel data. This research is instructive to the technology of industry. In the future, the authors would continue to investigate the intelligent diagnosis using multi-channel data in different types so as to further enhance the diagnostic accuracy and efficiency.

**Author Contributions:** Q.Y. conceived and designed the research, wrote the manuscript; S.L. validated and revised the manuscript; C.L. review and edited. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant No. 51974036), the Scientific Research Project of Education Department of Hubei Province (Grant No. Q20181301), Yangtze Youth Found of Yangtze University (Grant No. 2016cqn33).

**Acknowledgments:** The authors would like to thank the National Natural Science Foundation of China (Grant No. 51974036), the Scientific Research Project of Education Department of Hubei Province (Grant No. Q20181301), Yangtze Youth Found of Yangtze University (Grant No. 2016cqn33) for their support. The authors would also like to thank the editors and anonymous reviewers for their constructive comments and suggestions.

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

## References

1. Fang, H.-W.; Ma, J. Analysis of transformation countermeasures of automobile manufacturing enterprises from production type to service type. *J. Chang'an Univ. (Nat. Sci. Ed.)* **2013**, *33*, 131–136.
2. Yao, L.-J.; Ding, J.-X. An On-line Vibration Monitoring System for Final Drive of Automobile. *Noise Vib. Control* **2017**, *27*, 54–57.
3. Ye, Q.; Liu, C.; Pan, H. Simultaneous Fault Diagnosis Method Based on Improved Sparse Bayesian Extreme Learning Machine. *J. Southwest Jiaotong Univ.* **2016**, *51*, 793–799.
4. Khazaei, M.; Ahmadi, H.; Omid, M.; Moosavian, A.; Khazaei, M. Classifier fusion of vibration and acoustic signals for fault diagnosis and classification of planetary gears based on Dempster–Shafer evidence theory. *Proc. Inst. Mech. Eng. Part E J. Process. Mech. Eng.* **2013**, *228*, 21–32. [[CrossRef](#)]
5. Zhao, R.; Yan, R.; Wang, J.; Mao, K. Learning to Monitor Machine Health with Convolutional Bi-Directional LSTM Networks. *Sensors* **2017**, *17*, 273. [[CrossRef](#)] [[PubMed](#)]
6. Liu, X.; Tian, Y.; Lei, X.; Liu, M.; Wen, X.; Huang, H.; Wang, H. Deep forest based intelligent fault diagnosis of hydraulic turbine. *J. Mech. Sci. Technol.* **2019**, *33*, 2049–2058. [[CrossRef](#)]
7. Lu, Y.; Tang, J.; Luo, H. Wind Turbine Gearbox Fault Detection Using Multiple Sensors With Features Level Data Fusion. *J. Eng. Gas Turbines Power* **2012**, *134*, 042501. [[CrossRef](#)]
8. Eren, L.; Ince, T.; Kiranyaz, S. A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1D CNN Classifier. *J. Signal Process. Syst.* **2018**, *91*, 179–189. [[CrossRef](#)]
9. Yang, Z.; Wong, P.K.; Vong, C.-M.; Zhong, J.; Liang, J. Simultaneous-Fault Diagnosis of Gas Turbine Generator Systems Using a Pairwise-Coupled Probabilistic Classifier. *Math. Probl. Eng.* **2013**, *2013*, 1–14. [[CrossRef](#)]
10. Guoji, S.; McLaughlin, S.; Yongcheng, X.; White, P. Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears. *Mech. Syst. Signal Process.* **2014**, *43*, 76–89. [[CrossRef](#)]
11. Jena, D.; Sahoo, S.; Panigrahi, S. Gear fault diagnosis using active noise cancellation and adaptive wavelet transform. *Measurement* **2014**, *47*, 356–372. [[CrossRef](#)]
12. Amarnath, M.; Krishna, I.P. Local fault detection in helical gears via vibration and acoustic signals using EMD based statistical parameter analysis. *Measurement* **2014**, *58*, 154–164. [[CrossRef](#)]
13. Ye, Q.; Liu, C. A Multichannel Data Fusion Method Based on Multiple Deep Belief Networks for Intelligent Fault Diagnosis of Main Reducer. *Symmetry* **2020**, *12*, 483. [[CrossRef](#)]
14. Zhou, C.; Zhou, J.-F. Direction-of-Arrival Estimation with Coarray ESPRIT for Coprime Array. *Sensors* **2017**, *17*, 1779. [[CrossRef](#)]

15. Wen, F.; Shi, J. Fast direction finding for bistatic EMVS-MIMO radar without pairing. *Signal Process.* **2020**, *173*, 107512. [[CrossRef](#)]
16. Zhou, C.; Gu, Y.; Fan, X.; Shi, Z.; Mao, G.; Zhang, Y.D. Direction-of-Arrival Estimation for Coprime Array via Virtual Array Interpolation. *IEEE Trans. Signal Process.* **2018**, *66*, 5956–5971. [[CrossRef](#)]
17. Liu, T.; Wen, F.; Zhang, L.; Wang, K. Off-Grid DOA Estimation for Colocated MIMO Radar via Reduced-Complexity Sparse Bayesian Learning. *IEEE Access* **2019**, *7*, 99907–99916. [[CrossRef](#)]
18. Zhou, C.; Gu, Y.; He, S.; Shi, Z. A Robust and Efficient Algorithm for Coprime Array Adaptive Beamforming. *IEEE Trans. Veh. Technol.* **2018**, *67*, 1099–1112. [[CrossRef](#)]
19. Wen, F.; Zhang, Z.; Wang, K.; Sheng, G.; Zhang, G. Angle estimation and mutual coupling self-calibration for ULA-based bistatic MIMO radar. *Signal Process.* **2018**, *144*, 61–67. [[CrossRef](#)]
20. Zhou, C.; Gu, Y.; Zhang, Y.D.; Shi, Z.; Jin, T.; Wu, X. Compressive sensing-based coprime array direction-of-arrival estimation. *IET Commun.* **2017**, *11*, 1719–1724. [[CrossRef](#)]
21. Shi, Z.; Zhou, C.; Gu, Y.; Goodman, N.A.; Qu, F. Source Estimation Using Coprime Array: A Sparse Reconstruction Perspective. *IEEE Sens. J.* **2017**, *17*, 755–765. [[CrossRef](#)]
22. Song, J.; Liu, Y.; Shao, J.; Tang, C. A Dynamic Membership Data Aggregation (DMDA) Protocol for Smart Grid. *IEEE Syst. J.* **2020**, *14*, 900–908. [[CrossRef](#)]
23. Vanra, J.; Dhami, S.S.; Pabla, B. Hybrid data fusion approach for fault diagnosis of fixed-axis gearbox. *Struct. Health Monit.* **2017**, *17*, 936–945. [[CrossRef](#)]
24. Li, C.; Sánchez, R.-V.; Zurita, G.; Cerrada, M.; Cabrera, D.; Vásquez, R.E. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mech. Syst. Signal Process.* **2016**, *76*, 283–293. [[CrossRef](#)]
25. Nembhard, A.D.; Sinha, J.K.; Yunusa-Kaltungo, A. Development of a generic rotating machinery fault diagnosis approach insensitive to machine speed and support type. *J. Sound Vib.* **2015**, *337*, 321–341. [[CrossRef](#)]
26. Yunusa-Kaltungo, A.; Sinha, J.; Nembhard, A.D. A novel fault diagnosis technique for enhancing maintenance and reliability of rotating machines. *Struct. Heal. Monit.* **2015**, *14*, 604–621. [[CrossRef](#)]
27. Lei, Y.; Lin, J.; He, Z.; Kong, D. A Method Based on Multi-Sensor Data Fusion for Fault Detection of Planetary Gearboxes. *Sensors* **2012**, *12*, 2005–2017. [[CrossRef](#)]
28. Safizadeh, M.; Latifi, S. Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell. *Inf. Fusion* **2014**, *18*, 1–8. [[CrossRef](#)]
29. Jing, L.; Wang, T.; Zhao, M.; Wang, P. An Adaptive Multi-Sensor Data Fusion Method Based on Deep Convolutional Neural Networks for Fault Diagnosis of Planetary Gearbox. *Sensors* **2017**, *17*, 414. [[CrossRef](#)]
30. Li, Z.; Yan, X.; Wang, X.; Peng, Z. Detection of gear cracks in a complex gearbox of wind turbines using supervised bounded component analysis of vibration signals collected from multi-channel sensors. *J. Sound Vib.* **2016**, *371*, 406–433. [[CrossRef](#)]
31. Duro, J.A.; Padget, J.; Bowen, C.R.; Kim, H.A.; Nassehi, A. Multi-sensor data fusion framework for CNC machining monitoring. *Mech. Syst. Signal Process.* **2016**, *66*, 505–520. [[CrossRef](#)]
32. Khazaei, M.; Ahmadi, H.; Omid, M.; Banakar, A.; Moosavian, A. Feature-level fusion based on wavelet transform and artificial neural network for fault diagnosis of planetary gearbox using acoustic and vibration signals. *Insight-Non-Destr. Test. Cond. Monit.* **2013**, *55*, 323–330. [[CrossRef](#)]
33. Yunusa-Kaltungo, A.; Cao, R. Towards Developing an Automated Faults Characterisation Framework for Rotating Machines. Part 1: Rotor-Related Faults. *Energies* **2020**, *13*, 1394. [[CrossRef](#)]
34. Liu, Z.; Guo, W.; Tang, Z.-C.; Chen, Y. Multi-Sensor Data Fusion Using a Relevance Vector Machine Based on an Ant Colony for Gearbox Fault Detection. *Sensors* **2015**, *15*, 21857–21875. [[CrossRef](#)]
35. Su, Z.; Tang, B.; Liu, Z.; Qin, Y. Multi-fault diagnosis for rotating machinery based on orthogonal supervised linear local tangent space alignment and least square support vector machine. *Neurocomputing* **2015**, *157*, 208–222. [[CrossRef](#)]
36. Li, C.; Sánchez, R.-V.; Zurita, G.; Cerrada, M.; Cabrera, D. Fault Diagnosis for Rotating Machinery Using Vibration Measurement Deep Statistical Feature Learning. *Sensors* **2016**, *16*, 895. [[CrossRef](#)]
37. Jia, F.; Lei, Y.; Lin, J.; Zhou, X.; Lu, N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process.* **2016**, *72*, 303–315. [[CrossRef](#)]
38. LeCun, Y.; Bengio, Y.; Hinton, G. Review: Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]

39. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [[CrossRef](#)]
40. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)]
41. Chandra, B.; Sharma, R.K. Fast learning in Deep Neural Networks. *Neurocomputing* **2016**, *171*, 1205–1215. [[CrossRef](#)]
42. Suk, H.-I.; Lee, S.-W.; Shen, D.; Initiative, T.A.D.N. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage* **2014**, *101*, 569–582. [[CrossRef](#)] [[PubMed](#)]
43. Sun, W.; Shao, S.; Zhao, R.; Yan, R.; Zhang, X.; Chen, X. A sparse auto-encoder-based deep neural network approach for induction motor faults classification. *Measurement* **2016**, *89*, 171–178. [[CrossRef](#)]
44. Tamilselvan, P.; Wang, P. Failure diagnosis using deep belief learning based health state classification. *Reliab. Eng. Syst. Saf.* **2013**, *115*, 124–135. [[CrossRef](#)]
45. Hinton, G.; Mohamed, A.-R.; Jaitly, N.; Vanhoucke, V.; Kingsbury, B.; Deng, L.; Yu, D.; Dahl, G.; Senior, A.W.; Nguyen, P.; et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Process. Mag.* **2012**, *29*, 82–97. [[CrossRef](#)]
46. Shao, H.; Jiang, H.; Zhao, H.; Wang, F. A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. *Mech. Syst. Signal Process.* **2017**, *95*, 187–204. [[CrossRef](#)]
47. Zabalza, J.; Ren, J.; Zheng, J.; Zhao, H.; Qing, C.; Yang, Z.; Du, P.; Marshall, S. Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing* **2016**, *185*, 1–10. [[CrossRef](#)]
48. Xiong, P.; Wang, H.; Liu, M.; Zhou, S.; Hou, Z.; Lin, C. ECG signal enhancement based on improved denoising auto-encoder. *Eng. Appl. Artif. Intell.* **2016**, *52*, 194–202. [[CrossRef](#)]
49. Shao, H.; Jiang, H.; Wang, F.; Zhao, H. An enhancement deep feature fusion method for rotating machinery fault diagnosis. *Knowl.-Based Syst.* **2017**, *119*, 200–220. [[CrossRef](#)]
50. Shikkenawis, G.; Mitra, S. A New Proposal for Locality Preserving Projection. *Comput. Vis.* **2012**, *7143*, 298–305.
51. Ding, X.; He, Q.; Luo, N. A fusion feature and its improvement based on locality preserving projections for rolling element bearing fault classification. *J. Sound Vib.* **2015**, *335*, 367–383. [[CrossRef](#)]
52. Fan, Q.; Gao, D.Q.; Wang, Z. Multiple empirical kernel learning with locality pre-serving constraint. *Knowl.-Based Syst.* **2016**, *105*, 107–118. [[CrossRef](#)]
53. Chen, L.; Zhou, M.; Su, W.; Wu, M.; She, J.; Hirota, K. Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction. *Inf. Sci.* **2018**, *428*, 49–61. [[CrossRef](#)]
54. Du, J.; Xu, Y. Hierarchical deep neural network for multivariate regression. *Pattern Recognit.* **2017**, *63*, 149–157. [[CrossRef](#)]
55. Guo, X.; Chen, L.; Shen, C. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Measurement* **2016**, *93*, 490–502. [[CrossRef](#)]
56. Lu, C.; Wang, Z.-Y.; Qin, W.-L.; Ma, J. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification. *Signal Process.* **2017**, *130*, 377–388. [[CrossRef](#)]
57. Shin, H.-C.; Orton, M.R.; Collins, D.J.; Doran, S.J.; Leach, M.O. Stacked Autoencoders for Unsupervised Feature Learning and Multiple Organ Detection in a Pilot Study Using 4D Patient Data. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1930–1943. [[CrossRef](#)]
58. Nousi, P.; Tefas, A. Deep learning algorithms for discriminant autoencoding. *Neurocomputing* **2017**, *266*, 325–335. [[CrossRef](#)]
59. Galdámez, P.L.; Raveane, W.; Arrieta, A.G. A brief review of the ear recognition process using deep neural networks. *J. Appl. Log.* **2017**, *24*, 62–70. [[CrossRef](#)]
60. Yu, D. Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models. *Mech. Syst. Signal Process.* **2011**, *25*, 2573–2588. [[CrossRef](#)]
61. Bao, S.; Luo, L.; Mao, J.; Tang, D. Improved fault detection and diagnosis using sparse global-local preserving projections. *J. Process. Control.* **2016**, *47*, 121–135. [[CrossRef](#)]

