Article

An Adaptive Fusion Convolutional Denoising Network and Its Application to the Fault Diagnosis of Shore Bridge Lift Gearbox

Rongqiang Zhao *[ID] **and Xiong Hu**

Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China; huxiong@shmtu.edu.cn
* Correspondence: zhaorongqiang@stu.shmtu.edu.cn

Abstract: Traditional fault diagnosis methods are limited in the condition detection of shore bridge lifting gearboxes due to their limited ability to extract signal features and their sensitivity to noise. In order to solve this problem, an adaptive fusion convolutional denoising network (AF-CDN) was proposed in this paper. First, a novel 1D and 2D adaptive fused convolutional neural network structure is built. The fusion of both 1D and 2D convolutional models can effectively improve the feature extraction capability of the network. Then, a gradient updating method based on the Kalman filter mechanism is designed. The effectiveness of the developed method is evaluated by using the benchmark datasets and the actual data collected for the shore bridge lift gearbox. Finally, the effectiveness of the proposed algorithm is proved through the experimental validation in the paper. The main contributions of this paper are described as follows: the proposed AF-CDN can improve the diagnosis accuracy by 1.5–9.1% when compared with the normal CNN methods. The robustness of the diagnostic network can be significantly improved.

Keywords: gearbox fault diagnosis; convolution fusion; state identification

1. Introduction

With the further development of globalization, automated container terminals (ACTs) are increasingly widespread. The automatic loading and unloading of containers by ACTs ensure the orderly flow of goods. This plays an important role in the globalization of the economy. As an integral part of the ACT, the reliable working condition of the port crane ensures the efficient operation of the entire terminal. The gearbox of the port crane, as an important power component, works for long periods of time and under heavy loads. A reliable condition monitoring and fault diagnosis system for the port crane gearbox is essential for a port crane [1,2]. The failure of a port crane can lead to port blockages and unnecessary economic losses or even cause injury or death. Therefore, it is essential to ensure that it works safely and securely. In practical scenarios, it is usually experienced experts or engineers who perform the maintenance of the equipment through their previous experience. For example, an experienced expert can determine the status of a device by tapping on it and locating faults according to the feedback signal characteristics. However, some critical equipment requires effective online monitoring so that faults can be detected and handled as soon as they occur.

The development of sensors such as vibration sensors, acoustic sensors, temperature sensors, pressure sensors, etc., can provide an effective means of obtaining information for such equipment [3–5]. This provides an effective means of detecting equipment in real time. Traditional analysis methods are mainly based on manual feature extraction of the collected signals. The methods of feature extraction for signals include time domain features, frequency domain features, and time–frequency domain features. Common features of the time domain include the mean value, standard deviation, root mean square value, peak value, shape indicator, skewness, kurtosis, crest indicator, clearance indicator, impulse indicator, etc., [6]. Frequency domain features usually refer to feature signals

extracted from the frequency spectrum, mainly including the mean frequency, frequency center, root mean square frequency, the standard deviation of frequency, etc. [7]. Frequency signals often better represent some of the hidden features of the signal than the time domain. The time–frequency domain features include energy entropy, which is usually extracted by wavelet transform (WT), wavelet package transform (WPT), or empirical model decomposition (EMD) [8–13].

With the development of artificial intelligence technology [14,15], machine learning methods are used to identify faults based on the features extracted, such as expert systems, ANN, and SVM [16,17]. The intelligent algorithms of the fault diagnosis model have a strong nonlinear fitting capability [18,19]. With a provided training target and an optimization algorithm, the intelligent algorithm often achieves a good diagnosis result after continuous iteration of the optimal search. However, the efficiency of signal feature extraction may have a significant impact on the diagnostic accuracy of these methods

In this paper, a 1D and 2D adaptive fusion convolutional neural network structure is proposed, while the parameters are integrated with a Kalman filter during the iterative training process. AF-CDN converts raw data into 2D data and uses the fast Fourier transform (FFT) technique to extract features from the signal. Then, the two signals are adaptively fused. At the same time, the use of Kalman filter technology can effectively eliminate the influence of noise in the raw data on the diagnostic results. The network has excellent diagnostic accuracy, while the robustness is greatly improved. Based on the historical data, we built an online condition monitoring system for port crane gearboxes. We also test our proposed algorithm on a public bearing dataset from Case Western Reserve University, and the results show that AF-CDN is well suited for different situations.

The main contributions of this paper are summarized as follows:

(1) A 1D and 2D adaptive convolutional approach is proposed, through which the feature extraction capability of the network can be greatly enhanced. We design a 1D and 2D fused convolutional signal extraction layer (perception layer). First, the FFT-processed 1D information is fed into the 1D convolution input, and then the sequence of the original signal is aligned and fed into the 2D convolution input.

(2) A Kalman filter-based method for updating network parameters is proposed. Improvements are made to the minibatch stochastic gradient descent (MSGD) method. The information within the minibatch is effectively integrated based on the Kalman filter mechanism.

The subsequent sections of this manuscript are organized as follows: Section 2 presents the preliminary work. Section 3 describes the method proposed in detail. Section 4 designs experimental validation for the proposed algorithms. Section 5 presents the conclusion and provides suggestions for future work

2. Preliminary Work

2.1. Convolutional Neural Networks

Convolutional neural networks are an important branch of neural networks [20]. However, unlike back propagation (BP) neural networks, convolutional neural networks have a strong feature extraction capability. After the convolutional operation, the network can perform feature extraction on the signal fed into the network.

A CNN has a convolutional layer, a pooling layer, and a full connection layer. Since the proposal of the convolutional neural network, a rich variety of CNN structures have been developed over the decades, including LeNET, AlexNET, VGG, GoogleLeNET, ResNET, DenseNET, etc. A typical convolutional neural network structure is shown in Figure 1.
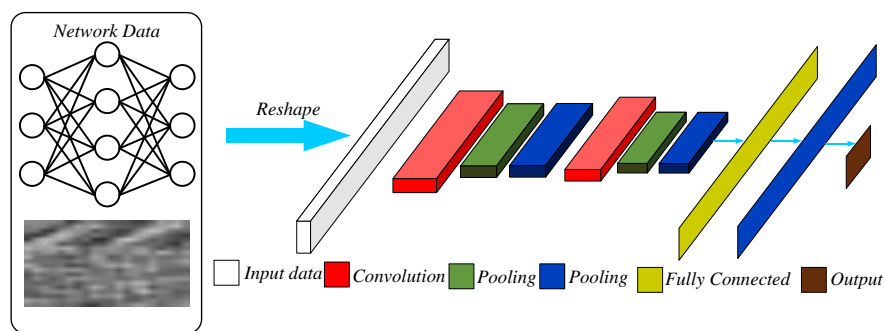
Figure 1. Typical convolutional neural network structure.

The forward computation of a CNN can be expressed as follows.

$$\mathcal{G}(X) \;=\; g^{(L)}(\ldots g^{(2)}(g^{(1)}(X,\theta^{(1)}),\theta^{(2)})\ldots,\theta^{(L)}). \tag{1}$$

$\mathcal{G}$ is the mapping equation of the network. g is the nonlinear function of each layer inside the network. θ is the connection parameters of each layer and L is the number of layers of the neural network. $X \;=\; \{x_1, x_2 \ldots x_p \ldots x_Q\}$ is the input to the network, which can be one- or two-dimensional. Q is the number of data in the dataset.

The convolutional layer of a CNN consists of a convolution core and a bias. After the input of the network has been convolved, the bias of the layer is added, and, finally, the output of the network is obtained by passing through the nonlinear layer. The equation for the convolution layer is shown as follows [20].

$$O_i^{(l)} \;=\; g^{(l-1)}\left(\sum_{j=1}^{n^{(l-1)}} w_{ij}^{(l)} * X_j^{(l-1)} + b_i^{(l)}\right). \tag{2}$$

In the formula, $O^{(l)}$ is the output of the lth layer. $i \;=\; 1,2\ldots n^{(l)}$, $n^{(l)}$ is the output size of the lth layer. $j \;=\; 1,2\ldots n^{(l-1)}$, $n^{(l-1)}$ is the output size of the $(l-1)$th layer. w is the value of the convolution core. b is the value of bias. The pooling layer is used to further extract the information from the convolutional output. The pooling operation can be max pooling, down pooling or average pooling. After the pooling operation, the representative features in the local area are further extracted. Taking down pooling as an example, it takes the smallest value inside the pooling size range and generates a new output. As shown in Equation (3), let the pooling size be $p \times p$.

$$x_{ijl} \;=\; \min(o_{i'j'l} : i \le i' < i + p, j \le j' < j + p). \tag{3}$$

The input information to the network is passed through a number of convolution, pooling, and nonlinear computations. Then, the output value of the last pooling layer is reshaped, and this value is fed into the full connecting layer. Finally, the diagnosis result is provided after a softmax layer. We assume that the network output has K classes. $Y \;=\; \{y_1, y_2 \ldots y_q \ldots y_Q\}$ is the output set of the dataset. Q is the number of data in the dataset. If $y_q \in \{1, 2 \ldots k \ldots K\}$, the predicted output of the network is shown as Equation (4):

$$\hat{y}(\hat{y}_q \;=\; k\big|o^{(L-1)}; w^{(L)}) \;=\; \frac{e^{w_k^{(L)} o^{(L-1)}}}{\displaystyle\sum_{i=1}^{K} e^{w_i^{(L)} o^{(L-1)}}}. \tag{4}$$

where $w^{(L)}$ is the network parameter for the softmax layer. $\hat{Y} \;=\; \{\hat{y}_1, \hat{y}_2 \ldots \hat{y}_q \ldots \hat{y}_Q\}$.

The network parameters are updated using the minibatch stochastic gradient descent method after the network has completed forward propagation. J is the loss function of the network, which can be mean square error (MSE) or cross-entropy, etc. After each

forward calculation is completed, the output of the network is updated by iterating backward derivation. The network parameters are updated by Equations (5) and (6). γ is the learning rate.

$$W = W - \gamma \frac{\partial J(W, B; X, \hat{Y})}{\partial W}. \tag{5}$$

$$B = B - \gamma \frac{\partial J(W, B; X, \hat{Y})}{\partial B}. \tag{6}$$

2.2. CNN-Based Fault Diagnosis

Convolutional neural networks are widely used in data-based fault diagnosis applications. Based on the different types of convolutional kernel operations, they can be divided into 1D-CNNs and 2D-CNNs. The 1D convolutional structure is proposed mainly in response to the fact that the neural network often requires manual feature extraction of the raw signal when performing recognition. One-dimensional convolutional neural networks can use the raw data directly as an input to the neural network. For example, Eren et al. proposed an adaptive 1D convolution method that can extract data features directly from the raw time-domain data [21]. An online diagnostic network based on 1D-CNN was designed for the effective diagnosis of a gearbox, where vibration sensors cannot be used, and the signal was collected by a rotary encoder [22]. A deep convolutional structure, Deep Inception Net with Atrous Convolution (ACDIN), was designed in [23] based on 1D-CNNs, which improved the feature extraction ability of the network by adding an inception layer. The 1D convolution was improved by Atrous convolution. This led to a significant increase in the diagnostic capability of the network. To address the problem of uneven distribution of samples in the dataset, Jia et al. proposed a 1D-CNN with normalized weights for one-dimensional input data [24]. Jiang et al. designed a multi-scale signal resolution method using one-dimensional convolution for signal feature extraction, which achieved a positive result [25]. Appana et al. proposed the extraction of the raw signal by CNN for the case of multiple faults and environmental influences [26]. One-dimensional CNNs have excellent environmental adaptability and can effectively resist interference.

The main issue that needs to be solved when 2D convolutional neural networks are used for fault diagnosis is how to convert the acquired 1D raw signal into 2D data that can be fed into the network. A number of approaches have been proposed to solve this problem. Guo et al. used the residual processed short-time Fourier transform (STFT)-transformed image of the original signal as the input into the CNN [27]. Long et al. proposed a signal to image conversion mechanism to transform the raw time domain signal into 2D grey images [28]. This enables feature extraction of the collected vibration signals in a similar manner to picture recognition. Han proposed a spatiotemporal convolutional neural network (ST-CNN), which extracts spatiotemporal features via the spatiotemporal pattern network (STPN) and then makes a diagnosis based on the CNN [29]. In [30], Yu et al. used a pseudo-color map to represent the data extracted by STFT and then fed the images into a CNN for training recognition. Sun et al. used the dual-tree complex wavelet transform method to extract features from the raw data, and the DTCWT wavelet sub-bands were used as multiple rows of a matrix so that a 2D signal was formed and sent to 2D convolution for processing [31]. Similarly, the Hilbert envelope demodulation spectra (HEDS) of reconstructed signals in each frequency band were also spliced into the 2D signal matrix [32]. The HEDS of the reconstructed signal for each frequency band were stitched into a 2D signal matrix to produce a 2D signal. The time domain signals were arranged row by row to form a 2D input matrix as the network input for diagnostics. Min et al. arranged the time domain signals row by row to form a 2D input matrix as the network input for diagnostics [33].

3. The Proposed Method

3.1. Adaptive Fusion Convolutional Denoising Network

　　The proposed AF-CDN structure diagram is shown in Figure 2. Firstly, the equipment data are collected by vibration sensors. Then, the 1D data are obtained by FFT and fed into the 1D channel of the perception layer. The raw data are arranged in order to obtain 2D data. These data are then fed into the 2D channel of the perception layer. After feature extraction in the perception layer, all feature values are flattened and then pooled. The features then go through two more inception and pooling layers. The output values are fed to the auxiliary classifier for classification after each pooling layer. The final pooling layer output is combined with the output of the auxiliary classifier to provide the final diagnosis result. The loss value is calculated from the output, and if the predetermined loss condition is satisfied, then the training is stopped; if not, then training continues.
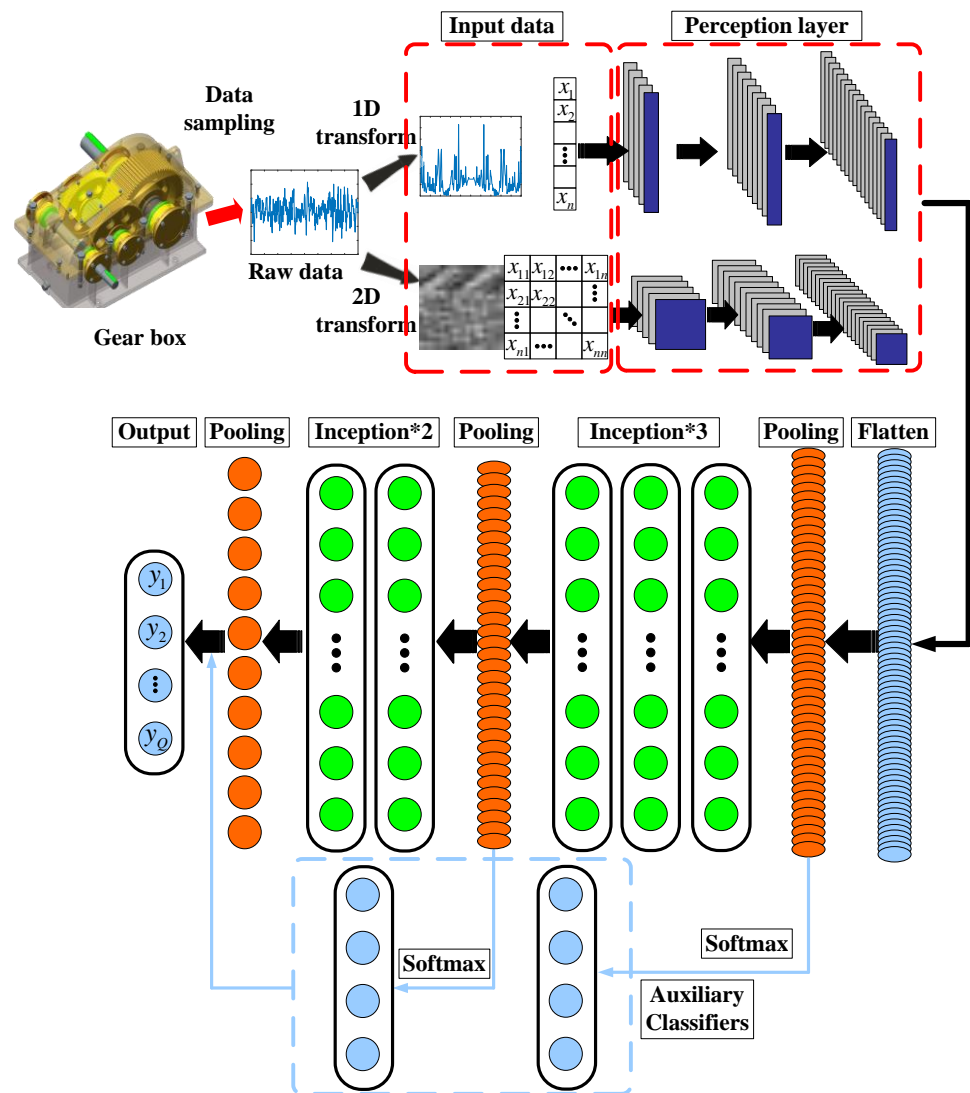

Figure 2. AF-CDN structure diagram.

　　Numerous research results in the field of CV have favorably illustrated the importance of 2D convolution for the accurate recognition of valuable information in 2D images. As shown in Section 2.2, the development of convolutional neural networks in the field of fault diagnosis also started with 1D-CNNs, and then researchers successively proposed various methods to convert 1D data into 2D data, thus enabling 2D-CNNs to be widely used in the field of fault diagnosis. The schematic diagram of the data processing is shown in Figure 3.

After the original data are FFT-transformed, the positive half-axis frequency data are taken and arranged sequentially to obtain 1D data. Two-dimensional data, on the other hand, are arranged with the raw data starting from the first row of the matrix, followed by the second row, until the entire matrix is filled. During the 2D conversion, the signals are arranged sequentially, so the time-series property of the signals is preserved.
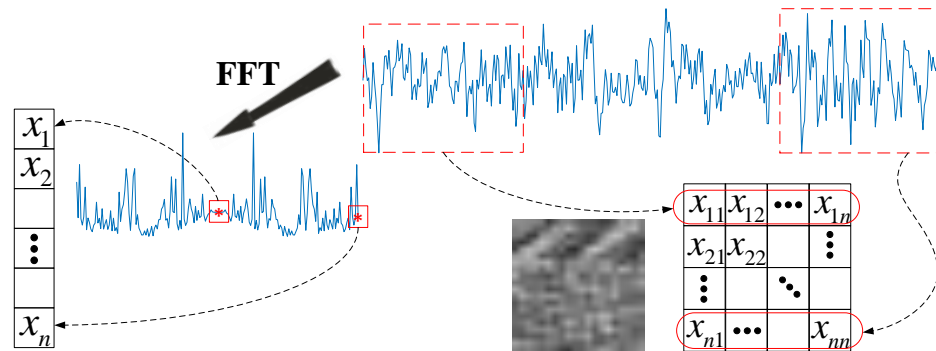

Figure 3. One-dimensional and two-dimensional data transform.

The algorithm flow chart of AF-CDN is shown in Figure 4. The main steps of the proposed algorithm are described as Algorithm 1.
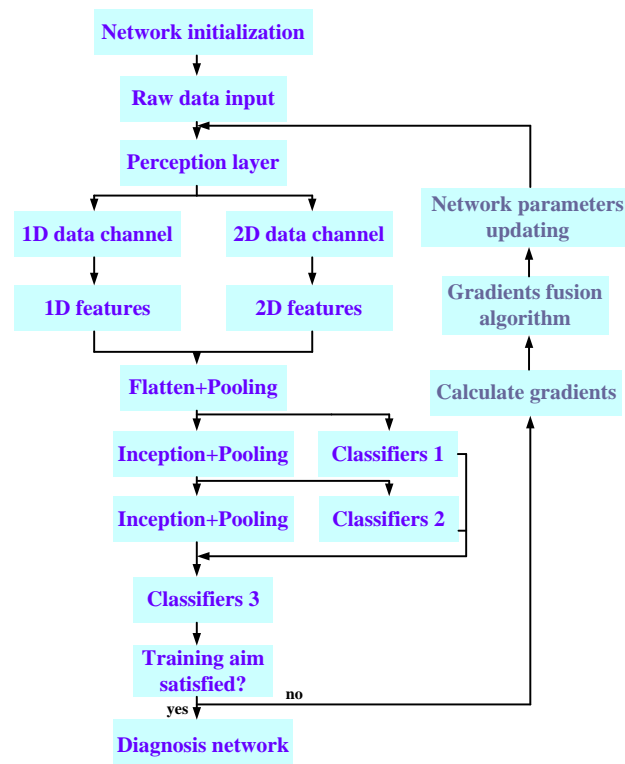

Figure 4. Algorithm flow chart of AF-CDN.

Algorithm 1: AF-CDN

Input: Network $\mathcal{G}$, Training epoch N, Input data $X = \{x_1, x_2 \ldots x_p \ldots x_Q\}$ and $Y = \{y_1, y_2 \ldots y_q \ldots y_Q\}$
Output: Trained $\mathcal{G}^{tr}$
1. Initialize network parameters
2. **For** $i = 1, 2, \ldots N$ **do**
3. Feed raw data into the network
4. Send the data to the perception layer to calculate 1D data and 2D data, respectively
5. Flatten and pool the 1D and 2D features of the perception layer output. The values after pooling are fed into classifier1 and inception layer, respectively
6. The values of the inception layer are pooled and fed into classifier2 and the next inception layer, respectively
7. The values of inception layer are pooled and fed into classifier3
8. Classifier3 combines the values of classifier1 and classifier2 to give the predicted output.
9. Calculate the loss value. Stop training if the training target is met, otherwise step forward.
10. Calculate the gradients and use the fusion algorithm to update the gradient value.
11. End

3.2. Gradient Fusion Algorithm

Ma et al. proposed a Kalman-filter-based fusion method for network parameters updating [6]. Based on this theory, we performed some improvements when we used the MSGD algorithm to calculate the updating gradients. Considering that after the sensor has been selected and the measurement position has been determined, noise in the acquired signal cannot be avoided. Therefore, we considered further analysis of the gradients during the gradient updating process. Signal noise is assumed to be hidden in the gradient information of each sample. Thus, the gradient information is fused inside the minibatch using a Kalman filter.

When calculating the gradient inside the batch, we fuse the gradient information using the Kalman filter on each of the batch size gradients.

We use k and $k-1$ moments as an example, and the Kalman filter-based gradient fusion process is described as follows.

$$w(k) = Fw(k-1) + \delta(k). \tag{7}$$

$$z(k) = H(k)w(k) + \gamma(k). \tag{8}$$

$w(k)$ is the network parameter at the moment k. F is the state transfer matrix. $\delta(k)$ is the state error, $\gamma(k)$ is the measurement error and $\delta(k)$, $\gamma(k)$ conform to a Gaussian distribution. Then according to the Kalman filter formula [34], the follow-up process in the iterative process is as follows.

First, an a priori estimate of the gradient is calculated.

$$\hat{w}^-(k) = F\hat{w}(k-1) + \delta(k). \tag{9}$$

$\hat{w}^-(k)$ is an a priori estimate of the moment k. $\hat{w}(k-1)$ is the optimal estimate at moment $k-1$. Next, we update the a priori estimated covariance.

$$P^-(k) = FP(k-1)F^T + Q(k). \tag{10}$$

$P^-(k)$ is the a priori estimated covariance, which will be used when calculating the Kalman gain. $P(k-1)$ is the posterior estimated covariance at moment $k-1$. $Q(k) = \delta(k)\delta^T(k)$ is the covariance of the state error. The measured value at the k moment is calculated according to Equation (8), and then the Kalman gain is updated [6].

$$K(k) = P^-(k)H(k)^T[H(k)P^-(k)H(k)^T + R(k)]^{-1}. \tag{11}$$

$R(k) = \gamma(k)\gamma^T(k)$ is the covariance of the measurement error. Then, the optimal estimate of the gradient at moment k can be found.

$$
\begin{aligned}
\hat{w}(k) &= \hat{w}^-(k) + K(k)[z(k) - H(k)\hat{w}^-(k)] \\
&= [1 - K(k)H(k)]\hat{w}^-(k) + K(k)z(k).
\end{aligned}
\tag{12}
$$

$[1 - K(k)H(k)]$ is the confidence level of the estimate. $K(k)$ is the confidence level of the measured value.

Finally, the posterior estimated covariance is updated.

$$
P(k) = [I - K(k)H(k)]P^-(k).
\tag{13}
$$

In specific applications, let the batch size be N_B, so $k = \{1, 2 \ldots N_B\}$.

$w(k)$ is the updated gradient value corresponding to the N_B samples. F is the state transfer matrix, it is set to I in the paper.

$H(k)$ is the measurement matrix with the measurement values matching the state values. $H(k)$ is set to I. We update the Kalman gain as in Equation (14).

$$
\begin{aligned}
K(k) &= P^-(k)[P^-(k) + R(k)]^{-1} \\
&= [P(k-1) + Q(k)][P(k-1) + Q(k) + R(k)]^{-1}.
\end{aligned}
\tag{14}
$$

In practical terms, the value of R can be set to a smaller value if the measurement error of the sensor is small, which means that we are more likely to believe that the gradient of each sample is the true value. Conversely, the value of $R(k)$ can be increased appropriately. The value of $Q(k)$ is adjusted according to the range of the gradient values during actual optimization.

4. Experimental Verification

In order to verify the effectiveness of the algorithm proposed in this paper, in Case One, data acquisition, fault classification, and diagnosis results are introduced in detail on the port crane built by NetCMAs. The feasibility of AF-CDN is verified. In Case Two, we verify the algorithm on the open-source rolling bearing fault dataset from Case Western Reserve University. AF-CDN can achieve excellent diagnostic results. The universality of the proposed algorithm is illustrated. The experiments in this paper were implemented on an Intel(R) Core (TM) i7-8550U CPU PC (1.80 GHz, 8 GB RAM) NVIDIA Geforce MX 150 GPU (4 GB) 64 Bit Windows 10 operating system in a Python environment.

4.1. Case One

4.1.1. Dataset Preparation and Parameter Settings

The NetCMAs system is installed with vibration sensors, stress sensors, temperature sensors, etc., which can effectively detect the status of the whole port crane system in real time. The whole system has 32 sampling channels, and sampling information points are distributed in the T frame of the upper beam area, beam rod load area, lifting motor, gearbox, car motor, etc. The collection positions of the gearbox are as follows, V-directional and H-directional vibration on the left side of the high-speed shaft, the temperature on both the left and right side of the high-speed shaft, and V-directional vibration on the left side of the low-speed shaft. The sampling frequency of the detection system is 2.5 kHz. The sampling time is 0.8 s, and the sampling interval is 10 s. In order to avoid continuous data transmission and storage consumption, the valid values of each sampling period are calculated and saved. Figure 5 shows the driving part of the port crane lifting mechanism. Figure 6 shows the reduction gearbox and bearings. Figure 7 shows the installation position of the vibration sensor and the damaged bearing.
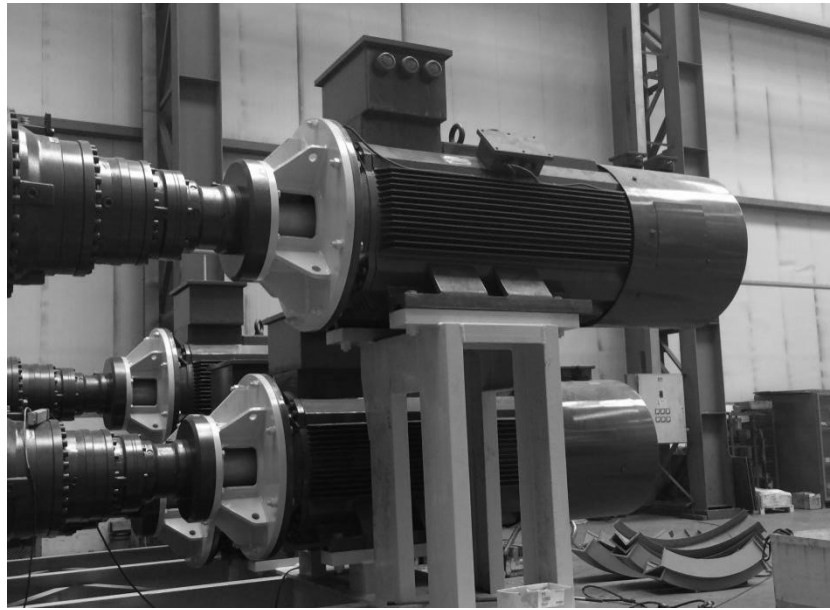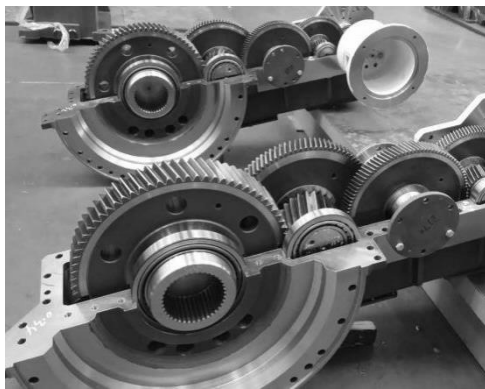
Figure 5. Lifting mechanism of port crane.
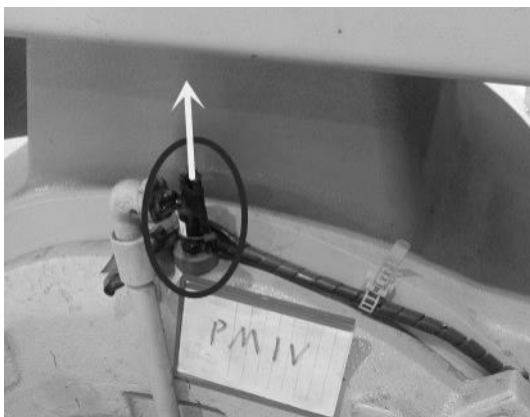

(**a**)

(**b**)

Figure 6. Gear box of lifting mechanism of port crane. (**a**) Drive gear set for installation. (**b**) Gear bearing.
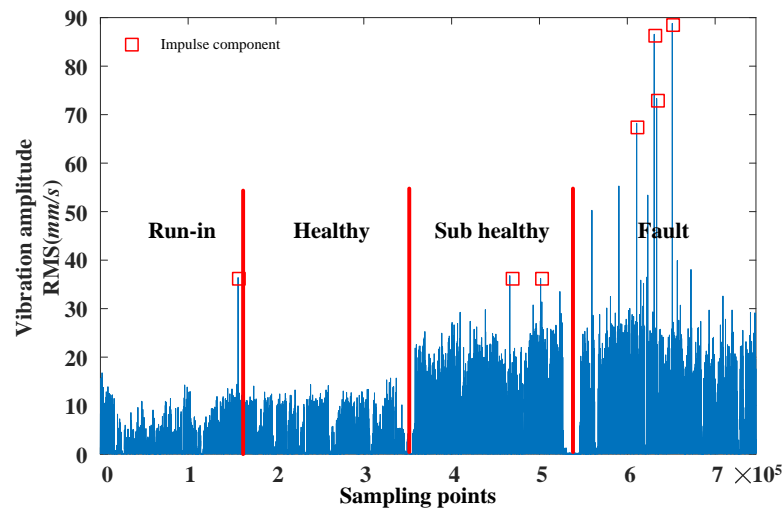

(**a**)

(**b**)

Figure 7. Experimental settings. (**a**) Installation position of vibration sensor. (**a**) Installation position of vibration sensor.

The dataset used in this case comprises four years of data on the No. 8114 lifting bridge of a port crane. The data were recorded from the time when the gearbox was first equipped until the time of failure. Due to extremely large volumes of data, only representative data are shown in Figure 8. It can be seen that there are some shock components in the wave. According to our practical application experience, all data are classified into four categories: healthy (H), sub-healthy (SH), failure (F), run-in period (R), and health (H).

Figure 8. V-direction vibration waveform of the gearbox.

After each new gearbox is re-installed, it will run through a run-in period of time before it enters into a healthy state. After a period of operation, the equipment will be in a sub-healthy state due to the occurrence of wear of the equipment. Eventually, the damage was so significant that the equipment entered a fault state. Table 1 describes in detail the label and quantity information of the four state data in the experiment. Sample labels are onehot encoded, and each state contains 100 samples. The dimension of each sample is 1600, so the dimensions after the 2D conversion are 40×40. In order to ensure the speed of iteration, the batch size is set to between 20 and 40.

Table 1. Four types of fault status information.

Fault Type	Fault Diameter (inch)	Label	Sample Size
H	Healthy	0001	100
SH	Sub-Healthy	0010	100
F	Failure	0100	100
R	Run-in	1000	100

4.1.2. Experiment and Analysis

In order to demonstrate the effectiveness of AF-CDN, a simulation comparison is performed. Table 2 shows the comparison between the algorithm proposed in this paper and the 1D-CNN algorithm extracted by the FFT signal, and the 2D-CNN algorithm based on raw data.

Table 2. Comparison of accuracy.

Number of Experiments	1D-CNN	Rawdata-CNN	AF-CDN
1	91.25%	96.75%	100.00%
2	89.58%	98.75%	99.25%
3	92.50%	98.5%	99.00%
4	90.83%	98.25%	100%
5	93.75%	98.75%	100%
6	89.58%	98.50%	99.50%
7	89.17%	98.00%	100.00%
8	93.75%	98.75%	99.75%
9	90.83%	98.75%	100.00%
10	88.33%	98.50%	99.00%
Mean accuracy	90.56%	98.35%	99.65%

As can be seen from the results of the 10 experiments in the table, AF-CDN provides better diagnostic results than both the algorithm that performs signal spectral analysis alone and the algorithm based on raw signal feature extraction. Figure 9 shows the statistical information of 10 experimental results. It can be seen that AF-CDN can not only effectively improve the diagnosis accuracy but also can greatly reduce the error of each diagnosis result. The performance of AF-CDN is 10.09% higher than that of FFT alone and is 1.3% higher than that of original signal feature extraction. The average execution time of AF-CDN was 0.196 s. The 1D-CNN was 0.083 s and Rawdata-CNN was 0.154 s. The improvement in accuracy compared to 1D-CNN is significant, so the computational consumption is worthwhile. Compared to Rawdata-CNN, the computational consumption is 0.042 s higher. In a practical scenario, the computational complexity can be significantly reduced by using a high-performance GPU device, so it is worthwhile to increase the computational consumption slightly.
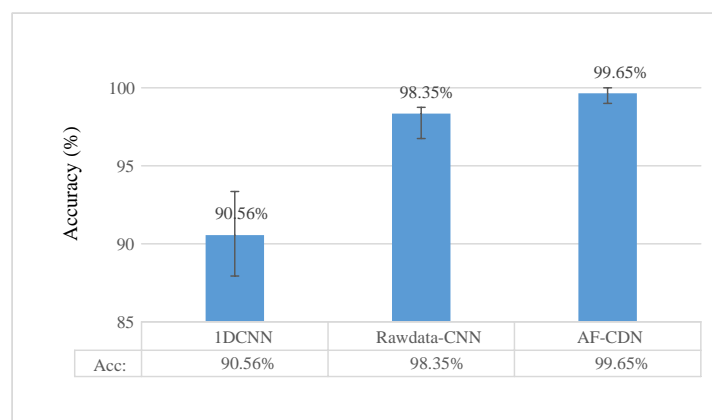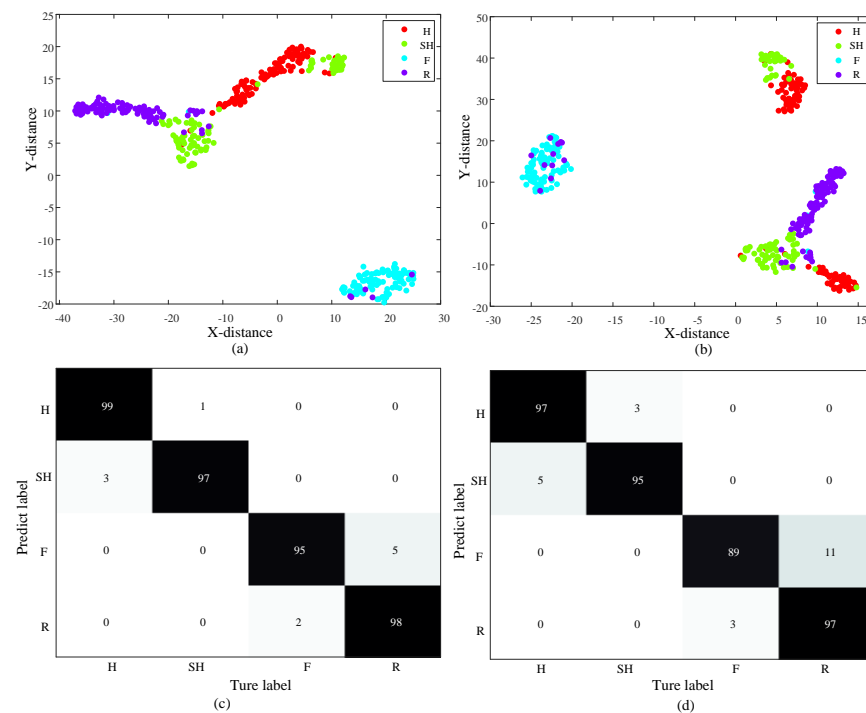

Figure 9. Results of the 10 experiments.

Figure 10 uses T-SNE visualization technology to visually display the experimental classification results. In Figure 10a,c the visualization and error matrix of the AF-CDN are presented, respectively. In Figure 10b,d the visualization and error matrix of 1D-CNN algorithm are presented, respectively.

Figure 10. Visualization of classification results and error matrix: (**a**) visualization of AF-CDN; (**b**) error matrix of AF-CDN; (**c**) 1D-CNN visualization; (**d**) 1D-CNN error matrix.

At the same time, it should be pointed out that, compared with the method of feature extraction based on raw data for diagnosis, the AF-CDN in this paper demonstrates great improvement in reducing the probability of two kinds of misdiagnosis. This is also intuitive in the visualization and error matrix. The main reasons for this result can be summarized as follows: (1) compared with the single FFT signal or the original 2D signal, and the AF-CDN fuses the two signals. (2) Compared with SGD, the gradient fusion algorithm based on the Kalman filter has a stronger parameter integration ability in feature fusion of gradient information. In this way, the diagnostic ability of the network can be better improved. As can be seen from the visual figures of the two algorithms, there are more cases of misdiagnosis between H and SH and between S and F. At the same time, it should be pointed out that compared with the method of feature extraction based on original data for diagnosis, the AF-CDN provides a great improvement in reducing the probability of misdiagnosis.

4.2. Case Two
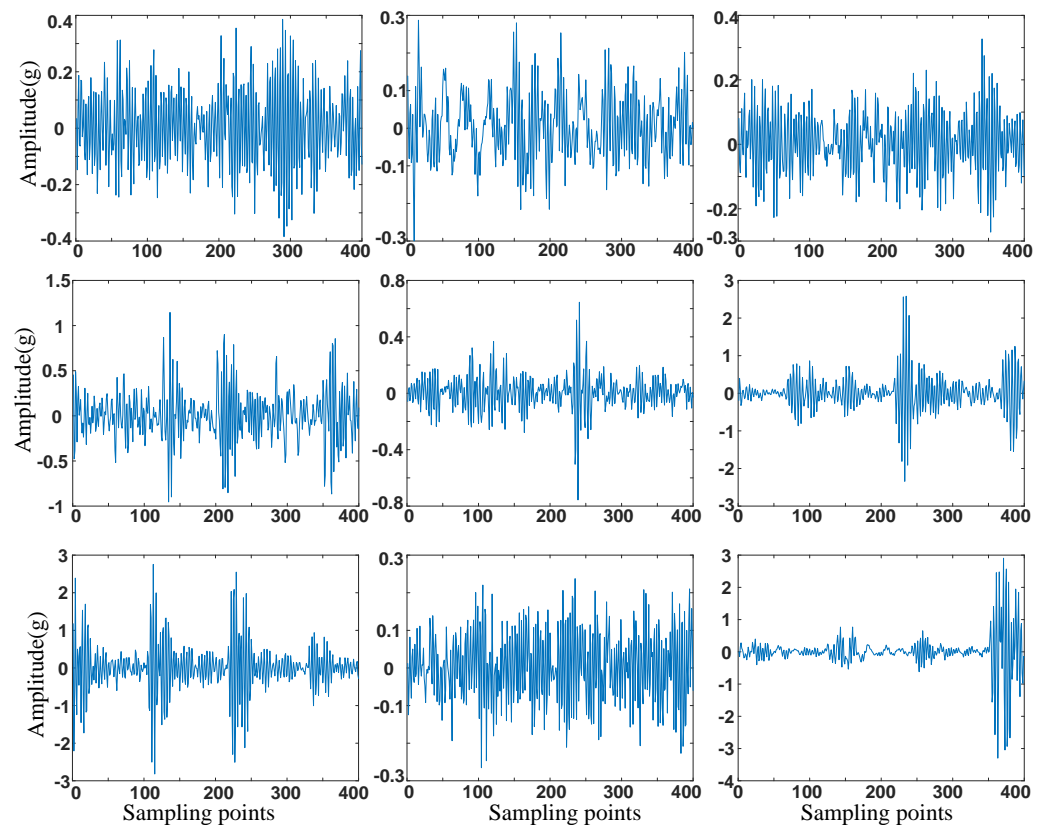
4.2.1. Dataset Preparation and Parameter Settings

The AF-CDN is also verified on the CWRU Bearing Dataset [33]. CWRU set up a variety of fault types on motor drive equipment. The vibration signals of the different positions of the drive end were collected by vibration sensors under different load conditions.

According to the different fault locations, the fault can be divided into three categories: the ball fault (BF), inner ring fault (IF), and outer ring fault (OF). Each type of fault can be further subdivided into three different fault levels according to different fault severity. Therefore, a total of nine types of faults were set up in this experiment, and the sampling method was carried out according to the paper [6,35]. For details about fault information, see Table 3. The size of each sample is 400, so the dimensions after the 2D conversion are 20×20. In order to ensure the speed of iteration, the batch size is set to between 20 and 40.

The continuous time signal collected by the vibration sensor is shown in Figure 11.

Table 3. Status information of 9 types of fault.

Fault Type	Fault Diameter(inch)	Label	Sample Size
BFI	0.007	000000001	300
BFII	0.014	000000010	300
BFIII	0.021	000000100	300
IFI	0.007	000001000	300
IFII	0.014	000010000	300
IFIII	0.021	000100000	300
OFI	0.007	001000000	300
OFII	0.014	010000000	300
OFIII	0.021	100000000	300

Figure 11. Continuous time signal diagram of 9 kinds of faults.

4.2.2. Experiment and Analysis

The AF-CDN is compared with some popular methods at present. As summarized in the previous chapter, the existing methods are mainly based on two signal extraction methods. One involves performing frequency domain transformation in the original time-continuous signal to obtain frequency domain features. The other mainstream approach is based on raw signals.

In this case, the existing methods and AF-CDN are compared in detail. The processing methods in the frequency domain mainly include wavelet transform, wavelet packet transform, statistical locally linear embedding, and other methods. Raw signal processing involves converting raw signals into 2D images and establishing Spectrogram methods.

Table 4 shows the comparison of experimental results between the existing mainstream methods and AF-CDN.

Table 4. Comparison of diagnostic accuracy of different methods.

Means of Classification	Features	Accuracy Rate
KNN [9]	HOCs and WT	91.2%
SVM [8]	WP	62.5–98.7%
ANN [10]	DWT (Morlet)	96.7%
ANN [10]	DWT (Daubechies 10)	93.3%
SVM [11]	Statistical Locally Linear Embedding	77.8–94.1%
2D-CNN [7]	Raw data	98.35%
2D-CNN [12]	Spectrogram	98.1–99.5%
AF-CDN	FFT + Raw data	99.44–99.78%

Compared with the traditional diagnosis methods, the proposed algorithm effectively integrates the frequency domain characteristics of vibration signals and their original signal characteristics. The accuracy of network diagnosis is further improved.

The main feature extraction methods used in the signal-based feature analysis approach include WP, WT, DWT, and statistical methods based on these. After obtaining these frequency features, KNN, SVM, ANN, etc., can be used for analysis. As shown in Table 4, such diagnosis methods based on "frequency domain signals + neural networks" have a diagnostic accuracy of 62.5–98.7%. The accuracy of the analysis method based on the raw signal is 98.1% to 99.5%. The diagnostic accuracy of the proposed method is 99.44–99.78%. The average execution time of AF-CDN is about 0.115 s. Compared with the traditional diagnosis methods, AF-CDN effectively integrates the frequency domain characteristics of vibration signals and their raw signal characteristics. The feature extraction capability of the network is excellent compared to the rest of the network structure. As analyzed in Case One, AF-CDN combines information on the frequency domain characteristics of the signal with the raw time domain information. Thus AF-CDN is able to have an excellent diagnostic result. The accuracy of network diagnosis is further improved.

5. Conclusions

This paper presents an adaptive fusion convolutional denoising network for the health monitoring of port crane speed gearboxes. At the same time, a Kalman filter is used to update the network parameters during the training process. Compared with traditional diagnosis methods, the main advantages of the method proposed in this paper are that the accuracy of diagnosis is improved greatly. The robustness of the diagnostic network can be significantly improved. Monitoring the status of port cranes' gearbox systems provides support for equipment health care. When the equipment is in the running-in state and sub-health state, workers can perform maintenance on bearing equipment in a timely manner so that the service time of the equipment can be effectively delayed. At the same time, once the equipment is in a sub-health state, it is necessary to monitor the system state at all times; once the system indicators reach the fault state, it is necessary to immediately shut down and perform the maintenance. This ensures the safety and reliability of the entire port machine. The results show that AF-CDN also has excellent diagnostic performance on public data sets. However, the shortcomings of this paper are that we have not integrated the other sensors well, including temperature sensors and vibration signals from other locations that the hardware system contains.

The system we have built in this paper is a four-stage bearing diagnosis system based on a single vibration sensor. In future work, we will further study the fusion of stress, temperature, and multi-directional vibration signals collected from multiple locations of the whole system so as to extract the accurate overall health status of the crane and fuse it with multi-sensor signals. This makes it possible to establish a whole health management system and enables effective real-time monitoring of the full lifecycle of a port crane. Secondly,

a whole life cycle inspection system for the equipment should be established based on the multi-sensor information fusion technology described above. Meanwhile, distributed learning techniques will be focused on in order to fuse the data from the multi-location port machines. Meanwhile, distributed learning techniques will be focused on in order to fuse the data from the multi-location port machines.

Author Contributions: Conceptualization, R.Z.; Data curation, X.H.; Formal analysis, R.Z.; Funding acquisition, X.H.; Methodology, R.Z.; Supervision, X.H.; Writing—original draft, R.Z.; Writing—review and editing, R.Z. and X.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: No applicable.

Informed Consent Statement: No applicable.

Data Availability Statement: The original data contributions presented in the study are included in the article; further inquiries can be directed to the corresponding authors.

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

References

1. Xu, L.; Chatterton, S.; Pennacchi, P. A Novel Method of Frequency Band Selection for Squared Envelope Analysis for Fault Diagnosing of Rolling Element Bearings in a Locomotive Powertrain. *Sensors* **2018**, *18*, 4344. [CrossRef] [PubMed]
2. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [CrossRef]
3. Li, H.; Liu, J.; Wu, K.; Yang, Z.; Liu, R.W.; Xiong, N. Spatio-Temporal Vessel Trajectory Clustering Based on Data Mapping and Density. *IEEE Access* **2018**, *6*, 58939–58954. [CrossRef]
4. Gao, K.; Han, F.; Dong, P.; Xiong, N.; Du, R. Connected Vehicle as a Mobile Sensor for Real Time Queue Length at Signalized Intersections. *Sensors* **2019**, *19*, 2059. [CrossRef] [PubMed]
5. Wu, M.; Tan, L.; Xiong, N. A Structure Fidelity Approach for Big Data Collection in Wireless Sensor Networks. *Sensors* **2014**, *15*, 248–273. [CrossRef]
6. Ma, X.; Wen, C.; Wen, T. An Asynchronous and Real-Time Update Paradigm of Federated Learning for Fault Diagnosis. *IEEE Trans. Ind. Inform.* **2021**, *17*, 8531–8540. [CrossRef]
7. Xia, M.; Li, T.; Xu, L.; Liu, L.; De Silva, C.W. Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks. *IEEE/ASME Trans. Mechatron.* **2017**, *23*, 101–110. [CrossRef]
8. Hu, Q.; He, Z.; Zhang, Z.; Zi, Y. Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble. *Mech. Syst. Signal Process.* **2007**, *21*, 688–705. [CrossRef]
9. Yaqub, M.F.; Gondal, I.; Kamruzzaman, J. Inchoate Fault Detection Framework: Adaptive Selection of Wavelet Nodes and Cumulant Orders. *IEEE Trans. Instrum. Meas.* **2011**, *61*, 685–695. [CrossRef]
10. Konar, P.; Chattopadhyay, P. Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs). *Appl. Soft Comput.* **2011**, *11*, 4203–4211. [CrossRef]
11. Wang, X.; Zheng, Y.; Zhao, Z.; Wang, J. Bearing Fault Diagnosis Based on Statistical Locally Linear Embedding. *Sensors* **2015**, *15*, 16225–16247. [CrossRef] [PubMed]
12. Verstraete, D.; Ferrada, A.; Droguett, E.L.; Meruane, V.; Modarres, M. Deep Learning Enabled Fault Diagnosis Using Time-Frequency Image Analysis of Rolling Element Bearings. *Shock Vib.* **2017**, *2017*, 5067651. [CrossRef]
13. Torrence, C.; Compo, G.P. A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 61–78. [CrossRef]
14. Chen, H.; Chen, Z.; Chai, Z.; Jiang, B.; Huang, B. A Single-Side Neural Network-Aided Canonical Correlation Analysis With Applications to Fault Diagnosis. *IEEE Trans. Cybern.* **2021**, 1–13. [CrossRef]
15. Chen, H.; Li, L.; Shang, C.; Huang, B. Fault Detection for Nonlinear Dynamic Systems With Consideration of Modeling Errors: A Data-Driven Approach. *IEEE Trans. Cybern.* **2022**, 1–11. [CrossRef]
16. Haykin, S. Neural networks expand SP's horizons. *IEEE Signal Process. Mag.* **1996**, *13*, 24–49. [CrossRef]
17. Guyon, I.; Weston, J.; Barnhill, S.; Vapnik, V. Gene Selection for Cancer Classification using Support Vector Machines. *Mach. Learn.* **2002**, *46*, 389–422. [CrossRef]
18. Huang, S.; Liu, A.; Wang, T.; Xiong, N.N. BD-VTE: A Novel Baseline Data Based Verifiable Trust Evaluation Scheme for Smart Network Systems. *IEEE Trans. Netw. Sci. Eng.* **2020**, *8*, 2087–2105. [CrossRef]
19. Yang, P.; Xiong, N.N.; Ren, J. Data Security and Privacy Protection for Cloud Storage: A Survey. *IEEE Access* **2020**, *8*, 131723–131740. [CrossRef]

20. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
21. Eren, L.; Ince, T.; Kiranyaz, S. A Generic Intelligent Bearing Fault Diagnosis System Using Compact Adaptive 1D CNN Classifier. *J. Signal Process. Syst.* **2018**, *91*, 179–189. [CrossRef]
22. Jiao, J.; Zhao, M.; Lin, J.; Zhao, J. A multivariate encoder information based convolutional neural network for intelligent fault diagnosis of planetary gearboxes. *Knowl. Based Syst.* **2018**, *160*, 237–250. [CrossRef]
23. Chen, Y.; Peng, G.; Xie, C.; Zhang, W.; Li, C.; Liu, S. ACDIN: Bridging the gap between artificial and real bearing damages for bearing fault diagnosis. *Neurocomputing* **2018**, *294*, 61–71. [CrossRef]
24. Jia, F.; Lei, Y.; Lu, N.; Xing, S. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization. *Mech. Syst. Signal Process.* **2018**, *110*, 349–367. [CrossRef]
25. Jiang, G.; He, H.; Yan, J.; Xie, P. Multiscale Convolutional Neural Networks for Fault Diagnosis of Wind Turbine Gearbox. *IEEE Trans. Ind. Electron.* **2018**, *66*, 3196–3207. [CrossRef]
26. Appana, D.K.; Prosvirin, A.; Kim, J.-M. Reliable fault diagnosis of bearings with varying rotational speeds using envelope spectrum and convolution neural networks. *Soft Comput.* **2018**, *22*, 6719–6729. [CrossRef]
27. Guo, D.; Zhong, M.; Ji, H.; Liu, Y.; Yang, R. A hybrid feature model and deep learning based fault diagnosis for unmanned aerial vehicle sensors. *Neurocomputing* **2018**, *319*, 155–163. [CrossRef]
28. 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]
29. Han, T.; Liu, C.; Wu, L.; Sarkar, S.; Jiang, D. An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems. *Mech. Syst. Signal Process.* **2018**, *117*, 170–187. [CrossRef]
30. Xin, Y.; Li, S.; Cheng, C.; Wang, J. An intelligent fault diagnosis method of rotating machinery based on deep neural networks and time-frequency analysis. *J. Vibroengineering* **2018**, *20*, 2321–2335. [CrossRef]
31. Sun, W.; Yao, B.; Zeng, N.; Chen, B.; He, Y.; Cao, X.; He, W. An Intelligent Gear Fault Diagnosis Methodology Using a Complex Wavelet Enhanced Convolutional Neural Network. *Materials* **2017**, *10*, 790. [CrossRef] [PubMed]
32. Cao, X.-C.; Chen, B.-Q.; Yao, B.; He, W.-P. Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification. *Comput. Ind.* **2019**, *106*, 71–84. [CrossRef]
33. Case Western Reserve University Bearing Data Center Website. Available online: http://csegroups.case.edu/bearingdatacenter/home (accessed on 15 March 2017).
34. Arulampalam, M.S.; Maskell, S.; Gordon, N.; Clapp, T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **2002**, *50*, 174–188. [CrossRef]
35. Chen, H.; Chai, Z.; Dogru, O.; Jiang, B.; Huang, B. Data-Driven Designs of Fault Detection Systems via Neural Network-Aided Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, 1–12. [CrossRef] [PubMed]