

Article

Fault Diagnosis Method of Waterproof Valves in Engineering Mixing Machinery Based on a New Adaptive Feature Extraction Model

Rui Zhang ¹, Jiyan Yi ², Hesheng Tang ¹ , Jiawei Xiang ¹  and Yan Ren ^{1,*}

¹ College of Mechanical and Electrical Engineering, Wenzhou University, Wenzhou 325035, China; 20451438015@stu.wzu.edu.cn (R.Z.); tanghesheng@wzu.edu.cn (H.T.); jwxiang@wzu.edu.cn (J.X.)

² Wenzhou Academy of Special Equipment Science, Plot 31, Phase I of China Shoes Capital, Fengmen Street, Lucheng District, Wenzhou 325035, China; yijian@126.com

* Correspondence: rentingting211@wzu.edu.cn

Abstract: Due to the complex working medium of oil in construction engineering, the waterproof valve in mixing machinery can easily cause different degrees of failure. Moreover, under adverse working conditions and complicated noise backgrounds, it is very difficult to detect the fault of waterproof valves. Thus, a fault diagnosis method is proposed, especially for the fault detection of waterproof valves as a key component in the construction of mixing machinery. This fault diagnosis method is based on a new adaptive feature extraction model, with multi-path signals to the improved deep residual shrinkage network–stacked denoising convolutional autoencoder (named DRSN–SDCAE). Firstly, the noisy vibration signals collected by the two vibration sensors are preprocessed, and then transmitted to the parallel structure improved DRSN–SDCAE for adaptive denoising and feature extraction. Finally, these results are fused through the feature fusion strategy to realize the effective fault diagnosis of the waterproof valve. The effectiveness of this method was verified through theory and experiments. The experimental results show that the proposed fault diagnosis method based on the improved DRSN–SDCAE model can automatically and effectively extract fault features from noise for fault diagnosis without relying on signal processing technology and diagnosis experiences. When compared with other intelligent fault diagnosis methods, the features extracted from multi-path inputs were more comprehensive than those extracted from single-path inputs, and contained more complete features of hidden data, which significantly improved fault diagnosis accuracy based on these fault features. The contribution of this paper is to learn fault features autonomously in signals with strong and complex noise through a deep network structure, which extends the fault diagnosis method to the field of construction machinery to improve the safe operation and maintainability of engineering machinery.

Keywords: engineering mixing machinery; waterproof valve; adaptive feature extraction; multi-path inputs; intelligent fault diagnosis



Citation: Zhang, R.; Yi, J.; Tang, H.; Xiang, J.; Ren, Y. Fault Diagnosis Method of Waterproof Valves in Engineering Mixing Machinery Based on a New Adaptive Feature Extraction Model. *Energies* **2022**, *15*, 2796. <https://doi.org/10.3390/en15082796>

Academic Editors: Hu Shi and Sérgio Cruz

Received: 16 February 2022

Accepted: 31 March 2022

Published: 11 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Mixing machines are commonly used in construction engineering to realize the mixing and stirring of building materials such as the cement, mortar and concrete. Due to the working process of the mixing machine itself, the working medium inevitably mixes complex building materials and the polluted fluid containing a large number of particles, which makes the hydraulic transmission system, including components in the mixing machine, vulnerable to varying degrees of wear and heating. This damage is irreversible, and the degree of damage will increase with working time. Moreover, the working conditions of mixing machinery are very harsh, and in a strong and complex noise environment, it is difficult to detect faults of the system or components in mixing machinery. Therefore,

effective fault diagnosis methods are of great significance for the safe and reliable operation of construction machinery.

In the fault diagnosis method, if the fault features can be extracted, it is easier to diagnose the fault. The working environments of hydraulic transmission systems are relatively bad, especially the working conditions of construction machinery, which often submerges the working state of the system or components in the background of strong and complex noise. In order to extract fault features under harsh working conditions, signal denoising has proven to be an effective method [1]. The common methods of processing noise signals mainly include empirical mode decomposition (EMD) [2,3], variational mode decomposition (VMD) [4,5], and Hilbert transform (HT) [6,7]. For example, Du et al. proposed a method combining EMD and the Wigner–Ville distribution to select the best intrinsic mode function to achieve the effective decomposition of noisy vibration signals from bearings in plunger pumps [8]. EMD is capable of linearizing and smoothing non-linear, non-smooth time-series data, thus preserving as many basic features of the original data as possible. Zhu et al. proposed a new method based on the VMD and relative entropy (RE) to extract the useful components of the signal, which could better extract the effective component of vibration signals with strong noise interference for a hydraulic axial piston pump under normal state, slipper wear, and slipper luxation [9]. VMD can decompose a raw signal into multiple subseries with different frequency scales with relative smoothness, avoiding the endpoint effect and spurious component problems encountered in the iterative process, with the advantages increased accuracy of complex data decomposition and better resistance to noise interference. Yu et al. addressed the problem of strong fault signal noise and weak fault information in the fault diagnosis of hydraulic pumps. An improved EWT and variance contribution ratio were proposed to dynamically fuse the vibration signals in different directions, and the Hilbert transform was used to extract weak fault features from the background noise and inherent impact [10]. However, these denoising methods rely on expert experience; thus, it is necessary to develop adaptive feature extraction to realize denoising by integrating feature extraction and classification into a single learning body.

With the development of artificial intelligence, deep learning as a complex machine learning algorithm can extract sensitive features from the original signal by learning the internal law and representation level of sample data. Multiple scholars have solved many complex problems of feature extraction, data mining or pattern recognition through deep learning techniques, such as convolutional neural networks (CNNs) [11,12], deep belief networks (DBNs) [13,14], recurrent neural networks (RNNs) [15,16], and autoencoders (AEs) [17,18]. For example, Wen et al. proposed a new LeNet-5-based CNN fault diagnosis method. By converting the signal into a two-dimensional image, the method could extract features from the converted two-dimensional image, which was applied to the fault diagnosis of axial piston pumps, obtaining good diagnosis results [19]. Mallak et al. provided a new combination of long short-term memory (LSTM) autoencoder architecture and deep learning methods for the fault diagnosis of hydraulic components, showing that the LSTM autoencoder can be effective in extracting features of fault signals [20]. Han et al. used generative adversarial networks (GANs) to compensate for imbalances in the sample data, and then used a stacked autoencoder (SAE) method to extract fault features of electrically driven feed pump signals; the comparison of the results showed that the GAN–SAE feature extraction strategy improved the robustness of the algorithm [21]. Liu et al. used a deep residual network to diagnose emulsion pump faults in order to improve the accuracy of emulsion pump fault diagnosis and to provide new ideas for intelligent fault diagnoses of emulsion pumps [22]. Huang et al. emphasized the power of deep learning in the application of feature engineering by using CNNs that did not rely on prior knowledge for feature learning and using LSTM layers to capture the time delay information of fault diagnoses of complex systems [23].

However, in strong and complex noise environments, the traditional deep learning models always exhibit weak results. Therefore, more and more new deep learning methods have emerged in recent years to address the interference of strong noise on the fault diag-

nosis of mechanical equipment, which enables end-to-end fault diagnosis in engineering applications. For example, Meng et al. proposed an enhanced denoising autoencoder for the fault diagnosis of rolling bearings, achieving appreciable accuracy [24]. Fu et al. proposed a fault diagnosis method combining a generative adversarial network (GAN) and stacked denoising autoencoder (SDAE) applied in bearings and gears. When the actual measurement data were limited, the data generated by GAN also achieved a diagnostic accuracy of more than 90% after denoising by SDAE, which demonstrated the effectiveness of SDAE for extracting fault information from signals [25]. SDAE is a deep learning network for processing noise-containing data based on an autoencoder. It is an unsupervised learning algorithm that adaptively extracts the main features in a signal as long as an arbitrary noise-containing signal is imputed, and is widely used in fault diagnosis. In general, SDAE needs the learned features to be as useful as possible and to be able to counteract the contamination and absence of raw data to some extent. However, in dealing with complex signals, especially in construction machinery, the features obtained are often limited. In recent years, improved residual networks with joint attention mechanism and soft thresholding have received increased attention. This provides new research solutions to mechanical equipment fault diagnosis. Zhao et al. proposed a deep residual shrinkage network (DRSN) to improve the feature learning capability of strong noisy vibration signals; the effectiveness of the proposed method was verified through experiments on different types of noise [26].

When a signal contains a lot of noise, it is difficult to accurately diagnose it by only improving the anti-noise performance, because the fault information contained in a single sensor is always limited. Thus, in addition to improving the noise immunity, it is also necessary to enrich the fault information as much as possible in order to achieve accurate fault diagnosis. Therefore, multi-path sensor information is necessary in strong noise working environments. Azamfar et al. implemented a novel two-dimensional convolutional neural network structure to fuse data obtained from multiple current sensors and use them directly for classification without manual feature extraction for the fault diagnosis of industrial gearbox test stands under different health states and different operating speeds [27]. Tao et al. proposed a fault diagnosis method based on multiple vibration signals and deep confidence networks (DBNs) in rolling bearing fault diagnosis, which can adaptively fuse multiple feature data to identify various bearing faults using the learning capabilities of DBNs. The effectiveness of multi-vibration signals in fault diagnosis is demonstrated through comparative experiments with individual sensors [28].

In the field of construction machinery, not only is the working environment very execrable, but the working state of the system itself in a polluted medium is also relatively complex. Therefore, it is difficult to extract the fault features of working components, especially the weak characteristics of early faults. Moreover, when the feature information is similar, it is difficult to identify correctly; thus, it is impossible to locate and identify faults in working components in the construction machinery quickly and accurately. Therefore, inspired by in-depth learning and multi-channel signals, this paper proposes a new adaptive feature extraction model to establish complex relationships between an original signal with complex and strong noise and the fault mode. Through the parallel structure—the improved deep residual shrinkage network and the superposition denoising convolution automatic encoder—the fault features can be extracted from different channels of the vibration signal of the waterproof valve, and the feature vectors extracted from the two channels are then combined into a single fusion feature vector, which is finally input into the fault classification.

2. Methodology

The fault diagnosis method of waterproof valves in engineering mixing machinery based on a new adaptive feature extraction model is shown in Figure 1. Firstly, the vibration signals collected by the two acceleration sensors are preprocessed, and then input into the parallel model: this is composed of improved DRSN and SDCAE. In the model, adaptive

feature learning is used to extract the basic features from the vibration signals polluted by strong noise. Then, the feature fusion strategy is used to fuse the two feature streams. Finally, the fully connected (FC) layer implements the mapping operation. Cross entropy is employed as the loss function to reflect the error between the two in the Softmax classifier, which is used to classify the fused features.

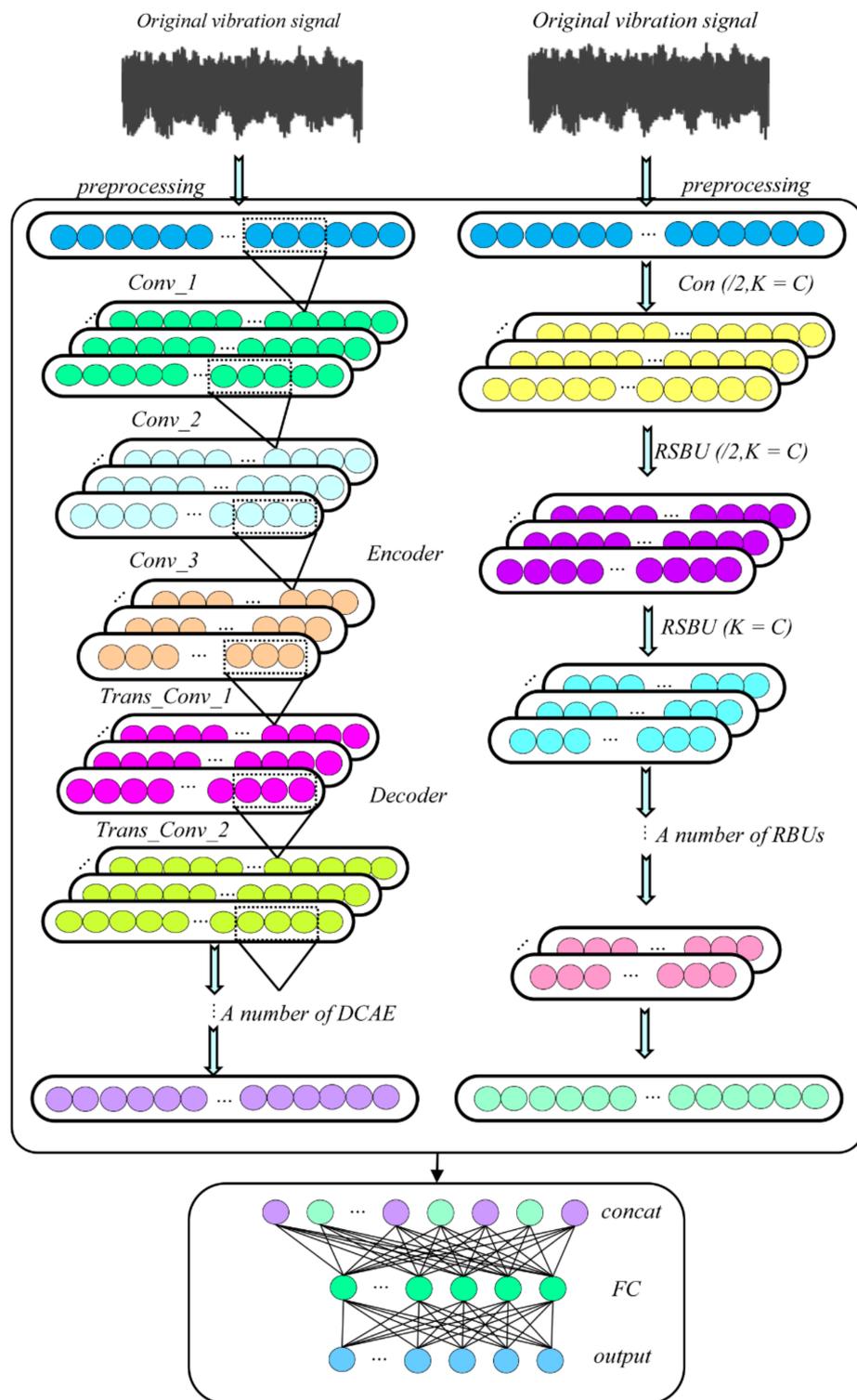


Figure 1. The fault diagnosis method based on a new adaptive feature extraction model.

2.1. Deep Residual Shrinkage Network Model

2.1.1. Basic Overview

The main aim of deep residual networks is to preserve as much original input information as possible during the training of deep convolutional neural networks, avoiding the gradient disappearance and degradation problems caused by increasing depth in the network. The core of the deep residual network lies in the residual building units (RBUs), which mainly consist of convolutional layers, batch normalization (BN) layers, and ReLUs.

The purpose of the convolutional layer is to extract different features in the input signal, which are called fault features in fault diagnosis, so that the parameters in the network training can be reduced and the occurrence of overfitting can be avoided, thus improving the accuracy of the network model. The convolution operation between the input feature map and the convolution kernel can be expressed as:

$$y_j = \sum_{i \in M_j} x_i * k_{ij} + b_j \tag{1}$$

where x_i and y_j are the i th and j th channels of the input and output feature maps, respectively, k is the convolution kernel, b is the bias, and M_j is the input feature vector.

The role of batch normalization (BN) is to normalize the input values with the aim of reducing the variation in the distribution of data in the hidden layers of the neural network. This has the advantage of helping to speed up the convergence of the neural network and alleviating the gradient disappearance problem in training. The calculation formula is as follows:

$$\mu = \frac{1}{N_{batch}} \sum_{n=1}^{N_{batch}} x_n \tag{2}$$

$$\sigma^2 = \frac{1}{N_{batch}} \sum_{n=1}^{N_{batch}} (x_n - \mu)^2 \tag{3}$$

$$\widehat{x}_n = \frac{x_n - \mu}{\sqrt{\sigma^2 + \epsilon}} \tag{4}$$

$$y_n = \gamma \widehat{x}_n + \beta \tag{5}$$

where x_n represents the input of the n th observation and y_n represents the output of the n th observation. N_{batch} represents the size of the mini batch. ϵ is a constant value which is close to zero. γ is a scaling parameter, and β is a bias parameter.

Both γ and β can be obtained through training.

The ReLU activation function is widely used because of its rapid convergence, the absence of a saturation interval, and the fact that the gradient is fixed to 1 in the part greater than 0, effectively solving the problem of disappearing gradients that exists in Sigmoid, which is expressed by the formula:

$$y = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{6}$$

where x represents the input feature and y represents the output feature.

Global average pooling (GAP) is an operation to average the feature maps, which can greatly reduce the parameters during neural network training and speed up the computation of neural networks.

2.1.2. Architecture of the Residual Shrinkage Building Unit

Deep residual shrinkage networks (DRSNs) accurately classify noise-polluted samples by introducing attention mechanisms and soft thresholding in ResNet. ‘‘Shrinkage’’ is the core contribution of deep residual shrinkage networks, referring to ‘‘soft thresholding’’, which is a key step in many signal denoising algorithms. The threshold required for soft thresholding in deep residual shrinkage networks is essentially set with the help of an

attention mechanism. The process of soft thresholding is shown in Figure 2a. The functions of the soft thresholding and its derivative form are expressed as follows:

$$y = \begin{cases} x - \tau, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ x + \tau, & x < -\tau \end{cases} \quad (7)$$

$$\frac{\partial y}{\partial x} = \begin{cases} 1, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ 1, & x < -\tau \end{cases} \quad (8)$$

where τ is the threshold, and is a positive number. x and y are the input feature and output feature, respectively.

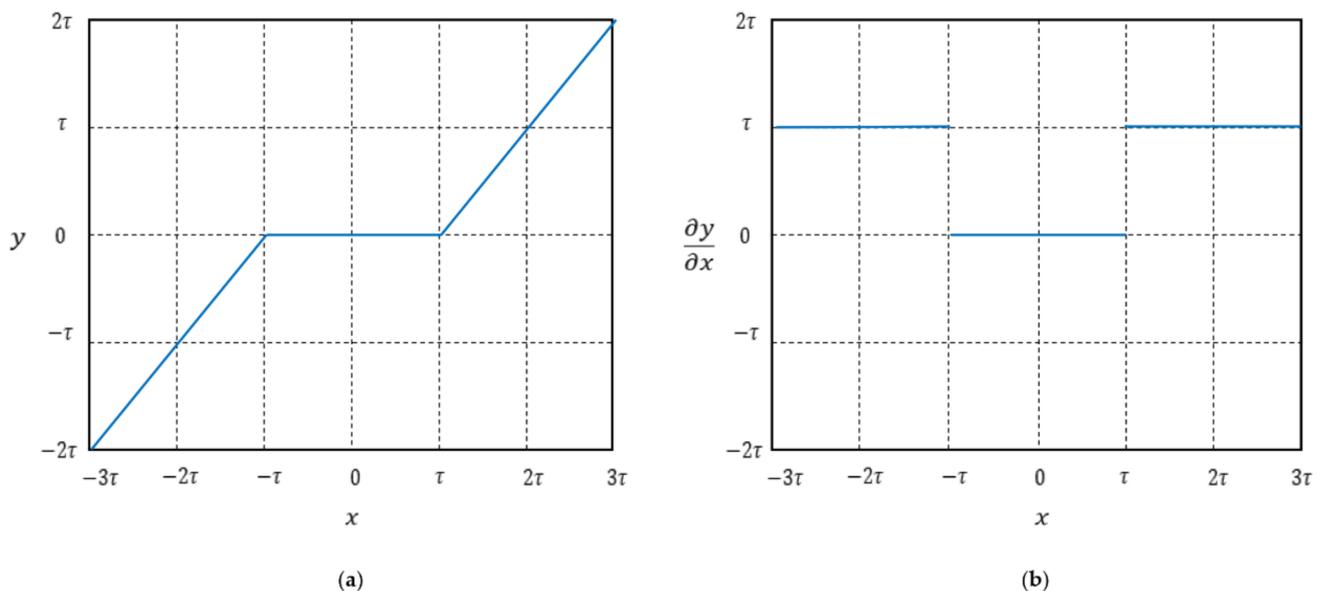


Figure 2. Soft thresholding and its derivative: (a) soft thresholding; (b) derivative of the soft threshold.

It can be seen from Equation (8) that the derivative of the soft threshold is either 1 or 0. The plotted image is shown in Figure 2b. This property is the same as the ReLU activation function. Therefore, soft thresholding can also reduce the risk of gradient disappearance and gradient explosion for deep learning algorithms. The soft thresholding function operates when both the threshold value is positive; the threshold value cannot be greater than the maximum value of the input signal.

The residual shrinkage building units (RSBUs) are the evolution of RBUs in the residual network, which is the most important module of the DRSN, as shown in Figure 3. Here, the number of feature map channels is C , the width is W , and the height is 1, and there is a red sub-network on the residual module in Figure 3. This sub-network is called the threshold module, and its role is to set the threshold adaptively by learning the features of the samples.

Multiple RSBUs are stacked, and the soft threshold is used as a systolic function to learn discriminative features by multiple nonlinear transformations to eliminate noise-related information.

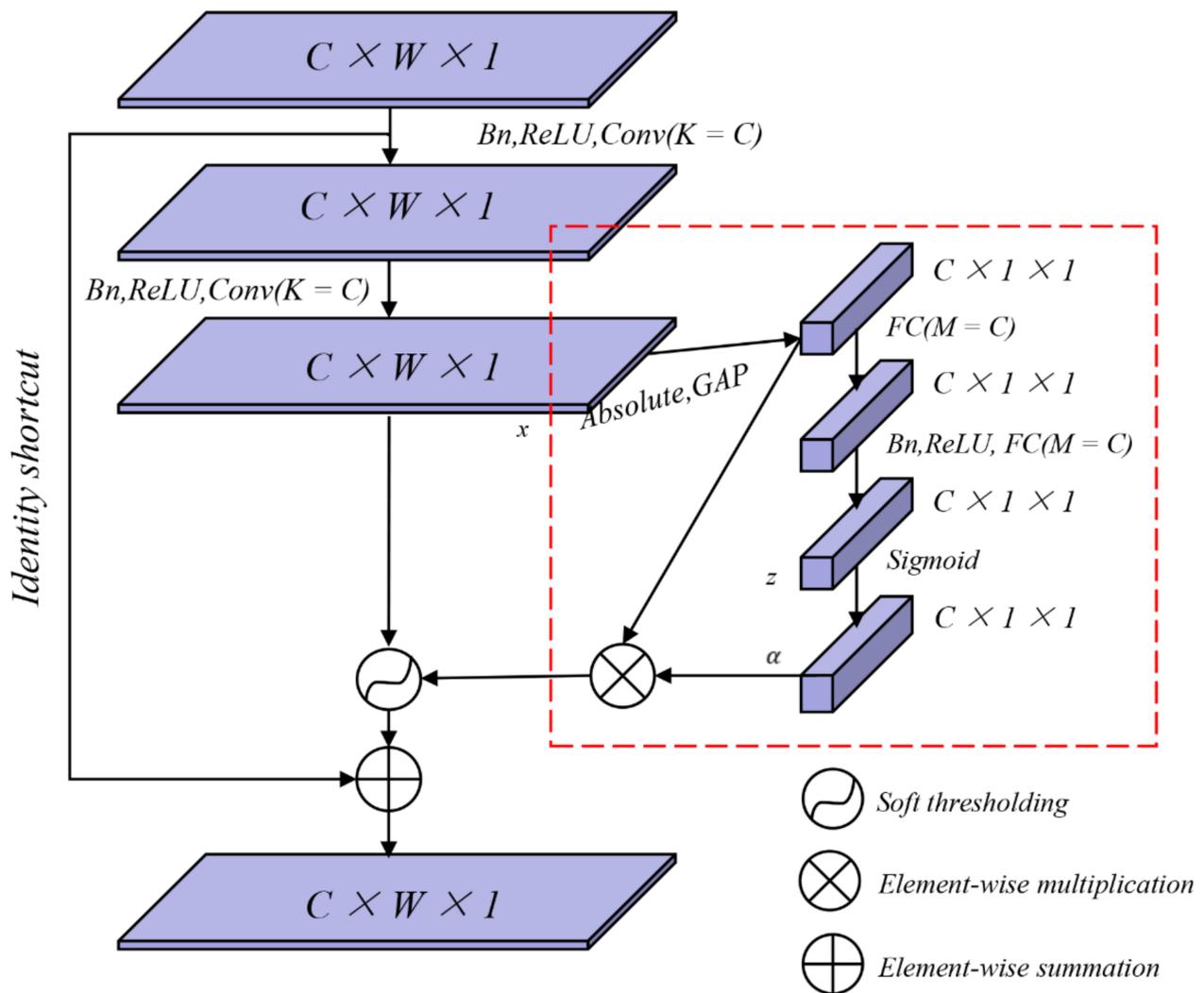


Figure 3. Residual shrinkage building unit.

2.2. Denoising Convolutional Autoencoder Model

The autoencoder is an unsupervised learning neural network composed of a three-layer neural network, whose learning process includes an encoding process and decoding process, and the basic structure is shown in Figure 4.

A denoising autoencoder (DAE) is an autoencoder that takes the original noise-containing data samples as inputs and trains them to output data samples with or without trace noise. This is not simply replicating the input of the original data, but can be understood as the process of removing noise. The hidden layer is also called the feature extraction layer. The input signal is to be $X = [x_1, x_2, \dots, x_n]$ and the coding function is f_θ . The formula for converting from the input layer to the hidden layer is:

$$Y = f_\theta(X) = s(WX + b) \tag{9}$$

where f_θ denotes the activation function of the encoder, $W = [w_1, w_2, \dots, w_n]$ is the encoder weights matrix, and $b = [b_1, b_2, \dots, b_n]$ is the offset coefficient.

After obtaining the hidden layer, the output layer is derived through a decoding process. In this process, the decode function is set to g_θ and the output signal to be $Z = [z_1, z_2, \dots, z_n]$. The output vector Z is reconstructed by Y according to the following equation.

$$Z = g_\theta(Y) = s(W'Y + b') \tag{10}$$

where g_θ is the activation function of the decoder, $W' = [w'_1, w'_2, \dots, w'_n]$ is the encoder weight matrix from the hidden layer to the output layer, and $b' = [b'_1, b'_2, \dots, b'_n]$ is the bias vectors for each layer.

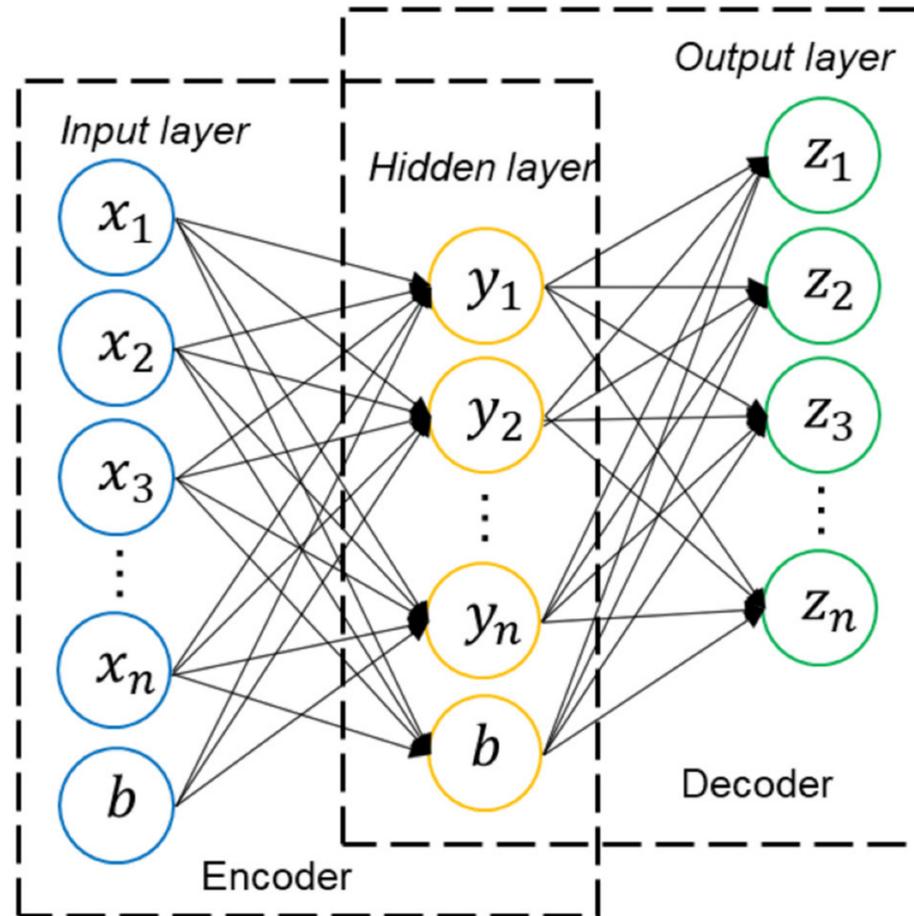


Figure 4. The network structure of autoencoder.

Compared with the conventional DAE, denoising convolutional autoencoder (DCAE) has the same basic structure of encoder and decoder, which is a fully connected DAE layer replaced by a convolutional layer. CNNs with a deep structure are easy to train; therefore, DCAE, as a type of CNN, can utilize the deep structure to improve the reconstruction ability. As shown in Figure 5, DCAEs consist of several convolutional and transposed convolutional layers, where the convolutional layer acts as an encoder and the transposed convolutional layer acts as a decoder. SDCAEs consist of multiple denoised convolutional autoencoders (DCAEs) by superposition. SDCAEs are capable of extensive and powerful fault feature extraction from a large amount of noise-laden vibration data in the original input.

2.3. Fusion Strategy

Feature fusion is a strategy between data fusion and decision fusion. It avoids a large amount of computation while avoiding a large amount of information detail loss. After denoising by DRSN and SDCAE, the feature vectors $v_1^T = \{v_{11}, v_{12}, \dots, v_{1h}\}$ and $s_2^T = \{v_{21}, v_{22}, \dots, v_{2h}\}$ are obtained, respectively. Fusing two multichannel feature vectors, the fusion strategy is defined as follows:

$$F(v_1^T, v_2^T) = \{v_{11}, v_{12}, \dots, v_{1h}, v_{21}, v_{22}, \dots, v_{2h}\} \quad (11)$$

where v_1 and v_2 are the feature vectors of DRSN and SDCAE after feature extraction, respectively.

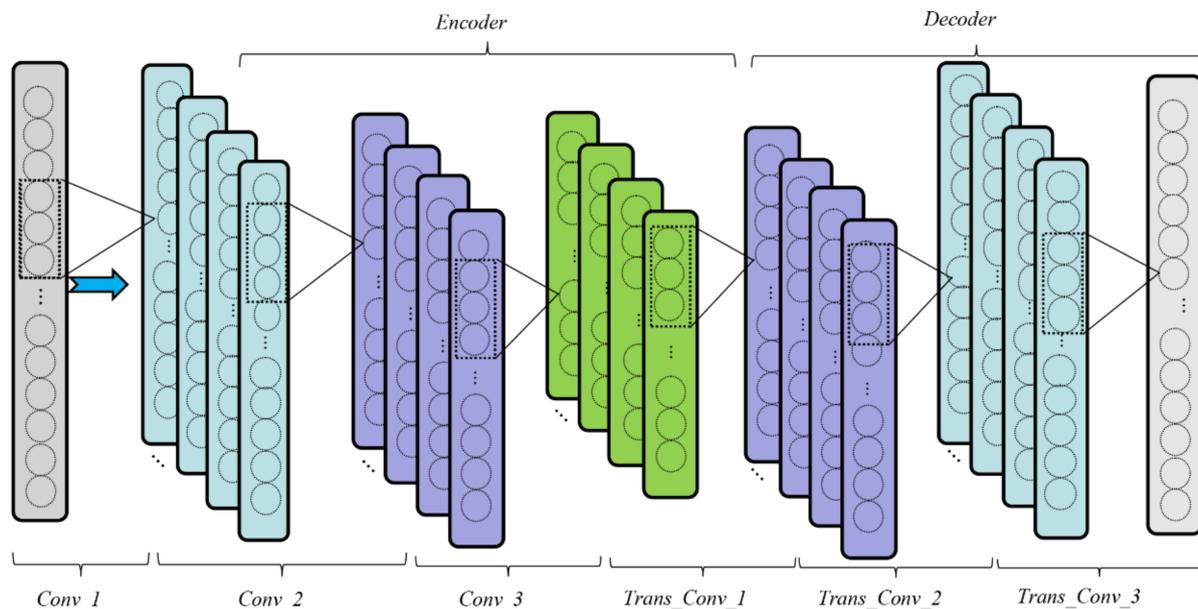


Figure 5. The structure of a DCAE.

This feature fusion strategy merges two feature vectors into one feature vector, and the resulting feature vector contains all the features contained in the two vibration signals. The feature extraction algorithms are different for the two channels, which minimizes the impact of redundant features on fault diagnosis. The fused feature vectors are transported to the FC layer and then passed through the final classifier to produce the results.

2.4. Fault Diagnosis Process

1. In this study, the proposed improved DRSN–SDCAE model was divided into two parallel paths, as shown in Figure 6. In simple terms, the signals are delivered to the DRSN sub-path and the SDCAE sub-path, and feature learning is performed adaptively; then, the fault feature vectors of the two sub-paths are obtained through a feature fusion strategy to accurately identify the fault category of the waterproof valve. The specific steps are as follows. The original vibration signals in the experiment are collected by the vibration signal acquisition device and vibration signal sensor, and randomly divided into several samples. The training set and testing set are used to train the improved DRSN–SDCAE model and verify the correctness of the improved DRSN–SDCAE model, respectively;
2. The training parameters are set, including the number of iterations, the learning rate, the number of training batches, the L2 regularization parameters, and the initialization of the network parameters, including the weights and offsets;
3. A supervised learning approach is used to train layer by layer and the network parameters are fine-tuned using the backpropagation algorithm;
4. When the number of iterations satisfies the set number N , the training of the improved DRSN–SDCAE model is completed, and it is then transferred to the next step;
5. The test sample set is fed into the improved DRSN–SDCAE model to obtain diagnostic results.

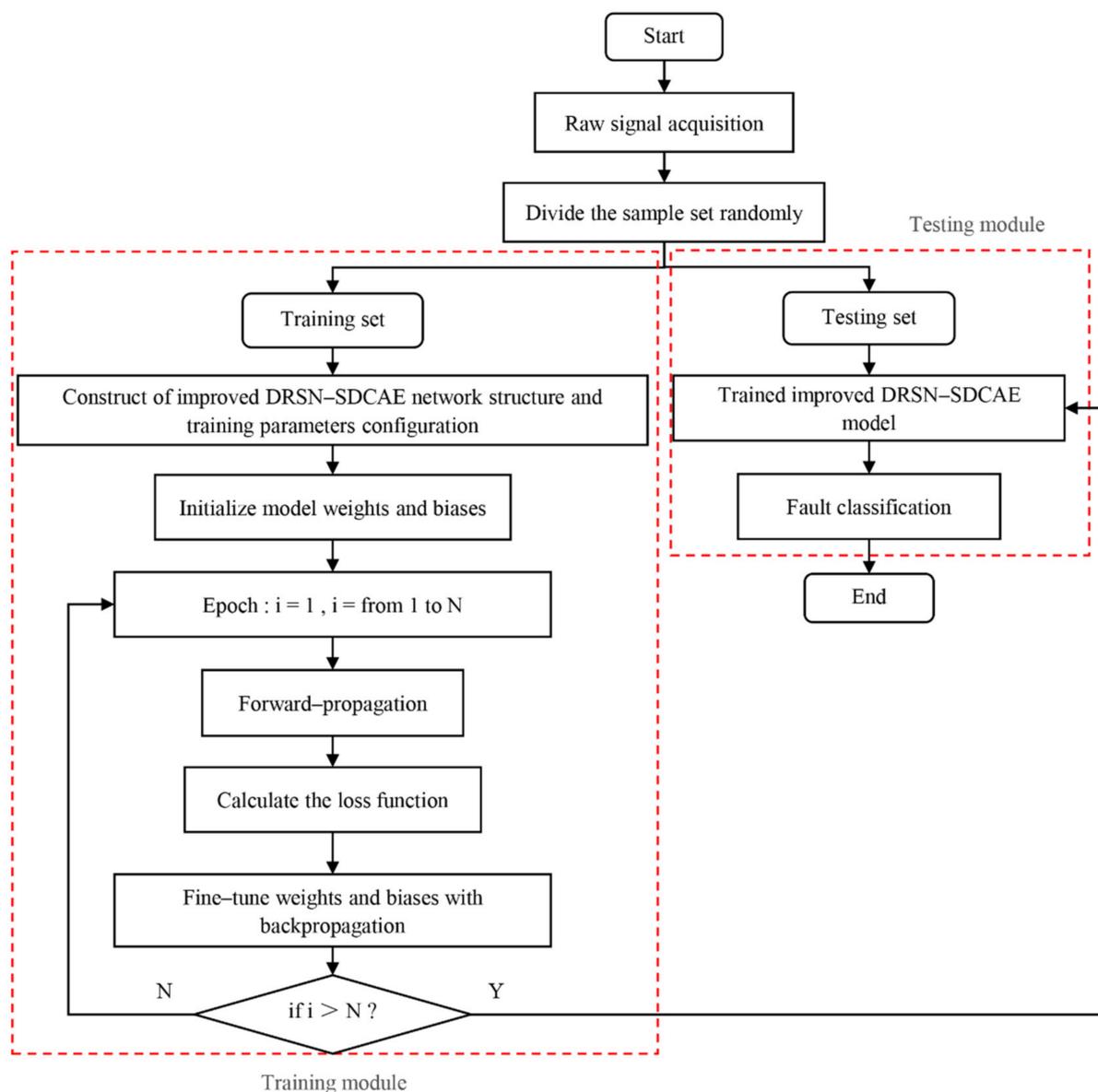


Figure 6. Flowchart framework of the improved DRSN-SDCAE fault diagnosis model.

3. Experimental Validation

3.1. Concept of the Waterproof Valve

The waterproof valve is a valve group composed of a safety valve, pressure-reducing valve and check valve. Waterproof valves are installed on the master cylinder of pump trucks, of which main function is to prevent the water in the water tank from adhering to the piston rod of the master cylinder and entering the cylinder barrel of the master cylinder, resulting in oil pollution and system emulsification. Figure 7 shows the internal oil circuit diagram of the waterproof valve. It can be seen that when the pressure in the sealing chamber of the master cylinder increases, the oil from the master cylinder will flow into the hydraulic waterproof valve through port A until it reaches the safety pressure set by the safety valve, and the oil will flow out of the safety port of the safety valve. At the project site, the main cylinder fluid contains pollutant particles, which will cause wearing of the safety valve to varying degrees. Under this working condition for a long time, the valve core will be stuck, which will seriously affect the normal operation of the hydraulic system.

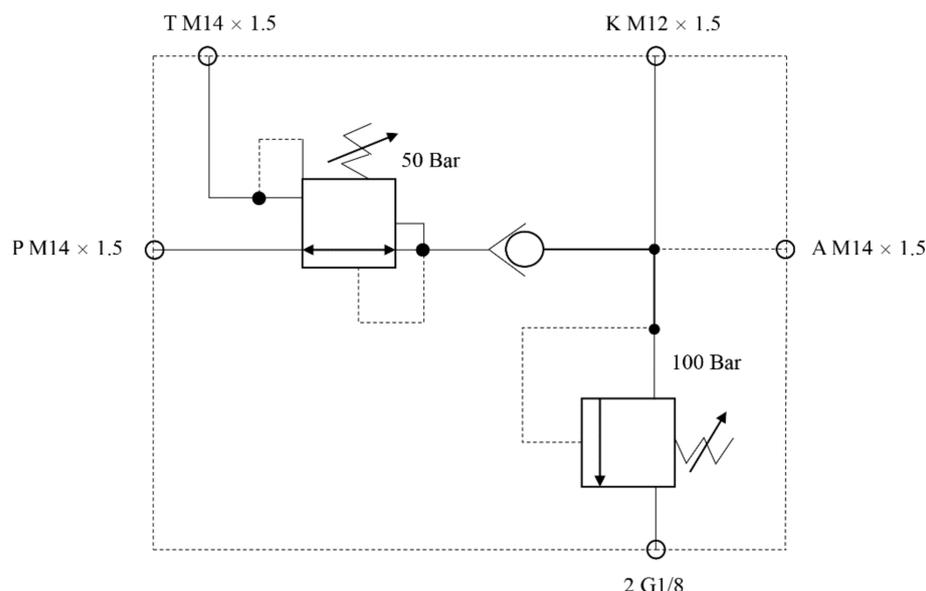


Figure 7. Waterproof valve internal oil circuit diagram.

During the working process of mixing machinery, mortar, concrete and other building particles will inevitably pollute the working medium. Working in this polluted medium, different degrees of scratches or irregular wear can easily occur between the valve core and the valve sleeve, at the acute angle of the shoulder or at the valve port of the waterproof valve, resulting in different degrees of internal leakage of the waterproof valve. Moreover, because the waterproof valve contains the conical valve structure, the waterproof valve is prone to be stuck and cannot work. In addition, due to wear, a cavity which easily resonates is formed in the front cavity of the conical valve, which affects the safe operation of the whole hydraulic system. In this experiment, the wear degree of the cone valve in the waterproof valve was defined as having two levels, namely, slight wear and serious wear. When the given pressure was 10 bar, the internal leakage under slight wear was 2.5 L/min; for severe wear, the leakage was 3.8 L/min. The locking degree of the spring inside the valve core was also defined as having two levels, i.e., slight locking and severe locking. The standard for defining slight locking was that at 120 bar, the given pressure reached the safety pressure of normal safety valve, and the safety flow for slight locking was 8.5 L/min; the pressure relief flow of severe locking was 6.3 L/min. The health valve was tested under the same conditions, and the pressure relief flow was 9.7 L/min. The fault definition of the waterproof valve is detailed in Table 1.

Table 1. Fault description of the waterproof valve.

Label	Fault Mode	Flow Description	Defined Conditions	Sample Points	Training Set	Testing Set	Experimental Conditions
0	Slight wear	Leakage rate is 2.5 L/min	A port pressure is 10 bar	2048	800	200	The maximum pressure of A port is 121 bar, the pressure of relief valve is set to 100 bar, the time of each experiment is 10 min, and the acquisition frequency of sensor is 6000 Hz.
1	Severe wear	Leakage rate is 3.8 L/min		2048	800	200	
2	Slight locking	Overflow rate of 8.5 L/min	A port pressure is 121 bar	2048	800	200	
3	Severe locking	Overflow rate of 6.3 L/min		2048	800	200	
4	Health	Overflow rate of 9.7 L/min		2048	800	200	

3.2. Experiment Description

As shown in Figure 8, the waterproof valve fault test bench mainly included the tested valve part, shown in Figure 8a, and the vibration signal acquisition part shown in Figure 8b, and the power source part, which was provided by the hydraulic pump station. The waterproof valve was fixed on the valve test bench to eliminate the influence of pressure impacts of the pump or the vibration of the system itself. Four acceleration sensors (marked as 1, 2, 3 and 4 in the figure) were installed near port A, port 2, port K, and the back of the safety valve to collect vibration signals. The hydraulic pump set a pulse pressure to deliver to port A; its pulse pressure setting curve is shown in Figure 9. It can be seen from the curve that the pulse pressure was set as a cycle for 6 s and decreased after reaching the maximum pressure of 140 bar to simulate the on-off-overflow process of the overflow valve. The sampling frequency of each experiment was 6 kHz, and the duration was 10 min.

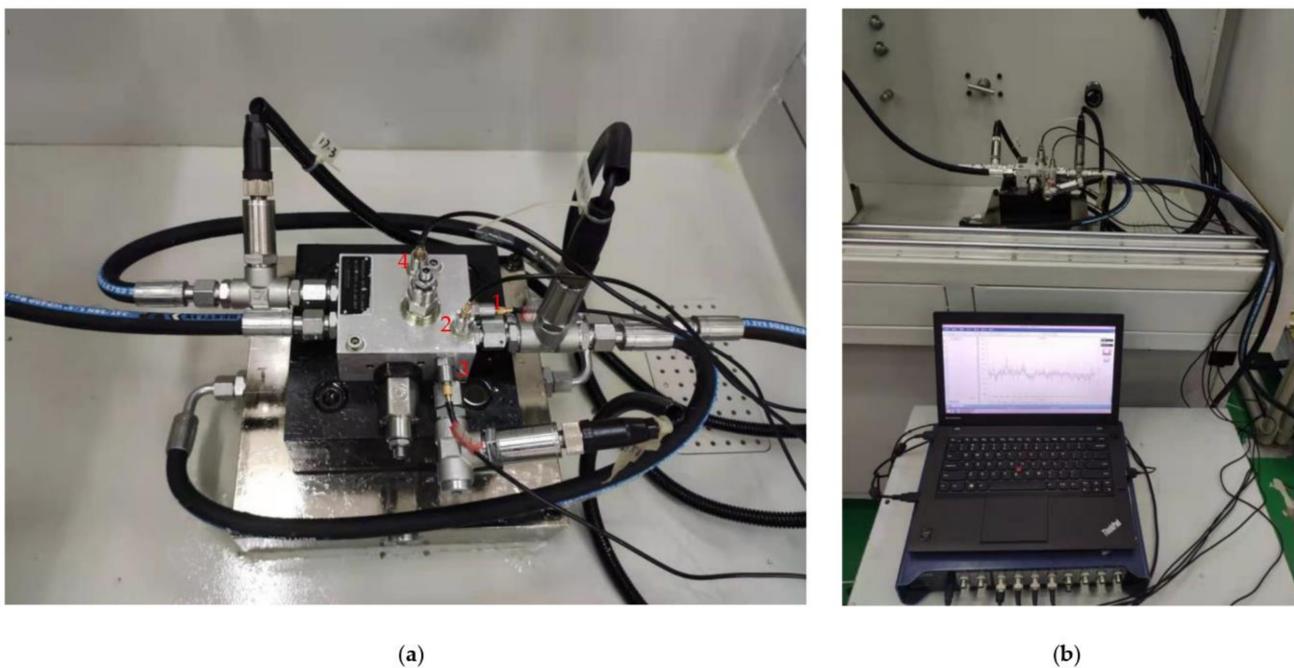


Figure 8. Waterproof valve failure test bench: (a) tested valve part; (b) vibration signal acquisition.

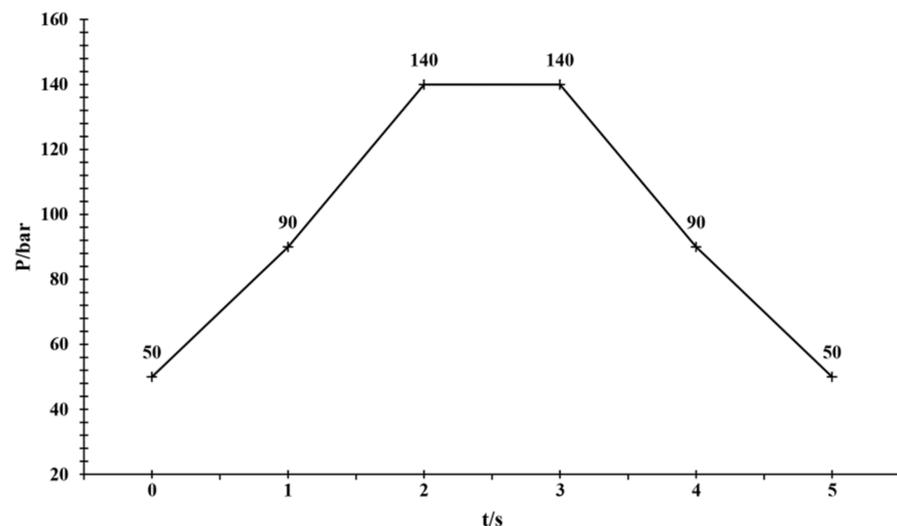


Figure 9. Pulse pressure curve for hydraulic pump settings.

3.3. Data Description

Figure 10 shows the vibration data collected by one of the acceleration sensors on the waterproof valve. The sliding time window method was used to obtain multiple signal samples (each sample contained 2048 data points). Table 1 shows the sample distribution of the fault diagnosis experimental dataset: 80% of the samples (800 training samples) were used for classifier training, and 20% of the samples (200 test samples) were used for classifier testing.

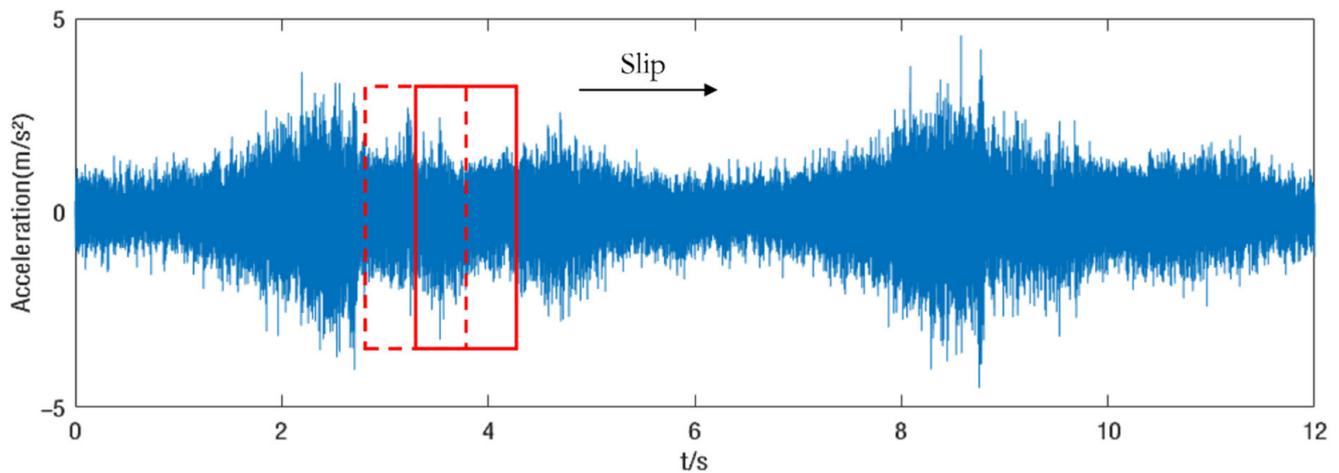


Figure 10. Vibration data collected by one of the acceleration sensors.

3.4. Parameter Setting

The hyperparameters had a significant impact on the performance of the proposed model, as shown in Table 2. In this study, the relevant hyperparameters chosen for the DRSN stage were based on popular suggestions, including the number of layers, the number of convolutional kernels, and the size of convolutional kernels. After pre-processing of the first part of the original signal, the size of the input to the DRSN tributary was 1×2048 , and the output size was 4×1024 after one convolutional layer processing. Next, denoising and feature extraction were performed in several RSBU: the kernel size, kernel number and stride in each RSBU were slightly different. The purpose was to extract features that are more suitable for the network and more different from the next SDCAE tributaries.

Table 2. Parameter settings for each layer of the adaptive feature extraction model.

Number of Components	The Name of the Layer	Kernel Size	Kernel Number	Stride	Output Size
DRSN phase					
1	Input 1				1×2048
1	Conv	3	4	2	4×1024
1	RSBU	3	4	2	4×512
3	RSBU	3	4	1	4×512
1	RSBU	3	8	2	8×256
3	RSBU	3	8	1	8×256
1	RSBU	3	16	2	16×128
3	RSBU	3	16	1	16×128
1	GAP				16
SDCAE phase					
1	Input 2				1×2048
3	Conv	2	16/32/64	2	64×256
3	Trans_Conv	2	32/16/1	2	1×2048

Table 2. Cont.

Number of Components	The Name of the Layer	Kernel Size	Kernel Number	Stride	Output Size
3	Conv	2	16/32/64	2	64×256
3	Trans_Conv	2	32/16/1	2	1×2048
3	Conv	2	16/32/64	2	64×256
3	Trans_Conv	2	32/16/1	2	1×2048
1	FC				16
Fusion and classification phase					
1	Fusion layer				32
1	FC layer				5
1	Dropout layer				5

In the SDCAE stage, because the vibration signal was a one-dimensional time series, the SDCAE used a denoising self-encoder with a one-dimensional convolution layer to process the input noise. Therefore, the model could take the original signal directly as an input without any additional processing. A single DCAE includes three convolutional layers as an encoder and three transposed convolutional layers as a decoder. The size of the same input signal is 1×2048 .

In the fusion layer, the arrays of eigenvalues of the two streams are stitched together. During the training process, we defined some optimization-related hyperparameters, used the Adam optimizer, and set the learning rate of 100 epochs to 0.001. The L2 regularization parameter was set to 0.0001, the minibatch size is set to 128, and the dropout rate was set to 0.5. Finally, the FC output layer had 5 neurons, equal to the number of trained categories (i.e., 1 healthy state and 4 faulty states).

4. Results and Discussion

4.1. Confusion Matrix

A confusion matrix is an error matrix used to evaluate the performance of supervised learning algorithms. In supervised learning, the confusion matrix is a square matrix and the coordinates of the square matrix are the predicted labels and the true labels, respectively. The predicted values are compared with the true values under the same characteristics, and if both are valid, they are placed in the corresponding matrix positions; if not, they are placed in the mismatched matrix positions. Figure 11 shows the confusion matrix of the proposed model compared with other models. Figure 11a shows the confusion matrix of the stacked autoencoder (SAE) fault diagnosis model, which is an unsupervised algorithm that automatically learns features from unlabeled data; Figure 11b is the confusion matrix of the stacked denoising convolutional autoencoder (SDCAE) fault diagnosis model, which is a variant of the SAE; Figure 11c is the confusion matrix of the empirical mode decomposition convolutional neural network (EMDCNN) conventional fault diagnosis model without data fusion; Figure 11d shows the confusion matrix of the EMDCNN–EMDCNN fault diagnosis model using data fusion, and the fault diagnosis model representing the two-way signal brings two different acceleration sensors to the convolutional neural networks after EMD processing, respectively; Figure 11e shows the confusion matrix of the model proposed in this paper. The comparison between (a) and (b) shows that the adaptive denoising had a lower error rate than the default denoising. By comparing (b) with (c), the performance of adaptive denoising algorithm was established to be better than that of the artificial denoising algorithm. The comparison between (c) and (d) shows that the multi-path data fusion algorithm was better than the single path signal algorithm in fault diagnosis. In addition, there are always some test samples that will produce large deviation misdiagnoses, as depicted in Figure 11c, which shows that the traditional manual feature extraction is still very random for fault diagnosis. In the proposed model, only individual samples are misdiagnosed, and the overall accuracy is high.

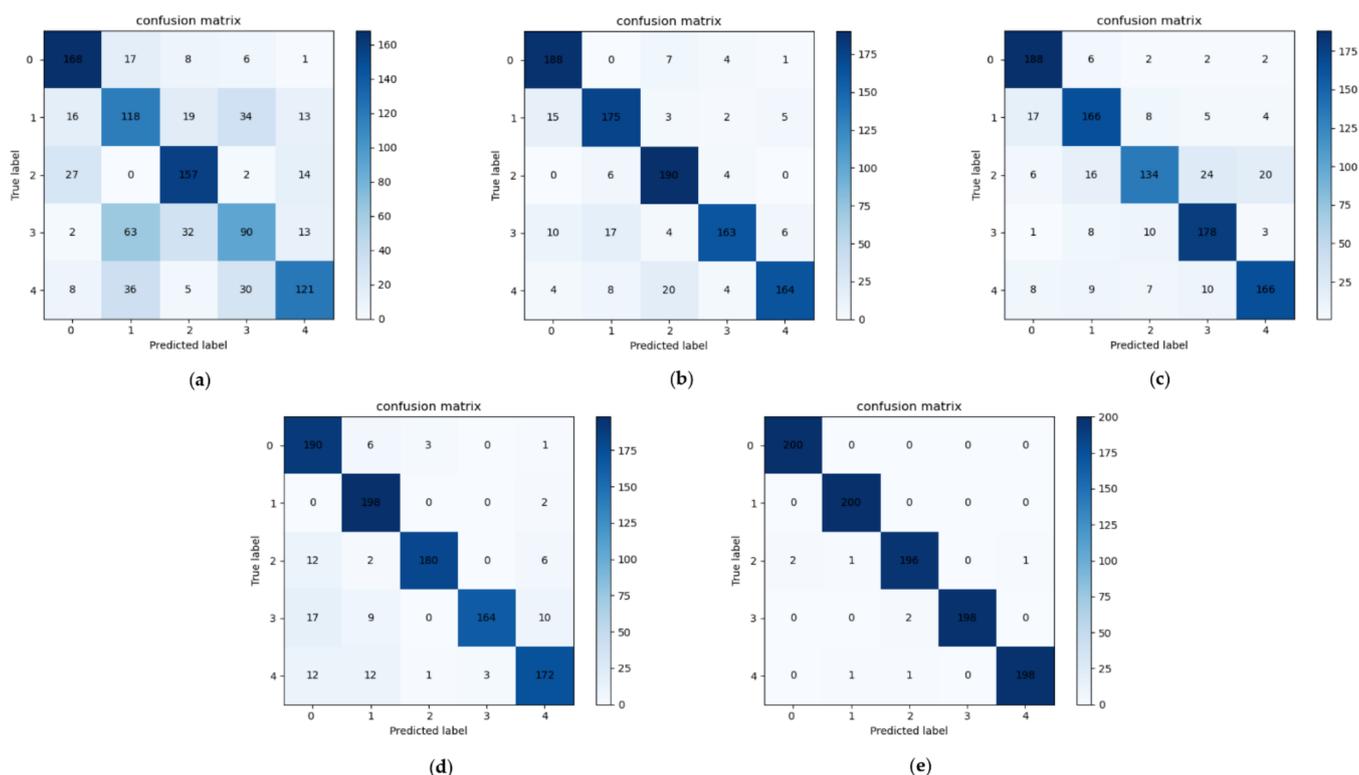


Figure 11. Confusion matrix for different methods: (a) SAE; (b) SDCAE; (c) EMDCNN; (d) EMDCNN-EMDCNN; (e) the proposed method.

Among the five methods, the method proposed in this paper had the lowest misdiagnosis rate and the highest overall accuracy rate. In addition, the stability of the diagnosis method was also very high, mainly due to: (1) The samples collected by two different acceleration sensors were imputed into different model branches for processing. Due to the different denoising model principles in each branch, the generated characteristic signal could avoid interference from redundant components; (2) Additional information obtained from multiple paths was helpful to improve the robustness of the proposed model.

4.2. Verification with Different Signal-to-Noise Ratios

Hydraulic systems work in complex environments with noise interference in practical engineering. In order to evaluate the noise immunity of the proposed fault diagnosis model, we artificially added Gaussian white noise to the measured vibration signal to further evaluate its performance under different signal-to-noise ratios (SNRs). The corresponding SNR is defined as:

$$SNR = 10 \log_{10} \left(\frac{P_{measured\ signal}}{P_{noise}} \right) \tag{12}$$

In order to conduct a fair comparison, the neural network models used for comparison were kept at as similar scale as possible, and were guaranteed to be the same as the optimizer of the proposed model; due to space limitations, we will not elaborate on the specific parameter settings in this paper. The compared networks were SAE, SDCAE, EMDCNN, and EMDCNN-EMDCNN.

In the analysis result of fault diagnosis, if only one experiment of the data sample is used as the final result, the influence of random factors is not sufficient to verify the performance of the fault diagnosis model. In order to accurately measure the model's ability to make correct decisions, we divided different testing samples 10 times, recorded the diagnostic accuracy of each time, and took the average accuracy of 10 tests as the final diagnostic result. Considering the interference of noise level on the experiment in actual engineering practice, this experiment applied noise disturbance to samples in the testing

set, and the testing sample dataset was divided randomly. Moreover, we used the standard deviation to describe the stability of the model. The smaller the standard deviation, the more stable the model. The average accuracy and standard deviation were calculated as shown in the following equations:

$$Accuracy = \frac{T}{T + F} \tag{13}$$

$$Std = \sqrt{\frac{1}{N} \sum_{i=1}^N (acc_i - acc_{avg})^2} \tag{14}$$

where T denotes the number of successfully diagnosed samples and F denotes the number of unsuccessfully diagnosed samples. The letter N denotes the number of diagnoses, which indicates the i th diagnostic accuracy, whereas acc_{avg} denotes the average diagnostic accuracy, here denoting the average diagnostic accuracy of 10 repetitions.

In this study, the fault diagnosis experiments of the proposed method and other methods were conducted under noise environments with different signal-to-noise ratios, and the noise resistance was evaluated by the average accuracy and standard deviation, as shown in Figure 12.

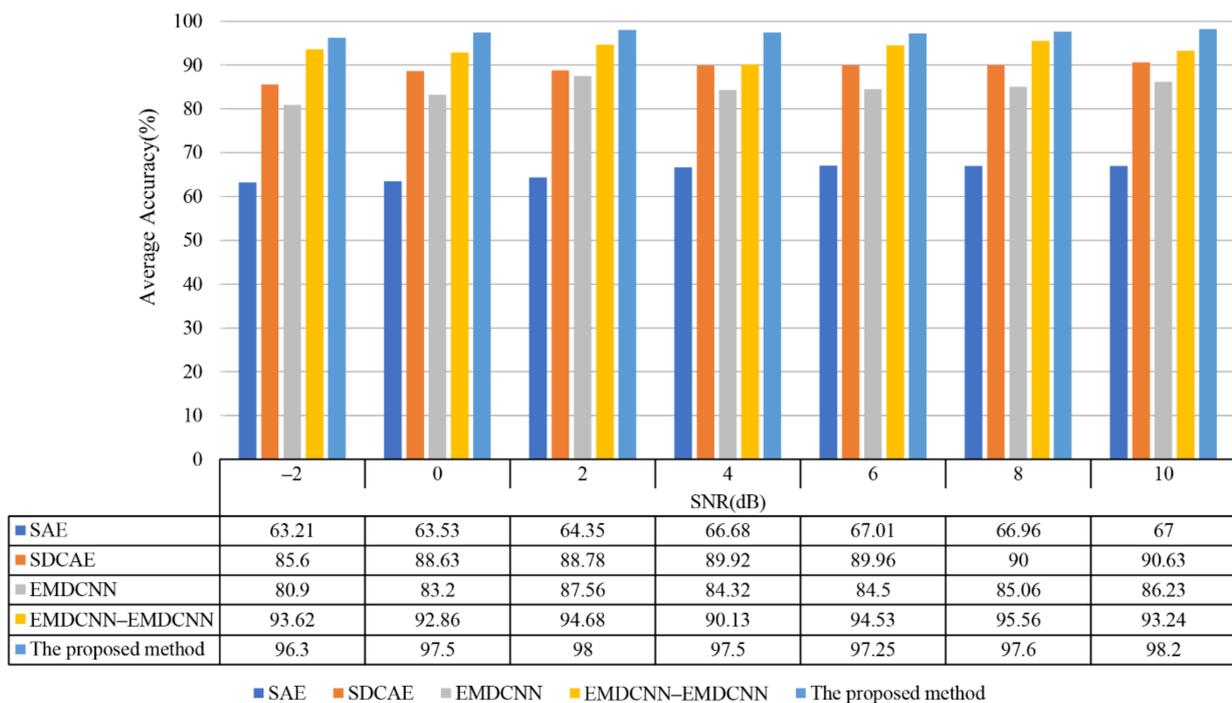


Figure 12. Average accuracy (%) of different methods for waterproof valve fault diagnosis at different signal-to-noise ratios.

As can be seen in Figure 12, the average accuracy of the methods proposed in this paper was above 96% at different signal-to-noise ratios. Notably, the diagnostic accuracy of manually denoised multi-sensor models, such as EMDCNN-EMDCNN, is generally higher than that of the single sensor. This is because the fault information obtained by multiple sensors is more abundant. Even so, in Table 3, the stability of the artificial denoising model is consistently worse than that of the adaptive denoising, regardless of a single sensor or multiple sensors, which indicates that there is a randomness in the artificial denoising method, and the adaptive denoising method is superior. In addition, we can see that the model proposed in this paper was more stable under different signal-to-noise ratios and adapted to the background of strong noise in engineering practice.

Table 3. Fault diagnosis standard deviation of different methods (%).

Method	SNR (dB)						
	−2	0	2	4	6	8	10
SAE	2.76	1.87	1.89	2.89	2.02	2.048	2.62
SDCAE	2.13	2.00	2.06	2.11	2.25	2.22	2.12
EMDCNN	2.89	3.18	2.95	2.84	2.94	3.01	2.95
EMDCNN–EMDCNN	2.22	2.12	2.02	2.35	2.34	2.24	2.25
The proposed method	0.32	0.43	0.39	0.33	0.51	0.28	0.57

5. Conclusions

A fault diagnosis method has been proposed, which was based on a new adaptive feature extraction model (the multi-path signals to improved DRSN–SDCAE), especially for the fault detection of waterproof valves as a key component in construction mixing machinery. Theoretical and experimental results show that compared with other intelligent fault diagnosis methods, the fault diagnosis method proposed in this paper can independently and effectively extract fault features from strong and complex noise to obtain accurate fault diagnosis results without relying on signal processing technology and diagnosis experience. When the signal-to-noise ratio (SNR) changed from −2 to 8 dB, the standard deviations of the proposed method were only 0.32%, 0.43%, 0.39%, 0.33%, 0.51%, and 0.28%, which are much smaller than those of other algorithms. Moreover, especially when the signal-to-noise ratio was 10 dB, the proposed diagnosis method could be used for adaptive feature extraction, and the average diagnosis accuracy of the model exceeded 98%.

Intelligent control in the field of mechanical engineering lags behind the development of other industrial fields. The main bottleneck lies in the intelligent fault diagnosis of hydraulic components in hydraulic systems. Therefore, expanding the fault diagnosis method to each key component of the hydraulic system is the development direction of our work in the future.

Author Contributions: Writing—original draft preparation, R.Z.; writing—review and editing, J.Y., H.T., J.X. and Y.R. 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 (No. 52175060) and the Zhejiang Provincial Natural Science Foundation of China (No. LY20E050028).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data is not publicly available due to privacy issues.

Acknowledgments: The authors are grateful to support from the National Natural Science Foundation of China (No. 52175060) and the Zhejiang Provincial Natural Science Foundation of China (No. LY20E050028).

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

References

- Zhang, T.; Liu, S.; Wei, Y.; Zhang, H. A novel feature adaptive extraction method based on deep learning for bearing fault diagnosis. *Measurement* **2021**, *185*, 110030. [[CrossRef](#)]
- Lu, C.; Wang, S.; Zhang, C. Fault diagnosis of hydraulic piston pumps based on a two-step EMD method and fuzzy C-means clustering. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2016**, *230*, 2913–2928. [[CrossRef](#)]
- Tang, H.B.; Wu, Y.X.; Hua, G.J.; Ma, C.X. Fault diagnosis of pump using EMD and envelope spectrum analysis. *J. Vib. Shock* **2012**, *31*, 44–48.
- Yang, T.; Yu, X.; Li, G.; Dou, J.; Duan, B. An early fault diagnosis method based on the optimization of a variational modal de-composition and convolutional neural network for aeronautical hydraulic pipe clamps. *Meas. Sci. Technol.* **2020**, *31*, 055007.
- Tang, X.; Hu, B.; Wen, H. Fault Diagnosis of Hydraulic Generator Bearing by VMD-Based Feature Extraction and Classification. *Iran. J. Sci. Technol. Trans. Electr. Eng.* **2021**, *45*, 1227–1237. [[CrossRef](#)]
- Wang, Y.; Li, Q.; Zhou, F.; Zhou, Y.; Mu, X. A New Method with Hilbert Transform and Slip-SVD-Based Noise-Suppression Algorithm for Noisy Power Quality Monitoring. *IEEE Trans. Instrum. Meas.* **2018**, *68*, 987–1001. [[CrossRef](#)]

7. Konar, P.; Chattopadhyay, P. Multi-class fault diagnosis of induction motor using Hilbert and Wavelet Transform. *Appl. Soft Comput.* **2015**, *30*, 341–352. [[CrossRef](#)]
8. Du, W.; Wang, Z.; Gong, X.; Wang, L.; Luo, G. Optimum IMFs Selection Based Envelope Analysis of Bearing Fault Diagnosis in Plunger Pump. *Shock Vib.* **2016**, *2016*, 1248626. [[CrossRef](#)]
9. Zhu, Y.; Wang, Q.; Wang, Y.; Yuan, S.; Tang, S.; Zheng, Z. A novel extraction method for useful component of vibration signal combining variational mode decomposition and relative entropy. *AIP Adv.* **2021**, *11*, 035129. [[CrossRef](#)]
10. 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](#)]
11. 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](#)]
12. Tang, S.; Zhu, Y.; Yuan, S.; Li, G. Intelligent Diagnosis towards Hydraulic Axial Piston Pump Using a Novel Integrated CNN Model. *Sensors* **2020**, *20*, 7152. [[CrossRef](#)]
13. Huang, J.; Wang, X.; Wang, D.; Wang, Z.; Hua, X. Analysis of Weak Fault in Hydraulic System Based on Multi-scale Permutation Entropy of Fault-Sensitive Intrinsic Mode Function and Deep Belief Network. *Entropy* **2019**, *21*, 425. [[CrossRef](#)] [[PubMed](#)]
14. Zhang, P.; Chen, X. Internal Leakage Diagnosis of a Hydraulic Cylinder Based on Optimization DBN Using the CEEMDAN Technique. *Shock Vib.* **2021**, *2021*, 8856835. [[CrossRef](#)]
15. Ji, X.; Ren, Y.; Tang, H.; Shi, C.; Xiang, J. An intelligent fault diagnosis approach based on Dempster-Shafer theory for hydraulic valves. *Measurement* **2020**, *165*, 108129. [[CrossRef](#)]
16. Jia, C.; Kang, K.; Gao, W.; Yang, D.; Chen, L.; Ai, C. Fault Prediction of Electro-hydraulic Servo Valve Based on CNN+ LSTM Neural Network. *Chin. Hydraul. Pneum.* **2020**, *12*, 173.
17. Gareev, A.; Protsenko, V.; Stadnik, D.; Greshniakov, P.; Yuzifovich, Y.; Minaev, E.; Nikonorov, A. Improved Fault Di-agnosis in Hydraulic Systems with Gated Convolutional Autoencoder and Partially Simulated Data. *Sensors* **2021**, *21*, 4410. [[CrossRef](#)] [[PubMed](#)]
18. Wang, C.; Zhang, Y.; Ma, J.; Tao, L.; Zhang, H.; Ding, Y. Modified Stacked Convolutional Auto-Encoder for Hydraulic Pump Fault Diagnosis with Unbalanced Data. In Proceedings of the 2021 Global Reliability and Prognostics and Health Management, Nanjing, China, 15–17 October 2021; pp. 1–7.
19. Wen, L.; Li, X.; Gao, L.; Zhang, Y. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method. *IEEE Trans. Ind. Electron.* **2017**, *65*, 5990–5998. [[CrossRef](#)]
20. Mallak, A.; Fathi, M. Sensor and component fault detection and diagnosis for hydraulic machinery integrating LSTM auto-encoder detector and diagnostic classifiers. *Sensors* **2021**, *21*, 433. [[CrossRef](#)]
21. Han, H.; Hao, L.; Cheng, D.; Xu, H. GAN-SAE based fault diagnosis method for electrically driven feed pumps. *PLoS ONE* **2020**, *15*, e0239070. [[CrossRef](#)]
22. Liu, R. Fault Diagnosis of Emulsion Pump Based on Depth Residual Network. *Coal Mine Mach.* **2021**, *42*, 177–180.
23. Huang, T.; Zhang, Q.; Tang, X.; Zhao, S.; Lu, X. A novel fault diagnosis method based on CNN and LSTM and its application in fault diagnosis for complex systems. *Artif. Intell. Rev.* **2021**, *55*, 1289–1315. [[CrossRef](#)]
24. Meng, Z.; Zhan, X.; Li, J.; Pan, Z. An enhancement denoising autoencoder for rolling bearing fault diagnosis. *Measurement* **2018**, *130*, 448–454. [[CrossRef](#)]
25. Fu, Q.; Wang, H. A Novel Deep Learning System with Data Augmentation for Machine Fault Diagnosis from Vibration Signals. *Appl. Sci.* **2020**, *10*, 5765. [[CrossRef](#)]
26. Zhao, M.; Zhong, S.; Fu, X.; Tang, B.; Pecht, M. Deep Residual Shrinkage Networks for Fault Diagnosis. *IEEE Trans. Ind. Inform.* **2019**, *16*, 4681–4690. [[CrossRef](#)]
27. Azamfar, M.; Singh, J.; Bravo-Imaz, I.; Lee, J. Multisensor data fusion for gearbox fault diagnosis using 2-D convolutional neural network and motor current signature analysis. *Mech. Syst. Signal Process.* **2020**, *144*, 106861. [[CrossRef](#)]
28. Tao, J.; Liu, Y.; Yang, D. Bearing Fault Diagnosis Based on Deep Belief Network and Multisensor Information Fusion. *Shock Vib.* **2016**, *2016*, 9306205. [[CrossRef](#)]