


Article

Failure Analysis and Intelligent Identification of Critical Friction Pairs of an Axial Piston Pump

Yong Zhu ^{1,2,3} , Tao Zhou ¹, Shengnan Tang ^{4,5,6,*} and Shouqi Yuan ¹¹ National Research Center of Pumps, Jiangsu University, Zhenjiang 212013, China² International Shipping Research Institute, Gongqing Institute of Science and Technology, Jiujiang 332020, China³ Leo Group Co., Ltd., Wenling 317500, China⁴ Institute of Advanced Manufacturing and Modern Equipment Technology, Jiangsu University, Zhenjiang 212013, China⁵ Saurer (Changzhou) Textile Machinery Co., Ltd., Changzhou 213200, China⁶ Wenling Fluid Machinery Technology Institute of Jiangsu University, Wenling 317525, China

* Correspondence: tangsn@ujs.edu.cn; Tel.: +86-188-5286-6635

Abstract: Hydraulic axial piston pumps are the power source of fluid power systems and have important applications in many fields. They have a compact structure, high efficiency, large transmission power, and excellent flow variable performance. However, the crucial components of pumps easily suffer from different faults. It is therefore important to investigate a precise fault identification method to maintain reliability of the system. The use of deep models in feature learning, data mining, automatic identification, and classification has led to the development of novel fault diagnosis methods. In this research, typical faults and wears of the important friction pairs of piston pumps were analyzed. Different working conditions were considered by monitoring outlet pressure signals. To overcome the low efficiency and time-consuming nature of traditional manual parameter tuning, the Bayesian algorithm was introduced for adaptive optimization of an established deep learning model. The proposed method can explore potential fault feature information from the signals and adaptively identify the main fault types. The average diagnostic accuracy was found to reach up to 100%, indicating the ability of the method to detect typical faults of axial piston pumps with high precision.



Citation: Zhu, Y.; Zhou, T.; Tang, S.; Yuan, S. Failure Analysis and Intelligent Identification of Critical Friction Pairs of an Axial Piston Pump. *J. Mar. Sci. Eng.* **2023**, *11*, 616. <https://doi.org/10.3390/jmse11030616>

Academic Editor: Leszek Chybowski

Received: 27 February 2023

Revised: 10 March 2023

Accepted: 12 March 2023

Published: 14 March 2023



Copyright: © 2023 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/>).

Keywords: axial piston pump; fault identification; deep learning; Bayesian algorithm

1. Introduction

An axial piston pump is one of the hydraulic components with the highest technical content in a hydraulic transmission system [1–3]. Axial piston pumps have been widely applied in the transmission system of equipment in many fields, including offshore drilling platforms, drilling machines, and diving equipment in deep sea, as depicted in Figure 1. The structure of an axial piston pump consists of many friction pairs, meaning friction and wear can occur during operation. Some critical friction pairs are illustrated in Figure 2. Any failure will lead to great loss of production and have an effect on the safety and validity of the system [4–6]. Therefore, the efficient fault diagnosis of axial piston pumps is greatly valuable and worth in-depth exploration.

Fault diagnosis aims to determine the cause, position, type, and level of faults and predict the present condition as well as evolutionary trends. It generally includes the detection, isolation, and identification of faults. Fault diagnosis methods are mainly based on signal, model, and knowledge [7–9]. Signal-based methods can decide the fault type and nature by analyzing the time domain, frequency domain, and time–frequency features of original signals. Model-based methods maintain inherent sensitivity to unknown faults, and an accurate mathematical model of the object needs to be constructed for its faults to be diagnosed. Methods based on knowledge achieve fault diagnosis by analyzing and

processing the raw signals. Considering the influence of noise interference on bearing vibration signals, the variational mode extraction method can be used for signal processing. An improved one-dimensional convolutional neural network (1D CNN) was developed for fault diagnosis of rolling bearings by introducing batch normalization, a self-attention layer, and global average pooling [10]. Yu et al. developed a fusion method of raw signals by combining modified empirical wavelet transform and the variance contribution rate [11]. The slight faults of hydraulic pumps can be accurately detected based on fused feature information. By fusing wavelet packet transformation and a new tangent hyperbolic fuzzy entropy measure, Zhou et al. proposed a new method to effectively achieve defect identification of axial piston pumps [12]. Owing to the insufficient analysis of the dynamic characteristics of piston pumps, Tang et al. investigated the effects of external loading and health status by constructing a virtual prototype model [13]. Furthermore, loose slipper failure was effectively detected for different loads. By acquiring the thermal images of a brushless direct current electric motor, three different faulty states were investigated using a new feature extraction method and deep learning methods. High accuracy of around 100% was obtained using the power of normalized image difference method and three deep models [14].

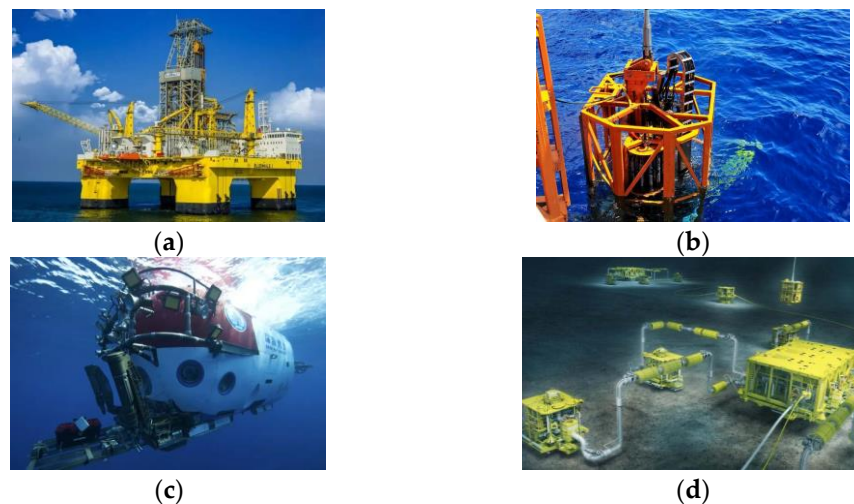


Figure 1. Applications of hydraulic piston pumps. (a) Offshore drilling platform, (b) deep-sea drill, (c) deep-sea manned submersible, (d) deep-sea oil and gas exploitation.

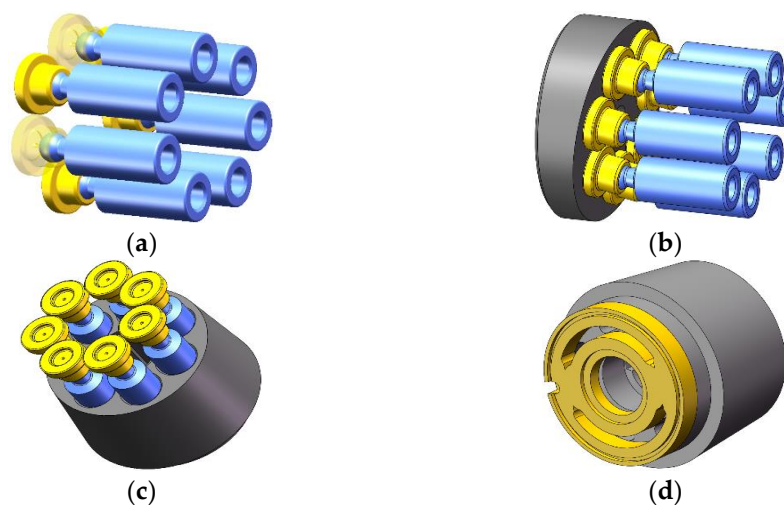


Figure 2. Main friction pairs in an axial piston pump. (a) Piston ball head and slipper, (b) slipper and swash plate, (c) piston and its hole, (d) cylinder and valve plate.

Artificial intelligence is known as a progress of knowledge-based methods and has boosted intelligence of mechanical fault diagnosis [15–17]. Kumar et al. used an artificial neural network to carry out the intelligent detection of a centrifugal pump under varying speeds. The method comprehensively learnt multisource feature information in acceleration, pressure, and motor line current signals and promoted the accuracy of detecting blockage failures [18]. To explore the causes of anomalies in wind turbines, Wang et al. proposed a two-stage decomposition strategy for fault detection and identification. Multi-head self-attention mechanism and position-wise feed-forward network were used. The method presented superior identification performance compared to autoencoder-based methods and had the attention of long short-term memory [19]. Yuan et al. explored an intelligent index on the basis of an enhanced weighted square envelope spectrum and solved the adaptive selection of multiwavelet basis functions. A support vector machine was applied for weight optimization. The results for two bearing datasets indicated that the method was effective for early machinery fault diagnosis [20]. Considering the large amounts of unlabeled data in project practice, Shi et al. constructed a novel fault diagnosis method based on deep hypergraph autoencoder embedding. The method combined the strengths of the extreme learning machine autoencoder in computation and the powerful capability of deep learning approaches in representation learning [21]. The method showed superiority in accuracy and computation of fault diagnosis of rolling bearings and rotors. Because of the severe distribution difference of samples in healthy and faulty conditions, Liu constructed a diagnosis method by integrating a hierarchical extreme learning machine and transfer learning. The method achieved fault identification of the gas path of an aero-engine and showed strong generalization ability [22]. Taking into account the potential complicated and heterogenous feature distribution difference, a strategy called interdomain decision discrepancy minimization was introduced to enhance the transfer learning method [23]. Through experimental validation, it was revealed to be more accurate and robust for fault diagnosis of bearings and gearbox datasets with small and massive samples. Considering the small faulty samples in nuclear power plants, a new method based on lightweight conditional generative adversarial network was developed for fault diagnosis via sample augmentation. The performance of the method was validated using three different public datasets [24]. Given the problems of sparse training data and the reliability of the conclusions in present intelligent diagnosis methods, a new unsupervised learning framework was established [25]. Contrastive learning was used to obtain the common features between the same fault types and distinguish differences between different fault types. A feature-assisted multibranch method was employed as a guidance for the model to weaken the possible interference from operational parameters. The proposed method was especially efficient for bearing fault diagnosis under different speeds. Zhao and Shen proposed a novel diagnosis method based on semisupervised domain generalization for unseen working conditions. Some new samples with pseudo labels were obtained on the strength of the samples with labels, and a sample purification strategy based on the entropy theory was introduced to enhance the restructured samples. Semisupervised fault diagnosis was accomplished for the bearings and gearbox [26]. The method was shown to be effective by analysis of the computation complexity and performance in noisy environments. However, the automatic sample selection rate and its generalization for different machines need to be further explored. To deal with unreliability from unseen faults, an ensemble diagnosis method was developed by combining five different CNNs, CNNs with different convolutional layers, ResNet, and inception network. The constructed method could identify unknown faults, and was found to be trustworthy for fault diagnosis of wind turbines and gearboxes [27]. To determine possible misdiagnosis, Zhou et al. constructed a probabilistic Bayesian CNN framework for more accurate fault identification. The method could complete diagnosis of visible bearing faults and was also verified to be valid for unseen gear faults under normal and noisy conditions [28]. Due to the multiple parameters and time-consuming model optimization, a multihierarchy compound network compression approach was used for bearing fault diagnosis [29]. Unlike traditional CNNs,

structured and unstructured pruning was carried out by removing unimportant filters in the convolutional layer and inconsequential connections in the fully connected layer. The parameters were reduced while achieving equal recognition accuracy. Based on denoising time–frequency images, Chao et al. explored the identification of cavitation severity on a high-speed aviation hydraulic pump using a CNN-based intelligent method [30]. Driven by acoustic signals, Kumar et al. investigated fault identification of a centrifugal pump employing a CNN for feature extraction and classification [31]. The experiments indicated increased accuracy and powerful generalization capability.

Most of the present literature concerns the use of fault diagnosis methods for bearing, gearing, and wind turbines. In an earlier study, a deep model named two-layer CNN was constructed for fault diagnosis of axial piston pumps based on vibration signals [32]. In this research, similar defect types and degree of damage of key friction pairs of axial piston pumps were considered. A new deeper CNN network structure with five convolutional layers was built for fault identification. The innovations are embodied in the following three points.

(1) The structure of a hydraulic axial piston pump is composed of many friction pairs and is very complex. This makes it difficult to achieve fault monitoring and diagnosis, and it is much harder to accomplish intelligent diagnosis process. This research provides a new intelligent method for fault diagnosis of the key friction pairs of axial piston pumps.

(2) Deep learning methods are employed for fault diagnosis of typical rotating machinery. However, the construction of the current deep models has shortcomings, such as its time-consuming nature and inefficient manual parameter tuning. This research utilizes the strengths of the Bayesian optimization algorithm and completes automatic learning of the model hyperparameters.

(3) Many fault diagnosis methods are used for machinery vibration. Some are implemented based on acoustic emission technology. Taking into account the pressure change associated with component failure and easy monitoring of the signal, this study probes into the monitored outlet pressure signals and dissects the hidden fault characteristics from a special angle.

The rest of the paper is structured as follows. A brief introduction of CNN and Bayesian optimization (BO) is provided in Section 2. Section 3 describes the implementation of the proposed fault identification method. The experimental setup and data acquisition are outlined in Section 4. In Section 5, the diagnosis accuracy and performance of the method are analyzed and discussed. Finally, a conclusion is drawn and prospects for future research are discussed in Section 6.

2. Theoretical Background

2.1. Convolutional Neural Network

A typical fully connected neural network can generally result in information loss, and lots of parameters lead to overfitting. With the progress of deep models, CNN is gradually exhibiting superiority in processing large-scale images [33–35]. In general, the basic structure of a CNN consists of an input layer, convolution layer, pooling layer, fully connected layer, and output layer. Convolution layer, pooling layer, and fully connected (FC) layer are collectively called hidden layers. The connection between the convolutional layer and the subsampled layer is a local connection. A fully connected mode is adopted in the fully connected layer. A CNN presents three distinguished characteristics: local connection, weight sharing, and downsampling. The local receptive field is a particular structure that endows CNN with the function of local connection and weight sharing.

The calculation of convolutional layer is as follows [36],

$$M_m = A(I \cdot W_m + b_m) \quad (1)$$

where A denotes a kind of nonlinear activation function; m denotes the m^{th} feature map; I is the input; \cdot is used as the convolution operator; and the new feature map, weight, and bias are denoted as M , W , and b , respectively.

The classification is achieved by the Softmax function, which can convert the output values of multiclass into probability distributions in the range $[0, 1]$ and 1. It can be expressed as follows [37]:

$$\text{Output} = \text{Softmax}(W_n x + b_n) = \frac{e^{(W_n x + b_n)}}{\sum_1^P e^{(W_o x + b_o)}} \quad (2)$$

where the input of the output is presented as x ; n denotes one class; and P is the number of output nodes, also known as the number of categories.

Figure 3 presents the convolutional operation. The convolutional kernel, known as a filter, conducts relative operation in the receptive field for feature extraction [38]. The pooling layer shown in Figure 4 can reduce dimension and parameters. It can complete downsampling using max-pooling or average pooling.

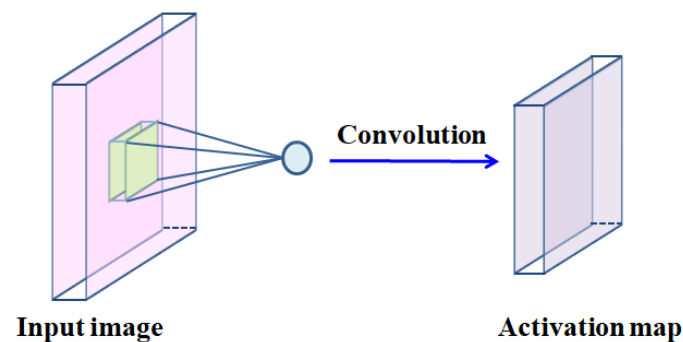


Figure 3. Convolutional layer.

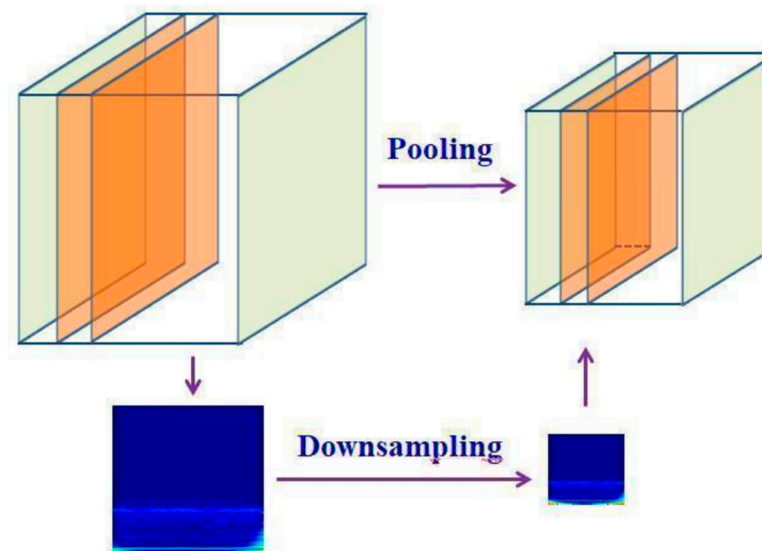


Figure 4. Pooling layer.

To reduce the interdependence of the parameters and achieve sparsity of the network, a rectified linear unit layer (ReLU) is usually used together with a convolutional layer [39]. The ReLU function is shown in Figure 5. The FC layer is located at the end of a CNN, and each node is fully interconnected with the nodes in the front layer. It integrates the features extracted from the previous layer and maps them to the label space. The structure of the FC layer is presented in Figure 6.

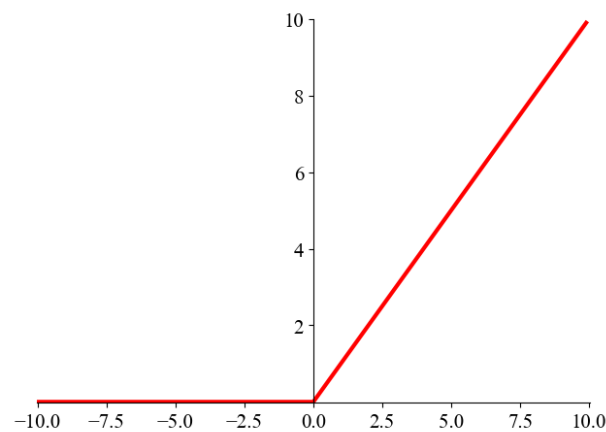


Figure 5. ReLU function.

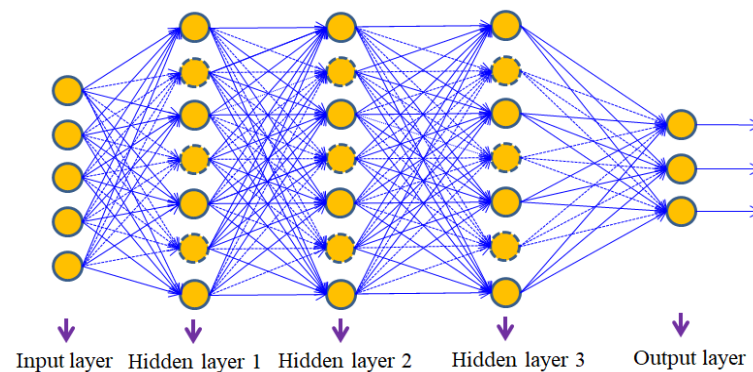


Figure 6. Fully connected layer.

2.2. Bayesian Algorithm

Machine learning algorithms contain numerous parameters. The ones that can be optimized through training are called parameters, such as the weight in neural networks, while those that cannot be optimized by training are called hyperparameters, such as the learning rate (LR), neurons per layer, and batch size. It should be noted that parameters are updated, while hyperparameters are always constant in training. It is essential to adjust the hyperparameters to acquire more efficient machine learning models [40,41].

There are two major challenges in hyperparameter optimization. First, it is a combinatorial optimization problem that is hard to complete using gradient descent. Second, it is time-consuming to evaluate a group of hyperparameters, which makes some evolutionary algorithms inapplicable. Grid search aims to search for a suitable configuration by trying all the possible combinations of a hyperparameter. Random search chooses the best configuration via random combinations of the hyperparameters. Compared to grid search, it can prevent unnecessary attempts at optimizing unimportant hyperparameters. The limitation is that the two search-based methods ignore the relationship between different hyperparameter combinations.

Bayesian optimization is an adaptive method for hyperparameter selection. BO can predict the next possible combination that can bring the greatest benefits according to the current tried groups [42]. BO includes prior information of parameters and shows fast convergence. BO is aimed at optimal hyperparameter combinations that can obtain the best model performance. BO uses two important functions: Gaussian process (GP) and acquisition function. GP can effectively deal with nonlinear tasks and is able to fit the objective function in optimization.

The following formula represents a multivariate Gaussian distribution [43,44]:

$$f(p_{1:k}) \sim N(a(p_{1:k}), c^\theta(p_{1:k}, p_{1:k})) \quad (3)$$

where f signifies a smooth function, a finite input collection is expressed as $p_{1:k}$, $a(p_{1:k})_j = a(p_j)$ presents a mean vector, $c^\theta(p_{1:k}, p_{1:k})_{ji} = c^\theta(p_j, p_i)$ denotes covariance matrix, and the parameterization is accomplished by θ .

The mean value and variance of samples are two critical statistics in Gaussian distribution. Exploitation refers to choosing points where the mean value is larger, while exploration refers to selection of points where the variance is larger. BO aims to balance exploitation and exploration. An appropriate point should be a tradeoff of exploitation and exploration. Acquisition function is considered as a suitable choice for sampling. Among many kinds of acquisition functions, expected improvement (EI) based on improvement is easy and effective. It does not have many parameters and has simple computation. The formula of EI is indicated as follows [44]:

$$\alpha_{EI}(D; \sigma, l) = E[M(0, y(D) - y(D^*))] \quad (4)$$

where l denotes a dataset, $y(D)$ presents the best objective that can be called the posterior expected value, and $y(D^*)$ shows the current best result or the maximum that has been encountered so far.

EI combined with GP is as follows:

$$\alpha_{EI}(D; \sigma, l) = \beta_t(D; \sigma, l)[x\Phi(x) + \phi(x)] \quad (5)$$

where $x = \frac{\mu_t(D; \sigma, l) - y(D^*)}{\beta(D; \sigma, l)}$, Φ denotes the cumulative distribution function, and ϕ presents the probability density function that conforms to the standard normal distribution.

EI with stochastic noise can be expressed as follows:

$$\alpha_{EI}(D; \sigma, l) = \beta(D; \sigma, l)[t\Phi(t) + \phi(t)] \quad (6)$$

where $t = \frac{\mu_t(D; \sigma, l) - \mu(\sigma^*)}{\beta(D; \sigma, l)}$, and $\mu(\sigma^*)$ signifies the most desired results that were obtained by calculating the mean value.

3. Proposed Diagnosis Method

A refined CNN model was constructed based on the traditional AlexNet (T-AlexNet) [45]. The strategies of data transform and dropout [46] and Adam algorithm were employed in the establishment of the model [47]. The proposed identification method combines a modified CNN model and Bayesian algorithm. The flowchart is displayed in Figure 7, while a more visualized account is depicted in Figure 8. The method can be divided into two steps.

The first step is called signal-to-image, where raw pressure signals are acquired with a pressure sensor. The obtained time series are converted into images using continuous wavelet transform. The time–frequency images are input into the deep model followed by data transform processing.

The second step is called feature extraction and identification, where a common CNN model is built based on T-AlexNet with initial hyperparameters. The important model hyperparameters are optimized using the Bayesian algorithm. LR, epoch, batch size, dropout ratio, convolutional kernels, and other features are included. The best hyperparameter groups are obtained through optimization, and the model achieves high performance. The Bayesian optimized AlexNet (B-AlexNet) is used for adaptive failure identification of an axial piston pump based on the above feature images.

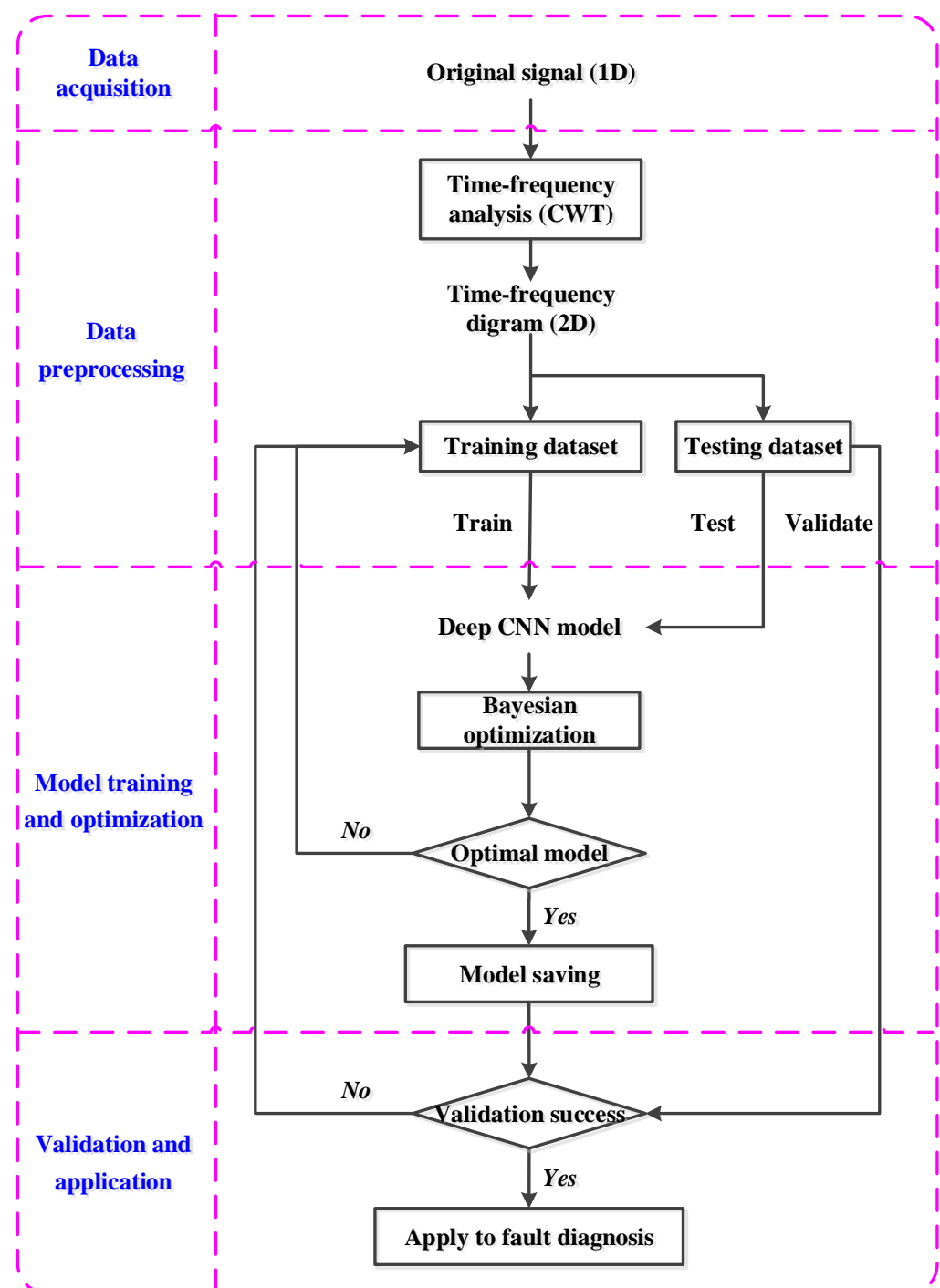


Figure 7. Flowchart of the proposed method.

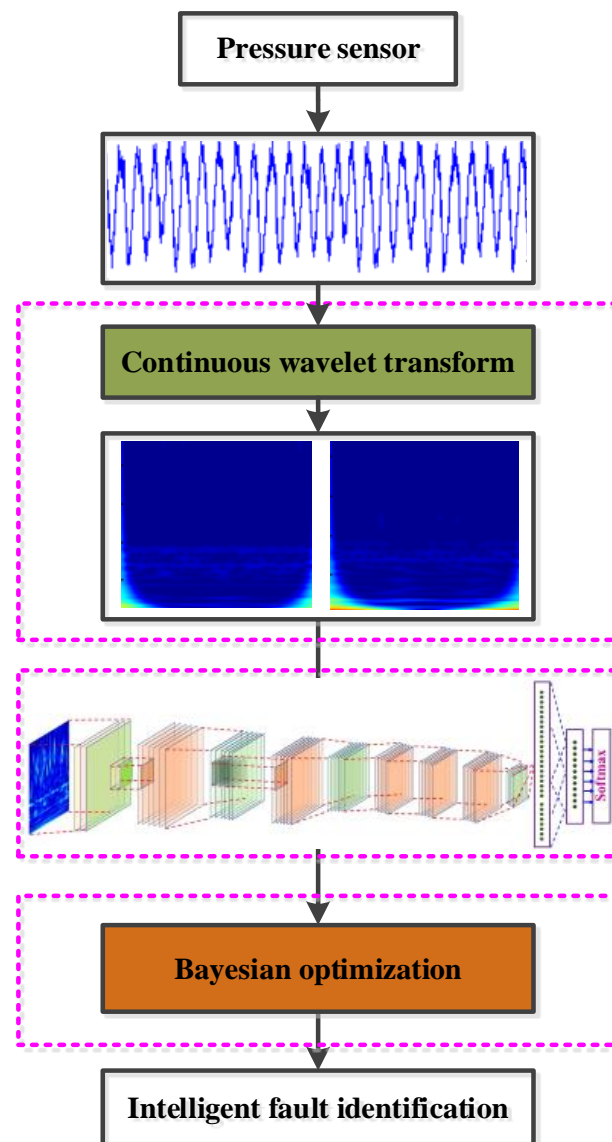


Figure 8. Proposed diagnosis method.

4. Design of the Experimental Bench

The bench of a swash plate plunger pump was used for fault simulation tests, as shown as Figure 9. It contained a Y132M-4 motor and an MCY14-1B axial piston pump with seven plungers. In the experiment, three types of sensors were installed: vibration sensors in three different directions, pressure sensor, and sound sensor. Data from different sensors were acquired. This research was conducted based on the analysis of pressure signals. We performed sampling with a frequency of 10 kHz.

The nominal pressure of a piston pump is 31.5 MPa. The motor has a constant rotating speed of 1470 r/min. The pressure sensor is an instrument of SYB-351 that can display parameters. It can measure from 0 to 25 MPa. This is mainly because the actual working pressure of the pump is generally lower than the nominal pressure. The pump outlet pipeline is equipped with a pressure sensor to monitor the outlet pressure signal of the pump. Five conditions were analyzed in the research: wear of the slipper (hx), wear of the swash plate (xp), failure of the loose slipper (sx), wear of the central spring (th), and no visible failure (zc). Different pressures and wear degrees were examined.

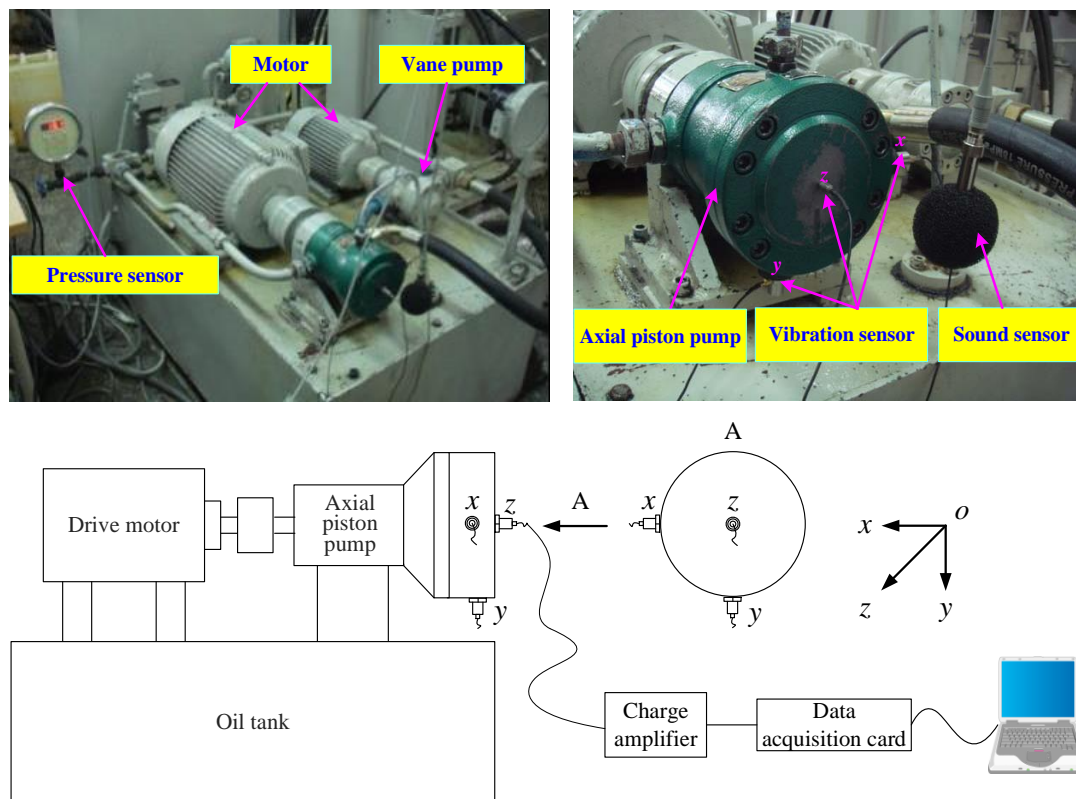


Figure 9. Test bed of an axial piston pump.

5. Results and Discussion

5.1. Signal Analysis

The sample length in the experiment was 1024. The complex wavelet was taken as the wavelet basis function. The scale sequence length was 256. The bandwidth and center frequency were both 3. Figure 10 shows the time–frequency diagrams of the pressure signal. The time–frequency images were used to build the dataset. Each fault type consisted of 1200 samples. The whole dataset was composed of 6000 images with the size of 256×256 . The images were resized into 224×224 , and random horizontal flip was performed. The training dataset accounted for 70%, and the remaining 30% was the test dataset. The frequency changed with time under each condition. The features were similar, and different fault modes were hard to identify. The visual information was almost the same in the states of slipper wear and loose slipper failure. Establishing a learning model to mine the valid information is important to complete fault classification of axial piston pumps.

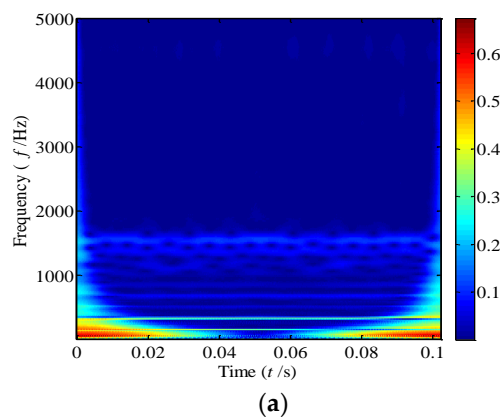


Figure 10. Cont.

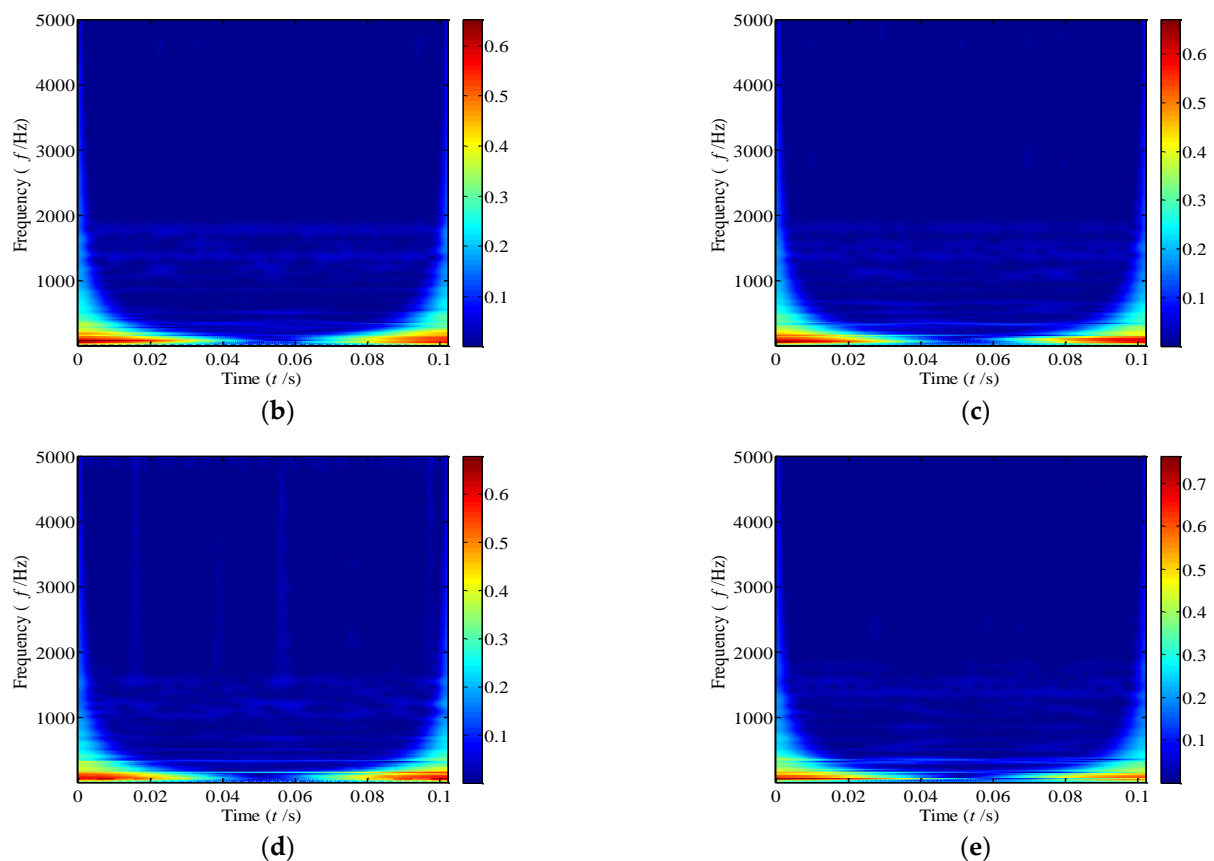


Figure 10. Time–frequency distributions of pressure signal via continuous wavelet transform. (a) Normal condition, (b) wear of the slipper, (c) failure of the loose slipper, (d) wear of the swash plate, (e) wear of the central spring.

5.2. Identification Results

5.2.1. Analysis of LR

The range of LR was set from 0.0001 to 0.0009. Then, optimization of the model was conducted according to the decreasing principle of LR, which was 0.0009, 0.0008, 0.0006, 0.0004, 0.0002, and 0.0001 in the experiments. A small loss was found in low LR, as shown in Figure 11. The lowest loss was found when LR was 0.0001. The rate of reduction became slow with the increase in LR. The loss decreased to the minimum value fastest when the LR was 0.0002. It fluctuated with the increase in epochs when the LR was 0.0009. Therefore, it can be seen that a large LR is detrimental to the training of this model. The change in test accuracy followed an opposite pattern compared to the loss. It reached lowest when the LR was 0.0009 and highest when it was 0.0001. The accuracy curve presented good convergence. Therefore, the initial LR was chosen as 0.0001.

The classification accuracy in 10 tests was analyzed for each LR. As can be seen from Figure 12, the improved AlexNet (I-AlexNet) presented relative stability with an LR of 0.0001. The accuracy fluctuated to varying degrees with small or large LR, especially when it was 0.0009, 0.0008, and 0.0004. Therefore, for pressure signals, the I-AlexNet showed good stability as the LR was 0.0001.

The average accuracy for 10 tests under different LR is shown in Figure 13. The average identification accuracy of 10 experiments reached the highest level of 99.99% when the LR was 0.0001. Combined with the above analysis on model convergence, the model showed great performance under this LR.

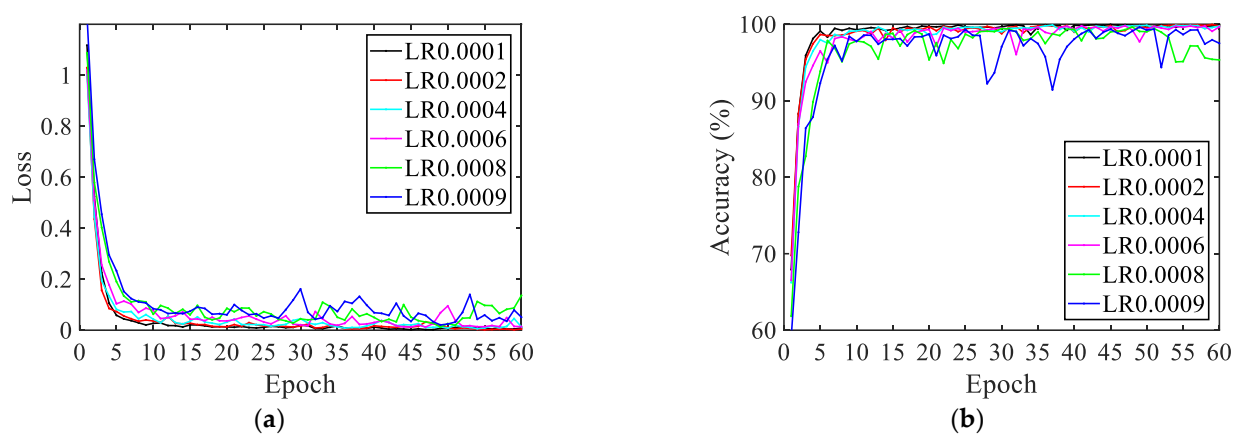


Figure 11. The curves of loss and test accuracy. (a) Loss, (b) accuracy.

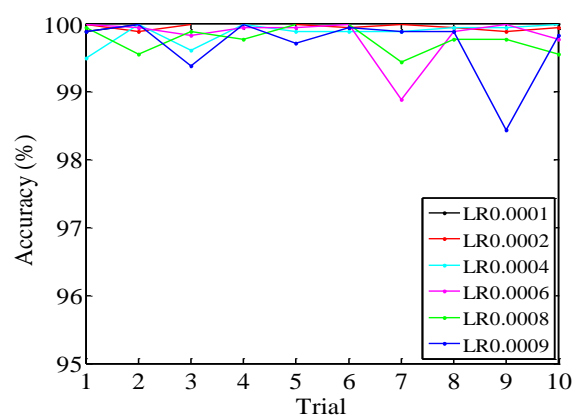


Figure 12. The accuracy in 10 trials.

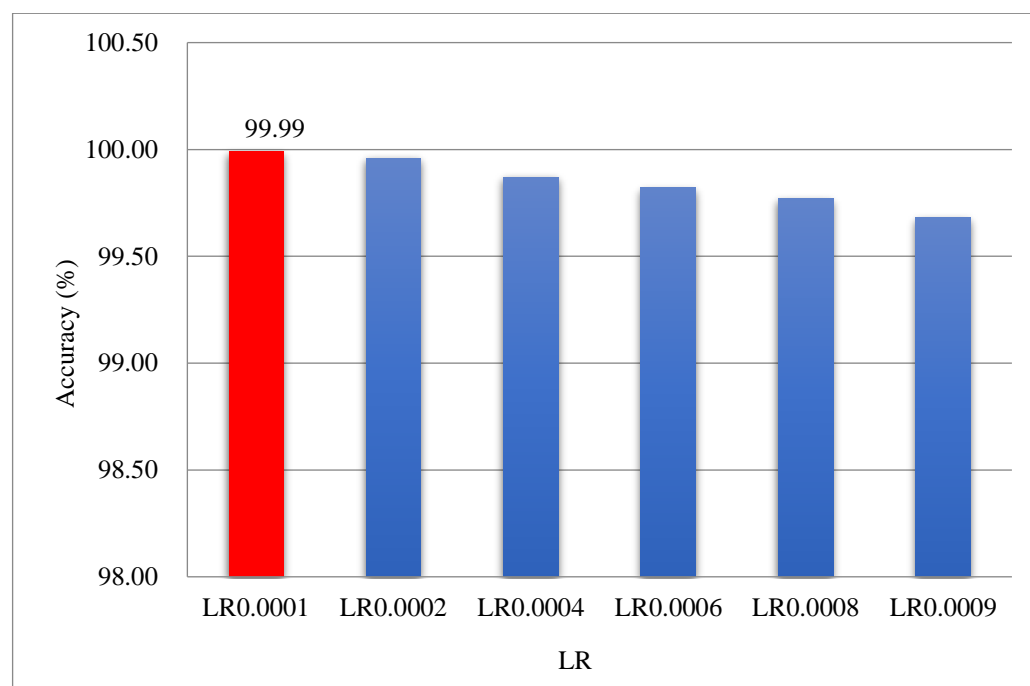


Figure 13. The average accuracy with different LRs.

5.2.2. Analysis of Epoch

When analyzing the influence of epochs on fault classification, only the epochs were changed and other parameters were fixed. The batch size was 32, and the initial value of the size and number of convolution kernels was the same as that in Section 5.2.1. The initial LR was selected as 0.0001. Ten independent trials were carried out by selecting the epoch as 60. The average accuracy is shown in Figure 14.

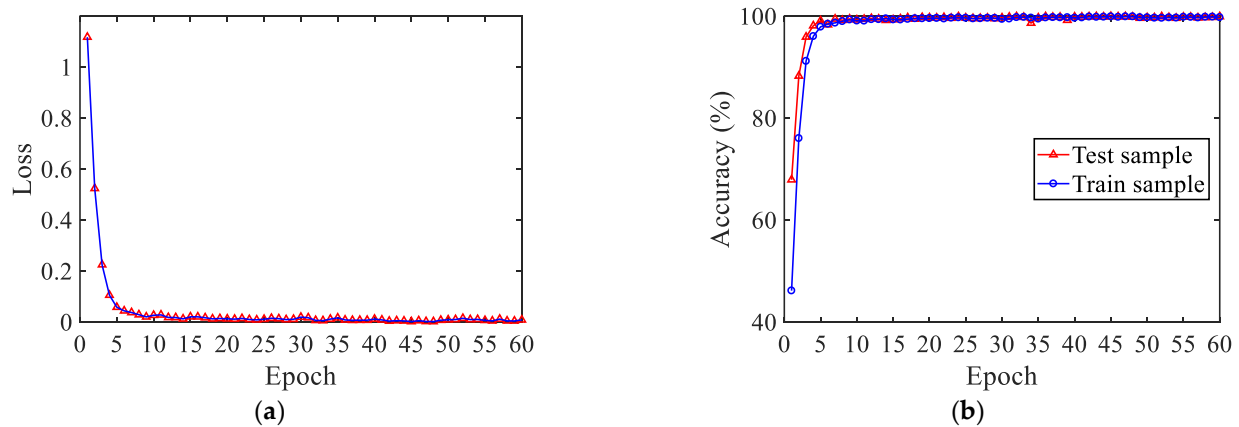


Figure 14. Tendency of loss and accuracy. (a) Loss, (b) accuracy.

The training loss gradually decreased and the accuracy showed increasing trend with increasing epochs. The test accuracy was slightly higher than that of the training samples in less than 10 epochs. The accuracy exceeded 99% after more than 10 epochs. The loss was stable after more than 30 epochs, indicating that the neural network had been trained to convergence. Therefore, the epoch of 30 was found to be suitable.

5.2.3. Analysis of Dropout Rate

In the model training stage, the dropout strategy was employed in the full connection layer. The range was set as 0.1 to 0.9 for the analysis of the model performance 0.1, 0.3, 0.5, 0.7, and 0.9, respectively. The accuracy curves under different dropout ratios are shown in Figure 15. The loss showed an increasing tendency as the dropout ratio increased. In contrast, the accuracy decreased gradually with the increase in dropout. The model represented better convergence under small ratios with the increase in epoch, and a small difference was found among the maximum accuracy of test samples under different ratios.

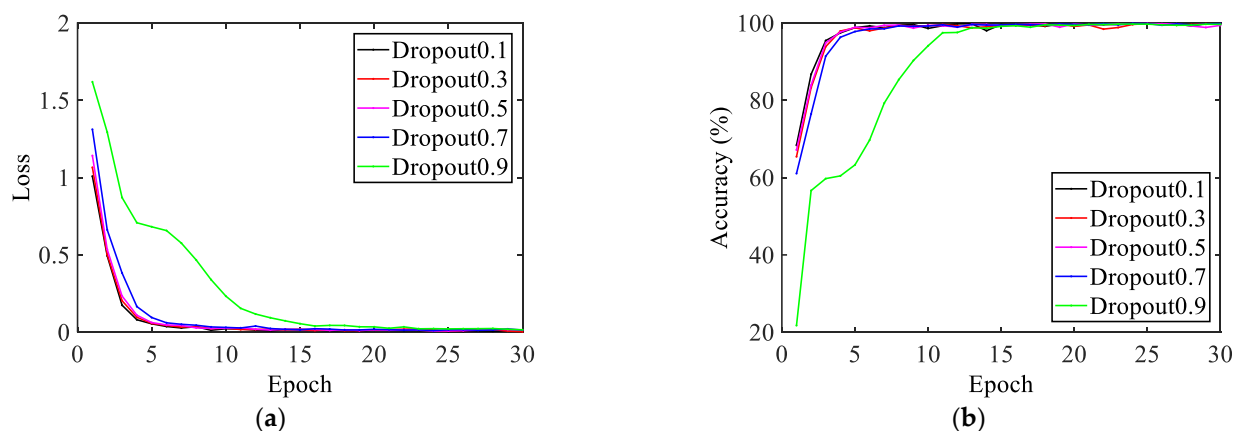


Figure 15. Trend of loss and accuracy. (a) Loss, (b) accuracy.

The average accuracy of 10 trials under different dropout ratios is displayed in Figure 16. There was no obvious difference in the classification accuracy with different dropout ratios. The average accuracy was slightly higher when the dropout was 0.5. Considering the convergence analysis and the most network structures to be obtained, the dropout ratio of 0.5 was selected.

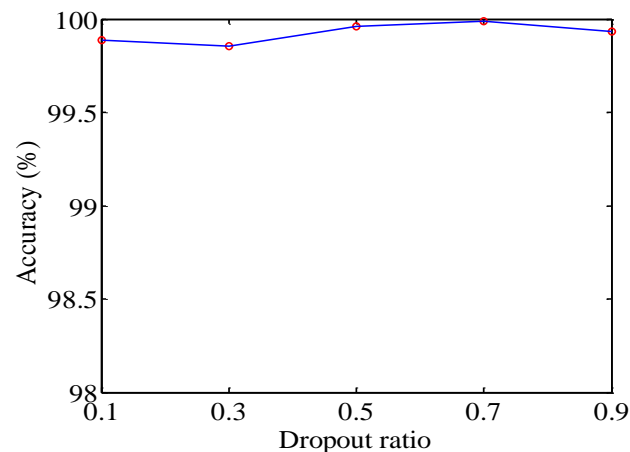


Figure 16. Classification accuracy under the different dropout ratios.

5.2.4. Effect of Batch Size on Diagnostic Accuracy

The epoch was set as 30, and the batch sizes were 8, 16, 21, 32, 42, 56, 64, and 84. The average accuracy of 10 independent repeated tests is shown in Figure 17. Little difference in accuracy was found among different batch sizes, and the average was more than 99.9%. When the batch size was 21, the classification accuracy reached 100%. The model achieved good recognition for different fault categories of the axial piston pump.

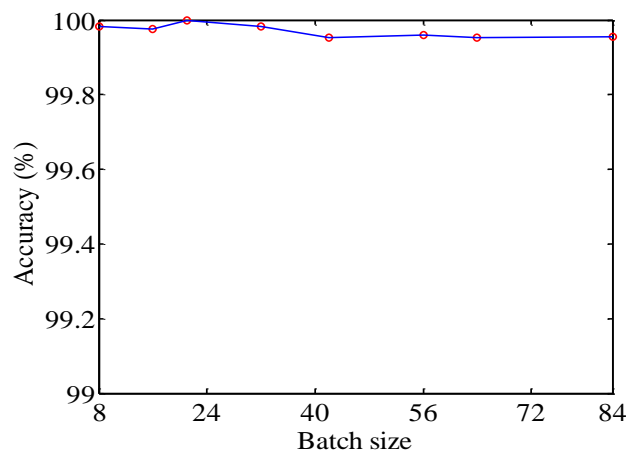


Figure 17. Relationship between classification accuracy and batch size.

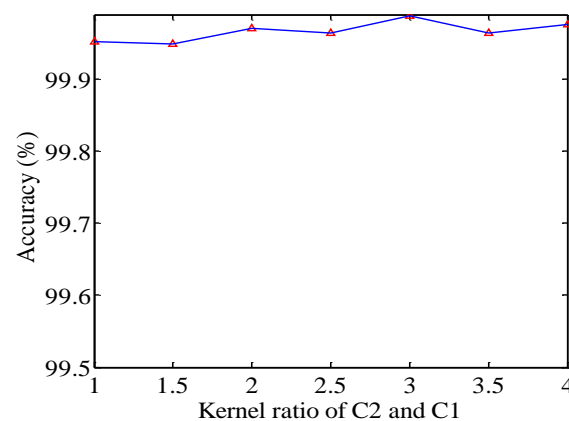
As there was no significant difference in accuracy of models at different batch sizes, the calculation time was analyzed. As can be seen from Table 1, when the batch size was small, the calculation time was long. When it increased to more than 32, the time difference of each epoch was smaller. For the present sample set, a smaller batch size was conducive to improving the accuracy of fault classification. It was more beneficial to the judgment of the final fault category. Therefore, the batch size of 32 was found to be suitable according to comprehensive analysis of classification accuracy and calculation cost.

Table 1. Computational time of different batch sizes in one epoch.

Batch Size	1	2	3	4	5	6	7	8	9	10	Time (s)
8	16.16	16.08	16.17	16.27	16.10	16.20	16.02	16.20	16.16	16.12	16.15
16	12.12	12.93	12.01	11.99	12.37	12.53	12.23	12.15	12.00	12.06	12.24
21	11.94	12.04	11.77	11.82	11.79	11.89	11.75	11.77	11.72	11.74	11.82
32	10.40	10.33	10.38	10.34	10.63	10.513	10.37	10.47	10.47	10.58	10.45
42	10.09	10.02	10.08	10.15	10.05	10.20	10.03	10.13	9.94	10.19	10.09
56	9.78	9.90	10.05	10.06	9.80	9.89	9.93	9.92	10.09	10.09	9.95
64	9.97	9.82	9.97	9.78	9.77	9.70	9.94	9.96	10.28	9.90	9.91
84	9.99	9.83	9.54	9.69	10.28	10.26	10.02	9.91	9.79	9.97	9.93

5.2.5. Effect of Kernel Number on Diagnostic Accuracy

The epoch was set as 30 and the batch size was set as 32 in the experiment. Meanwhile, the other parameters remained unchanged. Figure 18 depicts the accuracy at different kernel ratios. The accuracy of the model was more than 99.9%. There was a slight increase in accuracy with the increase in ratio. The accuracy achieved up to 99.97% when the ratio was 2. There was no obvious change when the ratio increased to more than 3 times. Due to the insignificant difference in accuracy with different kernel rates, the time consumption was analyzed. As shown in Table 2, a small change in average time was found with the increase in kernel number of C2.

**Figure 18.** Classification accuracy by different kernel number.**Table 2.** Computational time of different kernel numbers in one epoch.

C2/C1	1	2	3	4	5	6	7	8	9	10	Time (s)
1	10.21	10.65	11.01	10.32	10.57	10.73	10.56	11.42	10.38	10.58	10.64
1.5	11.22	11.17	11.01	11.14	11.15	11.00	11.17	11.12	11.03	11.03	11.10
2	11.12	11.19	10.99	11.26	11.09	11.06	10.91	10.93	11.05	11.21	11.08
2.5	10.66	10.91	10.98	11.00	10.90	10.95	10.88	10.65	10.85	10.75	10.85
3	11.89	11.98	12.05	10.89	10.96	10.95	11.38	11.55	11.03	11.02	11.37
3.5	11.20	11.14	11.05	11.22	11.042	11.08	10.96	11.034	10.91	11.16	11.08
4	11.41	11.14	11.23	11.09	11.02	11.10	10.90	11.03	11.109	11.35	11.14

5.2.6. Effect of Kernel Size on Diagnostic Accuracy

The influence of kernel size on accuracy was also examined, as shown in Figure 19. The highest test accuracy was obtained when the combination is k(9,5). The kernel size was selected as 9×9 for C1 and 5×5 for C2.

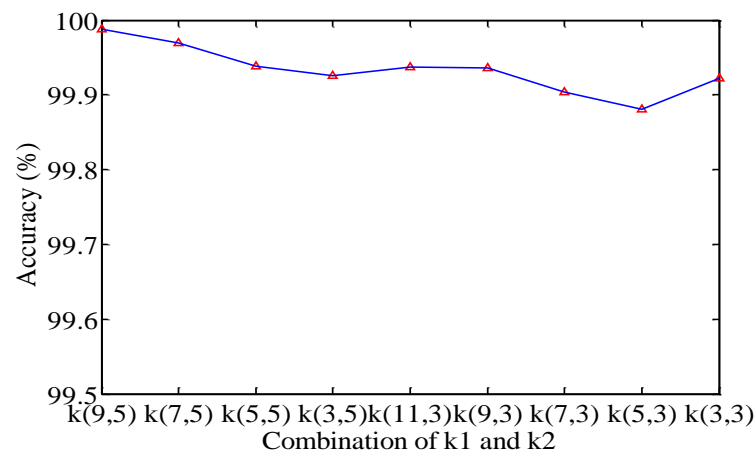


Figure 19. Classification accuracy by different kernel size.

As can be seen from the above analysis, different hyperparameters had different effects on the fault diagnosis result. An obvious difference could be found from the analysis of the LR and the kernel size. As for the other parameters, such as the epoch, the batch size, and the kernel number, there was no significant influence on the diagnosis performance. In particular, only a slight difference was observed with different dropout ratios.

5.2.7. Validation of Diagnostic Model

The optimal parameters obtained above were as follows: convolutional kernel, C1 and C2; number, 48 and 96; size of k1, 9×9 ; size of k2, 5×5 . The batch size was 32, the epoch was 30, the dropout ratio was 0.5, and the initial LR was 0.0001. To verify the reliability and stability of parameters, 10 independent tests were carried out. As shown in Figure 20, the difference between the test and training accuracy was not significant, and the test accuracy was above 99.9%. The stable results proved the effectiveness of the model structure and parameters.

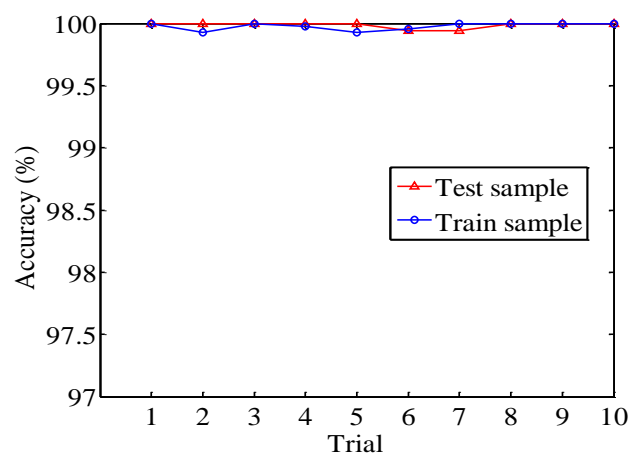


Figure 20. Curves of training and testing accuracy in 10 trials.

Dimensionality reduction was achieved using t-stochastic neighbor embedding (*t*-SNE). The visualization results are shown in Figure 21. The features of the five states in the original input signal were evenly distributed, and it was difficult to complete identification of the fault categories at this stage. The characteristics of different fault classes became much easier to distinguish from the initial convolution layer to the final full connection layer.

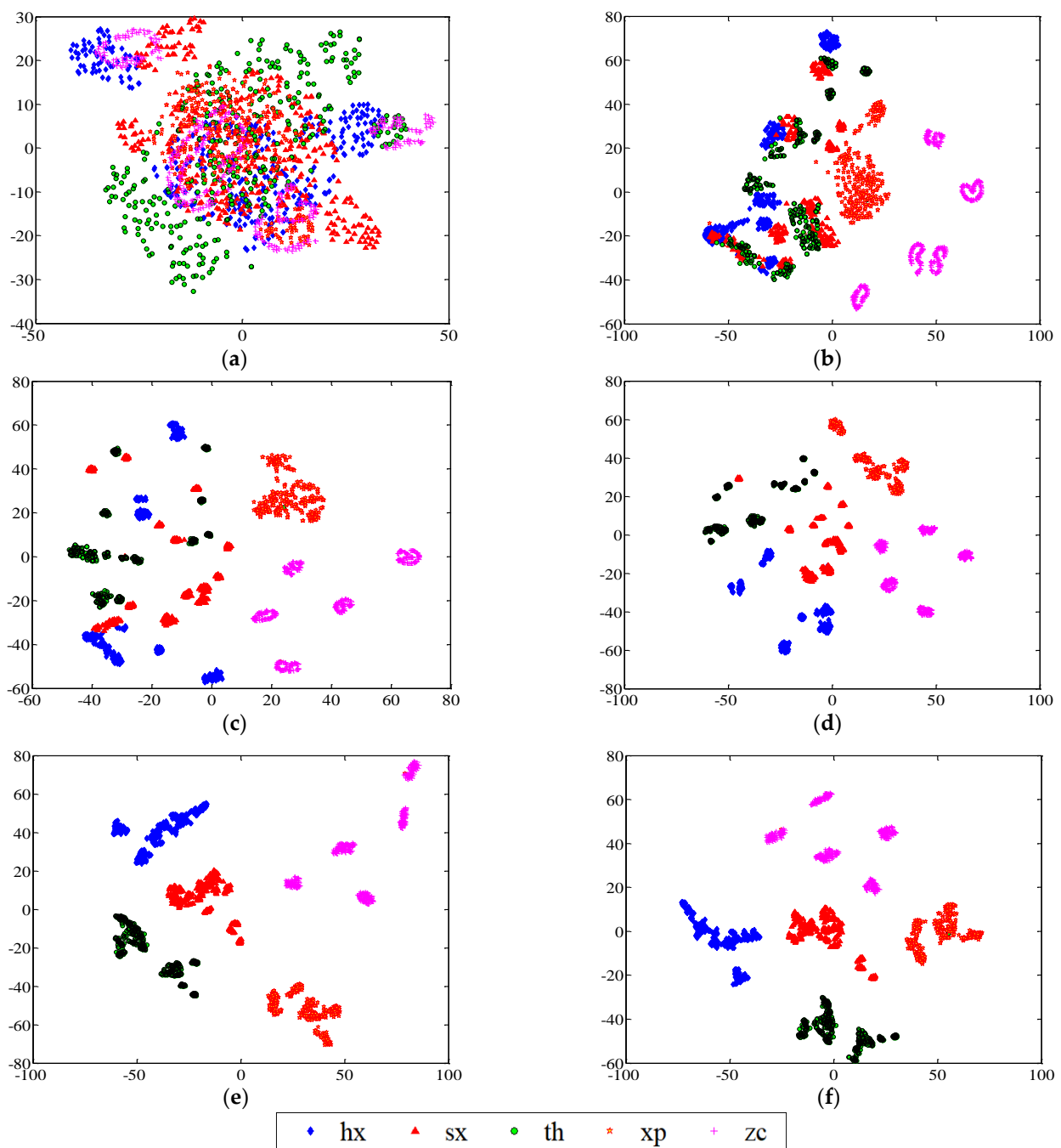


Figure 21. Visualization of feature representation via *t*-SNE. (a) Input data, (b) C1, (c) C3, (d) C5, (e) FC2, (f) classifier layer.

There was some overlapping among the feature distributions of different fault types when the model was shallow. The mixing phenomenon was more obvious through the learning of C1, and it could be seen that misclassification of the different fault types existed at this stage. The features of th gradually gathered after learning of C3, while the features of sx and hx were mixed with each other. The features of each fault type showed an aggregated trend by the extraction of C5. There was still some overlapping between the feature distributions of th and sx. The features of the different categories in the FC layer could be clearly identified. Although there was some distance among the features of the same fault type, the features of one type were collected in the same area. The feature representations of low level were transformed into those of high level. The identification

and classification of the signal were effectively accomplished for an axial piston pump via the feature learning of different network layers.

To explore the diagnostic performance of the proposed model, 10 independent tests were conducted and different CNN models were used for comparative analysis. As can be observed from Figure 22, the I-AlexNet converged better than other models. In the initial stage of feature extraction, the accuracy of the LeNet-based model was lower. When the epochs were more than 10, the accuracy of I-AlexNet was more than 99%, but the accuracy of T-LeNet was still less than 90%. After 30 epochs of training, I-AlexNet, T-AlexNet, 3-CNN, and 3-CNN all achieved high classification accuracy.

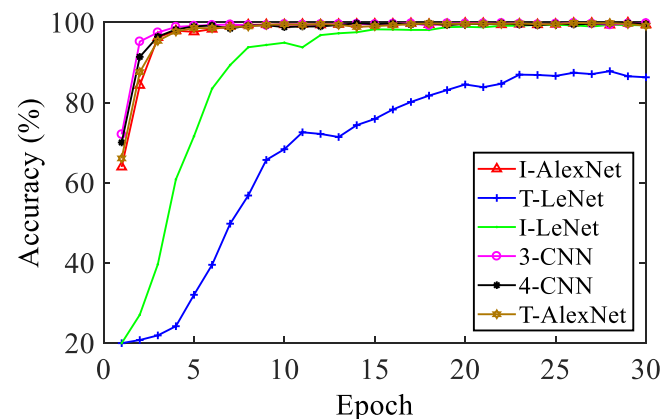


Figure 22. Curve of testing accuracy for different CNN models.

As can be seen from Table 3, the average accuracy of I-AlexNet reached 99.99%. The average accuracy of T-LeNet was only 94.06%, while the other models achieved more than 99%. Comparing VGG11 and deep network structure, the average accuracy and standard deviation were the same. This further confirmed the recognition performance of I-AlexNet. I-AlexNet showed good classification effect and stability in implementing fault identification of an axial piston pump.

Table 3. Average classification accuracy of different CNN models.

Model	Average Accuracy (%)	Standard Deviation
T-LeNet	94.06	0.08442
I-LeNet	99.61	0.001624
3-CNN	99.92	0.001342
4-CNN	99.93	0.001135
VGG11	99.99	0.0002530
T-AlexNet	99.94	0.001386
I-AlexNet	99.99	0.0002530

The confusion matrix of one trial is shown in Figure 23. As for I-LeNet and T-LeNet, the categories of recognition errors were mainly concentrated in hx, sx, and th. The identification of xp was obviously improved for I-LeNet. The misrecognition for sx and th was found in 3-CNN and 4-CNN. As a much deeper model, VGG11 achieved accurate classification. Misclassification was observed for sx and th faults using the T-AlexNet model. Among them, six samples in sx were misclassified as hx, and two samples in th were misclassified as sx. The I-AlexNet model had no misclassification phenomenon for any of the fault categories. The classification accuracy of each category is displayed in Table 4. The precision of I-AlexNet for the five states of an axial piston pump reached 100%. The classification precision of sx and hx was improved by the model improvement compared to the traditional model.

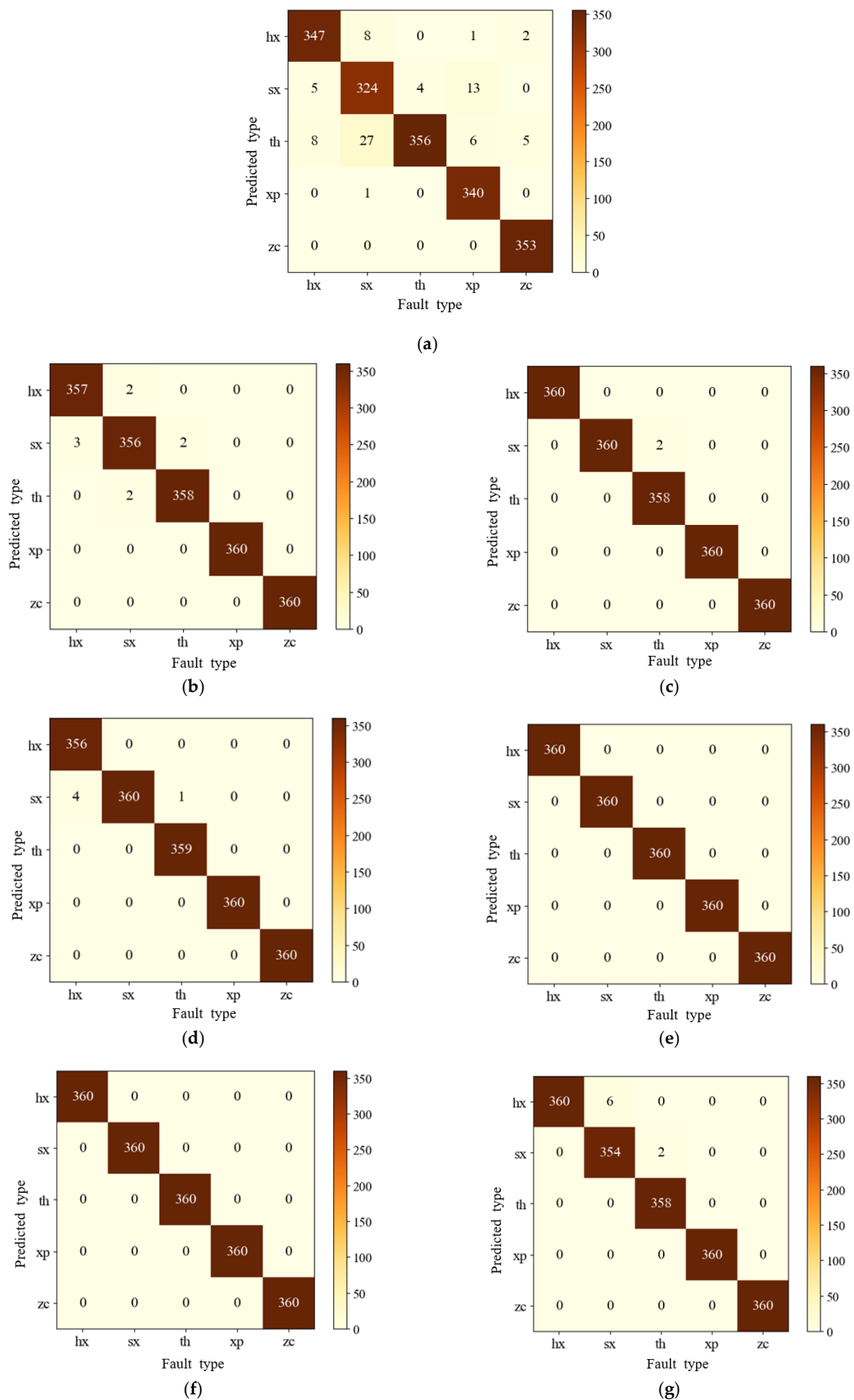
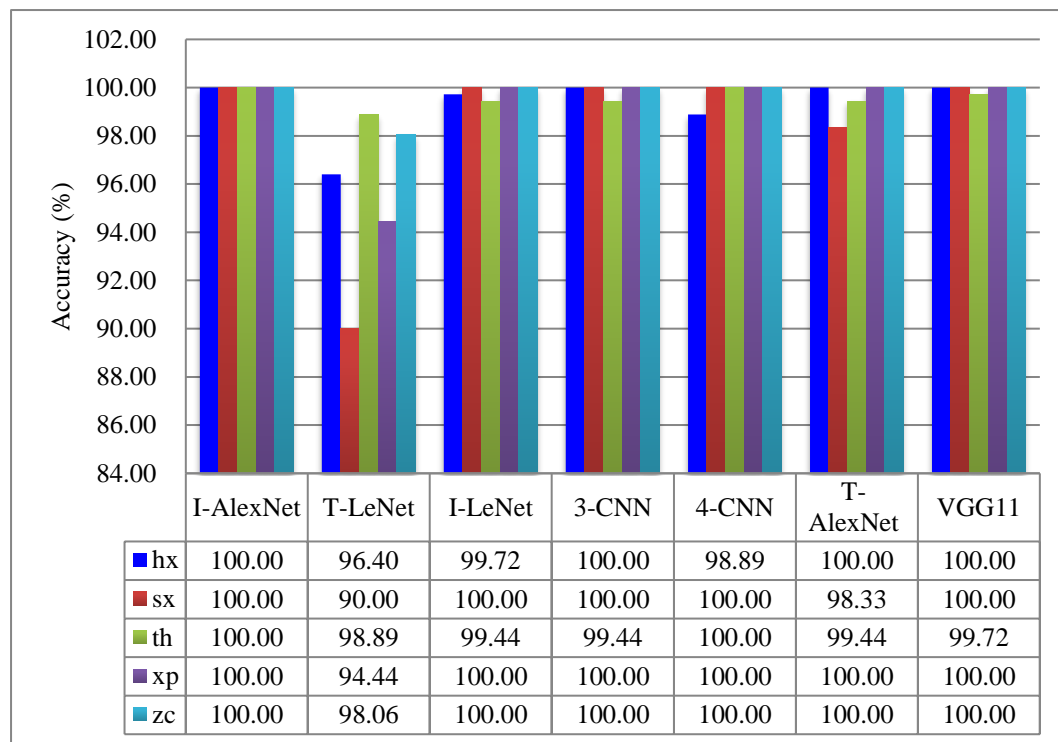


Figure 23. Confusion matrix in one trial. (a) I-LeNet, (b) I-LeNet, (c) 3-CNN, (d) 4-CNN, (e) VGG11, (f) I-AlexNet, (g) T-AlexNet.

Table 4. Precision of I-AlexNet and T-AlexNet for different fault types.

Fault Type	I-AlexNet (%)	T-AlexNet (%)
zc	100.0	100.0
xp	100.0	100.0
sx	100.0	99.4
hx	100.0	98.4
th	100.0	100.0

To investigate the classification effect of models on each fault category, different CNN models were used for comparative analysis. As can be seen from Figure 24, the performance of T-LeNet was significantly lower than that of other models. The other four models had no significant difference in the classification results of zc, xp, and sx. However, for the categories of hx and th, the classification effect of the I-AlexNet model showed a slight improvement and was better than other models.

**Figure 24.** Histogram of accuracy of different CNN models for each fault type.

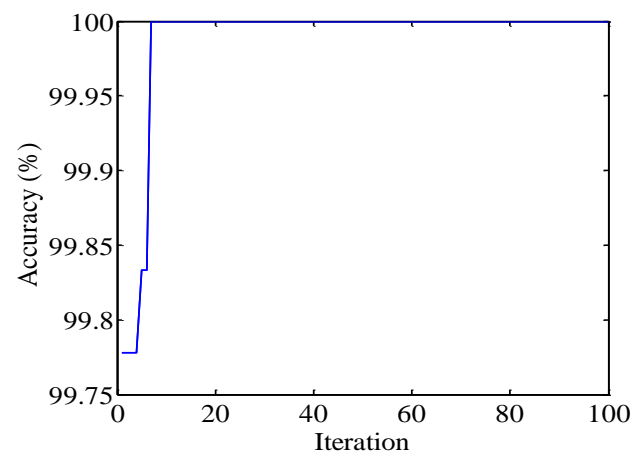
The computing time of one independent test was taken for analysis, and the results are depicted in Table 5. The mean computing time of the T-AlexNet model was 11.04 s, while the time consumption of the I-AlexNet model was slightly more. Compared to 3-CNN, 4-CNN, and VGG11, the average time consumption of the I-AlexNet model was significantly lower.

5.2.8. Optimization of Diagnostic Model

In the process of BO, some presetting hyperparameters of the I-AlexNet model were the same as above. The selection of the range for each hyperparameter was based on previous manual tuning. The parameter optimization is depicted in Figure 25. Table 6 shows the optimization results. The convergence speed of the improved AlexNet was faster with BO. The classification accuracy reached 100% after more than 10 iterations.

Table 5. Computational time of different CNN models in one epoch.

Model	1	2	3	4	5	6	7	8	9	10	Time (s)
I-AlexNet	11.09	11.06	11.22	12.23	11.14	11.12	11.19	11.09	11.09	11.16	11.24
T-AlexNet	11.15	11.05	11.22	11.15	11.08	10.90	10.95	10.91	11.05	10.96	11.04
T-LeNet	4.14	4.38	4.13	4.08	4.28	4.18	4.18	4.28	4.16	4.06	4.19
I-LeNet	4.78	4.63	4.89	4.83	7.72	4.72	4.80	4.75	4.88	4.80	5.08
3-CNN	12.99	13.05	12.97	13.24	13.20	13.26	13.03	12.99	12.89	13.01	13.06
4-CNN	13.38	13.45	13.60	13.68	13.67	13.56	13.79	13.52	13.67	13.43	13.58
VGG11	31.85	32.05	32.12	31.25	31.27	31.67	31.43	31.19	31.02	32.18	31.60

**Figure 25.** Classification accuracy with BO.**Table 6.** The range and results of hyperparameter optimization.

Serial Number	Hyperparameter	Range	Optimal Result
1	LR	[0.0001, 0.001]	0.00012
2	Batch size	[24, 56]	51
3	Epoch	[20, 50]	33
4	Size of convolutional kernel (C1)	[5, 9]	5
5	Size of convolutional kernel (C2)	[3, 7]	5
6	Number of convolutional kernel (C1)	[30, 60]	54
7	Number of convolutional kernel (C2)	[80, 140]	97
8	Neurons of FC1	[1000, 1600]	1457
9	Neurons of FC2	[400, 800]	541
10	Dropout ratio	[0.1, 0.9]	0.31

The hyperparameter combinations obtained by BO were used for model verification. As can be seen from the training loss curve in Figure 26, the model had good convergent results. The accuracy trend of the train and test samples showed a high degree of coincidence. The small difference indicated good stability of the B-AlexNet model. The confusion matrix of one experiment and the classification precision are depicted in Figure 27 and Table 7, respectively. The identification effects of I-AlexNet and B-AlexNet were superior to that of T-AlexNet, and the precision reached up to 100%. The B-AlexNet especially showed an improvement for sx and hx compared to T-AlexNet.

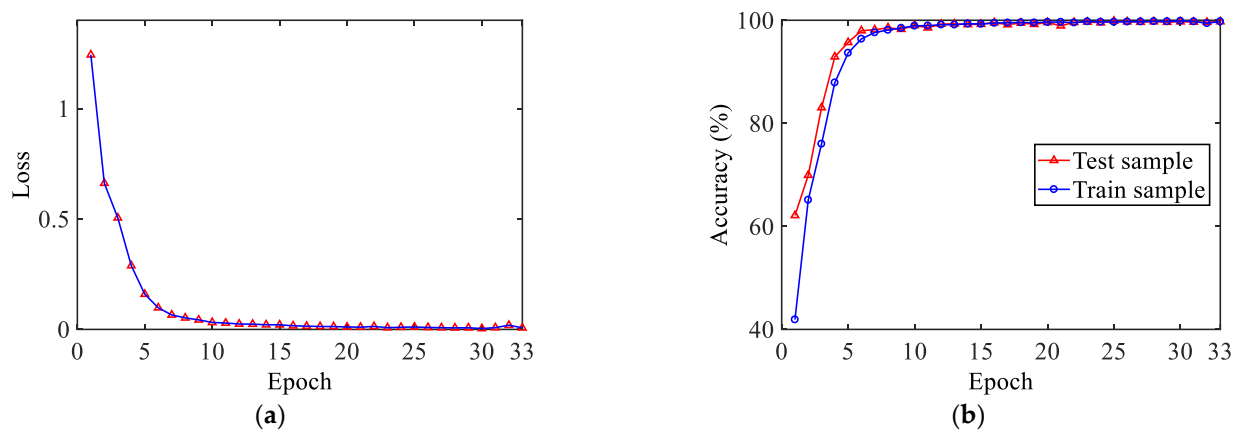


Figure 26. Trend of loss and identification accuracy. (a) Loss, (b) accuracy.

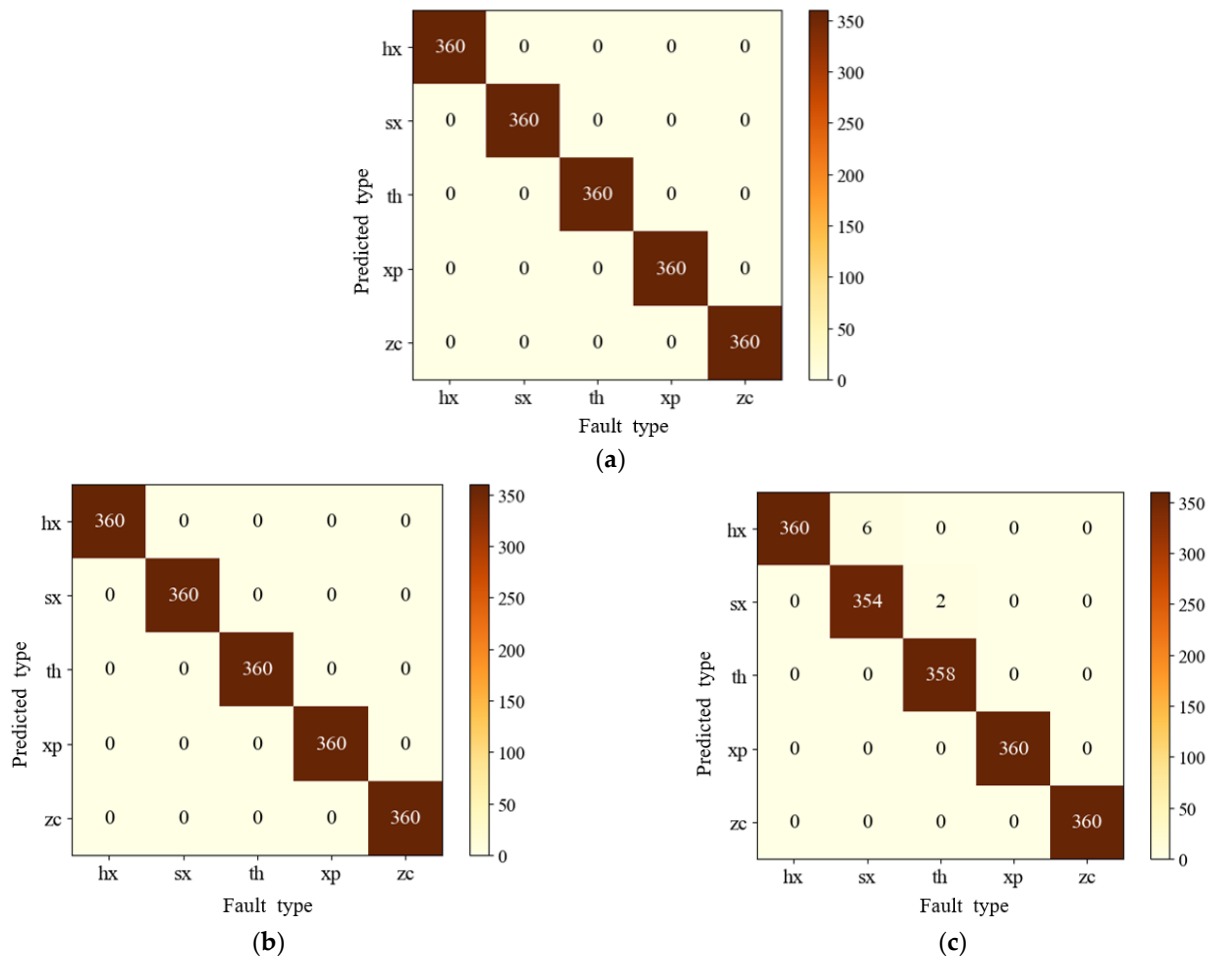


Figure 27. Confusion matrix of a trial. (a) B-AlexNet, (b) I-AlexNet, (c) T-AlexNet.

Table 7. Precision of each condition with different models.

Fault Type	B-AlexNet	I-AlexNet	T-AlexNet
zc	100.0	100.0	100.0
xp	100.0	100.0	100.0
sx	100.0	100.0	99.4
hx	100.0	100.0	98.4
th	100.0	100.0	100.0

The feature learning of B-AlexNet was investigated by employing *t*-SNE, and the results are shown in Figure 28. The features in raw input are presented with scattered points. It was difficult to find a certain principle between features of different faults. Gradually, the features of the same fault type got together in clusters due to the learning of convolutional layers. Five categories were clearly observed through the subsequent fully connected layers and classifier layer. B-AlexNet could gain useful knowledge in the time–frequency representations and achieve the classification of five typical states of an axial piston pump.

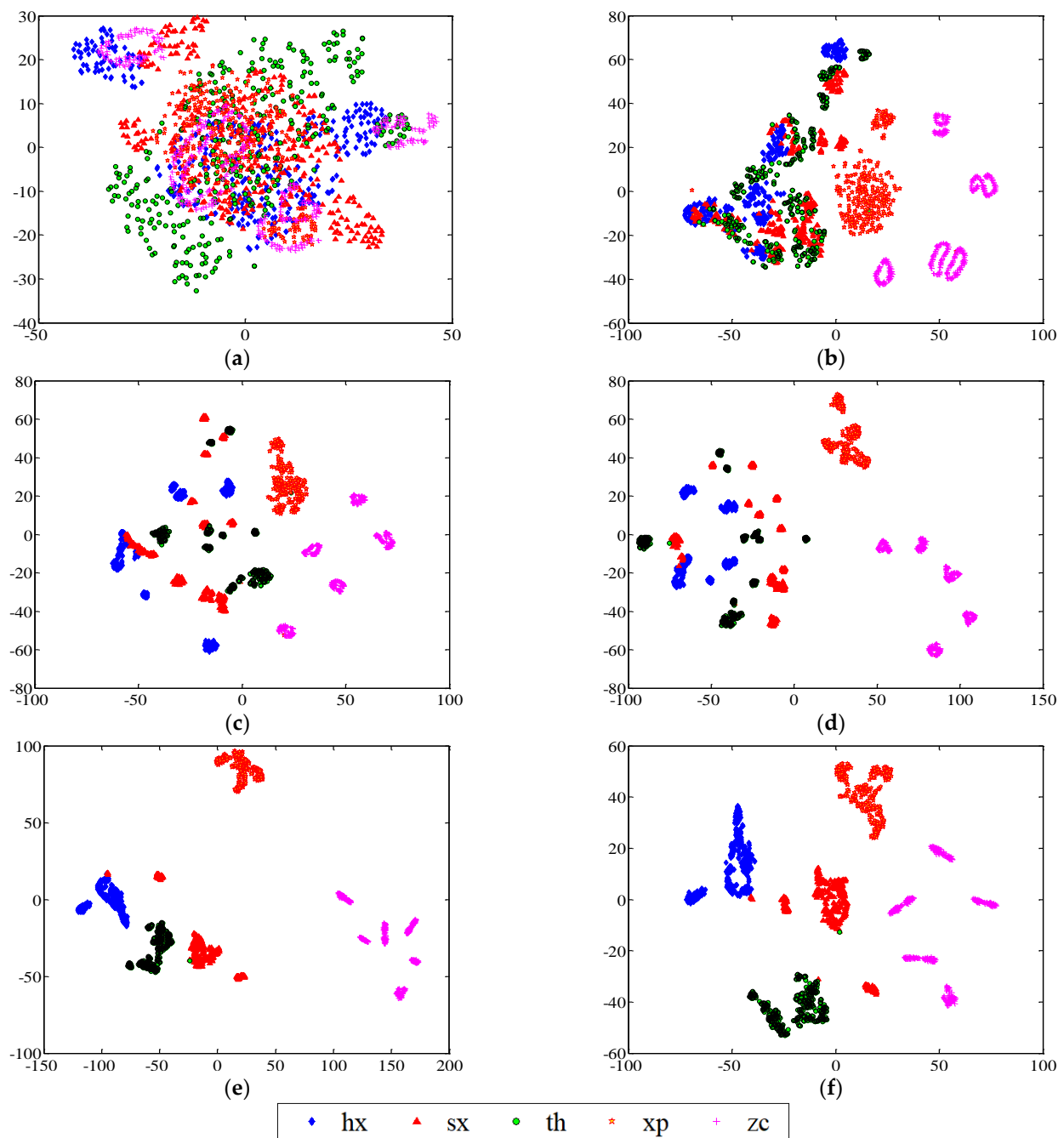


Figure 28. Visualization of feature representation via *t*-SNE. (a) Input data, (b) C1, (c) C3, (d) C5, (e) FC2, (f) classifier layer.

6. Conclusions

An integrated method was constructed based on a deep learning model and Bayesian algorithm. The feature information hidden in the pressure signal was fully mined, and the identification of typical faults of an axial piston pump was automatically accomplished. Continuous wavelet transform was used for signal analysis and to maintain information on both time and frequency. Bayesian algorithm adaptively learnt the critical hyperparameters of the deep model by absorbing the merits of the Gaussian process to improve the model. The following conclusions were obtained:

(1) Among the main selected hyperparameters, the learning rate showed the greatest effect on the diagnosis results. The kernel number and kernel size also had considerable influence. Compared to the above parameters, the epoch, dropout ratio, and batch size did not have a remarkable effect.

(2) The experiments showed that improved AlexNet with manual tuning showed better performance than the other contrastive models.

(3) By adopting Bayesian optimization algorithm, the proposed method named B-AlexNet had a higher accuracy and stronger robustness. The average accuracy reached up to 100% for five health states of an axial piston pump, and the precision for each fault type was also enhanced.

(4) The identification performance of the diagnostic model was further validated by feature distributions learned by different network layers. The constructed model can fuse feature extraction and fault classification, and it requires less professional knowledge and experience on signal processing.

In the future, further improvement to the Bayesian optimization algorithm will be investigated taking into account the influence of different kernel functions in the Gaussian process, and an improvement strategy based on noise addition will be developed. Moreover, the generalizability of the diagnosis model will be explored by mining more comprehensive information from multiple signals.

Author Contributions: Validation, writing—review and editing and funding acquisition, Y.Z.; formal analysis and writing—original draft preparation, T.Z.; conceptualization, methodology, investigation, and writing—original draft preparation, S.T.; supervision, S.Y. 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 (52205057, 52175052), National Key Research and Development Program of China (2020YFC1512402), Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJB460002), China Post-doctoral Science Foundation (2022M723702), Taizhou Science and Technology Plan Project (22gyb42), and the Youth Talent Development Program of Jiangsu University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author upon reasonable request.

Acknowledgments: The authors are grateful to Wanlu Jiang and Siyuan Liu from Yanshan University for their support in the experiments.

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

References

1. Lan, Y.; Li, Z.; Liu, S.; Huang, J.; Niu, L.; Xiong, X.; Niu, C.; Wu, B.; Zhou, X.; Yan, J.; et al. Experimental investigation on cavitation and cavitation detection of axial piston pump based on MLP-Mixer. *Measurement* **2022**, *200*, 111582. [[CrossRef](#)]
2. Kumar, A.; Gandhi, C.P.; Tang, H.; Vashishtha, G.; Kumar, R.; Zhou, Y.; Xiang, J. Adaptive sensitive frequency band selection for VMD to identify defective components of an axial piston pump. *Chin. J. Aeronaut.* **2022**, *35*, 250–265. [[CrossRef](#)]
3. Haidak, G.; Wei, X.; Li, F.; Larbi, A.; Wang, D. Heat effects modelling on the efficiency loss of the lubricating interface between piston and cylinder in axial piston pumps. *Tribol. Int.* **2022**, *175*, 107846. [[CrossRef](#)]

4. Ying, P.; Tang, H.; Ye, S.; Ren, Y.; Xiang, J.; Kumar, A. Dynamic modelling of swashplate with local defects in axial piston pump and coupled vibration analysis. *Mech. Syst. Signal Process.* **2023**, *15*, 110081. [\[CrossRef\]](#)
5. Tang, S.; Zhu, Y.; Yuan, S. A novel adaptive convolutional neural network for fault diagnosis of hydraulic piston pump with acoustic images. *Adv. Eng. Inform.* **2022**, *52*, 101554. [\[CrossRef\]](#)
6. Haidak, G.; Wang, D. Analysis of damage and failure mechanism under a lubricated slipper/swashplate interface in axial piston machines. *Procedia Struct. Integr.* **2022**, *35*, 124–131. [\[CrossRef\]](#)
7. Nie, S.; He, H.; Ji, H.; Nie, S.; Yan, X.; Yin, F. Failure analysis of auxiliary support bearing/shaft tribopair in seawater hydraulic axial piston pump. *Eng. Fail. Anal.* **2023**, *146*, 107069. [\[CrossRef\]](#)
8. Lv, Y.; Zhao, W.; Zhao, Z.; Li, W.; Ng, K.K.H. Vibration signal-based early fault prognosis: Status quo and applications. *Adv. Eng. Inform.* **2022**, *52*, 101609. [\[CrossRef\]](#)
9. Gu, L.; Ma, Z.; Tian, Q.; Sun, Y. Application of instantaneous speed fluctuation signal in fault diagnosis of axial piston pump. *J. Drain. Irrig. Mach. Eng.* **2021**, *39*, 740–746.
10. Ye, M.; Yan, X.; Chen, N.; Jia, M. Intelligent fault diagnosis of rolling bearing using variational mode extraction and improved one-dimensional convolutional neural network. *Appl. Acoust.* **2023**, *202*, 109143. [\[CrossRef\]](#)
11. Yu, H.; Li, H.; Li, Y. Vibration signal fusion using improved empirical wavelet transform and variance contribution rate for weak fault detection of hydraulic pumps. *ISA Trans.* **2020**, *107*, 385–401. [\[CrossRef\]](#)
12. Zhou, Y.; Kumar, A.; Parkash, C.; Vashishtha, G.; Tang, H.; Glowacz, A.; Dong, A.; Xiang, J. Development of entropy measure for selecting highly sensitive WPT band to identify defective components of an axial piston pump. *Appl. Acoust.* **2023**, *203*, 109225. [\[CrossRef\]](#)
13. Tang, H.; Fu, Z.; Huang, Y. A fault diagnosis method for loose slipper failure of piston pump in construction machinery under changing load. *Appl. Acoust.* **2021**, *172*, 107634. [\[CrossRef\]](#)
14. Glowacz, A. Thermographic fault diagnosis of shaft of BLDC motor. *Sensors* **2022**, *22*, 8537. [\[CrossRef\]](#)
15. Qian, P.; Pu, C.; Liu, L.; Li, X.; Zhang, B.; Gu, Z.; Meng, D. Development of a new high-precision friction test platform and experimental study of friction characteristics for pneumatic cylinders. *Meas. Sci. Technol.* **2022**, *33*, 065001. [\[CrossRef\]](#)
16. Tang, S.; Zhu, Y.; Yuan, S. An adaptive deep learning model towards fault diagnosis of hydraulic piston pump using pressure signal. *Eng. Fail. Anal.* **2022**, *138*, 106300. [\[CrossRef\]](#)
17. Zhu, Z.; Lei, Y.; Qi, G.; Chai, Y.; Mazur, N.; An, Y.; Huang, X. A review of the application of deep learning in intelligent fault diagnosis of rotating machinery. *Measurement* **2023**, *206*, 112346. [\[CrossRef\]](#)
18. Kumar, D.; Dewangan, A.; Tiwari, R.; Bordoloi, D.J. Identification of inlet pipe blockage level in centrifugal pump over a range of speeds by deep learning algorithm using multi-source data. *Measurement* **2021**, *186*, 110146. [\[CrossRef\]](#)
19. Wang, A.; Pei, Y.; Qian, Z.; Zareipour, H.; Jing, B.; An, J. A two-stage anomaly decomposition scheme based on multi-variable correlation extraction for wind turbine fault detection and identification. *Appl. Energy* **2022**, *321*, 119373. [\[CrossRef\]](#)
20. Yuan, J.; Luo, L.; Jiang, H.; Zhao, Q.; Zhou, B. An intelligent index-driven multiwavelet feature extraction method for mechanical fault diagnosis. *Mech. Syst. Signal Process.* **2023**, *188*, 109992. [\[CrossRef\]](#)
21. Shi, M.; Ding, C.; Wang, R.; Song, Q.; Shen, C.; Huang, W.; Zhu, Z. Deep hypergraph autoencoder embedding: An efficient intelligent approach for rotating machinery fault diagnosis. *Knowl. Based Syst.* **2023**, *260*, 110172. [\[CrossRef\]](#)
22. Liu, J. Gas path fault diagnosis of aircraft engine using HELM and transfer learning. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105149. [\[CrossRef\]](#)
23. Su, Z.; Zhang, J.; Tang, J.; Wang, Y.; Xu, H.; Zou, J.; Fan, S. A novel deep transfer learning method with inter-domain decision discrepancy minimization for intelligent fault diagnosis. *Knowl. Based Syst.* **2023**, *259*, 110065. [\[CrossRef\]](#)
24. Qian, G.; Liu, J. Fault diagnosis based on conditional generative adversarial networks in nuclear power plants. *Ann. Nucl. Energy* **2022**, *176*, 109267. [\[CrossRef\]](#)
25. Shi, Z.; Chen, J.; Zhang, X.; Zi, Y.; Li, C.; Chen, J. A reliable feature-assisted contrastive generalization net for intelligent fault diagnosis under unseen machines and working conditions. *Mech. Syst. Signal Process.* **2023**, *188*, 110011. [\[CrossRef\]](#)
26. Zhao, C.; Shen, W. Mutual-assistance semi-supervised domain generalization network for intelligent fault diagnosis under unseen working conditions. *Mech. Syst. Signal Process.* **2023**, *189*, 110074. [\[CrossRef\]](#)
27. Han, T.; Li, Y. Out-of-distribution detection-assisted trustworthy machinery fault diagnosis approach with uncertainty-aware deep ensembles. *Reliab. Eng. Syst. Saf.* **2022**, *226*, 108648. [\[CrossRef\]](#)
28. Zhou, T.; Han, T.; Droguett, E.L. Towards trustworthy machine fault diagnosis: A probabilistic Bayesian deep learning framework. *Reliab. Eng. Syst. Saf.* **2022**, *224*, 108525. [\[CrossRef\]](#)
29. Sun, J.; Liu, Z.; Wen, J.; Fu, R. Multiple hierarchical compression for deep neural network toward intelligent bearing fault diagnosis. *Eng. Appl. Artif. Intell.* **2022**, *116*, 105498. [\[CrossRef\]](#)
30. Chao, Q.; Zhang, J.; Xu, B.; Wang, Q.; Lyu, F.; Li, K. Integrated slipper retainer mechanism to eliminate slipper wear in high-speed axial piston pumps. *Front. Mech. Eng.* **2022**, *17*, 1. [\[CrossRef\]](#)
31. Kumar, A.; Gandhi, C.P.; Zhou, Y.; Kumar, R.; Xiang, J. Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images. *Appl. Acoust.* **2020**, *167*, 107399. [\[CrossRef\]](#)
32. Tang, S.; Zhu, Y.; Yuan, S. Intelligent fault diagnosis of hydraulic piston pump based on deep learning and Bayesian optimization. *ISA Trans.* **2022**, *129*, 555–563. [\[CrossRef\]](#)

33. Zhu, Y.; Li, G.; Tang, S.; Wang, R.; Su, H.; Wang, C. Acoustic signal-based fault detection of hydraulic piston pump using a particle swarm optimization enhancement CNN. *Appl. Acoust.* **2022**, *192*, 108718. [CrossRef]
34. Chao, Q.; Gao, H.; Tao, J.; Wang, Y.; Zhou, J.; Liu, C. Adaptive decision-level fusion strategy for the fault diagnosis of axial piston pumps using multiple channels of vibration signals. *Sci. China Technol. Sci.* **2022**, *65*, 470–480. [CrossRef]
35. Zhu, Y.; Li, G.; Wang, R.; Tang, S.; Su, H.; Cao, K. Intelligent fault diagnosis of hydraulic piston pump combining improved LeNet-5 and PSO hyperparameter optimization. *Appl. Acoust.* **2021**, *183*, 108336. [CrossRef]
36. Xie, Y.; Sun, W.; Ren, M.; Chen, S.; Huang, Z.; Pan, X. Stacking ensemble learning models for daily runoff prediction using 1D and 2D CNNs. *Expert Syst. Appl.* **2023**, *217*, 119469. [CrossRef]
37. Ali, D.; Mohammad, E.; Reza, H.; Biuok, E. A hybrid fine-tuned VMD and CNN scheme for untrained compound fault diagnosis of rotating machinery with unequal-severity faults. *Expert Syst. Appl.* **2021**, *167*, 114094.
38. An, F.; Wang, J. Rolling bearing fault diagnosis algorithm using overlapping group sparse-deep complex convolutional neural network. *Nonlinear Dyn.* **2022**, *108*, 2353–2368. [CrossRef]
39. Zhao, M.; Fu, X.; Zhang, Y.; Meng, L.; Tang, B. Highly imbalanced fault diagnosis of mechanical systems based on wavelet packet distortion and convolutional neural networks. *Adv. Eng. Inform.* **2022**, *51*, 101535. [CrossRef]
40. Tang, S.; Zhu, Y.; Yuan, S. Intelligent fault identification of hydraulic pump using deep adaptive normalized CNN and synchrosqueezed wavelet transform. *Reliab. Eng. Syst. Saf.* **2022**, *224*, 108560. [CrossRef]
41. Venkatesh, B.; Anuradha, J. A review of feature selection and its methods. *Cybern. Inf. Technol.* **2019**, *19*, 3–26. [CrossRef]
42. Tang, S.; Zhu, Y.; Yuan, S. An improved convolutional neural network with an adaptable learning rate towards multi-signal fault diagnosis of hydraulic piston pump. *Adv. Eng. Inform.* **2021**, *50*, 101406. [CrossRef]
43. Wu, J.; Chen, X.; Zhang, H.; Xiong, L.; Lei, H.; Deng, S. Hyperparameter optimization for machine learning models based on Bayesian optimization. *J. Electron. Sci. Technol.* **2019**, *17*, 26–40.
44. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; Freitas, N.D. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **2016**, *104*, 148–175. [CrossRef]
45. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]
46. Baldi, P.; Sadowski, P. The dropout learning algorithm. *Artif. Intell.* **2014**, *210*, 78–122. [CrossRef]
47. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. *Comput. Sci.* **2014**. Available online: <http://de.arxiv.org/pdf/1412.6980> (accessed on 1 January 2023).

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.