

# Frequency adaptive wavelet pyramid for noisy machinery fault diagnosis with multiple sensors

Aosheng Tian<sup>1</sup>, Ye Zhang<sup>1</sup>, Chao Ma<sup>1</sup>, Huiling Chen<sup>1</sup>, Shilin Zhou<sup>1</sup>, and Weidong Sheng<sup>1</sup>

<sup>1</sup>National University of Defense Technology College of Electronic Science and Technology

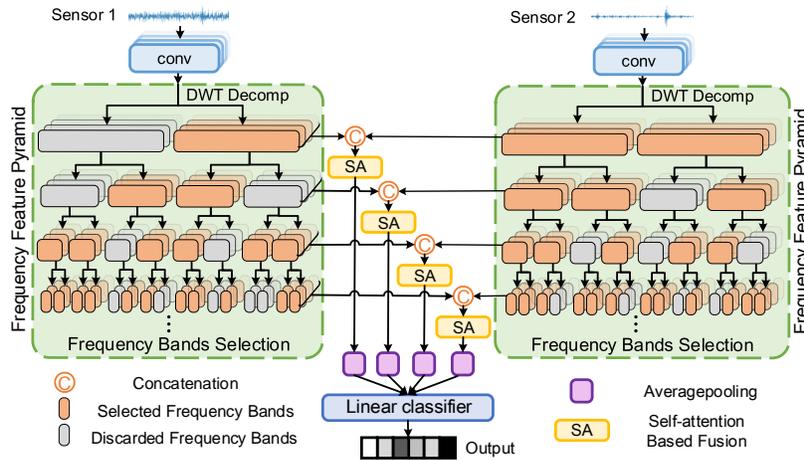
September 6, 2022

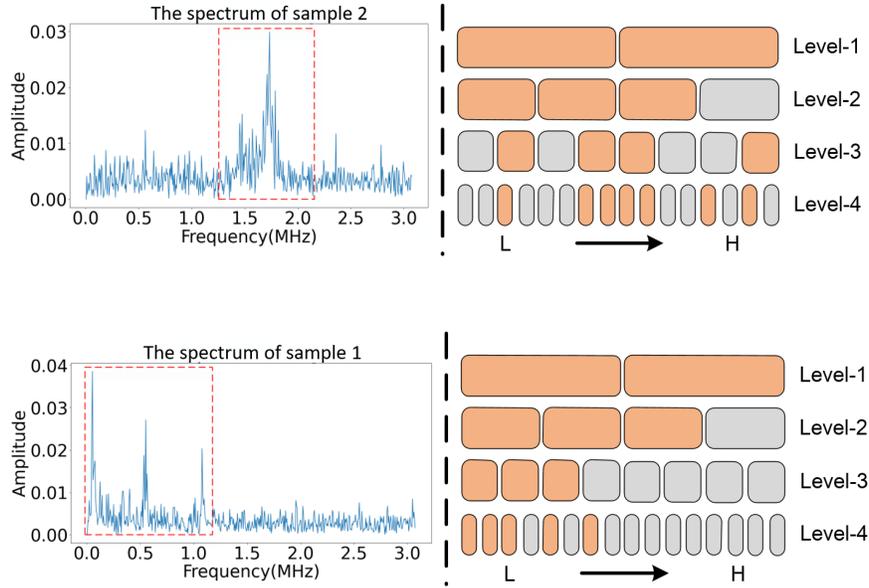
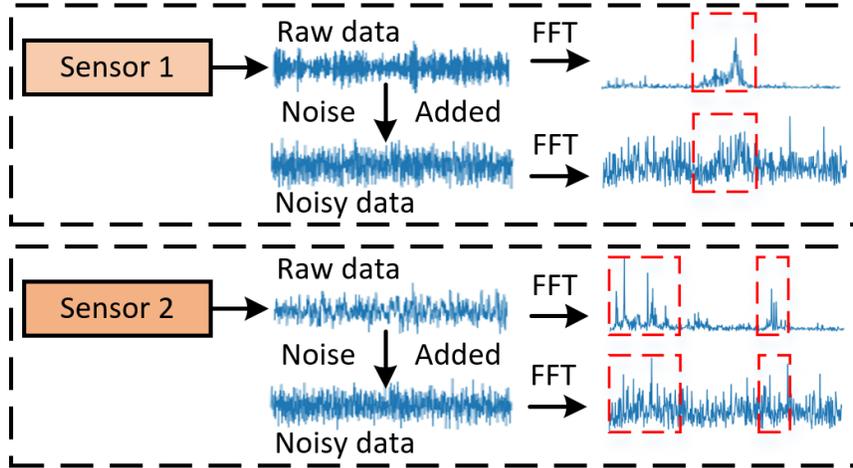
## Abstract

The fusion of multiple monitoring sensors is crucial to improve the accuracy and robustness of machinery fault diagnosis. However, existing fault diagnosis methods may underestimate the interference of noise in the multi-sensor fusion process, leading to unsatisfied performance. To handle this problem, this paper proposes a deep model based on the frequency adaptive wavelet pyramid. First, an adaptive frequency selection strategy is designed to prune the seriously polluted frequencies and only retain some key frequencies. Then, the self-attention mechanism is used to perform information fusion on the selected frequency bands of different sensors. Finally, a wavelet fusion pyramid is adopted by repeating the fusion process at multiple wavelet decomposition levels. In this way, different sensors can be fused in a more fine-grained manner. The experimental results on two multi-sensor-based fault diagnosis datasets demonstrate the anti-noise capability of our proposed method.

## Hosted file

IET-ELL-sample.tex available at <https://authorea.com/users/505967/articles/584882-frequency-adaptive-wavelet-pyramid-for-noisy-machinery-fault-diagnosis-with-multiple-sensors>





# Frequency adaptive wavelet pyramid for noisy machinery fault diagnosis with multiple sensors

Aosheng Tian,<sup>1</sup> Zhang Ye,<sup>1</sup> Chao Ma,<sup>1</sup> Huilin Chen,<sup>1</sup> Shilin Zhou,<sup>1</sup> and Weidong Sheng<sup>1</sup>

<sup>1</sup>*School of Electronic Science, National University of Defense Technology, Changsha, China*

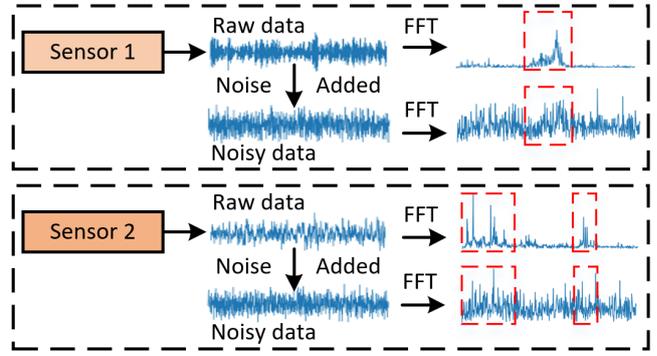
Email: machao0408@nudt.edu.cn

The fusion of multiple monitoring sensors is crucial to improve the accuracy and robustness of machinery fault diagnosis. However, existing fault diagnosis methods may underestimate the interference of noise in the multi-sensor fusion process, leading to unsatisfied performance. To handle this problem, this paper proposes a deep model based on the frequency adaptive wavelet pyramid. First, an adaptive frequency selection strategy is designed to prune the seriously polluted frequencies and only retain some key frequencies. Then, the self-attention mechanism is used to perform information fusion on the selected frequency bands of different sensors. Finally, a wavelet fusion pyramid is adopted by repeating the fusion process at multiple wavelet decomposition levels. In this way, different sensors can be fused in a more fine-grained manner. The experimental results on two multi-sensor-based fault diagnosis datasets demonstrate the anti-noise capability of our proposed method.

**Introduction:** To ensure the stability of rotating machinery in industrial environments, the information of multiple monitoring sensors is usually integrated for fault diagnosis [1, 2]. However, the amount of noise that exists in complex industrial environments brings difficulties to the fusion of different sensors, leading to poor fault diagnosis performance [3]. As shown in Figure 1, the overall shapes of the sensor signals are severely corrupted by noise. Specially, some key frequency bands (see the red blocks in Figure 1) are hidden in the noisy frequency bands, bringing difficulties to the multi-sensor fusion process. Therefore, it is urgent to develop effective anti-noise multi-sensor fusion methods.

In recent years, deep learning has been widely explored to perform multi-sensor fusion for fault diagnosis. Existing multi-sensor fusion methods can be classified into the convolutional neural network based (CNN-based), recurrent neural network based (RNN-based), and hybrid-model-based categories. Early works explored CNN-structured networks for multi-sensor fusion. In [4], different CNNs were applied to extract features from each sensor. The extracted features were then fed into the support vector machine for feature fusion. In [5], a 1D-CNN-based network was proposed to fuse the concatenated multi-sensor data for fault diagnosis. In [6], a multi-branch CNN was designed to extract temporal and spatial features from different sensors, respectively. Then, the attention mechanism was applied to enhance the features of important sensors. To capture the long-term dependencies, some works have explored RNN-based models. In [7], the signals of different sensors were split into multiple fixed-length segments. These segments were then horizontally concatenated and input to an LSTM network to construct both temporal and spatial correlations. In [8], the multi-sensor data was first decomposed to multiple frequency bands through discrete wavelet transform (DWT). Then, different frequency bands were concatenated and fed into an LSTM network for fusion. Furthermore, to integrate the advantages of CNNs in feature fusion and RNNs in capturing long-term correlations, hybrid models composed of CNN and RNN subnets have also been developed for multi-sensor based fault diagnosis [9]. As can be concluded, although existing multi-sensor fusion methods have obtained promising performance, these methods may lack fully consideration in the interference of noise during the multi-sensor fusion process.

This letter proposes an anti-noise multi-sensor fusion method, named the **Frequency Adaptive Wavelet Pyramid Network (FAWPNet)**. First, the input features are decomposed into multiple frequency bands through DWT, with an adaptive frequency selection strategy based on the Gumbel softmax trick [10] to prune the severely polluted frequency bands harmful to the multisensor fusion process. The Gumbel softmax trick is applied to handle the non-derivable problem in the frequency selection process. Second, the self-attention mechanism [11] is leveraged to per-



**Fig 1** The signal samples of the same fault category with and without noise. FFT denotes the Fast Fourier Transform. The waveforms on the right side represent the spectrums of the signals. The red blocks denote the key frequency regions of these samples.

form multi-sensor fusion on top of the selected frequency bands. Finally, since it is difficult to select a suitable wavelet decomposition level, a wavelet pyramid is constructed to perform multi-sensor fusion at multiple decomposition levels. In this way, different sensors can be fused in a more fine-grained manner.

The main contributions of this letter can be summarized as follows:

- 1) An adaptive frequency selection strategy is proposed to select key frequency bands, and discard the frequency bands seriously polluted by noise.
- 2) A wavelet pyramid is constructed to perform multi-sensor fusion on top of the selected frequency bands at multiple wavelet decomposition levels.
- 3) Experimental results demonstrate that our method obtains superior anti-noise performance on two multi-sensor-based fault diagnosis datasets under different signal-to-noise ratios (SNRs).

**Method:** The overall framework of FAWPNet is shown in Fig 2 (with the fusion process of two sensors as an example). As shown, the FAWPNet can be divided into three processes, including the adaptive frequency selection, the multi-sensor fusion and the multi-level fusion based on wavelet pyramid. These processes are described as follows.

**Adaptive Frequency Selection:** The input signals of different sensors are first passed through different convolutional branches, respectively. Then, the extracted features of each sensor are decomposed into multiple frequency bands through DWT. The DWT decomposition process is presented as follows:

$$x_L[t] = \sum_k x[2t - k] \cdot l[k], \quad (1)$$

$$x_H[t] = \sum_k x[2t - k] \cdot h[k], \quad (2)$$

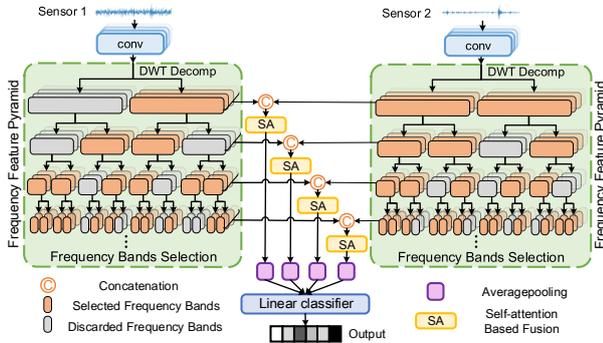
where  $t$  denotes the timestamps of the signal,  $k$  denotes the timestamps of the filter,  $l$  denotes the low-pass filter,  $h$  denotes the high-pass filter,  $x$  denotes the raw signal. Finally, the decomposed frequency bands are screened to select the key frequency bands, and discard the redundant ones (1 for “selected” band, and 0 for “discarded” band). Since the frequency selection operation is non-derivable, the Gumbel softmax trick is introduced to handle this problem. Concretely, in the training phase, to make the binary selection operation learnable, the Gumbel softmax trick is leveraged to approximate the one-hot distribution. In the testing phase, the Gumbel softmax operation is replaced with “argmax” operation to obtain the selected key frequency bands.

The process of Gumbel softmax trick can be illustrated with the following formula:

$$M_{i,j}[n] = \frac{\exp((F'_{i,j}[1, n] + G_{i,j}[1, n])/\tau)}{\sum_{m=0,1} \exp((F'_{i,j}[m, n] + G_{i,j}[m, n])/\tau)}, \quad (3)$$

$$F_{i,j}^S = M_{i,j} \cdot F_{i,j}, \quad (4)$$

where  $i$  is the wavelet decomposition level,  $j$  is the number of sensors,  $n$  is the indice of decomposed frequency bands,  $F'_{i,j} \in \mathbb{R}^{2 \times N}$  is the binary



**Fig 2** An overall framework of FAWPNet.

selection matrix ( $N$  represents the number of frequency bands in one decomposition level),  $G_{i,j}$  is a Gumbel noise for all frequency bands,  $M_{i,j} \in \mathbb{R}^N$  is the “argmax” vector,  $F_{i,j} \in \mathbb{R}^{C \times N \times L}$  is decomposed frequency bands ( $C$  is the channel number of frequency bands and  $L$  is the length of frequency bands),  $F_{i,j}^s \in \mathbb{R}^{C \times N' \times L}$  is the selected frequency bands ( $N'$  is the number of the selected frequency bands),  $\tau$  is a temperature hyperparameter. When  $\tau$  is closer to 0, the Gumbel softmax distribution becomes one-hot. When  $\tau$  is closer to infinity, the Gumbel softmax distribution becomes uniform.

**Multi-sensor Fusion:** The self-attention mechanism is used to construct correlations between different sensors in the fusion process. First, the selected frequency bands ( $F_{i,j}^s$ ) of different sensors are concatenated as  $F_i^a \in \mathbb{R}^{C \times S' \times L}$ , where  $S'$  denotes the number of selected frequency bands of all sensors. Then, the self-attention mechanism is used to construct correlations between different sensors on top of  $F_i^a$ . Finally, the correlated frequency bands are passed through an averagepooling operation for multi-sensor fusion.

The process of self-attention operation can be illustrated as follows:

$$F_i^{att} = \text{softmax}\left(\frac{f(F_i^a) \cdot g(F_i^a)}{\sqrt{d}}\right) \cdot h(F_i^a), \quad (5)$$

where  $f$ ,  $g$  and  $h$  denote the linear transformations,  $F_i^a$  is the concatenated selected frequency bands of different sensors,  $\sqrt{d}$  is the dimensions of  $f(F_i^a)$ , which is introduced for normalization,  $F_i^{att}$  is the weighted frequency bands of multiple sensors.

**Multi-level Fusion Based on Wavelet Pyramid:** A problem still remains to be tackled is that, it is difficult to choose a suitable wavelet decomposition level for multi-sensor fusion. Considering different wavelet decomposition levels produce frequency bands with different fine granularity, we repeat frequency selection and multi-sensor fusion process at multiple decomposition levels (see Figure 2). In this way, the multi-sensor fusion process can be more sufficient. Subsequently, the fused features of different decomposition levels are concatenated and fed into a linear layer for classification.

#### Performance Evaluation:

**Dataset:** Two multi-sensor based machinery fault diagnosis datasets are used to evaluate the compared methods. The first dataset is the Case Western Reserve University Bearings (CWRU) dataset [12]. In our experiments, only three types of measurement signals are used, including the signals recorded by accelerations placed at the motor drive end, fan end and supporting base plate, respectively. Moreover, ten kinds of bearing state categories are used for evaluation, following the suggestions in [13].

The second dataset is the Southeast University Gearbox (SEU) dataset [14]. The gearbox data are acquired by accelerations embedded in the Drivetrain Dynamics Simulator (DDS). This dataset comprises the signals with regard to the motor vibration of the planetary gearbox, motor torque and vibration of the parallel gearbox. In our experiments,

ten kinds of bearing state categories are used for evaluation, following the suggestions in [13].

**Experimental Settings:** Different models are implemented in Pytorch 1.7.1 on a PC with a single Nvidia RTX 1080Ti GPU. These models are optimized using the cross-entropy loss and the Adam optimizer. The learning rate was initialized as  $1e-3$  and decreased according to the training loss. Besides, the batch size and the training epochs were set to 64 and 200, respectively. Each dataset is separated into the training dataset and testing dataset with a ratio of 4 to 1. Moreover, the final testing results can be obtained by averaging the testing results of 5-fold cross-validation.

**Comparisons of Different Methods:** The proposed model is compared with three multi-sensor fusion based fault diagnosis methods, including DCNN [5], LSTM [7] and MS-PACNN [6]. To evaluate the anti-noise performance of different models, we add random Gaussian white noise with different SNRs to each data channel of different sensors. The formula of SNR is listed as follows:

$$\text{SNR} = 10 \log(P_{\text{Signal}}/P_{\text{Noise}}), \quad (6)$$

where  $P_{\text{Signal}}$  represents the power of the original signal,  $P_{\text{Noise}}$  represents the power of the added Gaussian noise. Table 1 presents the quantitative performance of FAWPNet and three compared methods under six different SNRs (including -10, -6, -2, 2, 6, 10). As can be observed from Table 1, FAWPNet performs obviously better than the compared methods. Therefore, it is demonstrated that the proposed method has superior anti-noise capabilities for multi-sensor based fault diagnosis.

**Table 1.** The classification results of the proposed model and the compared methods under different intensities of Gaussian noise.

Dataset	Methods	SNR(dB)					
		-10	-6	-2	2	6	10
CWRU	DCNN	80.15	89.22	98.09	99.78	99.81	99.98
	LSTM	70.76	67.59	63.81	58.40	72.25	70.88
	MB-PACNN	94.65	95.21	96.39	98.26	99.06	99.61
	FAWPNet	<b>98.65</b>	<b>99.33</b>	<b>99.88</b>	<b>99.98</b>	<b>99.98</b>	<b>1</b>
SEU	DCNN	71.08	65.11	72.13	85.62	89.80	92.29
	LSTM	53.11	51.54	83.99	96.37	97.06	97.20
	MB-PACNN	98.24	94.52	95.89	95.93	97.22	97.75
	FAWPNet	<b>99.50</b>	<b>99.90</b>	<b>99.94</b>	<b>99.98</b>	<b>1</b>	<b>1</b>

**Ablation Experiments:** This section provides the ablation experiments of FAWPNet. To be specific, the ablation experimental results of the adaptive frequency selection strategy and the wavelet fusion pyramid are presented, respectively.

To demonstrate the effectiveness of the adaptive frequency selection strategy, the models with and without the frequency selection process (FAWPNet-W) are compared. It is worth mentioning that, instead of selecting the key frequency bands, FAWPNet-W uses the self-attention mechanism to weight all of the frequency bands. The experimental results are presented in Table 2. As can be observed, FAWPNet shows a clear advantage compared with FAWPNet-W under all different SNRs. The experimental results indicates that, it is difficult to handle the interference of noise through the weighting operations, while the frequency selection strategy can effectively improve the anti-noise capability of the network by discarding the polluted frequency bands.

To demonstrate the effectiveness of the wavelet fusion pyramid, the pyramids with different decomposition levels are compared. In Table 2, FAWPNet shows better performance than the model with only two decomposition levels (FAWPNet-2) under all different SNRs. These experimental results illustrate that, the multi-level wavelet pyramid fuses the information of multiple sensors more sufficiently.

**Discussion:** To further illustrate the process of frequency selection, we show the spectrums of two samples (selected from the CWRU dataset)

Table 2. The classification results of ablation experiments under different intensities of Gaussian noise.

Dataset	Methods	SNR(dB)					
		-10	-6	-2	2	6	10
CWRU	FAWPNNet-W	97.77	98.95	99.50	99.83	99.97	99.93
	FAWPNNet-2	97.92	98.61	99.36	99.65	99.83	99.90
	FAWPNNet	<b>98.65</b>	<b>99.33</b>	<b>99.88</b>	<b>99.98</b>	<b>99.98</b>	<b>1</b>
SEU	FAWPNNet-W	94.70	98.79	98.94	99.12	99.64	99.73
	FAWPNNet-2	95.49	98.24	98.98	<b>99.98</b>	99.98	<b>1</b>
	FAWPNNet	<b>99.50</b>	<b>99.90</b>	<b>99.94</b>	<b>99.98</b>	<b>1</b>	<b>1</b>

in Figure 3. The key frequencies of the first sample are mainly located in the low-frequency region (see the red block in the left part of Figure 3 (a)). Differently, the key frequencies of the second sample are mainly located in the mid-high frequency region (see the red block in the left part of Figure 3 (b)).

The distributions of the selected frequency bands corresponding to two samples are presented in the right part of Figure 3. It can be observed that, the selected frequency bands of the first sample (see the right half part of Figure 3 (a)) are mostly distributed in the low-frequency region, which is consistent with the spectrum distribution in the left part of Figure 3 (a). Moreover, similar phenomena can also be observed from the second sample in Figure 3 (b). Through this toy example, we demonstrate that the adaptive frequency selection strategy can effectively select the key frequency bands of different sensors.

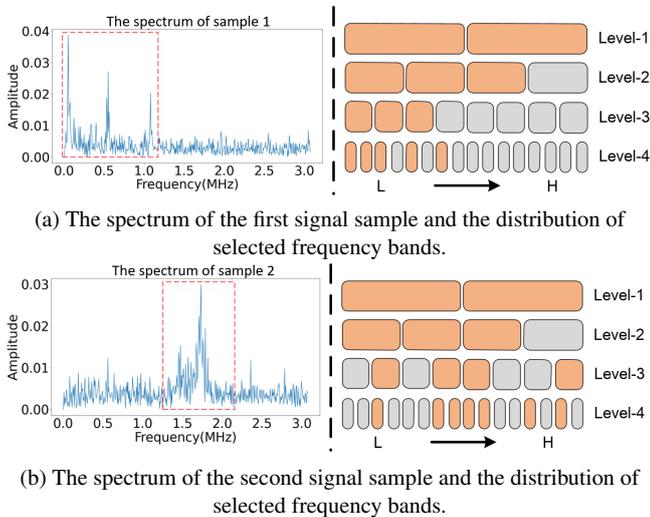


Fig 3 The spectrums of these two samples and the distributions of the selected frequency bands of two samples.

**Conclusion:** In this letter, we propose a frequency adaptive wavelet pyramid network for noisy multi-sensor fault diagnosis. First, the signals of different sensors are decomposed into multiple frequency bands through DWT, with designed frequency selection strategy to adaptively prune seriously polluted frequencies of each sensor. Then, the self-attention mechanism is applied to fuse different sensors on top of the preserved key frequency bands. Finally, the wavelet fusion pyramid repeats the fusion process at multiple decomposition levels. Experimental results on two multi-sensor based fault diagnosis datasets demonstrate the superior anti-noise performance of our method.

**Acknowledgments:** This work was supported by the National Natural Science Foundation of China under Grant (No. 62002372)

© 2022 The Authors. *Electronics Letters* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Received: 10 January 2021 Accepted: 4 March 2021

doi: 10.1049/ell2.10001

## References

- Rai, A., Upadhyay, S.H.: A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribology International* pp. 289–306. (2016)
- Song, Q., Zhao, S., Wang, M.: On the accuracy of fault diagnosis for rolling element bearings using improved dfa and multi-sensor data fusion method. *Sensors* 20(22) (2020)
- Jin, Y., et al.: Fault diagnosis of brake train based on multi-sensor data fusion. *Sensors* 21(13), 4370 (2021)
- Liu, Y., et al.: An ensemble convolutional neural networks for bearing fault diagnosis using multi-sensor data. *Sensors (Basel, Switzerland)* 19(23) (2019)
- Jing, L., et al.: An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors (Switzerland)* 17(2), 414 (2017)
- Xing, Z., et al.: Multi-sensor signals with parallel attention convolutional neural network for bearing fault diagnosis. *AIP Advances* 12(7), 075020 (2022)
- Lei, J., Liu, C., Jiang, D.: Fault diagnosis of wind turbine based on long short-term memory networks. *Renewable Energy* 133(APR.), 422–432 (2019)
- Gu, K., et al.: Dwt-lstm-based fault diagnosis of rolling bearings with multi-sensors. *Electronics* (17) (2021)
- Hao, S., et al.: Multisensor bearing fault diagnosis based on one-dimensional convolutional long short-term memory networks. *Measurement* 159, 107802 (2020)
- Jang, E., Gu, S., Poole, B.: Categorical reparameterization with gumbel-softmax. *arXiv e-prints* (2016)
- Vaswani, A., et al.: Attention is all you need (2017)
- Wade, et al.: Rolling element bearing diagnostics using the case western reserve university data: A benchmark study. *Mechanical Systems and Signal Processing* 64–65, 100–131 (2015)
- Zhao, Z., et al.: Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study. *ISA transactions* 107, 224–255 (2020)
- Shao, S., et al.: Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Transactions on Industrial Informatics* 15(4), 2446–2455 (2019)
- Li, Q., et al.: Wavelet integrated cnns for noise-robust image classification pp. 7243–7252. (2020). doi:10.1109/CVPR42600.2020.00727
- Gong, et al.: A novel deep learning method for intelligent fault diagnosis of rotating machinery based on improved cnn-svm and multichannel data fusion. *Sensors* 19(7) (2019)
- Han, D., et al.: A novel intelligent fault diagnosis method based on dual convolutional neural network with multi-level information fusion. *Journal of Mechanical Science and Technology* pp. 1–15.
- Li, X., et al.: Bearing fault diagnosis method based on attention mechanism and multilayer fusion network. *ISA Transactions* (2021). doi:10.1016/j.isatra.2021.11.020
- Wang, X., Mao, D., Li, X.: Bearing fault diagnosis based on vibro-acoustic data fusion and 1d-cnn network. *Measurement* 173(6), 108518 (2021)
- Shan, P., et al.: A multisensor data fusion method for ball screw fault diagnosis based on convolutional neural network with selected channels. *IEEE Sensors Journal* 20(14), 7896–7905 (2020)