

Research Article

A Multirate Sensor Information Fusion Strategy for Multitask Fault Diagnosis Based on Convolutional Neural Network

Xiangkai Ma ¹, Pei Wang ², Bozhou Zhang ³ and Ming Sun ²

¹School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

²School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

³Chengdu Science and Technology Development Research Center, Chengdu 610041, China

Correspondence should be addressed to Ming Sun; sunm@uestc.edu.cn

Received 22 March 2021; Revised 28 April 2021; Accepted 8 May 2021; Published 1 June 2021

Academic Editor: Esteban Jove

Copyright © 2021 Xiangkai Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In complicated mechanical systems, fault diagnosis, especially regarding feature extraction from multiple sensors, remains a challenge. Most existing methods for feature extraction tend to assume that all sensors have uniform sampling rates. However, complex mechanical systems use multirate sensors. These methods use upsampling for data preprocessing to ensure that all signals at the same scale can cause certain time-frequency features to vanish. To address these issues, this paper proposes a Multirate Sensor Information Fusion Strategy (MRSIFS) for multitask fault diagnosis. The proposed method is based on multidimensional convolution blocks incorporating multisource information fusion into the convolutional neural network (CNN) architecture. Features with different sampling rates from the raw signals are run through a multichannel parallel fault feature extraction framework for fault diagnosis. Additionally, time-frequency analysis technology is used to reveal fault information in the association between time and frequency domains. The simulation platform's experimental results show that the proposed multitask model achieves higher diagnosis accuracy than the existing methods. Furthermore, manual feature selection for each task becomes unnecessary in MRSIFS, which has the potential toward a general-purpose framework.

1. Introduction

In many complicated systems, researchers tend to take the multisource data measured in the manufacturing process as the features of deep learning (DL) algorithms [1–3]. However, existing studies guarantee that all sensors operate at the same rate [4], which is often unrealistic in multisampling rate signal fusion systems. Therefore, the issue of multirate sensor information fusion is of great significance in the actual industrial environment and has received extensive attention, especially in the recent ten years [5, 6].

In recent studies, a model based on convolution takes advantage of its excellent ability to extract features from multisource signals, which has achieved excellent performance in multitask fault detection [7]. CNN can effectively extract fault features from the raw signal with its weight-sharing strategy, spatial pooling layer, local connection mechanism [8, 9], and ability in handling periodic signals. It had been proved that CNN is suitable to learn potential features

hidden in rotating mechanical signals because of its ability in handling periodic signals [10]. With the 1D-to-2D conversion of vibration signals or 1D convolutional structure, the 2D CNN models have been successfully applied in fault diagnosis directly using raw signals. In recent years, some CNN-based deep learning models have been developed for mechanical fault diagnosis.

For specific fault detection problems, some researchers have made different improvements based on the original convolutional layer, as follows: Jia et al. [11] proposed a framework called deep normalized convolutional neural network (DNCNN) for imbalanced fault classification of machinery to overcome the weakness of imbalanced distribution of machinery health conditions. Apart from extracting potential features hidden in signals, CNNs can detect local features in a deep network. Recently, Peng et al. [12] proposed a novel deeper 1D convolutional neural network (Der-1DCNN), which includes the idea of residual learning and can effectively learn high-level and abstract features

while effectively alleviating the problem of training difficulty and performance degradation of a deeper network. All these works prove that CNN is capable of mechanical data analysis. However, these works require that all signals for training and testing CNN must be acquired at the same sensor sampling frequency, which limits its further application.

Besides, a key disadvantage of 1D CNN is to detect the local correlation of signals that is deficient. In the design of the 2D CNN applied to the fault diagnosis algorithm of a multisensor fusion system, some signal preprocessing transformation procedures [13] and time-frequency analysis technology are needed. 2D CNN based on signal processing techniques (fast Fourier transform, short-time Fourier transformation, wavelet transform, etc.) has many successful applications in the field of mechanical fault diagnosis. In some works, the raw sensor time series data has been preprocessed by some methods such as frequency transformation and time-frequency transformation before being input to the 2D CNN [14, 15].

Although researchers have tried to combine multisensor signal fusion and deep learning, different sampling rates of the mechanical system's multiple components are not considered in the existing articles [16, 17]. The multirate sensor problem has become an urgent problem to be solved. However, in the previous paper [18], only the raw signals from the low-rate sensor are transferred through the upsampling network. As mentioned earlier, the existing deep learning model cannot solve the multirate sensor problem well in a complicated mechanical system with multiple components.

This paper is aimed at developing an end-to-end Multirate Sensor Information Fusion Strategy (MRSIFS), which is dedicated to improving the feature fusion of multirate sampled signals. This strategy can automatically extract sensitive fault features from multirate sensor signals for fault detection and diagnosis. Specifically, the strategy consists of three sequential phases: multirate sensor feature extraction stage, feature fusion stage, and regression stage. A key advantage of the proposed strategy is that the fault features of different scales can be automatically learned from the vibration signals of different components through parallel multichannels in complex mechanical systems. As we know, the different dimensions of the filters have different frequency resolutions. The sensitive frequency of the signal may exist in different frequency bands. The proposed multidimensional parallel convolution kernel can be used as filters with different frequency resolutions for recognition, thus effectively enhancing the frequency domain classification information of raw signals. Thus, it is effective to add a multidimensional convolution block (MDCB) in the multirate sensor feature extraction layer to extract the different scale fault classification information from the raw signals. Since the sensitive frequency bands of the multidimensional convoluted signal are included in the frequency component of the sequence signal, to combine the fault features of different scale signals, the feature fusion layer is used for cascading processing of the fault features. Besides, 1D CNN and DNN implement low sampling rate signal feature extraction and dimensionality reduction, while 2D CNN based on short-time Fourier transform (STFT) extracts higher sampling rate features [19]. The

strategy combines the advantages of the three network structures, and the three networks supplement the feature information neglected by the other side.

The proposed strategy is tested on the hydraulic system condition monitoring dataset, which is available from the UC Irvine Machine Learning Repository [20]. In the experimental part, MRSIFS are compared with the existing fault diagnosis methods in classification accuracy. The comparison results show that MRSIFS can extract fault features from multisampling sensor signals. The main contributions of our work are listed as follows:

- (1) For the complicated mechanical system with multirate sampling, the designed multirate sensor feature extraction layer can extract fault features from the multirate sensor and fuse the fault feature automatically
- (2) To improve the capability of the fault detection model to learn fault feature information from different frequency band signals, multidimensional convolutional blocks are used to learn rich and complementary fault information from multirate sensor signals in parallel. MRSIFS is an end-to-end deep learning model, which takes the original signal as input directly without time-consuming feature selection. Therefore, it has the potential to be extended to other industrial systems
- (3) For the complicated mechanical system, the contribution of the present work is the implementation of a multitask classification framework, whereas the existing studies on the dataset of tag classification are all to propose and train a model for each task

The rest of this article is organized as follows. Section 2 discusses the related theories and the proposed framework. The experimental results and discussion are presented in Section 3. Finally, Section 4 is the conclusion.

2. Related Works

2.1. End-to-End Multisensor Model. Despite existing methods based on CNN have achieved the breakthrough performance in detection and diagnosis [21–23], there are still some shortcomings that can be improved: multirate information feature fusion can extract fault sensitive and complementary features, which are not contained in a single sensor signal, thus achieving higher accuracy and stability of the complex mechanical system [24, 25]. Nevertheless, most existing methods only used a single sensor, and few researchers attempted the multirate sensor information fusion for diagnosis. Furthermore, in previous studies about multisensor feature fusion, it is often assumed that all sensors have uniform sampling rates. In [16], the original AE signals of four independent sensor groups are first preprocessed by time-frequency analysis technology. Then, the feature matrices are converted into grey images. Finally, grey images are subsequently fed to the fine-tuned transfer learning (FTL) for fault diagnosis of different components and prediction of bearing degradation degree. In [17], the multisequence signal data collected by multiple sensors are converted into

multichannel feature matrices, and a parallel convolutional neural network (PCNN) is designed to fuse the fault information extracted from the transformed feature matrix.

Besides, the signals of various sensors are difficult to be measured at the same sampling rate, since the complicated mechanical system and design principles between different sensors [26–28]. Thus, the issue about multirate feature extraction inevitably occurs in the process of sensor fusion for complicated mechanical systems using multirate sensors [29]. However, most existing work in the area of CNN for multisensor feature fusion usage upsampling may make the vanish of certain time-frequency features, unless all sample rates are sufficiently close, and thus degrades the accuracy of fault diagnosis. In the multisensor system with different sampling rates, Li et al. [18] present an improved information fusion framework based on the atrous convolution. Specifically, to avoid tedious preprocessing, the model extracts fault features from multisource signals by constructing a convolution kernel of adaptive size matching the data source channels. Finally, the proposed method is compared with the existing research, as shown in Table 1.

2.2. Existing Works Based on the Same Dataset. There have been some researches about the hydraulic system condition monitoring dataset based on time-frequency analysis techniques and artificial feature extraction. Helwig et al. [31, 32] convert the time domain data into frequency domain using fast Fourier transform and generate statistical features. They then calculated features for fault label correlation and selected the n features by ranking or sorting the correlation (CS) for the fault classification. Prakash and Kankar [33] also utilized statistical features of frequency domain data and applied XGBoost [34] to define feature importance (XFI) and select half of the highest correlations along with a deep neural network for the classification model.

These previous approaches include several drawbacks. First, the proposed handcrafted feature selection methods, such as CS and XFI, may suffer from utilizing redundant features, which will disrupt the learning of the model. Second, statistical feature extraction, PCA, and other feature engineering methods are not suitable for real-time detection. These techniques must be applied to each new incoming sample, thereby consuming more time and computation power. Third, to ensure the quality of the feature extraction, a manual design and suitable features are needed based on the characteristics of the different types of faults. Furthermore, feature extraction usually turns out to be a computationally costly operation, but the existing studies on the dataset of tag classification based on UCI are all to propose and train a model for each tag. The quality of the features directly determines the system performance. Therefore, the system feature extraction is not automatic.

With the aforementioned open issues, we propose the end-to-end deep learning model and directly take the raw vibration signals as input. Finally, the multilabel classification problem is transformed into the regression prediction problem, and the accuracy rate is higher than the existing methods [31–34]. Unlike traditional methods relying on manually defined or extracted features, the proposed design

does not require any additional expert knowledge, which has great potential toward a general-purpose framework for intelligent fault diagnosis. Thus, it can be easily extended to deal with fault diagnosis problems of other industrial systems.

3. Methods

Although the most satisfactory level of anomaly diagnosis accuracy was reported [20, 32–34], most of these previous approaches had to design and train different feature extraction models for each task. In contrast, the multitask model is suitable for the complex mechanical system. In addition, manual feature selection often results in wasteful computational costs, but the existing studies on the dataset of tag classification based on UCI are all to propose and train a model for each tag. However, the stage of manual feature selection is difficult and time-consuming [35]. This process incurs considerable computational costs, which eventually may impede the use of existing approaches in real-time fault diagnosis applications. The MRSIFS model proposed in this paper transforms the multirate problem into a unified multidimensional convolutional neural network feature fusion strategy. Compared with the traditional methods about condition monitoring of hydraulic system dataset, MRSIFS has better robustness in noise environment and does not require predetermined feature selection. In particular, it proposes a feature extraction layer based on a multidimensional convolutional neural network and a concatenated feature fusion layer. Finally, the multitask classification problem is transformed into the regression prediction problem, and the accuracy rate is higher than the existing methods [20, 32–34].

The framework of the proposed MRSIFS is shown in Figure 1. The inputs to the model consist of 3 segments of raw temporal vibration signal at different sampling rates. The output Y is a vector made up of the labels of five tasks, where each value represents the predicted regression value for a task. The fault detection is defined to predict the vector Y based on the fault feature extraction from the raw temporal vibration signal X using the MRSIFS model. The MRSIFS model has three parts: the multidimensional convolution feature learning layer, the multirate feature concatenate layer, and the regression layer.

3.1. The Framework of MRSIFS. In the multirate sensor feature extraction stage, the fault feature extracted by the input layer is fed into the multidimensional convolution feature learning block. The kernels of the convolution layer have different dimensions. As we know, convolution kernels of different dimensions act as filters of different resolution scales to extract fault features in the raw signals and simultaneously extract the features of input signals in different frequency bands. As shown in Figure 2, features of the convolution kernel dimension are combined through the concatenation layer to form a multirate feature map. As a fault information collector, the concatenation layer can aggregate features of different scales to form a multirate feature set. It can be observed that the data from multiple sensors are transformed into multiple channels through the input layer, and then, the fault features are obtained through the feature

TABLE 1: Comparison with related works.

Framework	MSFTFI [16]	PCNN [17]	FAC-CNN [18]	MRSIFS
Data preprocessing	STFT	—	—	STFT
CNN dimension	2D	2D	1D	1D, 2D
Feature extraction	FTL, LeNet-5	LSTM, DRN	Atrous CNN	FPN [30]
Activation function	ReLU	ReLU	PReLU	PReLU
Sampling frequency	Fixed	Fixed	Different	Different

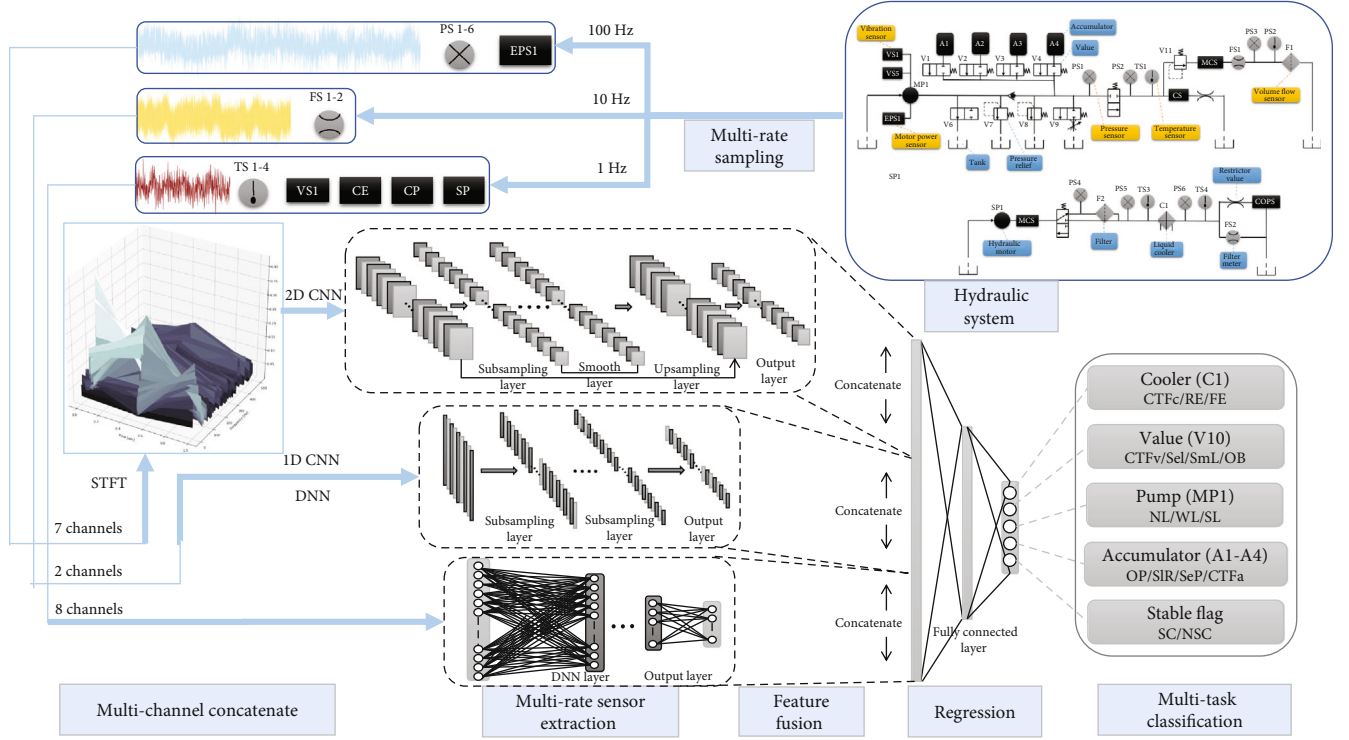


FIGURE 1: Flowchart of the proposed approach.

extraction layer and feature fusion layer sequentially. Specifically, the input layer of the proposed method is a three-dimensional matrix, and the prediction results are taken as the output of the multitask model. The specific size and other parameters of the filter are shown in Tables 2–4.

3.2. CNN Details and Regression. The 2D CNN is composed of six parts, including STFT, input layer, downsampling layer, smooth layer, and upsampling layer. The downsampling layer is composed of K prevention blocks, and each of them contains three convolutional layers. The upsampling layer is implemented based on bilinear. And the smooth layer contains a single convolutional layer. To extract the hidden fault information in the signal as far as possible, four downsampling layers, four smooth layers, and three upsampling layers were successively used in the experimental model with the dataset. The features are fed into the output layer and straighten the output of the last layer as the fault feature extracted by the 2D convolution component. The 2D CNN model summary is mentioned in Table 2. Table 2 describes the type of layers, in channels, out channels, kernel size, etc.

It must be noted that in the 2D convolution component, we use a deep convolutional network structure. In this way, although deep-level fault features can be extracted, according to some existing studies [30], the deep-level convolutional network may lose some shallow-level fault features. In this paper, the downsampling layer, smooth layer, and upsampling layer are used to cooperate, so that the features extracted from the shallow convolutional network can be integrated into the subsequent deep-level feature information. The proposed network structure is inspired by [36].

Compared with the two-dimensional convolution structure, the one-dimensional convolution structure only retains the part of the lower sampling layer, and some modifications are made in the convolutional network parameter settings. The DNN network block is composed of four full connection layers. Similarly, the regression layer consists of three full connection layers, with ReLU as an activation function. Parameter details and design implementation of DNN and regression layer are shown in Tables 3 and 4, respectively. After each convolutional layer in the input layer and downsampling layer, and the fully connected layer in DNN, batch

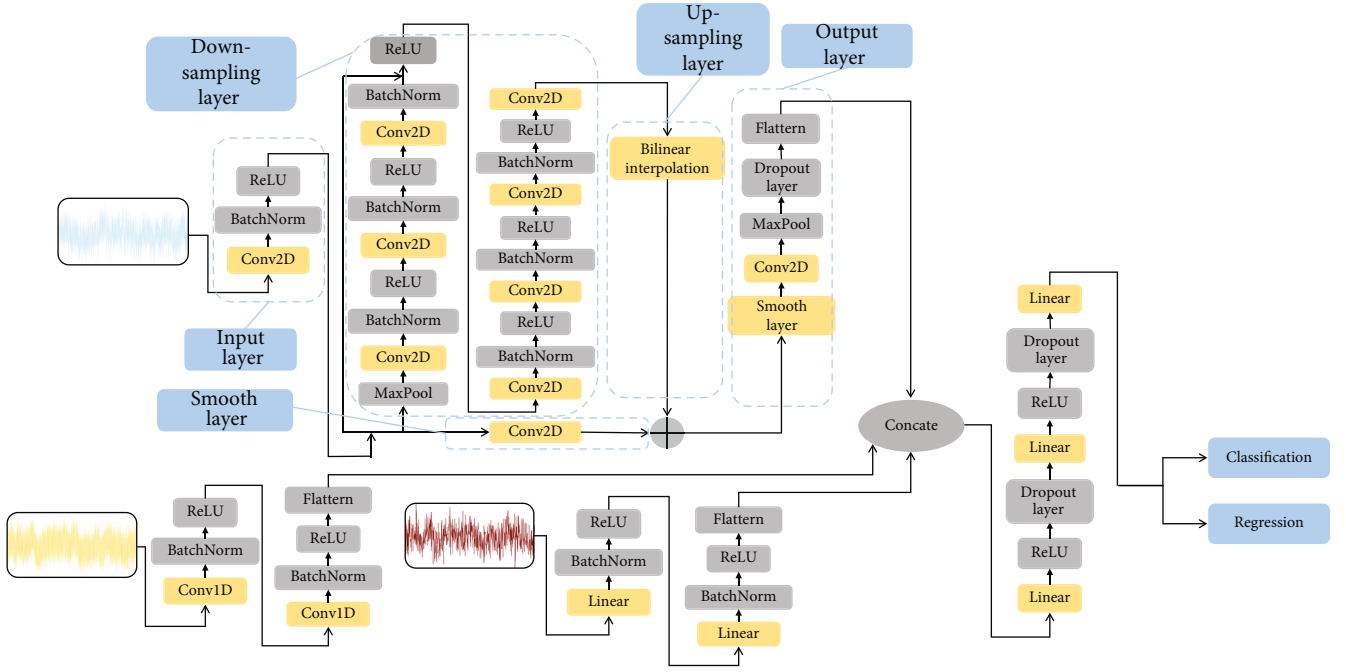


FIGURE 2: Detailed structure of the proposed method.

TABLE 2: Summary of the submodel of MRSIFS using 2D CNN for feature extraction.

Network structure	Layers	In_channels	Out_channels	Kernel_size	Stride	Padding
Input layer	Conv2D	7	16	7	2	3
	MaxPool			3	2	1
	Conv2D	16	16	1	1	0
	Conv2D	16	16	3	1	1
Downsampling layer	Conv2D	16	64	1	1	0
	Conv2D	64	16	1	1	0
	Conv2D	16	16	3	1	1
	Conv2D	16	64	1	1	0
Smooth layer	Conv2D	64	64	1	1	0
Upsampling layer	"Bilinear"					
	Conv2D	64	32	1	1	0
Output layer	MaxPool			4	4	0

Dropout ($p = 0.2$)

TABLE 3: Summary of the submodel of MRSIFS using DNN for feature extraction.

Network structure	Layers	In_features	Out_features	Bias
Fully connected layer	Linear	480	400	True
	Linear	400	256	True
	Linear	256	128	True
	Linear	128	64	True

TABLE 4: Summary of the regression layer of MRSIFS.

Network structure	Layers	In_features	Out_features	Bias
Regression	Linear	1280	512	True
	Dropout ($p = 0.2$)			
	Linear	512	128	True
	Dropout ($p = 0.2$)			
	Linear	128	5	True

normalization [37] is used to accelerate the training process of MRSIFS.

3.3. Training Details. In MRSIFS training, we adopt the mean-square error (MSE) between the predictive label and

the real label as the loss function. To reduce the computational overhead in the training, the Adam algorithm [38] is employed to optimize the parameters of the model. A critical task for MRSIFS training is the adjustment of hyperparameters, and this paper takes the batch size and learning rate

as hyperparameters. The batch size defines the number of samples to be processed in one batch. The learning rate determines the convergence speed of weight and bias in the neural network during training. The learning rate of model training is set to 0.001, and the batch size is set to 32. To prevent overfitting, the proposed method uses dropout technology [39] between the full connection layers in the regression prediction stage. With the dropout technique, every parameter in the full connection layer has a certain probability of being randomly removed during each training epoch. In the initialization stage of the model, all parameters of the neural network are initialized by a zero-mean standard uniform distribution. Particularly, in this model, the biases of all neurons are set to zero when initialized. In addition, all the original vibration signals are randomly used for training the model and testing the model. Specifically, 20% of the signals are selected as test samples, and the remaining 80% are used for testing models. To obtain a relatively stable experimental result, 10 trials of MRSIFS were repeated on the condition that each group of models had the same parameters.

Since its different scales and different depths of feature extraction layers (i.e., pairs of downsampling layers, smooth layer, and upsampling layers), MRSIFS architecture has the advantages of general purpose and flexibility. Furthermore, MRSIFS can effectively learn sensitive diagnostic information by using multidimensional feature extraction structure and capture complementary and useful fault features at different scales for fault diagnosis and detection. As we know, more robust and abstract fault information is expected to improve the diagnosis performance. Therefore, MRSIFS with more scales and layers can extract useful diagnostic features to adapt to the complicated mechanical system. In addition, in practical applications, the simple rate of the input samples and the depth of the layers have a restriction relationship. In MRSIFS, researchers can select the appropriate kernel size and layer depth based on the length of the input signal. More details can be found in III.

The proposed MRSIFS algorithm is conducted on torch 1.6.0+cu101. The hardware configuration for training and testing is Intel(R) MKL-DNN v1.5.0 + TITAN X (Pascal), while the software environment is Linux + Python 3.7 Version 3.7.9 [GCC 7.3.0].

4. Experiments and Discussion

4.1. Data Description. The proposed strategy is tested on the hydraulic system condition monitoring dataset, which is available from the UC Irvine Machine Learning Repository [10]. The system cyclically repeats constant load cycles (duration 60 seconds) and measures process values such as pressure, volume, flow, and temperature, while the conditions of the four hydraulic components (cooler, valve, pump, and accumulator) and the stable flag of the hydraulic system are quantitatively varied. As shown in Table 5, the dataset consists of the measurement signals of 17 sensors, including 14 physical sensor components and 3 virtual sensors. The measurement period of each sensor is 60 seconds, and the sampling rate range is between 1 Hz and 100 Hz. The dataset consists of 2205 samples with a sampling period of 60

TABLE 5: Hydraulic system sensors. *CE, CP, and SE are virtual sensors.

Sensor	Physical quantity description	Unit	Sampling rate	Sensor signal length
PS1-6	Pressure	Bar	100 Hz	6000
EPS1	Motor power	W	100 Hz	6000
FS1-2	Volume flow	L/min	10 Hz	600
TS1-4	Temperature	°C	1 Hz	60
VS1	Vibration	mm/s	1 Hz	60
CE (virtual)	Cooling efficiency	%	1 Hz	60
CP (virtual)	Cooling power	kW	1 Hz	60
SE (virtual)	Efficiency factor	%	1 Hz	60

seconds. Each sample contains a component status label that reflects the fault condition of the 5 components.

As shown in Table 6, the training results of each sample include five types of tags. Therefore, what this paper deals with is a multitask classification problem. Nevertheless, because the value of each tag has a specific physical meaning, therefore, the multitask classification problem can also be transformed into a regression problem to solve. The problem proposed by the dataset is actually using the regression vector with the predicted length of 5, and to set the threshold, convert the regression value to the final classification result. Ideally, the output of the network for any input should be limited to our given range (for example, the labelled hydraulic accumulator is estimated to be only 130, 115, 100, and 90), but actually, the output value is often in between these values. In this paper, the value with the minimum absolute value error is selected as the predicted class label.

4.2. Compared Models. The purpose of the experiment is to verify the higher classification accuracy of MRSIFS proposed in this paper compared with the existing MSFTFI, PCNN, and FAC-CNN. Besides, when changing the convolutional layer depth of the MRSIFS feature extraction block, MRSIFS still has good performance, which indicates that MRSIFS is not sensitive to the setting of super parameters, indicating that the strategy proposed in this paper has wide adaptability. The main comparisons of the proposed model and existing methods are listed as follows:

- (1) MSFTFI: in [16], the original AE signals of four independent sensor groups are first preprocessed by time-frequency analysis technology. Then, the feature matrices are converted into grey images. Finally, grey images are subsequently fed to the fine-tuned transfer learning (FTL) for fault diagnosis of different components and prediction of bearing degradation degree
- (2) PCNN: firstly, the multisequence signal data collected by multiple sensors are converted into multichannel feature matrices, and a parallel CNN is designed to concatenate the fault information extracted from the

TABLE 6: Summary of components and their simulated fault conditions of the hydraulic dataset. A load cycle with 60 s duration is repeated 2205 times with the distribution of instances as indicated by the examples.

Component	Fault conditions	States	Abbreviation	Examples
Cooler (C1)	Cooling power decrease	3%: close to total failure	CTFc	732
		20%: reduced efficiency	RE	732
		100%: full efficiency	FE	741
		73%: close to total failure	CTFv	360
Valve (V10)	Switching degradation	80%: severe lag	Sel	360
		90%: small lag	SmL	360
		100%: optimal behaviour	OB	1125
Pump (MP1)	Internal leakage	0: no leakage	NL	1221
		1: weak leakage	WL	492
		2: severe leakage	SL	492
		130 bar: optimal pressure	OP	599
Accumulator (A1-A4)	Gas leakage	115 bar: slightly reduced	SlR	399
		100 bar: severely reduced	SeP	399
		90 bar: close to total failure	CTFa	80 8
Stable flag (S1)	---	0: conditions were stable	SC	1449
		1: nonstatic conditions	NSC	756

transformed feature matrix. More details of PCNN can be found in [17]

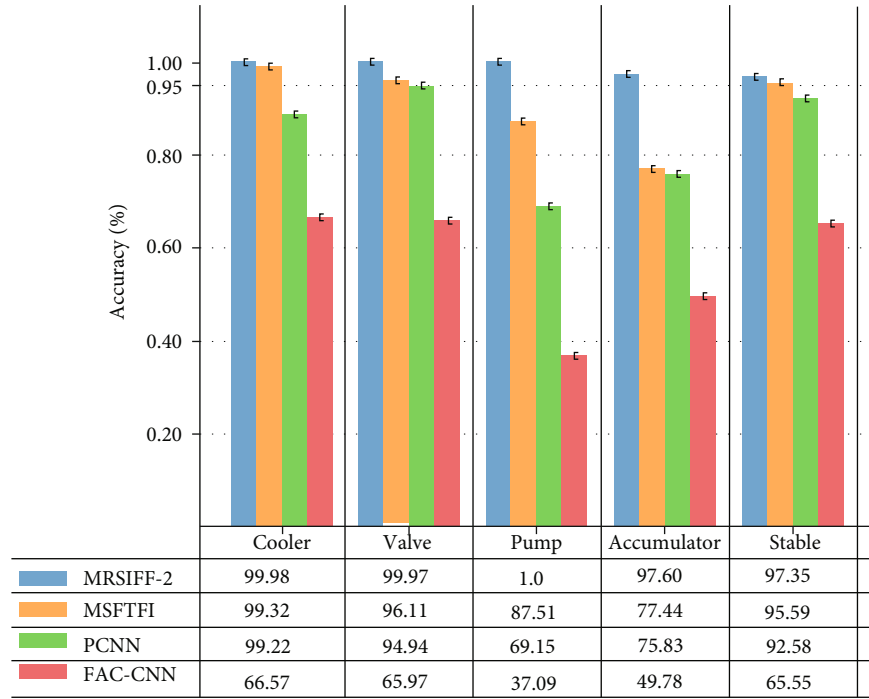
- (3) FA-CNN: researchers construct a convolution kernel of adaptive size matching data source channel to capture multiscale data without time-frequency analysis technology. In addition, to extract the diagnosis information of the fused data effectively, one-dimensional CNN and global average pooling methods are adopted to improve the domain adaptation of the network. More details of FA-CNN can be found in [18]
- (4) MRSIFS: in comparison experiments, MRSIFS were tested on layers 1 to 4 to test the depth of influence of MRSIFS on diagnostic performance. In addition, to provide a fair comparison, all comparison models have the same model depth as the MRSIFS proposed

To enable the fairness of comparison and improve the persuasiveness of the experiment results, all existing models have the same parameters and structure as the proposed MRSIFS. In addition, the same input form, number of training, batch size, and parameter optimization algorithm were adopted for all models.

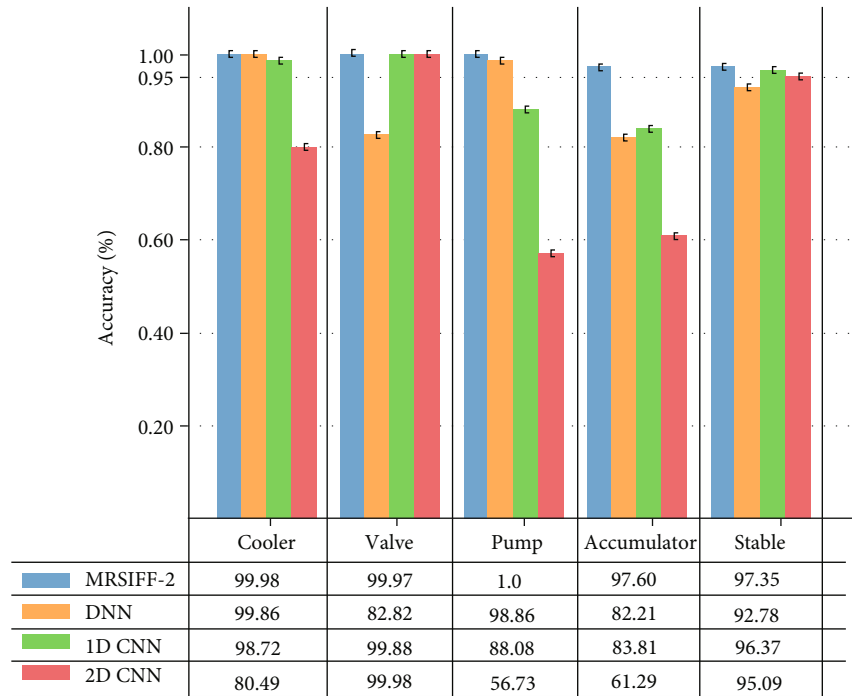
4.3. Diagnosis Results and Performance Comparison. To reduce the effect of randomness and improve the persuasiveness of experiment results, the average accuracies of ten replicate experiments for multitask fault diagnosis using different models are shown in Figure 3(a). The standard deviations of the cumulative experiments act as the error bars, which reflect the robustness of the method. In the figure, the mean represents the average accuracy of the model when solving the assigned tag task, and the standard deviation reflects the stability of the model.

As shown in Figure 3(a), the proposed MRSIFS achieves optimum performance on five tasks (coolers, valves, pumps, accumulators, and stabilizers). Specifically, the fault diagnosis accuracies of MRSIFS for each label are more than 97%, so MRSIFS has good multirate information extraction capability on multitask classification. Besides, since the standard deviations of MRSIFS on all tasks are very small, the proposed method has better robustness. The MSFTFI, PCNN, and FAC-CNN perform slightly less well than MRSIFS in the tag cooler, valve, and pump but have significantly worse performed than MRSIFS in tag accumulator and stable. Besides, as shown in Figure 3(b), MRSIFS perform better than the original CNNs and DNN in all five tasks. Therefore, we can draw two advantages, the first advantage is that MRSIFS performs better than the original CNNs. Another advantage of the proposed MRSIFS is that using multidimensional convolution kernels can extract more useful multiscale fault information than the traditional multisensor feature extraction model. Additionally, the multiscale feature connection layer fuses the fault feature information extracted from the original signals of different sampling frequencies by each submodel. Although FAC-CNN, PCNN, and MSFTFI can achieve excellent performance in task cooler, valve, and pump, they have lower than 80% accuracy in other tasks, which show that the existing models have low multirate feature extraction ability. The convergence performances of four different models are shown in Figure 3(b). We can see that the proposed MRSIFS converges faster than other models, which can be concluded that multidimensional cascade signals can provide more specific information than the raw signals.

To reveal that the existing models lack sufficient feature extraction capability, the losses of the four models on the condition monitoring of the hydraulic system dataset are shown in Figure 4. First of all, the loss of MRSIFS-2 (the depth of the multidimensional convolution feature learning



(a)



(b)

FIGURE 3: Testing fault diagnosis accuracy about four models (a) and the performance of the proposed MRSIFS in five tasks compared to all submodels (b).

layer is two) converges to zero. However, the loss of PCNN converges to about 0.1. Secondly, when more than 80 epochs were trained, the loss of MRSIFS-2 converges to zero, whereas the MSFTFI model needs to train more than 110 epochs to obtain similar results. In addition, since all labels of the raw signals have been normalized and preprocessed,

the losses of MRSIFS-2 and FAC-CNN are highly overlapping. To further explore the effectiveness of the proposed MRSIFS framework, the logarithmic function can be introduced to enlarge the difference between MRSIFS-2 and FAC-CNN. The logarithmic function value of the loss of MRSIFS is one order of magnitude lower than that of the

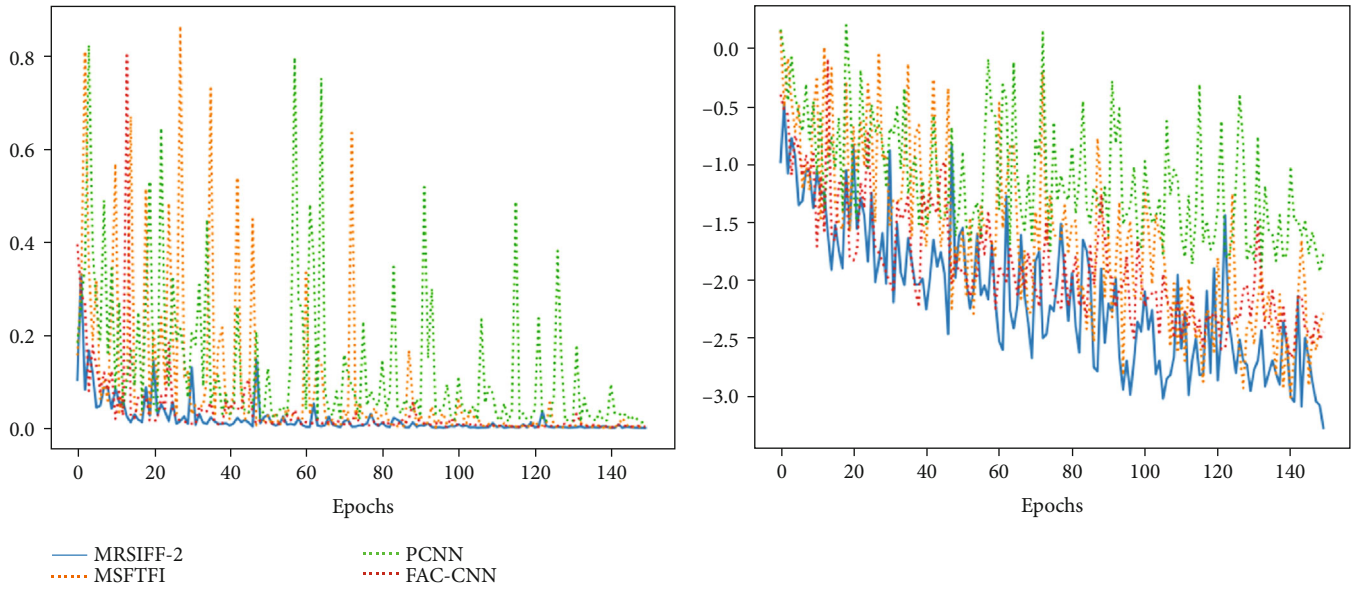


FIGURE 4: The losses of the four models based on the condition monitoring of the hydraulic system dataset.

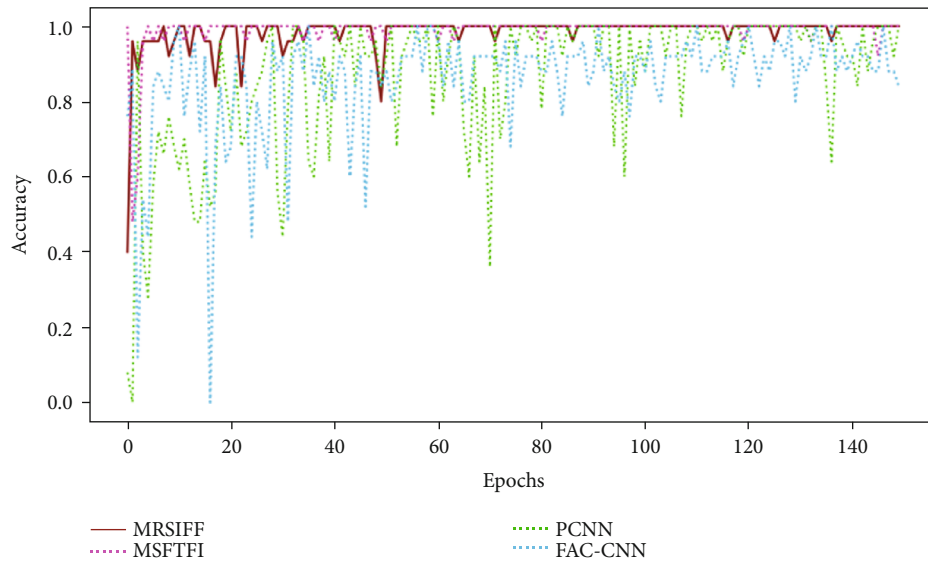


FIGURE 5: Training performance curve in terms of overall accuracy.

three existing models, which illustrates that the proposed MRSIFS model can learn the fault features and diagnosis information robustly from the raw vibration signals, and has fault discriminative ability. Thus, it can be proved that the multidimensional convolution feature learning stage actually makes the diagnosis information of the signal complementary by concatenating the useful components that come from multirate sensor systems.

It shows the overall accuracy curves for both MRSIFS, MSFTFI, PCNN, and FAC-CNN over 150 epochs in Figure 5 during the entire training process. As shown in Figure 5, compared with existing models, the proposed MRSIFS has stabilized over 95 accuracies after training 50 epochs. MRSIFS is more adaptable to complicated mechanical systems since it requires less time for training.

4.4. Feature Visualization via *t*-SNE and Confusion Matrix.

To further prove that multidimensional feature learning can improve the feature extraction capability of the model in complicated mechanical systems, the *t*-SNE method [40] is adopted to realize the visualization of the feature maps learned from the multidimensional convolution block. The feature map obtained through the feature extraction stage of multidimensional convolution is shown in Figure 6, which uses different colours to distinguish the features of samples with different tag types. Only consider the classification results of the original data of the four models when the hydraulic accumulator is used as the label because when the other four hydraulic components are used as the label, the accuracy of the classification results obtained by the MRSIFS model has reached 99.5%.

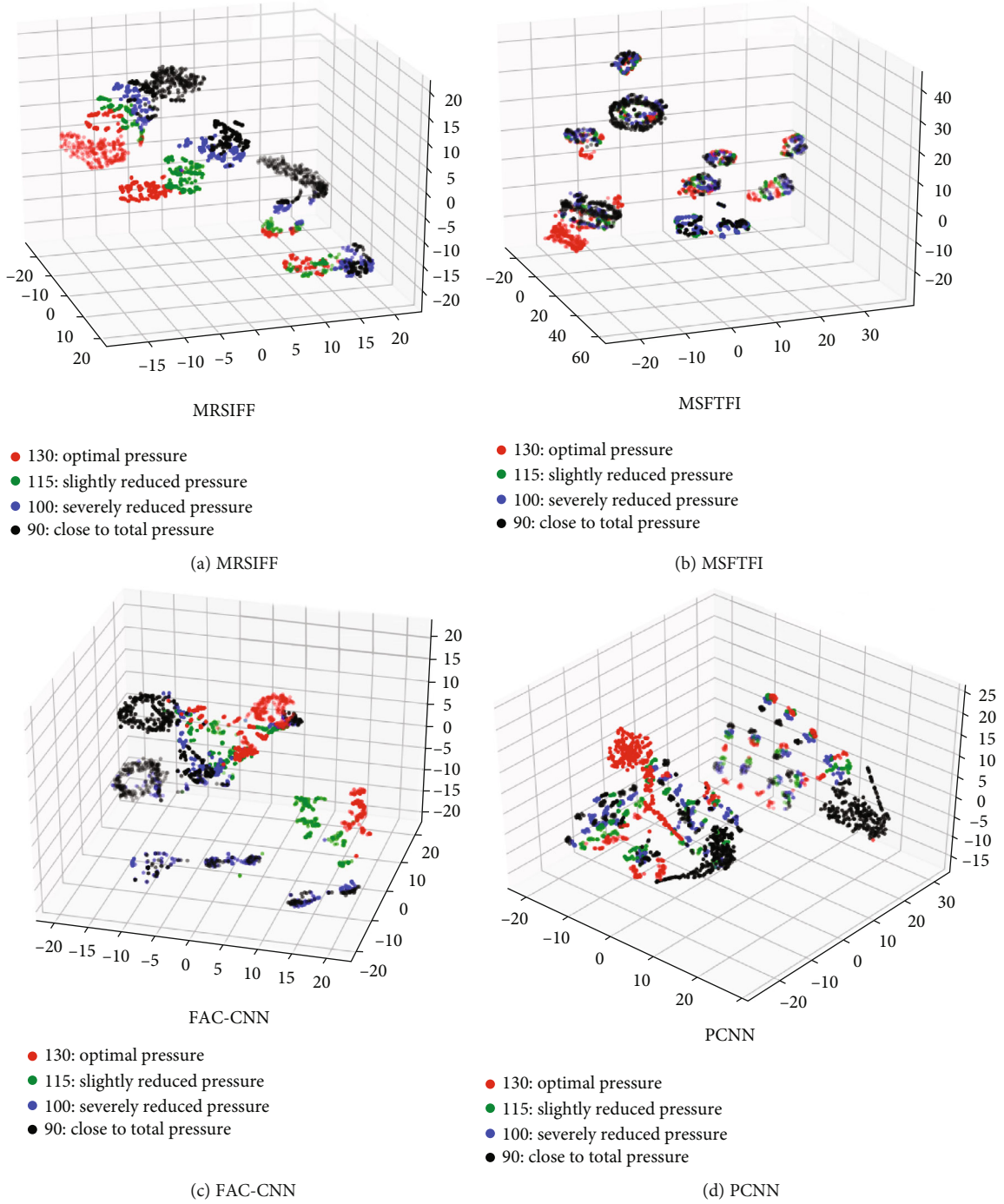


FIGURE 6: The feature visualization of different models.

Specifically, for the multitask classification problem, t-SNE is used to display the classification effect of the model better. As can be seen from Figure 6(a), samples are grouped into four categories under other labels, and each category is further divided into four categories, with accumulator samples as a label. However, in Figures 6(b)–6(d), the clustering results of the features of samples with the same tag type are worse, which indicates that the existing models cannot extract fault features from a multirate sensor. On the contrary, the proposed MRSIFS model has good domain adaptability and diagnosis information extraction ability

under complex working conditions, so it can finally learn fault discriminative features robustly.

Figure 7 gives the probability of the model making the correct classification and the probability of making the wrong classification for a given health condition. The X -axis represents the accurate label, and the Y -axis represents the predicted label.

4.5. Time Consumption. It is worth noting that, to make the comparison results more convincing, all the existing models and the proposed MRSIFS have the same convolutional layer

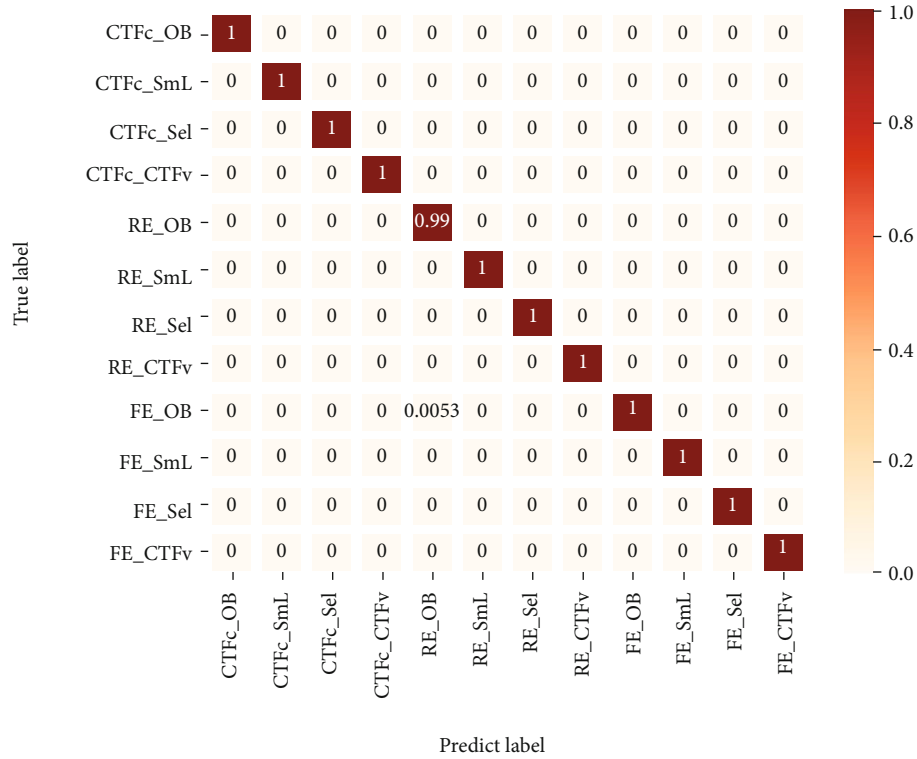


FIGURE 7: Confusion matrix results of cooler and valve.

depth. The training time and testing time spent by different models on the hydraulic system dataset are shown in Table 7. As shown in Table 7, MRSIFS, MSFTFI, and PCNN all require more time consuming than the original CNNs, which can be explained by the introduction of more parameters in multidimensional convolution blocks and result in more computing time. The testing time is a determining factor in the performance of the online fault diagnosis and detection model since the model is trained offline. Even if MRSIFS-2 needs more time to test, the testing time of the proposed method on the hydraulic system dataset is just 14.2161 ms, which shows the possibility about using MRSIFS-2 for online fault detection and diagnosis.

4.6. Discussions on the Effects of Depth. The depth of the convolutional layer will affect the diagnostic performance of the model based on deep learning. The proposed MRSIFS model can automatically adjust the depth of the model (i.e., the number of convolutional layer and pool layer pairs) according to the characteristics of the dataset. This section explores the influence of depth on feature extraction and fault diagnosis. The abstraction level of the feature is determined by the depth of MRSIFS and the scale of the convolution kernel. In complicated mechanical systems, the accuracies of classification results depend largely on the abstraction level of fault features. To test the influence of model depth on fault diagnosis, the accuracy of MRSIFS with one to four layers is recorded and compared with existing models in Figure 8.

As we know, the abstraction level of the extracted features can be determined by the depth of the MRSIFS. In addition, speed or load variations and background noise may suffer

TABLE 7: Cost time for MRSIFS for different scales and other models.

Method	Training time (s)/epoch	Testing time (ms)/one sample
MRSIFS-1	1.7278	2.1811
MRSIFS-2	2.0080	2.8432
MRSIFS-3	2.0813	3.7681
MRSIFS-4	2.2856	4.4364
MSFTFI	2.0809	1.7704
PCNN	1.7067	2.5053
FAC-CNN	1.4600	1.1219
DNN	1.3744	0.8290
1D CNN	1.5845	1.7763
2D CNN	1.7792	1.8995

low-level features obtained from the raw signal. Since the abstraction level of fault features can significantly impact the classification results, MRSIFS with layer depth of 1 to 4 are tested in this paper to investigate the influence of depth on diagnostic performance. The result is shown in Figure 8. In these experiments, the average accuracies of ten replicate experiments for each condition are shown in Figure 8, where the standard deviations of the cumulative experiments act as the error bars, which reflect the robustness of the method. A conclusion from the figure is that, for each tag, all MRSIFS outperform MSFTFI, PCNN, and FA-CNN in the fault diagnosis accuracy index. In general, as the depth increases to equal to two, MRSIFS can achieve the best and most reliable

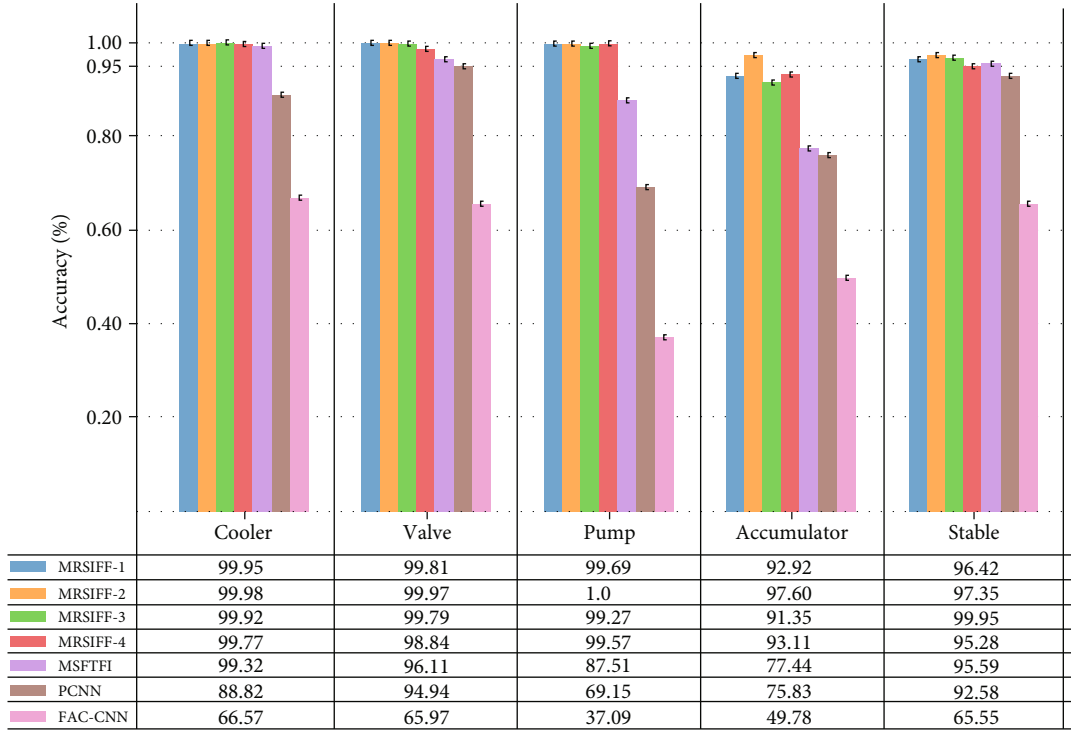


FIGURE 8: Diagnosis performance on the testing set using the MRSIFS with different scales from 1 to 4.

performance. Specifically, for both MRSIFS, MSFTFI, PCNN, and FAC-CNN, since MRSIFS with multidimensional convolution kernels can extract more useful and abstract fault information, which help in classification at higher levels, the classification performance improves with increasing depth. As the depth continues increasing, the accuracy of MRSIFS begins to decline. In addition, the standard deviation of each condition is smaller than these existing models, especially MRSIFS in the fourth task, which showed more excellent fault diagnosis performance than existing literatures. The result proves that MRSIFS can learn more abstract and fault sensitive features from multirate sensor systems. The time consumption of MRSIFS for training and testing at different convolutional layer depths was calculated, and the results are shown in Table 7. Obviously, the time overhead for optimizing model parameters increases with increasing depth. Therefore, to reduce the computational cost, MRSIFS with two layers are selected for feature extraction. For complex diagnostic tasks in practical applications, MRSIFS can be modified to improve the performance further.

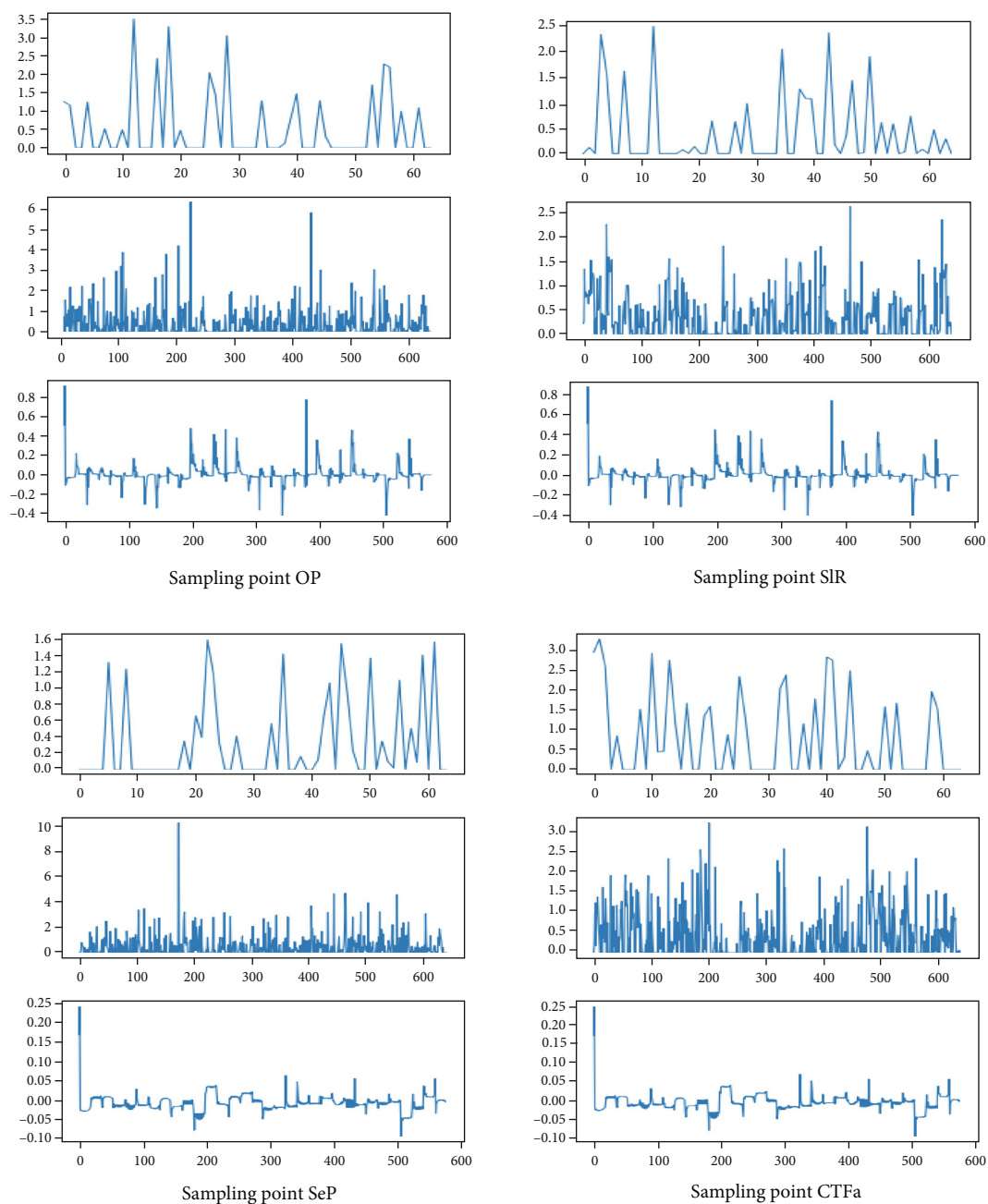
Another advantage of the proposed model is that, for a compound and changeable industrial environment, MRSIFS does not require a complex parameter adjustment process and can effectively extract fault information. Therefore, the model introduced in this paper is not sensitive to the setting of neural network parameters.

4.7. The Details of MRSIFF. To reveal the function of a multidimensional convolution feature learning layer, the evolution of the neurons in the MRSIFF, the learned kernels, multirate sensor cascade signals, and the output of each layer in the MRSIFF are displayed in this section. It can be seen

from Figure 9 that the length of multiscale feature fusion signals has been increased because of the series connection of signals convoluted by kernels with different dimensions. In addition, the waveform shapes of the fusion signals have been changed. In Figure 9(a), there is little difference between the signal wave shapes (SWS) of ball and normal. However, in Figure 9(b), the fusion-SWS of them are totally different, which can provide more useful features for the classifier.

To reveal how the raw signals change in the MRSIFF and what is the input of the classifier, the evolution of the inputs in MRSIFF and the features of Multilayer Perceptron (MLP) are shown in Figure 10, respectively. It can be seen from Figure 10 that the outputs of the C3 (the third convolutional layer) are similar to the wave shape of fusion signals and they can be distinguished obviously in the four conditions. The MLP features are constituted by the features in C3, and it can be seen from Figure 10 that the features before putting into the classifier are linear separable, which can increase the classification accuracy. Furthermore, the fully connected features of MRSIFF, MSFTFI, PCNN, and FAC-CNN are mapped into three-dimension features using t-SNE which are shown in Figure 6. Obviously, the mapped features of the MRSIFF cluster are better than other models based on CNN even though other models can differentiate most samples. Thus, the t-SNE error of MRSIFF is the lowest among the three models.

4.8. Summary of Experimental Results. In this paper, another deep convolution structure is proposed, which is different from the traditional CNN structure that only utilizes features in the last convolution layer. In the feature extraction stage, the upsampling layer is combined with the



(a)

FIGURE 9: Continued.

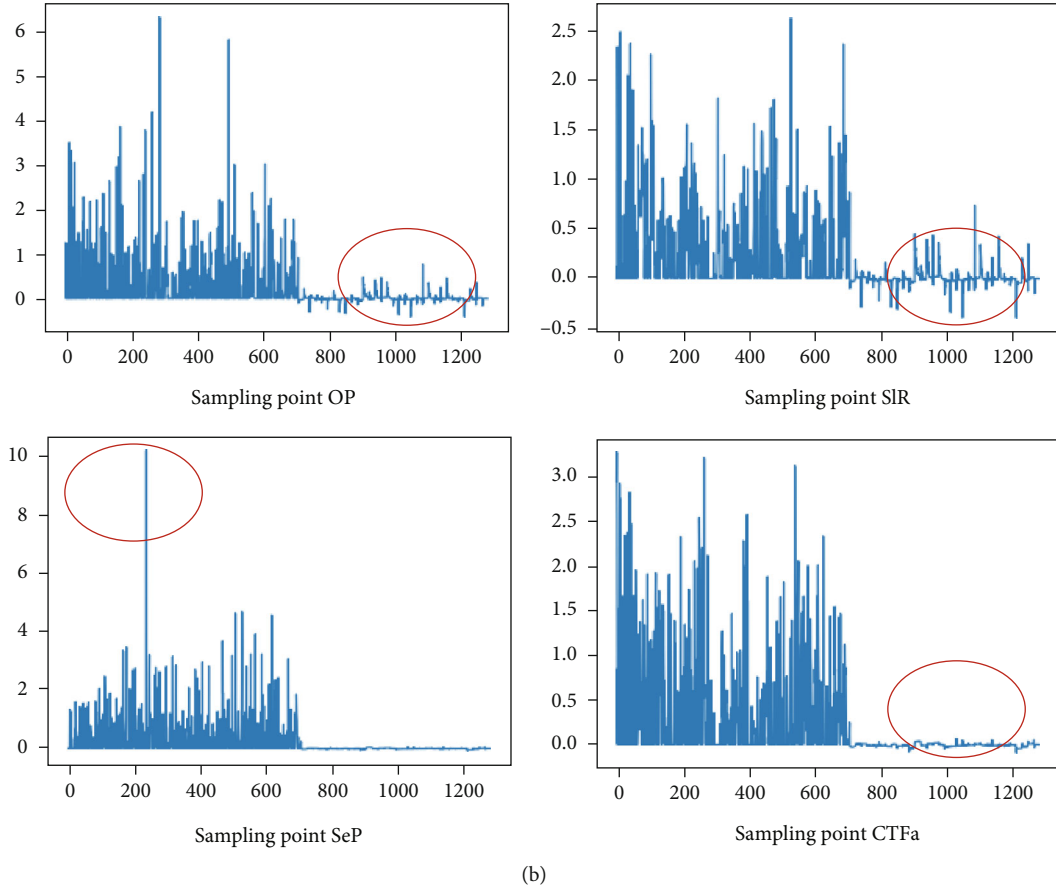


FIGURE 9: Raw signal (a) and fusion signal (b) of each healthy condition of the hydraulic system.

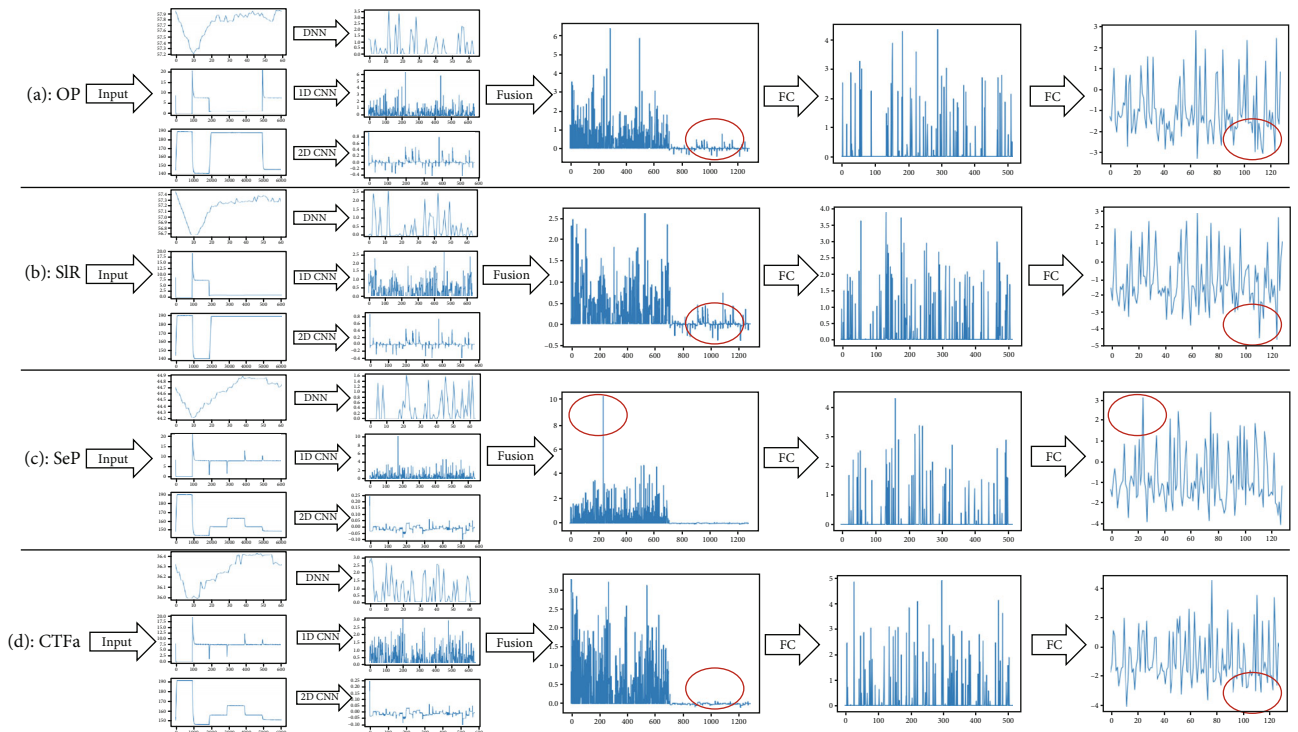


FIGURE 10: The whole evolution of signals of each healthy condition in MRSIFF.

convolutional layer to form the compound convolutional block, while maintaining the global and local features and improving the network capacity. In this structure, the multidimensional convolution layer and a full connection layer are used to extract features from multiscale input signals. In this method, the extracted features are concatenated to improve the accuracy of fault diagnosis, since the receptive fields between 1D convolution and 2D convolution are different.

The proposed MRSIFS, which is put forward, is tested on the hydraulic system condition monitoring dataset. The experiment results show that the fault diagnosis performance of the model based on 1D CNN is better than the model based on 2D CNN. However, when 1D CNN, 2D CNN, and DNN are combined, the accuracy of MRSIFS for fault diagnosis is improved to 97–99.5%, which indicates that multidimensional convolutional blocks can learn complementary and rich fault features. In addition, the same parameters and structures are used in MSFTFI, PCNN, and FAC-CNN, and the results show that the performance of the proposed Multirate Sensor Information Fusion Strategy is much better than all existing models.

Compared with the traditional intelligent fault detection models, the proposed MRSIFS has better performance in both feature extraction and detection accuracy. Therefore, the proposed method outperforms the existing methods.

4.9. Feature Work. In the future, the normal samples are far more than the fault samples due to the fault that will cause damage to the mechanical system, and existing research methods assume that the original signal dataset has sufficient balanced sample types. Therefore, it is necessary to develop a fault diagnosis model combining Generative Adversarial Networks (GAN), which is of great significance for the practical engineering environment and is also the research direction of the author in the future.

5. Conclusions

This paper introduced a Multirate Sensor Information Fusion Strategy (MRSIFS). The proposed method is based on multidimensional convolution block and time-frequency analysis technology, which implement multichannel parallel fault feature extraction, and the features from raw signals with different sampling rates are used for fault diagnosis. The multirate sensor feature extraction is a novel multitask feature extraction unit using a multidimensional convolution block and Adam loss function, which significantly improves the feature extraction capability. Finally, the simulation platform's experimental results show that the proposed multitask model achieves higher diagnosis accuracy than the existing methods. Besides, the manual feature selection for each task is unnecessary in MRSIFS, which has the potential toward a general-purpose framework. The main conclusions can be listed as follows:

- (1) The hierarchical learning structure of multiple convolution blocks can be used to learn advanced fault features effectively

- (2) To obtain fault sensitive and complementary detection features, this paper proposes a multidimensional convolution feature extraction model to adapt signal sources at different sampling frequencies
- (3) Another valuable characteristic of the proposed strategy is the ability to work directly with raw sensor data, thus providing an end-to-end model to perform feature extraction and classification simultaneously

Data Availability

The hydraulic system condition monitoring dataset is available from the UC Irvine Machine Learning Repository at <https://archive.ics.uci.edu/ml/datasets/Condition+monitoring+of+hydraulic+system>.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (grant no. 61976047) and the Science and Technology Department of Sichuan Province of China (grant no. 2021YFG0331).

References

- [1] C. Sun, Z. Zhang, Z. He, Z. Shen, and B. Chen, "Manifold learning-based subspace distance for machinery damage assessment," *Mechanical System & Signal Processings*, vol. 70–71, pp. 637–649, 2016.
- [2] Z. du, X. Chen, H. Zhang, and R. Yan, "Sparse feature identification based on union of redundant dictionary for wind turbine gearbox fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6594–6605, 2015.
- [3] G. Jiang, H. He, J. Yan, and P. Xie, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3196–3207, 2019.
- [4] B. Muruganatham, M. A. Sanjith, B. Krishnakumar, and S. A. V. Satya Murthy, "Roller element bearing fault diagnosis using singular spectrum analysis," *Mechanical System & Signal Processing*, vol. 35, no. 1–2, pp. 150–166, 2013.
- [5] B. Cai, Y. Liu, Q. Fan et al., "Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network," *Applied Energy*, vol. 114, no. 2, pp. 1–9, 2014.
- [6] P. Imperatore, G. Persichetti, G. Testa, and R. Bernini, "Continuous liquid level sensor based on coupled light diffusing fibers," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 4, pp. 1–8, 2020.
- [7] L. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomedical Optics Express*, vol. 8, no. 5, pp. 2732–2744, 2017.

- [8] J. Heaton, "Ian Goodfellow, Yoshua Bengio, and Aaron Courville: deep learning," *Genetic Programming and Evolvable Machines*, vol. 19, no. 1-2, pp. 305–307, 2018.
- [9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [10] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2018.
- [11] F. Jia, Y. Lei, N. Lu, and S. Xing, "Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization," *Mechanical System and Signal Processing*, vol. 110, pp. 349–367, 2018.
- [12] D. Peng, Z. Liu, H. Wang, Y. Qin, and L. Jia, "A novel deeper one-dimensional CNN with residual learning for fault diagnosis of wheelset bearings in high-speed trains," *IEEE Access*, vol. 7, pp. 10278–10293, 2019.
- [13] Z. Q. Chen, C. Li, and R.-V. Sanchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock and Vibration*, vol. 2015, Article ID 390134, 10 pages, 2015.
- [14] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals," *Sensors*, vol. 17, no. 2, p. 425, 2017.
- [15] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep ConvNet for intelligent spindle bearing fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 1926–1935, 2017.
- [16] M. J. Hasan, M. M. M. Islam, and J.-M. Kim, "Multi-sensor fusion-based time-frequency imaging and transfer learning for spherical tank crack diagnosis under variable pressure conditions," *Measurement*, vol. 168, p. 108478, 2021.
- [17] X. Xu, Z. Tao, W. Ming, Q. An, and M. Chen, "Intelligent monitoring and diagnostics using a novel integrated model based on deep learning and multi-sensor feature fusion," *Measurement*, vol. 165, p. 108086, 2020.
- [18] S. Li, H. Wang, L. Song, P. Wang, L. Cui, and T. Lin, "An adaptive data fusion strategy for fault diagnosis based on the convolutional neural network," *Measurement*, vol. 165, p. 108122, 2020.
- [19] H. Tao, P. Wang, Y. Chen, V. Stojanovic, and H. Yang, "An unsupervised fault diagnosis method for rolling bearing using STFT and generative neural networks," *Journal of the Franklin Institute*, vol. 357, no. 11, pp. 7286–7307, 2020.
- [20] N. Helwig, E. Pignatelli, and A. Schütze, "Condition monitoring of a complicated hydraulic system using multivariate statistics," in *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, pp. 210–215, Pisa, Italy, 2015.
- [21] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [22] H. Qiao, T. Wang, P. Wang, L. Zhang, and M. Xu, "An adaptive weighted multiscale convolutional neural network for rotating machinery fault diagnosis under variable operating conditions," *IEEE Access*, vol. 7, pp. 118954–118964, 2019.
- [23] H. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2727–2736, 2018.
- [24] J. Wu, Y. Su, Y. Cheng, X. Shao, C. Deng, and C. Liu, "Multi-sensor information fusion for remaining useful life prediction of machining tools by adaptive network based fuzzy inference system," *Applied Soft Computing*, vol. 68, pp. 13–23, 2018.
- [25] X. Ma, P. Wang, B. Zhang, and M. Sun, "An improved multi-channels information fusion model with multi-scale signals for fault diagnosis," in *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, pp. 552–555, Nanchang, China, March 2021.
- [26] E. Jove, J. L. Casteleiro-Roca, H. Quintián, D. Simić, J. A. Méndez-Pérez, and J. Luis Calvo-Rolle, "Anomaly detection based on one-class intelligent techniques over a control level plant," *Logic Journal of IGPL*, vol. 28, no. 4, pp. 502–518, 2020.
- [27] J. L. Calvo, S. H. Tindemans, and G. Strbac, "Risk-based method to secure power systems against cyber-physical faults with cascading impacts: a system protection scheme application," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 5, pp. 930–943, 2018.
- [28] R. Bernini, G. Persichetti, E. Catalano, L. Zeni, and A. Minardo, "Refractive index sensing by Brillouin scattering in side-polished optical fibers," *Optics Letters*, vol. 43, no. 10, pp. 2280–2283, 2018.
- [29] E. Jove, J. L. Casteleiro-Roca, H. Quintián, J. A. Méndez-Pérez, and J. L. Calvo-Rolle, "A new method for anomaly detection based on non-convex boundaries with random two-dimensional projections," *Information Fusion*, vol. 65, pp. 50–57, 2020.
- [30] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *2017 IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 936–944, Honolulu, HI, USA, 2017.
- [31] N. Helwig, E. Pignatelli, and A. Schütze, "Condition monitoring of a complex hydraulic system using multivariate statistics," in *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, pp. 210–215, Pisa, Italy, 2015.
- [32] T. Schneider, N. Helwig, and A. Schütze, "Industrial condition monitoring with smart sensors using automated feature extraction and selection," *Measurement Science and Technology*, vol. 29, no. 9, p. 094002, 2018.
- [33] J. Prakash and P. Kankar, "Health prediction of hydraulic cooling circuit using deep neural network with ensemble feature ranking technique," *Measurement*, vol. 151, p. 107225, 2020.
- [34] B. Krishnapuram, M. Shah, A. Smola, C. Aggarwal, D. Shen, and R. Rastogi, "Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining," in *Acm Sigkdd International Conference on Knowledge Discovery & Data Mining ACM*, San Francisco California USA, 2016.
- [35] M. Zhao, M. Kang, B. Tang, and M. Pecht, "Multiple wavelet coefficients fusion in deep residual networks for fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 6, pp. 4696–4706, 2019.
- [36] Y. Li, Y. Chen, N. Wang, and Z. Zhang, "Scale-aware trident networks for object detection," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 6053–6062, Seoul, Korea (South), 2019.
- [37] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," 2015, <https://arxiv.org/abs/1502.03167>.

- [38] D. Kingma and J. Ba, “Adam: a method for stochastic optimization,” *Computer Science*, vol. 12, 2014.
- [39] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [40] L. J. P. Van Der Maaten, “Accelerating t-SNE using tree-based algorithms,” *The Journal of Machine Learning Research*, vol. 15, no. 93, pp. 3221–3245, 2014.