A ROLLING BEARING FAULT DIAGNOSIS METHOD USING MULTI-SENSOR DATA AND PERIODIC SAMPLING

Jianbo Zheng[†a], *Chao Yang*[†], *Fangrong Zheng*[‡] *and Bin Jiang*[†a*]

[†]College of Computer Science and Electronic Engineering, Hunan University, China;
[a]Key Laboratory for Embedded and Network Computing of Hunan Province;
[‡]School of Computational Science and Electronics, Hunan Institute of Engineering, China;
[*]jiangbin@hnu.edu.cn.

2022 IEEE International Conference on Multimedia and Expo (ICME) | 978-1-6654-8563-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICME52920.2022.9859658

ABSTRACT

In recent years, bearing fault diagnosis based on deep learning has gradually become the mainstream. However, the existing studies still have some defects, such as unreasonable sampling and incomplete utilization of bearing data, limiting the further improvement of the performance of the fault diagnosis model. This paper proposes a fault diagnosis method using multi-sensor data and periodic sampling to solve the problems above. First, the vibration data of different bearing positions are fused into multi-channel fusion data to improve the defect of insufficient data utilization. Second, based on the sampling length and sampling stride, periodic sampling is carried out for the fusion data to solve the problem of unreasonable sampling. Third, the traditional convolutional neural network is adjusted to extract more detailed fault features and obtain the best recognition effect. Finally, the experimental results verify the effectiveness of the proposed method.

Index Terms— Fault diagnosis, Convolutional neural networks, Data fusion, Periodic sampling

1. INTRODUCTION

With the gradual implementation of intelligent manufacturing and the industrial Internet of things, lots of low-cost acceleration sensors and acoustic sensors are gradually deployed in the field of modern industrial manufacturing to obtain a large number of operating state parameters of industrial equipment at different positions. Acceleration and acoustic data are waveform data, and the features of waveform data at different positions (multi-sensor data) are quite different. The explosive growth of multi-sensor waveform data not only opens up new opportunities for fault diagnosis of industrial equipment, but also brings many application challenges such as data idle, data misuse, and so on. According to the literature, bearing failures are the most common type of equipment failure, accounting for 30 to 40 percent of total failures and as high as 45 to 55 percent for rotating machinery [1]. And rotating machinery composed of stator, rotor, shaft, and bearing plays an essential role in the industrial system. In the past decades, fault diagnosis of bearing based on waveform data, especially acceleration data, has been the research frontier of intelligent fault diagnosis field.

According to the failure components, the bearing fault can be divided into four categories: outer ring (OR), inner ring (IR), roller, and cage. These faults are caused by changes in the internal spatial structure of the bearings due to long-term wear and tear. And the faulty bearing will vibrate violently during operation, resulting in the abnormal state of the whole equipment. This situation not only reduces the performance of the equipment and aggravates the degradation of the entire kit, but causes severe economic losses and even endangers the safety of personnel in extreme cases. High-precision bearing fault diagnosis based on industrial data can effectively reduce financial loss, protect the security of employees, enhance the stability of equipment and extend the overall service life of the equipment. Moreover, it's well known that sound (acoustic data) comes from vibrations (acceleration data). Therefore, it is necessary to diagnose the fault state of the bearing according to acceleration data.

Based on acceleration data, current approaches for fault diagnosis can be divided into three categories: signal processing (SP) methods, traditional machine learning (TML) methods, and deep learning (DL) methods [2], [3]. Without sampling, SP methods directly extract and analyze fault features in time-domain, frequency-domain, and time-frequency-domain from the original data to diagnose the fault. SP methods contain wavelet transformation (WT), fast Fourier transform (FFT), wavelet packet transform (WPT), empirical mode decomposition (EMD), etc. SP methods need a lot of domain knowledge, and it is difficult to extract the hidden relationship between abnormal signals. Therefore, many experts and scholars have applied some TML methods to diagnosis bearing faults from vibration data.

This work was supported by National Natural Science Foundation of China under Grant No.62072169, National Key R&D Program of China under Grant No.2020YFB1713003 and The Key R&D Program of Changsha under Grant No.KQ2004005.

To diagnosis faults, TML methods consist of four steps. First, the raw data needs to be sampled to obtain a sample set containing features and labels. Second, the feature representation of bearing fault needs to be designed manually based on domain prior knowledge. Third, the hand-feature extraction is carried out. Finally, the extracted features are input into classifiers such as support vector machines (SVM) or naive Bayes (NB) to complete the fault classification of bearing data. The use of TML is helpful for the early diagnosis of bearing faults. However, TML methods still have the shortcoming that hand-features are challenging to contain most of the high discrimination features, which causes the fault diagnosis algorithms based on deep learning to increase rapidly in recent years [4].

DL methods have only two steps. The first step is to sample the vibration data like TML. The second step is to classify the samples by using the high-dimensional representation ability of deep neural network, which integrates feature design, extraction and classification. DL methods include Auto-Encoder, Deep Belief Network, Deep Boltzmann Machines, Convolutional Neural Network (CNN), Recurrent Neural Network, etc [5] [6] [7]. The main advantages of DL methods are as follows: (1) The DL methods realize the end-to-end solution of fault diagnosis (mapping the original data to the output), thus alleviating the loss of abnormal information in the original data caused by multiple diagnosis steps. (2) DL can automatically extract the abnormal high-dimensional features from original data and its labels without the hand-feature design. (3) The overall framework of DL is straightforward, including input, neural network and output, which has good universality and scalability [8]. (4) A deep neural network trained in an environment with strong expressive ability and high performance can be transferred to other similar environments.

Although DL methods have achieved a series of research achievements, it still suffers from the following shortcomings. (1) The sampling method of samples is highly random. Generally, a simple random sampling strategy or sliding window strategy with different sample lengths are used to sample the original data, and the diagnostic model is difficult to efficiently, reasonably, and completely extract abnormal features from the data [9], [10], [11]. (2) Most of them only use the original data of one position or fuse the vibration data of different positions into a matrix to diagnose faults, ignoring the difference of fault features between vibration data of different positions, so that the model can not thoroughly learn the fault features with high discrimination [12], [13], [14]. (3) Most networks are complex and have too many network layers, resulting in a waste of computing resources. (4) There are few open-source codes for bearing fault diagnosis based on DL, limiting the further development and expansion of the research direction of bearing fault diagnosis based on DL [15].

To overcome the aforementioned weakness, this paper proposes a fault diagnosis method using multi-sensor data and periodic sampling (MDPS). The main contributions of this paper are summarized as follows:

(1) The multi-channel fusion of multi-sensor data is realized, which makes the multi-channel fusion data have richer fault features.

(2) Based on sampling length and sampling stride, we reasonably model the sampling process of samples and propose a sampling method based on periodic sampling.

(3) The traditional convolution neural network structure is adjusted to realize more comprehensive feature extraction of multi-channel fusion data, and we will release our code[1] after the publication of this paper.

2. PROPOSED METHOD

2.1. Architecture of the proposed MDPS model

In this part, the proposed MDPS framework shown in Fig.1 is elaborated.

Firstly, the vibration data of sensors at different positions are obtained from the environment. In order to make full use of hidden abnormal features, the multi-sensor data is fused according to the channel fusion scheme, and the fused data X is obtained. Then, the fusion data X are periodically sampled based on the sampling length and stride to obtain the sample set, including the feature x and label y. Finally, the feature of samples x is input into the CNN model, and the label y of samples is used to supervise the learning process of the CNN model or test the performance of trained CNN.

In the CNN model, the input (size: $N \times m$) is first convolved through the filter K_f, and the ReLU operation is performed on the convolution results to extract the fusion features of channel fusion data and form a feature map $M1$. What is more, batch-normalization is used to prevent the problem of gradient disappearance and over-fitting, and the convergence speed of network model is accelerated. Then, a max-pooling layer is followed to sub-sample the feature maps $M1$. Repeat the above steps to extract more high-dimensional abnormal features and form the feature map $M2$. Next, a fully connected layer is used to model the relationship between features and results, and a softmax layer is used to classify the output of the fully connected layer.

2.2. Channel fusion of multi-sensor data

First, the bearing has K operating states (normal state and several fault states). Under a certain load, the acceleration data of N sampling points at m positions of rotating machinery are described as the following formula.

$$Data = \{(X_i, y_i)\}, i \in (1, K) \tag{1}$$

$$\begin{cases} X_i = \left(X_i^1, X_i^2, \cdots, X_i^j\right), j \in (1, m) \\ y^i \in \{label1, label2, \cdots, labelK\} \end{cases} \tag{2}$$

[1]https://github.com/IWantBe/MDPS

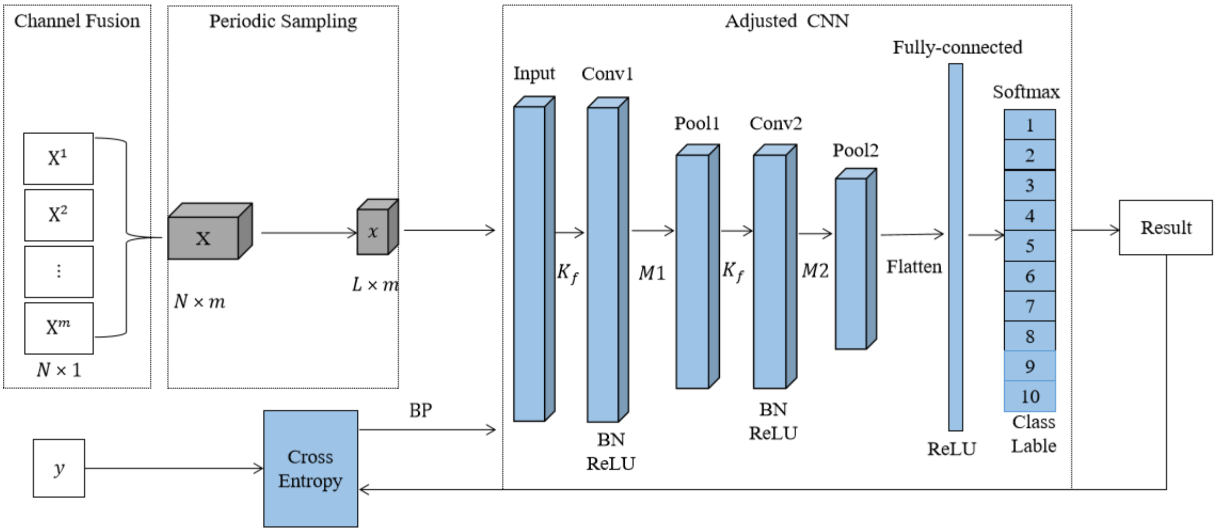

Fig. 1. Frame diagram of MDPS. It consists of multi-sensor data fusion, periodic sampling of fused data, and a multi-channel convolutional neural network.

$$X_i^j = \left[X_{i_1}^j, X_{i_2}^j, \cdots, X_{i_N}^j\right] \tag{3}$$

where $Data$ contains the vibration data and status labels of the bearing in K operating states. X_i is the monitoring data under state i, X_i^j is the acceleration data with the total number of sampling points N recorded at the position j, $X_{i_N}^j$ represents the Nth sampling point of the acceleration monitoring data at position j, and y^i represents the status label of the bearing in state i.

Second,inspired by the idea of using RGB three color channels to describe image features in image classification task comprehensively, we intuitively believe that bearing fault features can also be represented by m position channels. Therefore, the monitoring data at different positions X_i^1, ..., X_i^m (size : N) is modeled as m single-channel data (size : $N \times 1$), and then m single-channel data are merged in channel dimension to obtain the multi-channel fusion data (size : $N \times m$). ultimately, the fusion features of multi-channel vibration data are extracted by one-dimensional convolution with m-channel. The schematic diagram of multi-sensor data fusion in channel dimension is shown in Fig.2.

Third, the definition of the fusion data after channel fusion is shown as the following formula.

$$Data_{fusion} = \{(F_i, y_i)\}, i \in (1, K) \tag{4}$$

$$F_i = \left[F_i^1, F_i^2, \cdots, F_i^N\right] \tag{5}$$

$$F_i^l = \left(X_{i_l}^1, X_{i_l}^2, \cdots, X_{i_l}^m\right), l \in (1, m) \tag{6}$$

where $Data_{fusion}$ is the fusion data after performing channel fusion, which contains the fusion vibration data and its status labels. F_i is the fused vibration data with m channels and N sampling points.

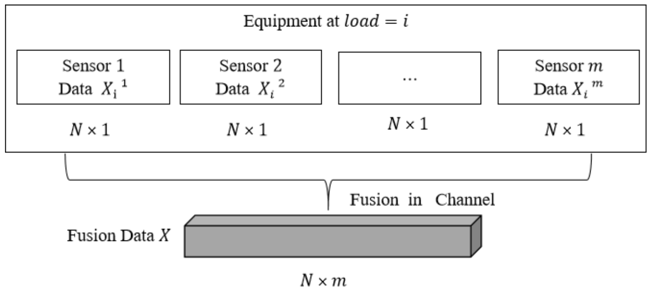

Fig. 2. Channel fusion of multi-sensor data

2.3. Periodic sampling of multi-channel fusion data

The revolutions per minute (RPM) of the bearing are variable under different loads. According to RPM, we can calculate the bearing's rotation period T (revolutions per second, RPS) with formula (7). Then, according to the sampling frequency f and rotation period T, we can calculate the number of sampling point n in a rotation period with formula (8). Since the vibration signal of the bearing is generated by rotation, when the bearing is abnormal, the vibration signal generated by rotation will fluctuate abnormally. Therefore, we reasonably assume that the abnormal signal of the bearing also has periodic characteristics, and the period is consistent with the rotation period T. Based on the above assumption, the length of each sample is set as the sampling points collected during one or more rotation cycles of the bearing, and the length L of each sample is defined with formula (9).

$$T = RPM/60 \tag{7}$$

$$n = f/T \tag{8}$$

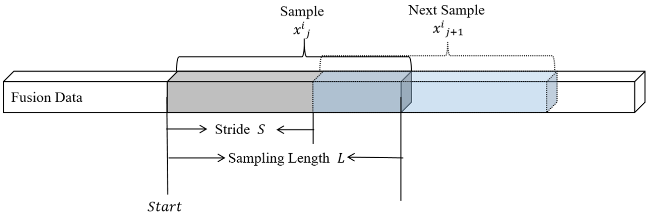

Fig. 3. Sampling process of multi-channel fusion data

Table 1. Table of fault data used in this paper

Load	RPM	Diameter	File(.mat)			
			IR	Ball	OR	Normal
0	1797	7 mils	105	118	130	97
		14mils	169	185	197	
		21mils	209	222	234	
1	1772	7 mils	106	119	131	98
		14mils	170	186	198	
		21mils	210	223	235	
2	1750	7 mils	107	120	132	99
		14mils	171	187	199	
		21mils	211	224	236	
3	1730	7 mils	108	121	133	100
		14mils	172	188	200	
		21mils	212	225	237	

$$L = \alpha \times n, \alpha \in N_+ \tag{9}$$

Considering the total number of sampling points in the dataset is small and fixed, and many parameters need to be trained in deep learning model, it is accessible to over-fitting the model if the training samples are insufficient. Therefore, in fault diagnosis, it is necessary to use data enhancement technology to alleviate the problem of insufficient training samples, and the sliding window is the most commonly used method. So, from the original vibration data, the samples are sampled according to the sample length L and sliding stride S. Finally, the formal definition of the sample set is defined as follows.

$$sample\_set = \{x^i, y_i\}, i \in (1, k) \tag{10}$$

$$x^i = \{x_1^i, x_2^i, \cdots\} \tag{11}$$

$$\left\{ \begin{array}{c} x_j^i = \left[F_i^{start}, \cdots, F_i^{start+L}\right] \\ x_{j+1}^i = \left[F_i^{start+S}, \cdots, F_i^{(start+L)+S}\right] \end{array} \right. \tag{12}$$

where $start$ is initialized to 0, stride S can be adjusted according to actual needs. And the sampling process is shown in Fig.3.

3. EXPERIMENTAL RESULTS

3.1. Dataset description

In order to verify the effectiveness of our MDPS method, several experiments were carried out based on the bearing dataset of the Case Western Reserve University (CWRU), which is used widely in bearing fault diagnosis. "12K drive end bearing fault data" is selected as the data studied in this paper (sampling frequency $f = 12000$), and after eliminating the incomplete or unavailable data, the fault diagnosis dataset as shown in Table 1 is obtained.

As shown from Table 1, The dataset contains the vibration data of the bearing under four different conditions marked by load 0, 1, 2, 3. Under any load, the vibration data includes normal operating state and three types of fault operating state, namely outer ring (OR), inner ring (IR) and roller. Each fault type has three levels of severity (the fault diameter is 7, 14 or 21 miles). In addition, each file (.mat) contains bearing vibration data at the drive end (DE) and fan end (FE). Then, according to the ratio of [0.7, 0.15, 0.15], the dataset is divided into training set, verification set, and test set.

3.2. Experimental setup

3.2.1. The setting of sample length and sampling stride

According to the sampling method in Part 2.3, the number of sampling points for a cycle of bearing rotation under loads 0, 1, 2, and 3 is calculated as $n_0 \approx 401$,$n_1 \approx 406$, $n_2 \approx 412$ and $n_3 \approx 416$ respectively. In order to make the number of sampling points n greater than the number of sampling points collected by one cycle of bearing rotation under all load conditions, this paper set $n = 420$. Within a specific limit, the longer the sampling length L of each sample (The greater the value of α), the better the fault diagnosis performance. However, considering the limited total number of sampling points in the dataset used in this paper. the number of rotation cycles α is set to 1 to obtain more training samples, and the length of each sample is $L = n = 420$.

Because our sampling method is sliding window sampling, another factor affecting the number of samples is the sampling stride S. For high-quality sampling, this paper define the degree of overlap between two samples (the repetition rate of sampling points) as the following formula.

$$repetition = 1 - S/L \tag{13}$$

If the stride S is small, more samples can be formed, but the $repetition$ between two adjacent samples is very high, which is not conducive to the training and learning of the model; On the contrary, the stride S is large and the $repetition$ of two adjacent samples is very low, but the number of samples that can be formed is small, which is challenging to help the model obtain the best recognition performance. Therefore, in this paper, in order to avoid the problem of high

Table 2. Average accuracy on load = 1 under different strides

Stride	Acc(%)	Stride	Acc(%)	Stride	Acc(%)
210	**99.7**	220	99.5	230	99.6
240	99.6	250	99.4	260	99.5
270	99.2	280	99.7	290	99.0
300	84.8	310	99.0	320	99.4
330	**99.7**	340	99.4	350	99.2
360	98.8	370	98.7	380	99.3
390	98.9	400	**84.4**	410	99.3
420	98.9	\	\	\	\

Table 3. Hyper-parameters table

Parameter	Value	Parameter	Value
Epoch	100	Filters	32
Batch_size	10	Kernel_size	20
Pool_size	2	Activation	Relu
Stride(Pool)	2	Stride(Conv)	1
Padding(Pool)	same	Padding(ConV)	same

$repetition$ but retain a promising recognition performance, we assume that stride S should be greater than or equal to 210 with the $repetition = 0.5$. In addition, under the conditions of $load = 1$ and $epoch = 80$, some experiments are carried out to study the effects of different strides on model performance. All experiments in this paper are carried out five times to reduce random interference, and the average accuracy of the model under different sampling stride S is finally obtained as shown in Table 2. It can be seen from Table 2 that the stride has a significant impact on the accuracy of the diagnostic model; When $S = 210$, the average accuracy is 99.7%, while $S = 400$, the accuracy is 84.4%. In order to get more samples and maintain a high accuracy, this paper set $S = 210$.

3.2.2. Parameter setting of the network model

In this part, the predefined hyper-parameters in CNN are introduced in detail, as shown in Table 3. And the deep learning framework tensorflow 2.6.0 is applied to build the model.

3.3. Result analysis

Because the existing methods have different sampling strategies, their sampling lengths vary greatly, and most of the existing work has not been open-source, so the classic 6-layer CNN architecture with accuracy similar to 12-layer network WDCNN is used as the baseline network architecture [16]. To demonstrate the advanced nature and effectiveness of our method, we set up the following six methods.

Table 4. Average accuracy of six methods under four loads

	Load 0	Load 1	Load 2	Load 3	Average
M0	70.5%	84.8%	87.2%	81.9%	81.9%
M1	99.6%	99.4%	**99.9**%	99.7%	99.7%
M2	98.0%	98.6%	99.0%	99.5%	98.8%
M3	99.5%	98.9%	99.2%	99.6%	99.3%
M4	99.5%	99.0%	99.6%	99.9%	99.5%
M5	**100**%	**99.6**%	99.7%	**100**%	**99.9**%

M0: SVM + single-position data (DE) + periodic sampling ($L = 420$, $S = 210$).

M1: CNN + single-position data + random sampling ($L = 2048$).

M2: CNN + single-position data + periodic sampling.

M3: CNN + two-position data (DE and FE) + periodic sampling.

M4: our network (Adjusted CNN) + single position data + periodic sampling.

M5: our network + two-position data + periodic sampling (MDPS).

After five experiments, the average accuracy of six methods under four loads is shown in Table 4. By analyzing the results in Table 4, the following conclusions can be drawn:

(1) The comparative experimental results of single-position and two-position show the effectiveness of the multi-sensor data fusion scheme: Based on M2, M3 uses the vibration data at two positions, and its performance is better than M2 under all loads. Moreover, the average accuracy of M3 is 99.3% and the average accuracy of M2 is 98.8%. It is confirmed that the vibration data at different positions hide different abnormal features, and the fusion of multi-sensor vibration data can effectively improve the diagnostic performance of the model.

(2) The comparative experimental results between CNN and our network structure show the effectiveness of our network structure: In the experiment using single-position data, M4 performed better than M2; and in the experiment using two-position data, M5 performed better than M3. This is because when the $stride_{convolutional} = 1$ and the $stride_{pooling} = 2$, the convolutional network can extract the abnormal features in the data constructed by periodic sampling more comprehensively.

(3) The comparative experimental results between random sampling and periodic sampling reveal the rationality of the periodic sampling scheme: M1 adopts random sampling with $L = 2048$, and the average accuracy is 99.7%. M2 to M5 adopt periodic sampling with $L = 420$, and the average accuracy of M5 is 99.9%. In addition, the main drawback of random sampling is high uncertainty, resulting in unstable performance of M1 (when load=3, the highest accuracy is 100%, and the lowest is 98.3% in five trials), while using periodic sampling has stable performance. So periodic sampling is more reasonable than random sampling.

To summarize, M5 has the best diagnostic performance in all comparative experiments, which shows that periodic sampling (more streamlined data), multi-sensor data fusion (more comprehensive abnormal feature), and the adjusted network structure of CNN (more detailed feature mining) can cooperate to help the model learn more comprehensive abnormal features. What's more, the average accuracy of M5 is 99.9%, which indicates that the samples obtained by periodic sampling contain almost all abnormal characteristics, and confirms the hypothesis that the abnormal features of the bearing have the same periodic characteristic as its rotation.

4. CONCLUSION

In this paper, the MDPS method is proposed for the bearing fault diagnosis task. MDPS performs multi-sensor data channel fusion, periodic sampling of fused data and fine extraction of fault features to obtain the best fault diagnosis performance. The experimental results under four different loads demonstrate the superiority of MDPS, and MDPS has a specific positive significance for the development of bearing fault diagnosis.

5. REFERENCES

[1] Arfat Siddique, GS Yadava, and Bhim Singh, "A review of stator fault monitoring techniques of induction motors," *IEEE transactions on energy conversion*, vol. 20, no. 1, pp. 106–114, 2005.

[2] Xiang Li, Wei Zhang, Qian Ding, and Jian-Qiao Sun, "Multi-layer domain adaptation method for rolling bearing fault diagnosis," *Signal Processing*, vol. 157, pp. 180–197, 2019.

[3] Yaguo Lei, Bin Yang, Xinwei Jiang, Feng Jia, Naipeng Li, and Asoke K Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mechanical Systems and Signal Processing*, vol. 138, pp. 106587, 2020.

[4] Syed Umar Amin, Ghulam Muhammad, Wadood Abdul, Mohamed Bencherif, and Mansour Alsulaiman, "Multi-cnn feature fusion for efficient eeg classification," in *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2020, pp. 1–6.

[5] Rui Zhao, Ruqiang Yan, Zhenghua Chen, Kezhi Mao, Peng Wang, and Robert X Gao, "Deep learning and its applications to machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 115, pp. 213–237, 2019.

[6] Dhiraj Neupane and Jongwon Seok, "Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review," *IEEE Access*, vol. 8, pp. 93155–93178, 2020.

[7] Wei Zhang, Chuanhao Li, Gaoliang Peng, Yuanhang Chen, and Zhujun Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mechanical Systems and Signal Processing*, vol. 100, pp. 439–453, 2018.

[8] Ye Yu, Qiang Jin, and Chang Wen Chen, "Ff-cmnet: A cnn-based model for fine-grained classification of car models based on feature fusion," in *2018 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2018, pp. 1–6.

[9] Weining Lu, Xueqian Wang, Chunchun Yang, and Tao Zhang, "A novel feature extraction method using deep neural network for rolling bearing fault diagnosis," in *The 27th Chinese Control and Decision Conference (2015 CCDC)*. IEEE, 2015, pp. 2427–2431.

[10] Min Xia, Teng Li, Lin Xu, Lizhi Liu, and Clarence W De Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME transactions on mechatronics*, vol. 23, no. 1, pp. 101–110, 2017.

[11] Zhiqiang Chen, Shengcai Deng, Xudong Chen, Chuan Li, René-Vinicio Sanchez, and Huafeng Qin, "Deep neural networks-based rolling bearing fault diagnosis," *Microelectronics Reliability*, vol. 75, pp. 327–333, 2017.

[12] Zhuang Zilong and Qin Wei, "Intelligent fault diagnosis of rolling bearing using one-dimensional multi-scale deep convolutional neural network based health state classification," in *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*. IEEE, 2018, pp. 1–6.

[13] Xiaoxi Ding and Qingbo 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.

[14] Panikos Heracleous, Pierre Badin, Gérard Bailly, and Norihiro Hagita, "Exploiting multimodal data fusion in robust speech recognition," in *2010 IEEE International Conference on Multimedia and Expo*. IEEE, 2010, pp. 568–572.

[15] Xiang Li, Wei Zhang, and Qian Ding, "Understanding and improving deep learning-based rolling bearing fault diagnosis with attention mechanism," *Signal processing*, vol. 161, pp. 136–154, 2019.

[16] Wei Zhang, Gaoliang Peng, Chuanhao Li, Yuanhang Chen, and Zhujun 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, pp. 425, 2017.