SCIENCE CHINA
Technological Sciences

•Article•

February 2022 Vol.65 No.2: 470–480
https://doi.org/10.1007/s11431-021-1904-7

Adaptive decision-level fusion strategy for the fault diagnosis of axial piston pumps using multiple channels of vibration signals

CHAO Qun[1,2], GAO HaoHan[1,2], TAO JianFeng[1,2*], WANG YuanHang[3], ZHOU Jian[3] & LIU ChengLiang[1,2]

[1] *State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China;*
[2] *MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, Shanghai 200240, China;*
[3] *China Electronic Product Reliability and Environmental Testing Research Institute, Guangzhou 510610, China*

Received April 20, 2021; accepted July 23, 2021; published online November 18, 2021

An axial piston pump is a key component that plays the role of the "heart" in hydraulic systems. The pump failure will lead to an unexpected breakdown of the entire hydraulic system or even economic loss and catastrophic safety consequences. Several vibration-based machine learning methods have been developed to detect and diagnose faults of axial piston pumps. However, most of these intelligent diagnosis methods use single-sensor vibration data to monitor the pump health states. Additionally, the diagnostic accuracy is unacceptable in most situations due to the complex pump structure and limited sensor information. Therefore, this study proposes a multi-sensor fusion method to improve the fault diagnosis performance of axial piston pumps. The convolutional neural network receives three channels of vibration data and makes the final diagnosis through information fusion at the decision level. The proposed decision fusion method is evaluated on the classification task of leakage levels of an actual axial piston pump. The experimental results show that the proposed method improves the classification accuracy by adjusting the probability distribution of classification according to the learned weight matrix.

axial piston pump, leakage, fault diagnosis, CNN, decision-level fusion

Citation: Chao Q, Gao H H, Tao J F, et al. Adaptive decision-level fusion strategy for the fault diagnosis of axial piston pumps using multiple channels of vibration signals. Sci China Tech Sci, 2022, 65: 470–480, https://doi.org/10.1007/s11431-021-1904-7

1 Introduction

An axial piston pump is a key component in hydraulic systems, and its performance greatly affects the efficiency and reliability of the entire hydraulic system. Severe operating conditions, such as high pressure, high speed, and high temperature, accelerate the deterioration of the pump performance, leading to the wear of friction pairs. Pump failure due to the worn friction pairs may cause unexpected breakdown of the hydraulic system or even considerable economic losses and catastrophic safety incidents. Therefore, it is necessary and urgent to develop effective fault detection methods and make maintenance decisions accordingly for the axial piston pump.

Generally, the fault diagnosis methods are divided into model-based, signal-based, and data-driven-based methods [1]. An axial piston pump is a complex mechanical-electrical-hydraulic coupling system. Thus, there are highly nonlinear relationships between parameters and strong couplings among various fault features [2]. Therefore, data-driven-based diagnosis methods are suitable for axial piston pumps since it is not easy to establish explicit models or signal symptoms. Machine learning is one of the emerging data-driven fault diagnosis methods for rotating machinery [3]. It includes two steps of manual feature extraction and fault classification. In the first step, signal processing techniques are used to design handcrafted features and extract them from sensor data. For the axial piston pump, the

*Corresponding author (email: jftao@sjtu.edu.cn)

common signals for reflecting its health conditions include vibration signal [4–6], discharge pressure signal [7], acoustic signal [8], and a combination of them [9]. In the second step, a certain classifier, such as support vector machine [5], decision tree [10], and Naïve Bayes [11], is trained by the extracted features and then recognizes the fault patterns using new data.

Traditional machine learning methods require carefully designed features and cannot handle raw data directly. Therefore, their performance depends on expert knowledge and the manual feature extraction process, which is laborious and time-consuming [12]. However, deep learning methods have a powerful end-to-end learning ability to automatically extract representative features from raw data and integrate feature extraction and pattern recognition into a single body. As one of the most popular deep learning methods, convolutional neural network (CNN) has achieved great success in computer vision [13] and speech recognition [14]. Recently, CNN has become a hot topic of research in the fault diagnosis of rotating machinery [15], such as bearing [16], motor [17], and gearbox [18].

In addition, some researchers have found applications of CNN in the fault diagnosis of hydraulic pumps. Yan et al. [19] performed fault diagnosis of an axial piston pump using a one-dimensional (1D) CNN framework and one-channel raw vibration data. Wen et al. [1] utilized the advantage of two-dimensional (2D) CNN in image recognition tasks. They converted the raw vibration data into grayscale images to classify the wear faults among the slipper, swash plate, and valve plate in an axial piston pump. Similarly, Tang et al. [20] applied continuous wavelet transform to convert raw time series vibration data into time-frequency images and took them as input of CNN to recognize five fault types of an axial piston pump. Wang and Xiang [21] proposed a vibration-based fault diagnosis method by combining CNN and minimum entropy deconvolution to detect four common faults in the axial piston pump. Sun et al. [22] compared the gear pump fault diagnosis performance with the same CNN framework between different input time-frequency images converted from vibration data by wavelet transform, Fourier transform, and Wigner-Will distribution. Xu et al. [23] designed a deep transfer CNN framework and verified it on the vibration dataset of a centrifugal pump under five health conditions. Kumar et al. [24] developed an improved CNN model using a modified cost function. They combined it with acoustic images to identify the defective components of a centrifugal pump.

From the literature review, CNN offers significant potential for the fault diagnosis of hydraulic pumps. In most cases, single-sensor vibration data are used to monitor the health state of hydraulic pumps [25]. However, it is a great challenge to obtain an acceptable fault diagnosis performance using limited information captured by single-sensor data due to the complex structure of axial piston pumps. Compared with single-sensor data, multiple-sensor data acquire redundant and complementary fault information that can be fused to achieve more accurate and reliable diagnosis results. Some multi-sensor fusion methods have been applied to the fault diagnosis of rotating machinery, such as bearing [26–28], motor [29], gear [26,30–32], and hydraulic pump [33–35]. The information fusion methods can be classified into three levels: data level [26,30–34], feature level [27,35], and decision level [28,29]. At the data level, the raw data from the same or different sensors are integrated directly, followed by feature extraction from the combined data. At the feature level, features are extracted from raw data, and the optimal ones are combined as an input to the decision layer. At the decision level, each sub-classifier receives single-sensor data to provide preliminary result separately, and multiple classifiers determine the final classification.

The typical decision fusion strategies include the voting method [36], Bayesian theory [37], and D-S evidence theory [38]. Although the voting method is a simple and widely used fusion strategy, it suffers from the drawback of the equal contribution of all classifiers during decision fusion. This is unreasonable because the classifiers have different diagnosis performances. Therefore, this study proposes an improved decision-level fusion (DF) strategy for the fault diagnosis of axial piston pumps based on CNN and multiple channels of vibration signals to enhance classification accuracy. The proposed method extracts features from multi-channel vibration data automatically and fuses the information from all CNN models at the decision level adaptively.

The remainder of the paper is organized as follows. Sect. 2 briefly introduces the CNN model. Sect. 3 presents a detailed CNN-based decision fusion method for identifying the health states and degradation levels of an axial piston pump. The experiments are presented in Sect. 4 to validate the effectiveness of the proposed method. Finally, Sect. 5 presents the conclusion.

2 Convolutional neural network

CNN is a feed-forward neural network biologically inspired by the visual system structure of the human brain. It presents a powerful ability to process 2D data, such as images, due to its three main characteristics: local connection, weight sharing, and spatial subsampling. A typical CNN consists of layers: convolutional, pooling, and fully connected layers. The convolutional layer applies a set number of kernels (also known as filters) to extract features from the input images or feature maps of the previous layer. The output of the convolutional layer can be described as [26]

$$\mathbf{X}_j^l = f\left(\sum_{i=1}^{M} \mathbf{W}_{ij}^l * \mathbf{X}_i^{l-1} + \mathbf{b}_j^l\right), \tag{1}$$

where $\mathbf{X}_j^l$ is the jth feature map at the l layer, $\mathbf{X}_i^{l-1}$ is the ith feature map at the (l–1) layer, $\mathbf{W}_{ij}^l$ denotes the weight matrix of filter connecting these two feature maps, * denotes the operator of 2D convolution, $\mathbf{b}_j^l$ is the bias matrix, M is the number of input feature maps, and f represents a nonlinear activation function. In this study, the rectified linear unit $f(x)$ = max(0, x) is selected as the activation function since it accelerates training and avoids overfitting [39].

The pooling layer is a subsampling layer, which usually follows a convolutional layer to reduce the dimensions of feature maps. The pooling operation is applied to each feature map. It fuses the feature values in a pooling window into one value, contributing to spatial invariance and robustness of the learned features. Typical pooling operations include max-pooling and average-pooling that calculate the maximum and average values among the neighboring elements in the pooling window, respectively. In this study, max-pooling is used at the pooling layer.

$$y_{mnk}^l = \max\left(y_{m'n'k}^{l-1} \middle| \begin{matrix}(m-1)H+1 \le m' \le mH, \\ (n-1)W+1 \le n' \le nW\end{matrix}\right), \tag{2}$$

where $y_{m'n'k}^{l-1}$ represents the kth feature map input to the pooling layer, y_{mnk}^l represents the corresponding output feature map after max-pooling operation; H and W are the height and width of the pooling window, respectively.

The fully connected layer follows several alternating convolutional and pooling layers. It works like a regular multilayer neural network with hidden and classification layers. Here the neurons are connected to all output elements of the previous layer, and the softmax regression plays a classifier role. The softmax classifier predicts the probability $\mathbf{p}_{\boldsymbol{\theta}}(\mathbf{x}^{(i)})$ of each class as follows:

$$\mathbf{p}_{\theta}(\mathbf{x}^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1|\mathbf{x}^{(i)}; \boldsymbol{\theta}) \\ p(y^{(i)} = 2|\mathbf{x}^{(i)}; \boldsymbol{\theta}) \\ \vdots \\ p(y^{(i)} = C|\mathbf{x}^{(i)}; \boldsymbol{\theta}) \end{bmatrix} = \frac{1}{\sum_{c=1}^{C} \exp\left(\boldsymbol{\theta}_c^{\mathrm{T}} \mathbf{x}^{(i)}\right)} \begin{bmatrix} \exp\left(\boldsymbol{\theta}_1^{\mathrm{T}} \mathbf{x}^{(i)}\right) \\ \exp\left(\boldsymbol{\theta}_2^{\mathrm{T}} \mathbf{x}^{(i)}\right) \\ \vdots \\ \exp\left(\boldsymbol{\theta}_C^{\mathrm{T}} \mathbf{x}^{(i)}\right) \end{bmatrix}, \tag{3}$$

where $\mathbf{x}^{(i)}$ is the ith input sample, $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_C]$ are the trainable parameters of the model, the lowercase letter c denotes the cth class, and the uppercase letter C denotes the total classification number. The superscript T denotes the matrix transpose operation.

The loss function of the softmax classifier is given by

$$J(\boldsymbol{\theta}) = -\frac{1}{S}\sum_{i=1}^{S}\sum_{c=1}^{C} I(y^{(i)} = c)\log\left(\frac{\exp\left(\boldsymbol{\theta}_c^{\mathrm{T}} \mathbf{x}^{(i)}\right)}{\sum_{c=1}^{C} \exp\left(\boldsymbol{\theta}_c^{\mathrm{T}} \mathbf{x}^{(i)}\right)}\right), \tag{4}$$

where S is the total number of samples, I is an indicator function; it is 1 for a true condition; otherwise, it is 0 for a false condition. The gradient descent method is used to minimize the loss function during the training process. In this study, RMSprop [40] is selected as the optimizer to update the learnable parameters.

LeNet-5 is a famous CNN architecture developed by Lecun et al. [41]. It contains two convolutional layers and two pooling layers alternatively, and a fully connected layer. In this study, the CNN model is an improved version of the LeNet-5 network (Figure 1). The input layer receives 128 × 128-pixel gray images converted from one channel of raw vibration data by the short-time Fourier transform (STFT). The feature maps are automatically extracted layer by layer and finally flattened into a vector to feed the fully connected layer for classification. Several batch normalization layers are added to the network to improve the convergence and

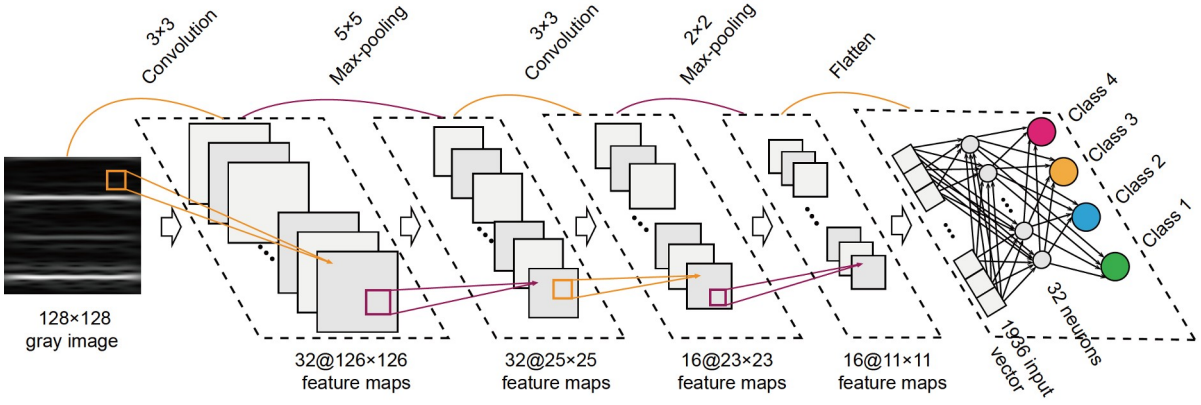

Figure 1 (Color online) Architecture of the convolutional neural network for fault diagnosis of the axial piston pump.

stability of the network. In addition, dropout layers are integrated to enhance the generalization ability of the CNN model.

3 Decision-level information fusion strategy

This section presents two strategies of decision-level information fusion, both of which include four steps: training, validation, testing, and fusion processes. The final result of fault diagnosis is jointly determined by m CNN models at the decision level. The basic idea behind the proposed method is that the contribution of each CNN model to the final decision depends on its diagnosis performance in the validation process.

Figure 2 shows the first strategy of decision-level information fusion. Assume that the acquired data are collected from m types or channels of sensors, and there are C fault types to be recognized. Each type or channel of sensor data is divided into training, validation, and testing data. After model training on the training data, a total number of m validation accuracies are obtained on the validation data. These validation accuracies form of the weight matrix $\mathbf{W}_{\text{val}}$ of all CNN models as follows:

$$\mathbf{W}_{\text{val}} = \left[W_{\text{val1}} \quad W_{\text{val2}} \quad \cdots \quad W_{\text{val}m}\right]^{\text{T}} = \left[\frac{\text{val}_1}{\sum_{i=1}^{m}\text{val}_i} \quad \frac{\text{val}_2}{\sum_{i=1}^{m}\text{val}_i} \quad \cdots \quad \frac{\text{val}_m}{\sum_{i=1}^{m}\text{val}_i}\right]^{\text{T}}, \tag{5}$$

where val_i and $W_{\text{val}i}$ (i = 1, 2, …, m) are the validation accuracy and weight coefficient of the ith model, respectively.

In the testing process, each trained CNN model receives the testing data to predict the probability vector $\mathbf{P}_{\text{test}i}$ that contains the probabilities of all fault types.

$$\mathbf{P}_{\text{test}i} = \left[p_i(y=1|x_{\text{test}}) \quad p_i(y=2|x_{\text{test}}) \quad \cdots \quad p_i(y=C|x_{\text{test}})\right]^{\text{T}}. \tag{6}$$

The probability matrix $\mathbf{P}_{\text{test}}$ for all CNN models on the testing data before the information fusion is expressed as

$$\mathbf{P}_{\text{test}} = \left[\mathbf{P}_{\text{test1}} \quad \mathbf{P}_{\text{test2}} \quad \mathbf{P}_{\text{test}m}\right] = \begin{bmatrix} p_1(y=1|x_{\text{test}}) & p_2(y=1|x_{\text{test}}) & \cdots & p_m(y=1|x_{\text{test}}) \\ p_1(y=2|x_{test}) & p_2(y=2|x_{test}) & \cdots & p_m(y=2|x_{\text{test}}) \\ \vdots & \vdots & \ddots & \vdots \\ p_1(y=C|x_{\text{test}}) & p_2(y=C|x_{\text{test}}) & \cdots & p_m(y=C|x_{\text{test}}) \end{bmatrix}. \tag{7}$$

The fusion process generates the final probability for all fault types by combining the probability matrix $\mathbf{P}_{\text{test}}$ and weight matrix $\mathbf{W}_{\text{val}}$.

$$\mathbf{P} = \left[P_1 \quad P_2 \quad \cdots \quad P_C\right]^{\text{T}} = \mathbf{P}_{\text{test}}\mathbf{W}_{\text{val}} = \begin{bmatrix} p_1(y=1|x_{\text{test}})\frac{\text{val}_1}{\sum_{i=1}^{m}\text{val}_i} + \cdots + p_m(y=1|x_{\text{test}})\frac{\text{val}_m}{\sum_{i=1}^{m}\text{val}_i} \\ p_1(y=2|x_{\text{test}})\frac{\text{val}_1}{\sum_{i=1}^{m}\text{val}_i} + \cdots + p_m(y=2|x_{\text{test}})\frac{\text{val}_m}{\sum_{i=1}^{m}\text{val}_i} \\ \vdots \\ p_1(y=C|x_{\text{test}})\frac{\text{val}_1}{\sum_{i=1}^{m}\text{val}_i} + \cdots p_m(y=C|x_{\text{test}})\frac{\text{val}_m}{\sum_{i=1}^{m}\text{val}_i} \end{bmatrix}, \tag{8}$$

where P_c (c = 1, 2, …, C) is the final probability for the cth fault type after the information fusion.

The first fusion method calculates the weight coefficients using the overall validation accuracy of each CNN model. It

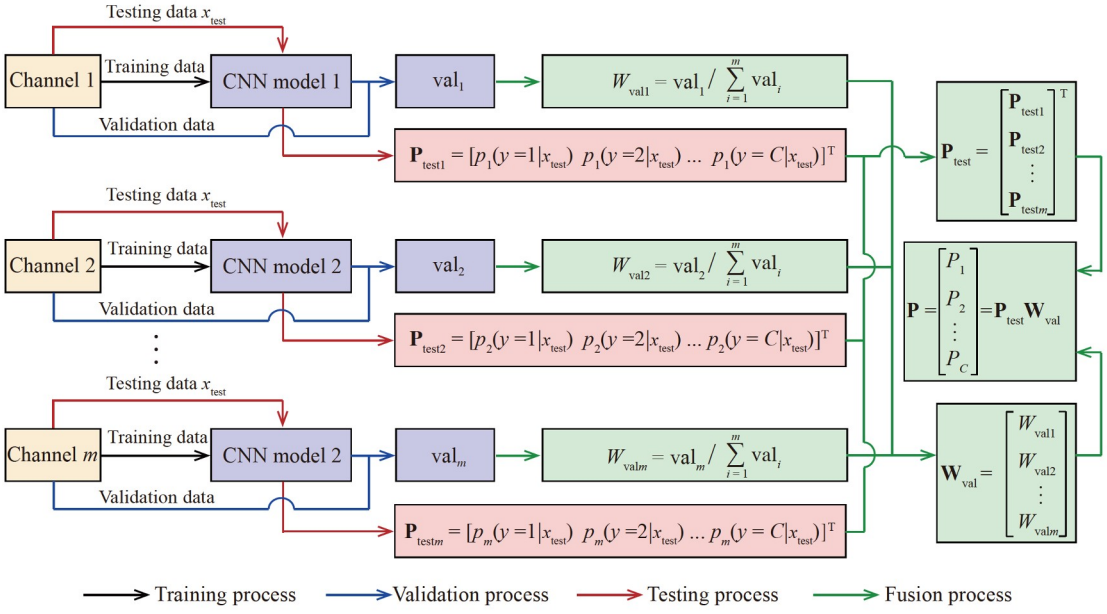

Figure 2 (Color online) Flowchart of the first decision-level information fusion method.

means that each CNN model has the same weight coefficient for the final classification probability regardless of the fault types. An alternative method for determining weight coefficients is to consider the validation accuracy of each fault type instead of the overall validation accuracy (Figure 3).

The new weight matrix of all CNN models is given by

$$\mathbf{W}_{\text{val}}=\begin{bmatrix}\mathbf{W}_{\text{val}1}\\ \mathbf{W}_{\text{val}2}\\ \vdots\\ \mathbf{W}_{\text{val}m}\end{bmatrix}=\begin{bmatrix}\text{val}_{11} & \text{val}_{12} & \cdots & \text{val}_{1C}\\ \text{val}_{21} & \text{val}_{22} & \cdots & \text{val}_{2C}\\ \vdots & \vdots & \ddots & \vdots\\ \text{val}_{m1} & \text{val}_{m2} & \cdots & \text{val}_{mC}\end{bmatrix}, \tag{9}$$

where val_{ij} is the validation accuracy of the ith CNN model for the jth fault type and $\mathbf{W}_{\text{val}j}$ (j = 1, 2, …, C) is the weight vector containing validation accuracies of all fault types.

The second information fusion method has the same testing process and probability matrix on the testing dataset as the first method. The final probability for all fault types is given by

$$\begin{aligned}\mathbf{P}&=\begin{bmatrix}P_1 & P_2 & \cdots & P_C\end{bmatrix}^{\mathrm{T}}=\mathbf{P}_{\text{test}}\otimes\mathbf{W}_{\text{val}}^{\mathrm{T}}\\ &=\begin{bmatrix}p_1(y=1|x_{\text{test}})\text{val}_{11}+\cdots+p_m(y=1|x_{\text{test}})\text{val}_{m1}\\ p_1(y=2|x_{\text{test}})\text{val}_{12}+\cdots+p_m(y=2|x_{\text{test}})\text{val}_{m2}\\ \vdots\\ p_1(y=C|x_{\text{test}})\text{val}_{1C}+\cdots+p_m(y=C|x_{\text{test}})\text{val}_{mC}\end{bmatrix},\end{aligned} \tag{10}$$

where the operator $\otimes$ is defined as

$$\begin{aligned}&\mathbf{A}_{m\times n}\otimes\mathbf{B}_{m\times n}\\ &=\begin{bmatrix}\text{A}_{11} & \text{A}_{12} & \cdots & \text{A}_{1n}\\ \text{A}_{21} & \text{A}_{22} & \cdots & \text{A}_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \text{A}_{m1} & \text{A}_{m2} & \cdots & \text{A}_{mn}\end{bmatrix}\otimes\begin{bmatrix}\text{B}_{11} & \text{B}_{12} & \cdots & \text{B}_{1n}\\ \text{B}_{21} & \text{B}_{22} & \cdots & \text{B}_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \text{B}_{m1} & \text{B}_{m2} & \cdots & \text{B}_{mn}\end{bmatrix}\\ &=\begin{bmatrix}\sum_{j=1}^{n}\text{A}_{1j}\text{B}_{1j} & \sum_{j=1}^{n}\text{A}_{2j}\text{B}_{2j} & \cdots & \sum_{j=1}^{n}\text{A}_{mj}\text{B}_{mj}\end{bmatrix}^{\mathrm{T}}.\end{aligned} \tag{11}$$

Thus, the proposed DF strategies make a final decision on the fault classification based on two sources of information, i.e., validation accuracies on the validation data and preliminary probability matrix on the testing dataset. The final probability matrix for the fault classification is adaptively adjusted by multiplying a weight matrix consisting of the validation accuracies. Figures 2 and 3 show that the main difference between the two information fusion strategies lies in the weight matrix calculated from the validation accuracies of all CNN models. The first method considers only the overall validation accuracy of each CNN model, and the CNN model with higher overall validation accuracy means a greater contribution to the final decision. In contrast, the second method further considers detailed information about the classification performance of each fault type when calculating the weight matrix. As a result, the latter method carries richer fused information in the weight matrix than the former.

4 Experiment and results

4.1 Experimental setup

Experimental data were acquired from an actual axial piston pump with nine pistons to evaluate the effectiveness of the proposed method. Figure 4 shows the experimental setup in the laboratory environment. The main specifications of the tested pump were as follows: theoretical displacement of 1.3 cm^3/r, rated rotational speed of 10000 r/min, rated outlet pressure of 21 MPa, rated inlet pressure of 0.3 MPa, and working medium of No. 15 aviation grade hydraulic oil. The tested pump was driven by an electric motor and operated under rated conditions. The pump load was simulated by a pressure relief valve. During the pump operation, the oil

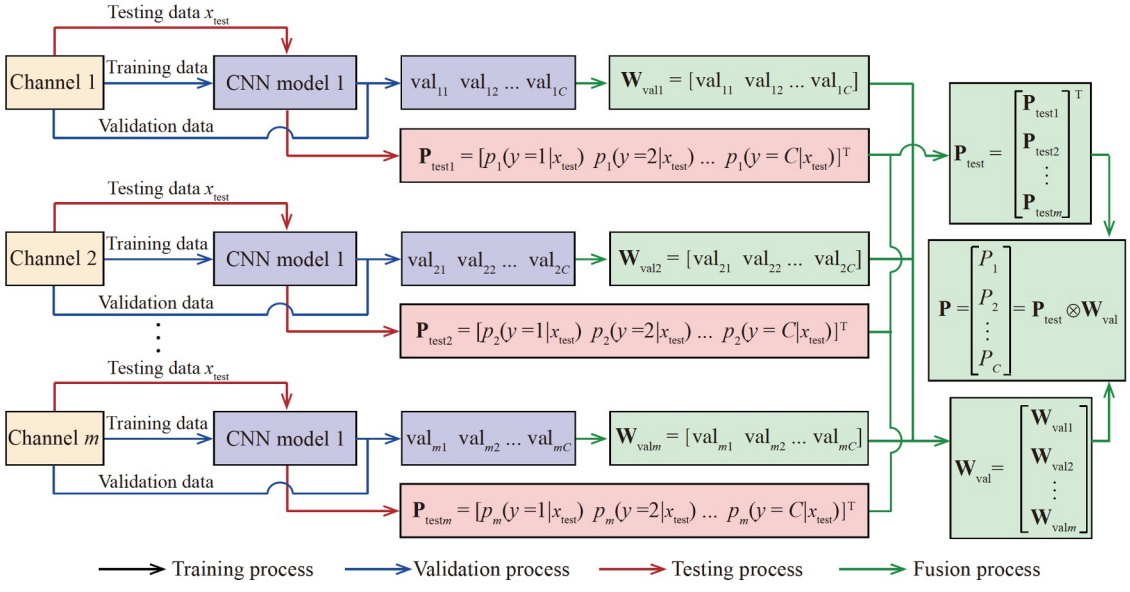

Figure 3 (Color online) Flowchart of the second decision-level information fusion method.

temperature of the inlet port was kept at 60°C ± 2°C to minimize its effects on pump performance. Additionally, a tri-axial accelerometer was mounted on the pump end cover to collect vibration signals at a sampling frequency of 10240 Hz. Note that the vibration signals in three orthogonal directions were simultaneously recorded for the information fusion at the decision level.

The experiments were conducted on the slipper pair of the pump under different health states: normal, slight leakage, medium leakage, and severe leakage (Table 1). The pump leakage levels were obtained by adjusting the gap height between the slippers and swash plate to save testing costs and ensure experimental safety. According to the law of Poiseuille flow, the leakage flow from the slipper pair increases with the gap height between the slippers and swash plate. An increase in gap height increases the discharge pressure fluctuation [42] and the slipper impact on the swash plate. As a result, high-frequency fluid and structure borne vibration was transmitted to the pump housing and end cover. Therefore, the leakage fault could be reflected in the vibration signals of the pump end cover. Figure 5 shows an example of the vibration signals under different leakage levels.

4.2 Data preprocessing

The 1D raw vibration data in each channel are converted into 2D spectrograms (Figure 6) because 2D images carry more powerful features that can be extracted by CNN. The 1D raw time series data collected from each channel are split into frames with equal lengths of 256 points and converted into grayscale spectrograms using STFT.

The discrete-time form of STFT is

$$\mathrm{STFT}\{x[n]\} = \sum_{m=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n}, \tag{12}$$

where $x[n]$ is the vibration signal, and $w[n]$ is the window function. In this study, the Hanning window is selected as the window function expressed as follows:

$$w[n] = \begin{cases} 0.5\left[1-\cos\left(\dfrac{2\pi n}{N-1}\right)\right], & 0 \le n \le N-1, \\ 0, & \text{otherwise}, \end{cases} \tag{13}$$

where the window size N is set to 128.

Figure 7 shows examples of the STFT-based spectrograms of different leakage levels converted by three channels of vibration data. It is found that the spectrogram energy is dominated at the frequencies of 1500, 3000, and 4500 Hz. The piston pass frequency of the tested pump is calculated as $Zn/60$. Here, Z is the piston number, and n is the rotational speed. For a nine-piston axial piston pump operating at a rotational speed of 10000 r/min, its piston pass frequency is 1500 Hz. The second-order and third-order harmonic frequencies are 3000 and 4500 Hz, respectively. For the same level of leakage, the spectrograms of individual data channels differ at dominating frequencies and regions between them. In addition, the spectrograms of the same data channel also exhibit different energy distributions for various leakage levels. Intuitively, it is easier to distinguish normal and slight leakage levels than medium and severe leakage levels.

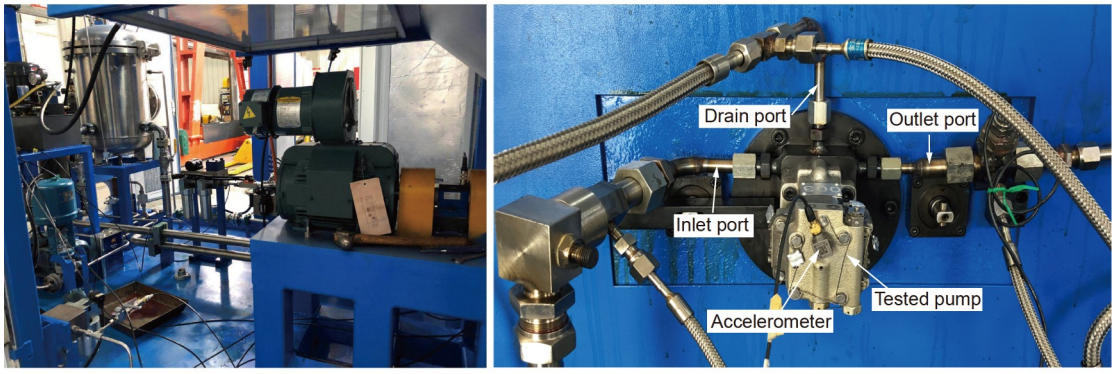

Figure 4 (Color online) Test rig of the axial piston pump.

Table 1 Pump health states in the dataset

Leakage level	Increased gap height (mm)	Leakage flow (L/min)	Training sample	Testing sample
Normal	0	0.40	192	48
Slight	0.05	0.57	192	48
Medium	0.15	0.65	192	48
Severe	0.20	1.10	192	48

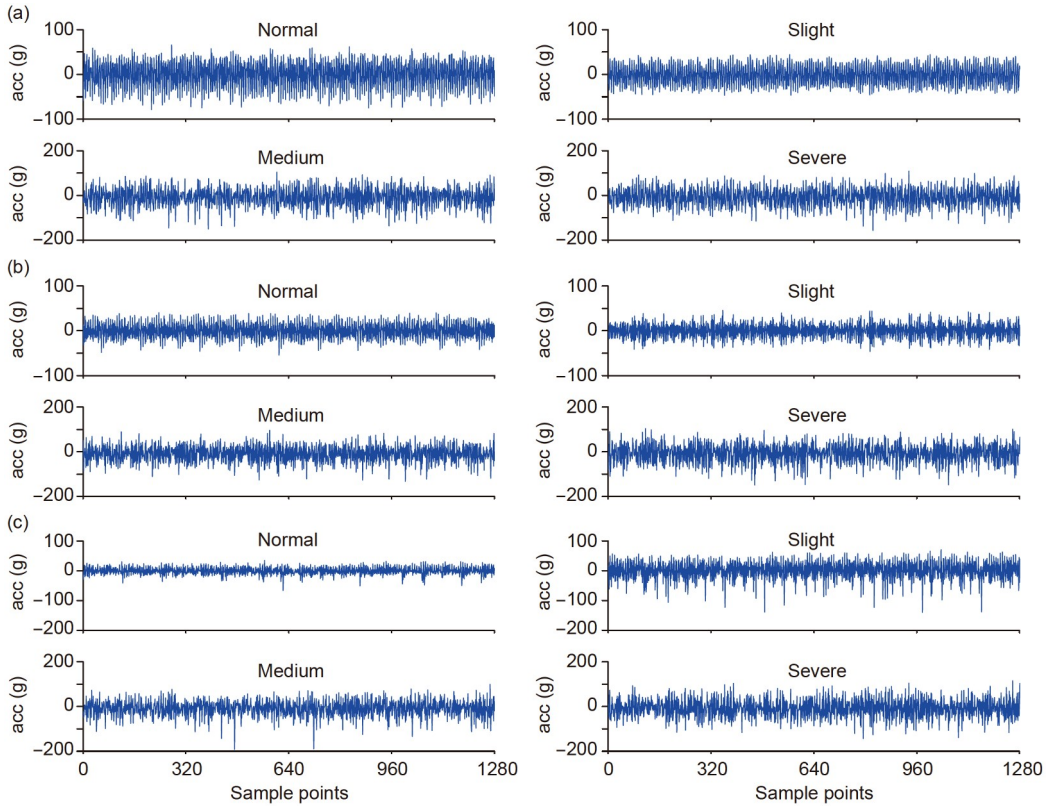

Figure 5 (Color online) Raw vibration signals in three orthogonal directions under different pump health states. (a) Vibration signals of channel 1; (b) vibration signals of channel 2; (c) vibration signals of channel 3.

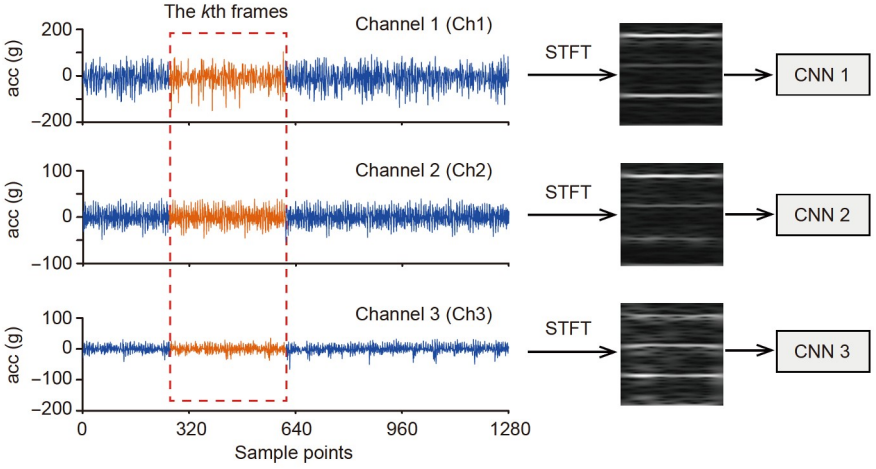

Figure 6 (Color online) Illustration of the data preprocessing.

4.3 Experimental results and discussion

Each channel of raw 1D vibration data is transformed into spectrograms to train a corresponding CNN model. Then, three trained CNN models are obtained with validation accuracy rates for each fault type. The validation accuracy rates form the weight matrix as described in Sect. 3. Each trained CNN model predicts a preliminary probability distribution when it receives new vibration data. Finally, the final probability distribution is obtained by summing up all channels of weighted probability through information fusion at the decision level. The maximum weighted probability represents the predicted pump leakage level.

Figure 8 compares the average accuracy and standard deviation of five trials between the same CNN models with and without DF. The proposed fusion strategy has the highest average accuracy of 99.2% and the lowest standard deviation of 0.23%. However, both average accuracy and standard

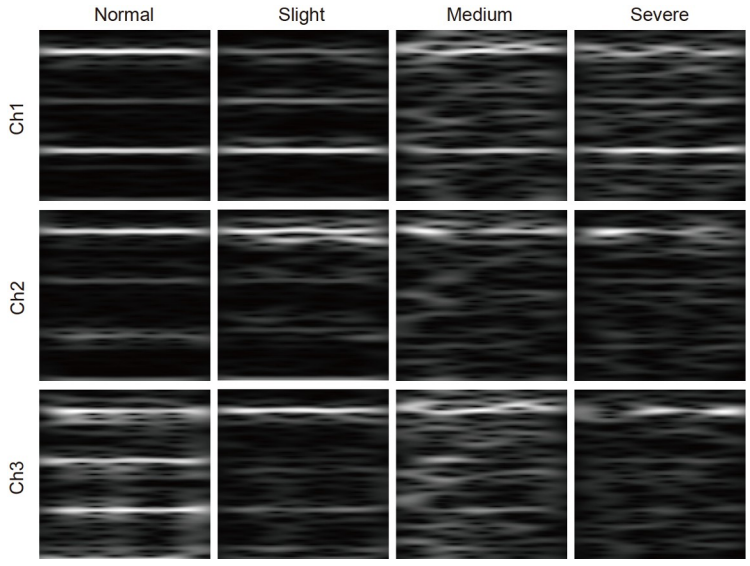

Figure 7 Comparison of grayscale spectrograms for different leakage levels.

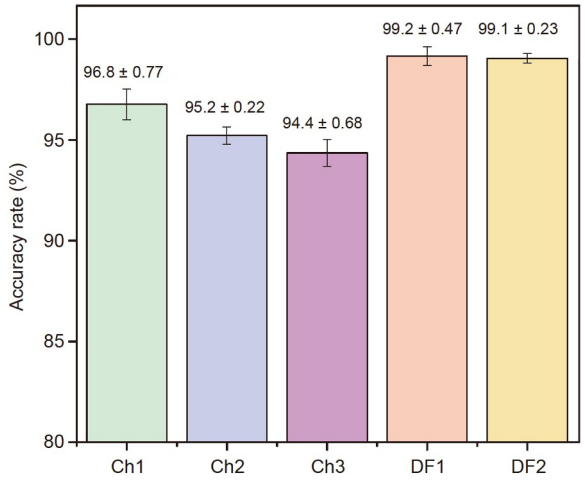

Figure 8 (Color online) Comparison of diagnostic accuracy of the same convolutional neural network model with and without decision-level fusion.

deviation differ from each other between individual single-channel data. This indicates that the vibration signals are differentially sensitive to the pump leakage in different directions, and single-channel vibration data are insufficient to capture the representative features effectively. In contrast, information fusion at the decision level utilizes the advantage of all data channels and has a more accurate fault diagnosis. It is worth noting that the proposed two DF methods have comparable diagnosis performance despite different weight matrices. A possible explanation is that the two DF methods have almost identical final classification predictions, although their probability distributions are different for all leakage levels.

Figure 9 further compares the confusion matrix of diagnostic accuracy between CNN models with and without DF. All diagonal elements of the confusion matrix represent the classification accuracy rates, whereas the non-diagonal ones represent the error rates. Figure 9(a)–(c) show that the CNN model can recognize all normal and slight leakages even without information fusion. However, the medium and severe leakages are easily misclassified to each other with low class accuracy rates. Specifically, the CNN model with an input of single-channel data has class accuracies of 95.8%, 91.7%, and 77.1% for medium leakage and 91.7%, 89.6%, and 97.9% for severe leakage. The poor diagnosis performance for the medium and severe leakage classification results from similar grayscale spectrograms (Figure 7). In contrast, the same CNN model has high class accuracy rates for all leakage levels when it integrates the DF (Figure 9(d) and (e)). For instance, the first fusion method increases the class accuracy rate of medium leakage by 2.1%, 6.2%, and 20.8%, and the class accuracy rate of severe leakage by 6.2%, 8.3%, and 0%. Similarly, the second fusion method improves the class accuracy rate. Here, the increase in the accuracy rate ranges from 4.2% to 22.9% for medium leakage and from 0% to 8.3% for severe leakage.

Figure 10 highlights how the decision-level information fusion improves the diagnosis performance. It shows only probability distributions for true labels of medium and severe leakages since DF improves their class accuracies. Each inset in Figure 10 presents the probability distributions of the four classes before and after decision fusion. Here, the first three probability distributions represent the cases before decision fusion, and the last two represent the cases after decision fusion. Thus, the decision fusion enables the class probability distribution to be changed so that the true class has the highest probability value. For example, the CNN model with the first data channel misclassifies medium leakage to severe

Figure 9 (Color online) Confusion matrix of the diagnostic accuracy rate with and without decision-level fusion. (a) Ch1; (b) Ch2; (c) Ch3; (d) DF1; (e) DF2.

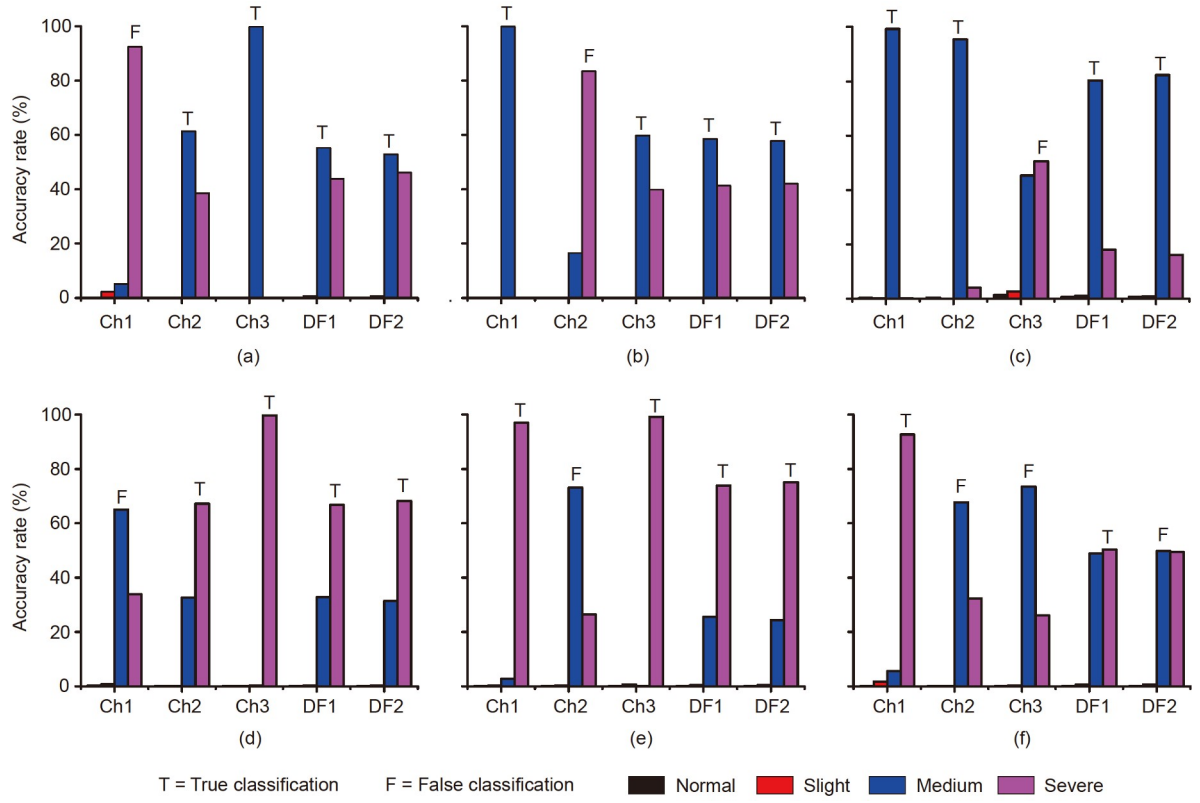

Figure 10 (Color online) Class probability distributions before and after decision fusion. (a)–(c) True label of medium leakage; (d)–(f) true label of severe leakage.

leakage (Figure 10(a)), whereas the decision fusion helps the CNN model to make a correct classification. Additionally, multi-channel data outperform the second or third data channel for the true label of medium leakage (Figure 10(b) and (c)). Similarly, the decision fusion improves the class accuracy of severe leakage by adjusting the class probability distribution (Figure 10(d)–(f)). However, there is a case that the second decision fusion method still misclassifies the severe leakage (Figure 10(f)). This is because two-thirds of sub-models have misclassifications before decision fusion.

5 Conclusions

This study presents a decision-level multi-sensor fusion method to improve the recognition accuracy of leakage levels in an axial piston pump. First, three channels of vibration data are fed separately to three identical CNN models to generate preliminary classification results. Then, the final prediction results are achieved by fusing these preliminary classification results at the decision level. The proposed decision fusion strategy considers different contributions of all sub-models by constructing a weight matrix that highlights the validation accuracy of each sub-model on historical data. Specifically, higher validation accuracy represents a greater contribution of the sub-model to final classification results. Experiments were conducted on four different levels of pump leakage. The experimental results show that the proposed decision fusion strategy can improve class accuracy by adjusting the class probability distribution. As a result, the classification accuracy rate is increased by about 2%, 4%, and 5% after decision fusion.

The proposed DF algorithm currently uses the same type of monitoring signals and identical CNN models for all data channels. The future study will extend the fusion algorithm by employing different monitoring signals or CNN models.

This work was supported by the National Key R&D Program of China (Grant No. 2020YFB2007202), the National Natural Science Foundation of China (Grant No. 52005323), the National Postdoctoral Program for Innovative Talents (Grant No. BX20200210), and the China Postdoctoral Science Foundation (Grant No. 2019M660086).

1 Wen L, Li X, Gao L, et al. A new convolutional neural network-based data-driven fault diagnosis method. IEEE Trans Ind Electron, 2018, 65: 5990–5998
2 Lan Y, Hu J, Huang J, et al. Fault diagnosis on slipper abrasion of axial piston pump based on extreme learning machine. Measurement, 2018, 124: 378–385
3 Manikandan S, Duraivelu K. Fault diagnosis of various rotating equipment using machine learning approaches—A review. Proc Institution Mech Engineers Part E-J Process Mech Eng, 2021, 235: 629–642
4 Du J, Wang S, Zhang H. Layered clustering multi-fault diagnosis for hydraulic piston pump. Mech Syst Signal Processing, 2013, 36: 487–504
5 Xia S, Zhang J, Ye S, et al. A spare support vector machine based fault detection strategy on key lubricating interfaces of axial piston pumps. IEEE Access, 2019, 7: 178177
6 Tang H, Fu Z, Huang Y. A fault diagnosis method for loose slipper failure of piston pump in construction machinery under changing load. Appl Acoustics, 2021, 172: 107634
7 Wang Y, Zhu Y, Wang Q, et al. Effective component extraction for hydraulic pump pressure signal based on fast empirical mode decomposition and relative entropy. AIP Adv, 2020, 10: 075103
8 Jiang W, Li Z, Li J, et al. Study on a fault identification method of the hydraulic pump based on a combination of voiceprint characteristics and extreme learning machine. Processes, 2019, 7: 894
9 Gao Q, Tang H S, Xiang J W, et al. A multi-sensor fault detection strategy for axial piston pump using the Walsh transform method. Int J Distrib Sens Networks, 2018, 14: 1550147718772531
10 Muralidharan V, Sugumaran V. Feature extraction using wavelets and classification through decision tree algorithm for fault diagnosis of mono-block centrifugal pump. Measurement, 2013, 46: 353–359
11 Muralidharan V, Sugumaran V. A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. Appl Soft Computing, 2012, 12: 2023–2029
12 Huang H R, Li K, Su W S, et al. An improved empirical wavelet transform method for rolling bearing fault diagnosis. Sci China Tech Sci, 2020, 63: 2231–2240
13 Lecun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521: 436–444
14 Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Process Mag, 2012, 29: 82–97
15 Tang S, Yuan S, Zhu Y. Convolutional neural network in intelligent fault diagnosis toward rotatory machinery. IEEE Access, 2020, 8: 86510–86519
16 Islam M M M, Kim J M. Automated bearing fault diagnosis scheme using 2D representation of wavelet packet transform and deep convolutional neural network. Comput Industry, 2019, 106: 142–153
17 Ince T, Kiranyaz S, Eren L, et al. Real-Time motor fault detection by 1-D convolutional neural networks. IEEE Trans Ind Electron, 2016, 63: 7067–7075
18 Liang P, Deng C, Wu J, et al. Compound fault diagnosis of gearboxes via multi-label convolutional neural network and wavelet transform. Comput Industry, 2019, 113: 103132
19 Yan J, Zhu H, Yang X, et al. Research on fault diagnosis of hydraulic pump using convolutional neural network. J Vibroeng, 2016, 18: 5141–5152
20 Tang S, Yuan S, Zhu Y, et al. An integrated deep learning method towards fault diagnosis of hydraulic axial piston pump. Sensors, 2020, 20: 6576
21 Wang S, Xiang J. A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps. Soft Comput, 2020, 24: 2983–2997
22 Sun S W, Zhang S, Jiang W L, et al. Study on the health condition monitoring method of hydraulic pump based on convolutional neural network. In: Proceedings of the 12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA). Phuket, 2020. 149–153
23 Xu G, Liu M, Jiang Z, et al. Online fault diagnosis method based on transfer convolutional neural networks. IEEE Trans Instrum Meas, 2020, 69: 509–520
24 Kumar A, Gandhi C P, Zhou Y, et al. Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images. Appl Acoustics, 2020, 167: 107399
25 Lázaro J G M, Pinilla C B. A methodology for detection of wear in hydraulic axial piston pumps. Int J Interact Des Manuf, 2020, 14: 1103–1119
26 Xia M, Li T, Xu L, et al. Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. IEEE/ASME

Trans Mechatron, 2018, 23: 101–110
27 Chen Z, Li W. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. IEEE Trans Instrum Meas, 2017, 66: 1693–1702
28 Hoang D T, Kang H J. A motor current signal-based bearing fault diagnosis using deep learning and information fusion. IEEE Trans Instrum Meas, 2020, 69: 3325–3333
29 Liu M K, Tran M Q, Weng P Y. Fusion of vibration and current signatures for the fault diagnosis of induction machines. Shock Vib, 2019, 2019: 7176482
30 Chen H, Hu N, Cheng Z, et al. A deep convolutional neural network based fusion method of two-direction vibration signal data for health state identification of planetary gearboxes. Measurement, 2019, 146: 268–278
31 Azamfar M, Singh J, Bravo-Imaz I, et al. Multisensor data fusion for gearbox fault diagnosis using 2-D convolutional neural network and motor current signature analysis. Mech Syst Signal Processing, 2020, 144: 106861
32 Di Z Y, Shao H D, Xiang J W. Ensemble deep transfer learning driven by multisensor signals for the fault diagnosis of bevel-gear cross-operation conditions. Sci China Tech Sci, 2021, 64: 481–492
33 Yu H, Li H, Li Y. Vibration signal fusion using improved empirical wavelet transform and variance contribution rate for weak fault detection of hydraulic pumps. ISA Trans, 2020, 107: 385–401
34 Wang H, Li S, Song L, et al. A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals. Comput Industry, 2019, 105: 182–190
35 Yu H, Li H, Li Y, et al. A novel improved full vector spectrum algorithm and its application in multi-sensor data fusion for hydraulic pumps. Measurement, 2019, 133: 145–161
36 Zeng F, Li Z, Zhou Z, et al. Fault classification decision fusion system based on combination weights and an improved voting method. Processes, 2019, 7: 783
37 Chen B, Varshney P K. A Bayesian sampling approach to decision fusion using hierarchical models. IEEE Trans Signal Process, 2002, 50: 1809–1818
38 Li S, Liu G, Tang X, et al. An ensemble deep convolutional neural network model with improved D-S evidence fusion for bearing fault diagnosis. Sensors, 2017, 17: 1729
39 Dahl G E, Sainath T N, Hinton G E. Improving deep neural networks for LVCSR using rectified linear units and dropout. In: Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, 2013. 8609–8613
40 Vani S, Madhusudhana Rao T V. An experimental approach towards the performance assessment of various optimizers on convolutional neural network. In: Proceedings of the 3rd International Conference on Trends in Electronics and Informatics (ICOEI). Tirunelveli, 2019. 331–336
41 Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proc IEEE, 1998, 86: 2278–2324
42 Huang J, Yan Z, Quan L, et al. Characteristics of delivery pressure in the axial piston pump with combination of variable displacement and variable speed. Proc Institution Mech Engineers Part I-J Syst Control Eng, 2015, 229: 599–613