sciendo

POLISH MARITIME RESEARCH 1 (117) 2023 Vol. 30; pp. 112-120
10.2478/pomr-2023-0011

DEEP LEARNING-BASED FAULT DIAGNOSIS FOR MARINE CENTRIFUGAL FAN

Congyue Li [1*]
Yihuai Hu [1]
Jiawei Jiang [2]
Guohua Yan [1]

[1] Shanghai Maritime University, College of Merchant Marine, Shanghai, China
[2] Shanghai Institute of Electronic Information Technology, School of Mechanical and Energy Engineering, Shanghai, China

* Corresponding author: *licongyue0126@163.com (Congyue Li)*

ABSTRACT

Marine centrifugal fans usually work in harsh environments. Their vibration signals are non-linear. The traditional fault diagnosis methods of fans require much calculation and have low operating efficiency. Only shallow fault features can be extracted. As a result, the diagnosis accuracy is not high. It is difficult to realize the end-to-end fault diagnosis. Combining the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and lightweight neural network, a fault classification method is proposed. First, the CEEMDAN can decompose the vibration signal into several intrinsic modal functions (IMF). Then, the original signals can be transformed into 2-D images through pseudo-colour coding of the IMFs. Finally, they are fed into the lightweight neural network for fault diagnosis. By embedding a convolutional block attention module (CBAM), the ability of the network to extract critical feature information is improved. The results show that the proposed method can adaptively extract the fault characteristics of a marine centrifugal fan. While the model is lightweight, the overall diagnostic accuracy can reach 99.3%. As exploratory basic research, this method can provide a reference for intelligent fault diagnosis systems on ships.

Keywords: CEEMDAN; Lightweight neural network; Marine centrifugal fan; Fault diagnosis

INTRODUCTION

Marine centrifugal fans are essential equipment for ships. They are mainly responsible for cabin ventilation, the boilers' induced air, ventilation of inert gas systems, etc. Damage to them will increase the cost of ship maintenance and affect the safety of ship operations. Therefore, conducting condition testing and fault diagnosis for the marine centrifugal fan is necessary. In the past, this work was accomplished through reactive maintenance and time-based maintenance, which depends heavily on human cognition and experience [1]. It is difficult to meet the development needs of large-scale, intelligent and unmanned ships [2].

Traditional machine learning fault diagnosis methods can only extract shallow features, and rely heavily on expert experience. In addition, they have poor robustness and generalization ability when dealing with massive data. With the continuous development of technology, deep learning fault diagnosis methods are developing rapidly. The convolutional neural network (CNN) is the most popular. Xie et al. [3] constructed an empirical mode decomposition (EMD)-CNN model, which enables the classification of faults in marine blowers. Guan et al. [4] combined EMD-sample entropy and a deep confidence network to realize rotating machinery fault diagnosis. However, EMD has defects, including mode mixing, false mode, and endpoint effects [5]. Rafia et al. [6] proposed the model of

ensemble EMD (EEMD)-continuous wavelet transform (CWT)-CNN, which realized the accurate classification of connecting rod bearings. EEMD is an improvement based on EMD. But the heavy computing load because of the added noise seriously reduces the executive efficiency of EEMD [7]. In [8], the images were obtained by short-time Fourier transform (STFT). Then, they were fed into the CNN to implement the fault diagnosis of bearings. Although using CNN for fault diagnosis improves the feature extraction ability and generalization ability, it also has some adverse effects. That is, with the increasing depth of the model, some current CNN models have problems such as a too complex structure, too large scale, and they have difficulty converging quickly [9]. In addition, the commonly used time-frequency analysis methods have many drawbacks. For example, STFT has difficulty in describing the local characteristics of the signals and it cannot be applied for the analysis of non-linear signals. CWT requires the selection of wavelet basis functions. These shortcomings do not apply to the fault diagnosis of marine centrifugal fans, however.

The lightweight CNN has been gaining popularity. MobileNetV1 [10] uses depth-separable convolution (DSC) to reduce model complexity. MobileNetV2 [11] improves the feature extraction ability by introducing the idea of a residual. ShuffleNetV1 [12] proposed group convolution (GC) and channel shuffle (CS), which can reduce the amount of computation. Ma et al. [13] designed ShuffleNetV2 based on summarizing four criteria affecting the network operation efficiency. Compared with the traditional CNN, the above lightweight CNN has fewer parameters and more efficient operation. However, when there is interference from irrelevant signals, such as noise, the performance of the model will be affected.

To address these problems, this paper combines CEEMDAN with CBAM-ShuffleNetV2 for fault classification of marine centrifugal fans. Referring to [14], the 1-D signals are converted to 2-D images. The main contributions of this paper are as follows:

(1) The vibration signal was decomposed by CEEMDAN. This can not only solve the problems of modal aliasing and noise transfer, but also preserve the main components of signals.
(2) Through the signal-to-image conversion, time-frequency maps can be obtained. These can describe the global and local characteristics of non-linear signals.
(3) Using the lightweight CNN, which can overcome the shortcomings of traditional network models, such as complex calculation, low diagnostic accuracy, etc., CBAM is introduced to improve the model's ability to extract critical feature information.

BACKGROUND THEORY

CEEMDAN

EEMD and complete EEMD (CEEMD) are improved algorithms based on EMD. However, these two improved algorithms cannot deal with residual noise, which is not conducive to subsequent signal analysis.

Torres et al. [15] proposed CEEMDAN based on EMD and EEMD. It improves this problem in two ways: (1) The noise is not directly fed into the native signal, but the noise component after EMD decomposition is added; (2) After CEEMDAN decomposes the signal to obtain the first IMF, the overall average is performed. CEEMDAN not only reduces the amount of computation but also improves the signal decomposition accuracy. The CEEMDAN decomposition steps are as follows:

After adding the signal $x(t)$ with white noise, it can be expressed as:

$$x(t) + (-1)^q \varepsilon v^i(t) \tag{1}$$

where q is equal to 1 or 2.

C_1^j is the first-order IMF obtained from the EMD decomposition of the new signal. It can be obtained as:

$$E[x(t) + \varepsilon v^i(t)] = C_1^j + r^j \tag{2}$$

The obtained IMF components are then averaged. This can be obtained as:

$$\overline{C_1(t)} = \frac{1}{N}\sum_{j=1}^{N} C_1^j(t) \tag{3}$$

where N denotes the number of IMFs.

The first residual is denoted as:

$$r_1(t) = x(t) - \overline{C_1(t)} \tag{4}$$

The same white noise is added to $r_1(t)$ to produce a new signal, which is decomposed by EMD. The obtained IMF components are then averaged. This can be obtained as:

$$\overline{C_2(t)} = \frac{1}{N}\sum_{j=1}^{N} D_1^j(t) \tag{5}$$

where D_1^j denotes the first IMF, and N denotes the number of IMFs.

The second residual is denoted as:

$$r_2(t) = r_1(t) - \overline{C_2(t)} \tag{6}$$

These steps are repeated until EMD cannot decompose the signal further. The original signal is decomposed into:

$$x(t) = \sum_{k=1}^{K} \overline{C_k(t)} + r_k(t) \tag{7}$$

where K is a positive integer.

GC AND CS

Standard convolution performs repeated convolution operations on the information of each channel of the input feature map. Assume that the convolution kernel size is $D \times D$ and the input feature channels are W. The number of convolution kernels is B. The number of standard convolution parameters is:

$$Q = D \times D \times W \times B \tag{8}$$

GC is different from standard convolution. It first groups the input feature channels and convolution kernels and then convolves the operation. Fig. 1 shows the GC process. The number of GC parameters is:

$$Q = D \times D \times \frac{W}{G} \times B \tag{9}$$

where G is the number of groups.

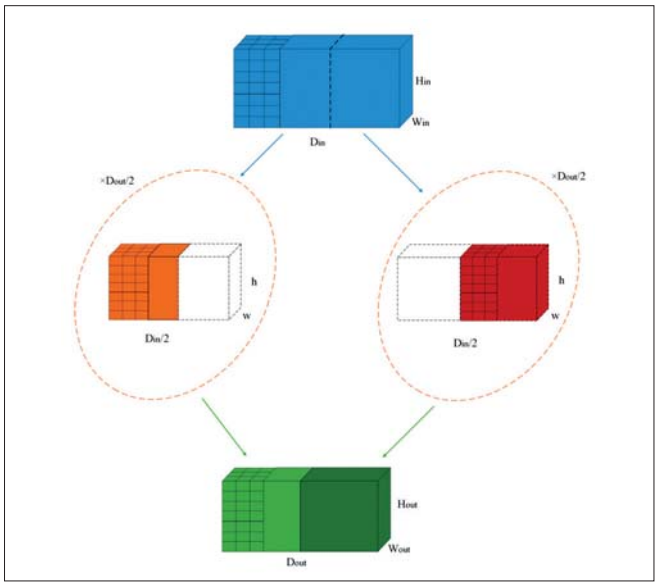

Fig. 1. The GC process

The number of parameters for GC is smaller. Krizhevsky et al. [16] pointed out that GC prevents overfitting and facilitates regularization. GC reduces the amount of computation. But the groups are independent of each other and there is no intermingling of information. By introducing CS, information blending between groups can be realized. Fig. 2 shows the CS process.

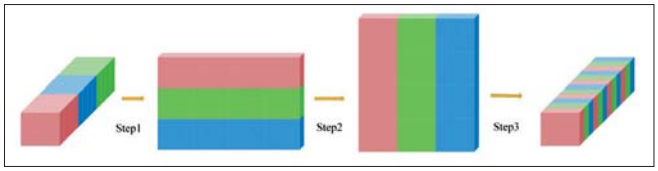

Fig. 2. The CS process

DSC

For standard convolution, each convolution kernel performs a convolution operation on all channels of the input. Unlike standard convolution, DSC is divided into depthwise convolution (DWC) and pointwise convolution (PWC) operations. DWC means that each convolution kernel handles only one channel of input. PWC is a *1×1* unit convolution. Fig. 3 shows the DSC process.

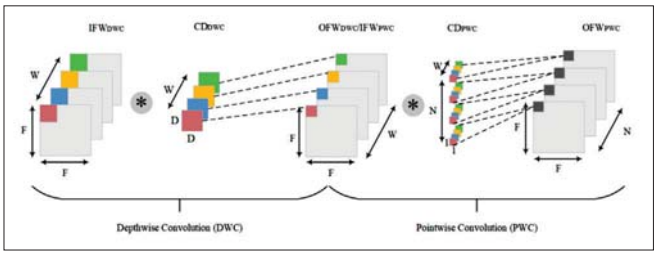

Fig. 3. The DSC process

SHUFFLENETV2 UNIT

Fig. 4 shows the ShuffleNetV2 basic units. In Unit 1, the input channels are divided equally. The left half does not perform any operations. The right half contains two 1×1 convolutions and one DWC. Then, the two branches are cascaded together. Finally, the CS operation is carried out to ensure the full integration of feature information. Unit 2 is the down-sampling module. Both branches perform a *3×3* deep convolution with step size 2. The length and width are reduced by half, and the output channels are doubled.

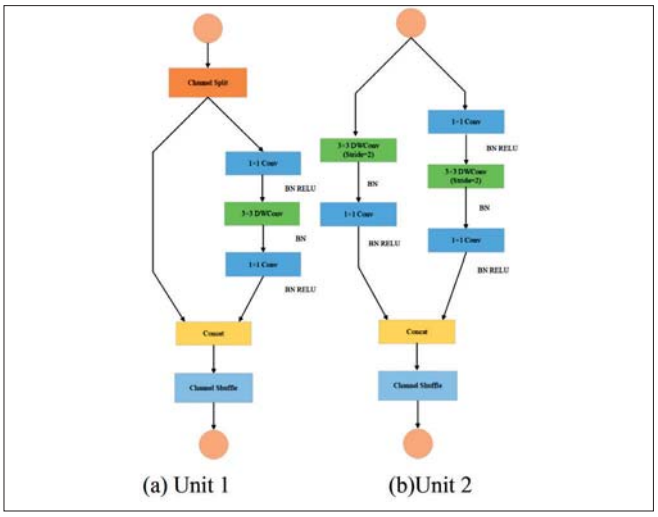

Fig. 4. Two basic units of ShuffleNetV2

CBAM

Woo et al. [17] proposed the CBAM. In contrast to SENet [18], CBAM focuses not only on the importance of each channel feature but also on the importance of the local features of each channel. Fig. 5 shows the CBAM structure, which is a simple one. It can be easily embedded in the CNN while remaining lightweight. It mainly includes the channel attention mechanism (CAM) and spatial attention mechanism (SAM). Their structures are shown in Fig. 6 and 7. The CAM performs average pooling and maximum pooling on the input feature maps. It yields two *C×1×1* features, which are fed into a multiple layer perceptron (MLP). Then the element-by-element summation is performed. The activation function *Sigmoid* is used to obtain the weight coefficient M_c, which is multiplied by F_c. The output of CAM is used as the input of SAM. After SAM, the spatial output features are obtained, which are multiplied by the CAM output. The final output is obtained.

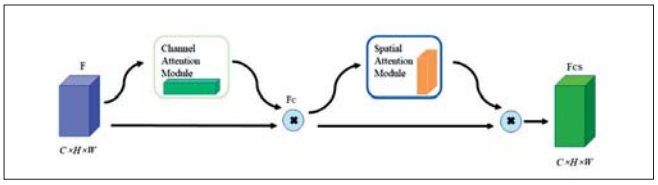

Fig. 5. The structure of CBAM

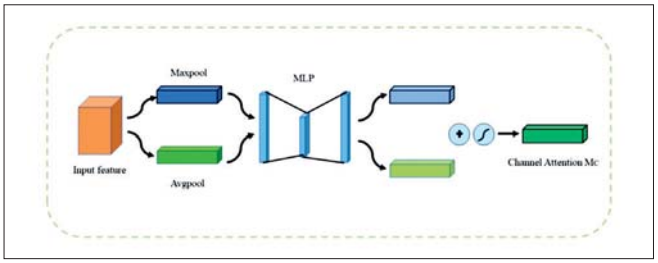

Fig. 6. The structure of CAM

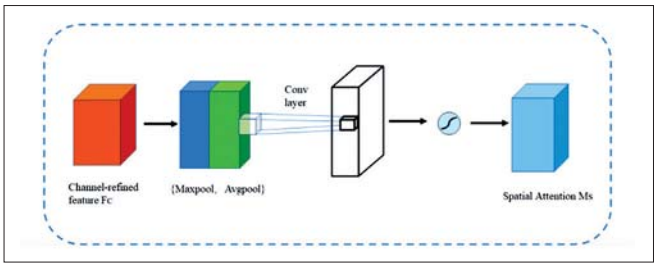

Fig. 7. The structure of SAM

THEORETICAL MODEL

Among the above two units of ShuffleNetV2, the more efficient Unit 1 is adopted. The CBAM is embedded in ShuffleNetV2. Fig. 8 is the CBAM-ShuffleNet unit. This module retains the maximum amount of light weight in the basic unit. This can enhance the extraction of crucial information and suppress irrelevant information, such as redundancy and noise.

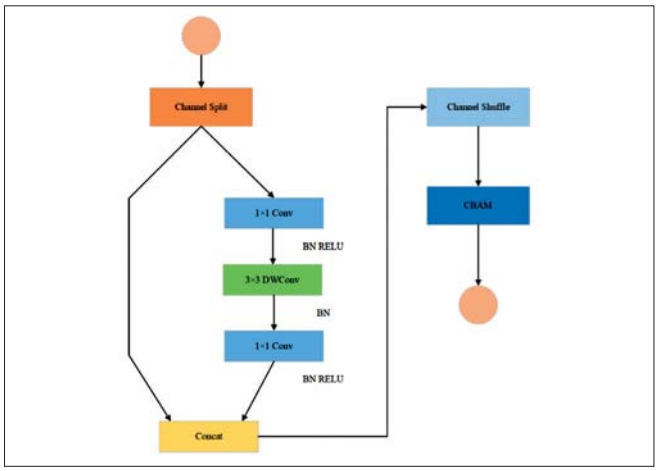

Fig. 8. Architecture of CBAM-ShuffleNet unit

EXPERIMENTAL ANALYSIS AND DISCUSSION

FAULT DIAGNOSIS PROCESS

The fault diagnosis process consists of four main stages: signal acquisition, image conversion, defining the lightweight CNN, and fault classification. Fig. 9 shows the fault diagnosis process. The detailed steps are as follows:

Step 1: Vibration signals of different health states are collected.

Step 2: CEEMDAN decomposes the vibration signal to get several IMF components, which are arranged in order from low to high frequency. The time–frequency matrix can be obtained, and the pseudo-colour coding is used to get the 2-D colour time–frequency maps.

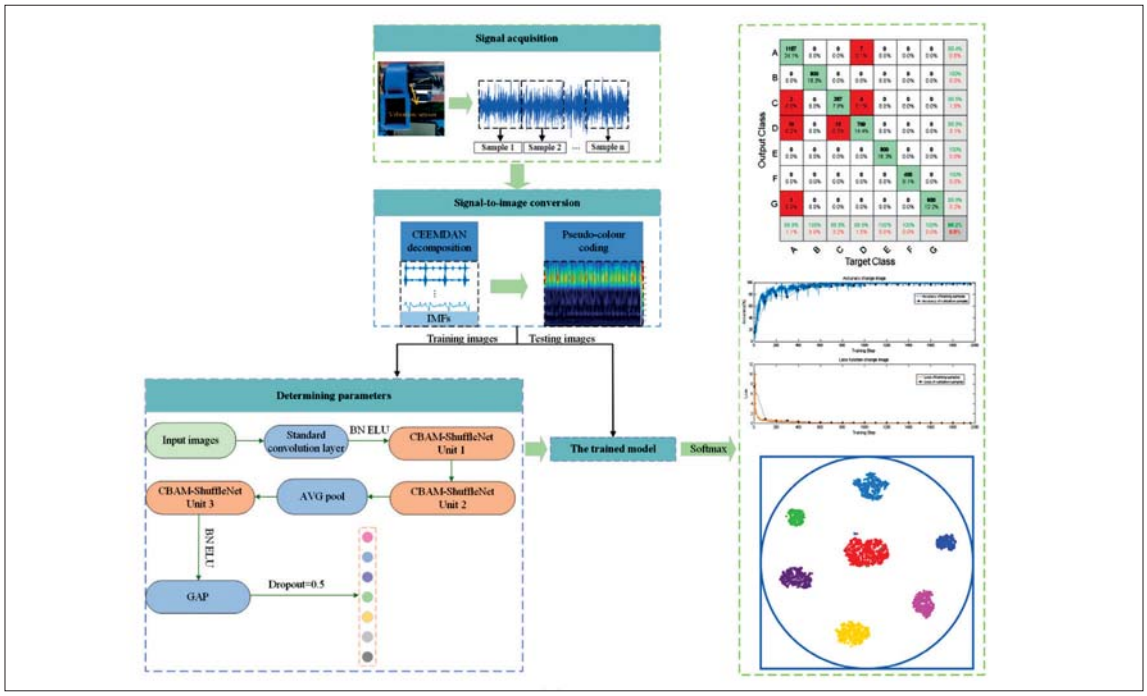

Fig. 9. Fault diagnosis process

Step 3: The data sets are divided into training and testing sets. Then, the training sets are fed into the network for training.

Step 4: The testing sets are fed into the lightweight CNN for adaptive fault feature extraction and classification.

DATA PRE-PROCESSING

The experimental marine centrifugal fan is driven by a three-phase asynchronous motor. An ICP vibration detector developed by Beijing East Vibration and Noise Technology Institute is adopted. The vibration sensor uses a magnetic suction sensor with a sensitivity of 10.20 mV/g and a bias voltage of 12.40 V. The sampling time is one second. The sampling frequency is 1652 Hz. Seven modes are set: normal, motor short circuit, fan blocking, loose bolts, impeller unbalanced, loose bolts and impeller unbalanced compound failure, and fan and motor coupling misalignment. The working conditions are numbered sequentially from A to G. 1200 groups of samples are collected for each working condition. 800 groups of samples are randomly selected for each working condition as training samples, and the rest are testing samples. The motor of the fan has a rated voltage of 380 V, rated current of 4.6 A, rated power of 1.1 kW, rated frequency of 50 Hz, and rated speed of 1400 r/min. Normal and short circuit faults are taken as examples for spectrum analysis. The 0~250 Hz spectrum amplification diagrams are shown in Fig. 10. The rotation frequency is 23.3 Hz (f_b). Under normal conditions, the vibration is mainly based on the fundamental frequency (23.3 Hz). The amplitude is small at the multiplier and rated frequency. The signal shock phenomenon is not obvious. When the motor is short-circuited, the magnetic field is distorted and the original magnetic field balance is destroyed. The vibration amplitude varies significantly at the fundamental frequency, multiplier frequency, and rated frequency. The amplitude is up to 0.137 m/s^2. It is a three-phase asynchronous motor with 2 pairs of poles. So, at triple and quadruple rotation frequencies, the amplitude changes significantly. The maximum deviation of the extracted fault characteristic frequency is 0.7 Hz. Fault characteristics are very obvious. This method is also applicable to other fault analysis of the centrifugal fan.

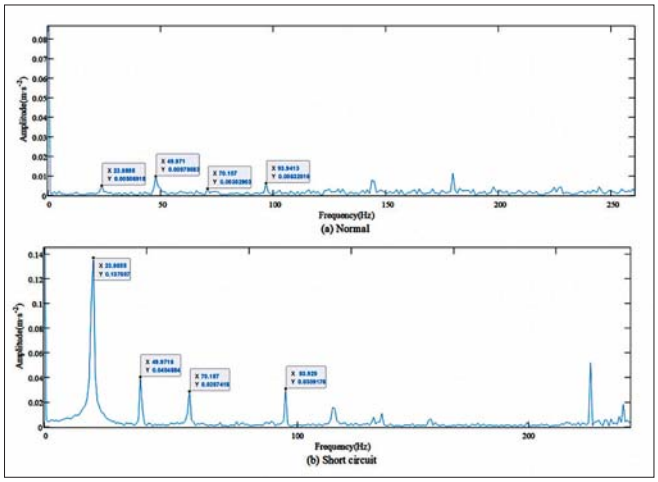

Fig. 10. The time domain waveforms for each operating condition

The original signals are decomposed by CEEMDAN. Fig. 11 provides the decomposition results of the first three working conditions. It can be seen that the decomposition process does not produce modal mixing and eliminates the interference of redundant information.

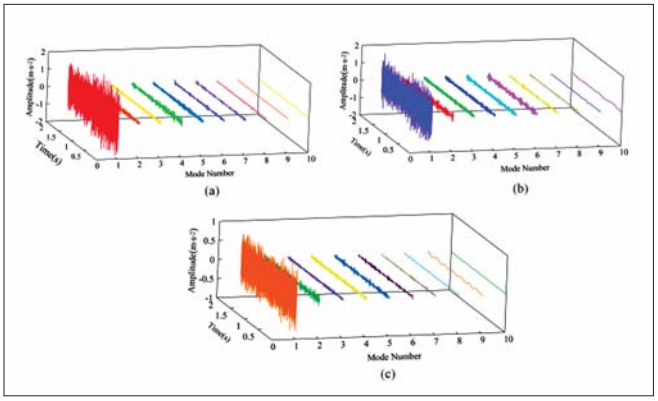

Fig. 11. Decomposition results for the first three working conditions

The time–frequency matrix can be obtained by CEEMDAN decomposition, and 2-D maps can be obtained by signal-to-image conversion. Fig. 12 shows the colour time–frequency maps for each working condition. Under normal conditions, the signal energy is essentially uniformly distributed in the time-scale plane. The analysis continues with the example of the short circuit fault. There is obvious periodic shock energy at 0~100 Hz and 400~500 Hz. The distribution of energy has a certain time interval. For example, there are some significant signal components occur at 0.0702 s, 0.1033 s, 0.1338 s, and 0.1608 s. The time interval is about 0.0318 s. Between the two energy peaks, there are two energy-low parts. That is, the energy change time interval is 0.0106 s. The frequency is about 94.3 Hz, which is close to the motor short-circuit fault frequency (93.929 Hz). The 2-D maps retain the fault characteristic information very well. There are significant differences in the energy distribution characteristics among the faults, i.e., their energy intensity and energy fluctuation time are significantly different. They fully demonstrate the time–frequency variation characteristics of the fault, facilitating subsequent feature extraction.

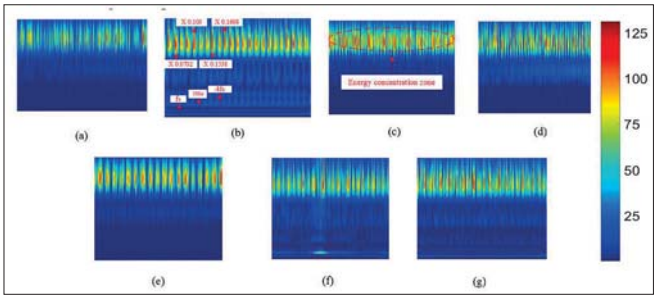

Fig. 12. 2-D time-frequency maps

MODEL TESTING AND FINE-TUNING

The deep learning is configured as comprising an AMD Ryzen 7 5800H CPU@3.2GHz, NVIDIA GeForce RTX 3060 Laptop GPU, and win11 operating system. The number of

CBAM-ShuffleNet units can affect the network performance. We use a five-fold cross-validation method to compare. The number of CBAM-ShuffleNet units is taken as 2, 3, 4, and 5, respectively. Noise with a signal-to-noise ratio (SNR) of -10dB is fed to the native signal. Table 1 shows the diagnostic accuracy for different numbers of units. Floating points of operations (FLOPs) are used to measure the network complexity. The larger the number of FLOPs, the more complex the network. With the increasing number of units, the numbers of parameters and FLOPs increase significantly, which makes the model less efficient. Although the efficiency of 3 units is slightly lower than that of 2 units, the highest diagnostic accuracy is achieved with 3 units. Therefore, the number of units selected is 3.

Tab. 1. Comparison of results of different numbers of units

Number of units	FLOPs/10^5	Number of parameters/10^5	Training time/s	Accuracy/%
2	0.65	0.32	1542	75.8
3	1.68	0.83	2285	88.7
4	8.39	4.18	3671	85.7
5	30.65	15.36	6087	85.4

Adam optimization parameters are used to update the network parameters. They are not only good at handling sparse gradients but can also handle non-smooth targets. First-order and second-order moment estimation preserve the adaptive learning rate for each parameter adaptation. After each iteration, the learning rate updates the interval, reducing the range of parameter fluctuations, speeding up convergence, and running efficiently. The starting parameters are shown in Table 2. The model computation is further reduced by global average pooling (GAP). The dropout layer can alleviate overfitting. Softmax is used for fault classification. Table 3 shows the optimal network parameters.

Tab. 2. Starting parameters

Name of parameter	Value
InitialLearnRate	0.001
LearnRateDropPeriod	10
LearnRateDropFactor	0.05
L2Regularization	0.004
MaxEpochs	20
MiniBatchSize	50

RESEARCH RESULTS

The training cycle is 20 rounds in total, the number of iterations in each round is 98, and the maximum number of iterations is 1960. Fig. 13 shows the diagnostic accuracy and the loss function images. The loss function gradually decreases until it tends to stabilize. At the same time, the accuracy of the training sets gradually increases and finally reaches 100%. This shows that this method is stable and converges quickly. The accuracy of the test sets can reach 100%. This network has

Tab. 3. Network Parameters

Sequence number	Structure name	Parameter values	Output
1	Input Layer	224×224	224×224
2	Standard convolution layer	24@3×3×3; stride=2	112××112
3	CBAM-ShuffleNet unit 1	Con (64@1×1×3; stride=1) DWC (64@3×3×3; stride=1) Con (64@1×1×3; stride=1)	56×56
4	CBAM-ShuffleNet unit 2	Con (64@1×1×3; stride=1) DWC (64@3×3×3; stride=1) Con (64@1×1×3; stride=1)	28×28
5	AVG pool	3×3; stride=2	16×16
6	CBAM-ShuffleNet unit 3	Con (128@1×1×3; stride=1) DWC (128@3×3×3; stride=1) Con (128@1×1×3; stride=1)	8×8
7	GAP	–	1×1
8	Dropout	0.5	–
9	FC	–	1×1×7
10	Softmax	–	1×1×7

no overfitting. The training time is about 56 seconds, and the testing time is about 18 seconds. It responds quickly.

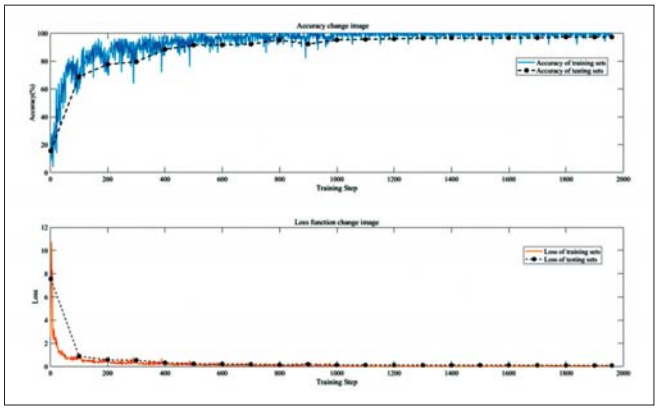

Fig. 13. Accuracy and loss function images

Fig. 14. The confusion matrix of the testing set

To reflect the diagnostic accuracy of the network model for different working conditions, Fig. 14 is the confusion matrix of

the testing sets. Green indicates the correct classification, and red means the opposite. The network model is slightly less sensitive to working conditions A, C, and D. But the diagnostic accuracy of the other working conditions can reach 100%. The validity of this method is verified based on the test sets.

T-SNE [19] can map high-dimensional nonlinear data into 2-D space. Based on this, the learning process within the network is visualized by T-SNE. Fig. 15 shows the learning process. At the input layer, different fault characteristic information intersects each other and it is not easy to separate them. After the first layer of convolution, there is a trend of clustering the feature information of the same working condition. With the deepening of the network model, it is possible to distinguish clearly between different working conditions. It follows that the proposed method can classify effectively.

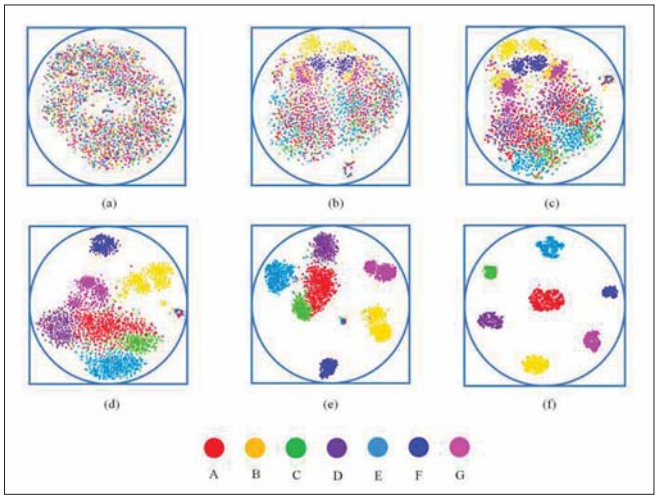

Fig. 15. Visualization results: (a) input layer; (b) convolution layer 1; (c) unit 1; (d) unit 2; (e) unit 3; (f) FC

COMPARISON OF OTHER MODELS

Further, the proposed method is compared with ShuffleNetV1/V2, MobileNetV2, and ResNet [20]. Noise with SNR of -20dB, -15dB, -10dB, and -5dB is added to the original signal, respectively. Then these samples are fed into the different network models for fault classification. Table 4 shows the comparison of the other methods. There is no significant change in the number of parameters and FLOPs, but the running time fluctuates slightly. MobileNetV2 has the simplest structure, so it has the highest efficiency. ShuffleNetV2 is more efficient than V1 because V2 removes fragmentation and element-by-element addition operations. The proposed method had slightly more FLOPs than ShuffleNetV2/V1 due to the embedding of CBAM in the model. But, the structure of CBAM has less impact on the model. The ResNet network model is deeper, so there are more FLOPs, resulting in the lowest operational efficiency.

Fig. 16 shows the diagnostic accuracy and average recall of different models. With the noise intensity decreasing, the classification accuracy and recall rate of the different network models gradually increased. Through comparative analysis, the proposed method has strong robustness while ensuring the model's light weight.

Tab. 4. Comparison of different methods

SNR/ dB	Network model	FLOPs/10^5	Number of parameters/10^5	Training time/s
-20	ResNet	50.27	25.07	11826
	ShuffleNetV1	1.38	0.68	6143
	ShuffleNetV2	1.52	0.76	2183
	MobileNetV2	1.41	0.57	2076
	The proposed method	1.68	0.83	2253
-15	ResNet	50.27	25.07	11843
	ShuffleNetV1	1.38	0.68	6128
	ShuffleNetV2	1.52	0.76	2164
	MobileNetV2	1.41	0.57	2018
	The proposed method	1.68	0.83	2241
-10	ResNet	50.27	25.07	11864
	ShuffleNetV1	1.38	0.68	6139
	ShuffleNetV2	1.52	0.76	2151
	MobileNetV2	1.41	0.57	2031
	The proposed method	1.68	0.83	2285
-5	ResNet	50.27	25.07	11907
	ShuffleNetV1	1.38	0.68	6157
	ShuffleNetV2	1.52	0.76	2162
	MobileNetV2	1.41	0.57	2043
	The proposed method	1.68	0.83	2261
Original signal	ResNet	50.27	25.07	11816
	ShuffleNetV1	1.38	0.68	6121
	ShuffleNetV2	1.52	0.76	2155
	MobileNetV2	1.41	0.57	2037
	The proposed method	1.68	0.83	2273

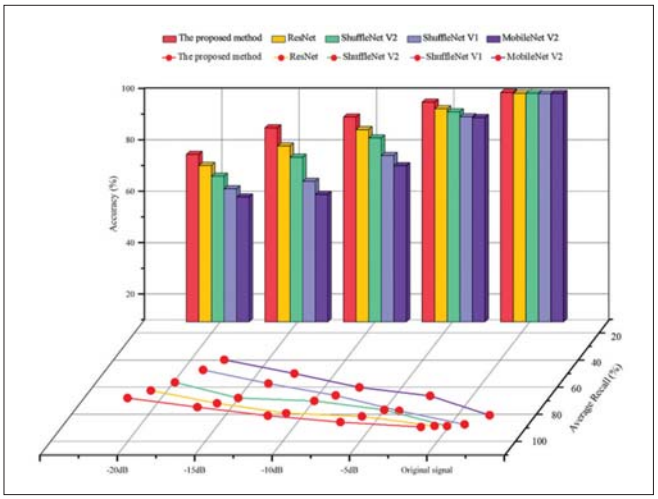

Fig. 16. Comparison chart of diagnostic accuracy and average recall rate

CONCLUSION

This paper proposes and validates a fault classification method for a marine centrifugal fan based on a combination of signal-to-image conversion and lightweight CNN. The main conclusions are as follows:

(1) CEEMDAN is used to decompose the original signal. The problems of mode aliasing and complex calculation are overcome effectively, which benefits the subsequent signal analysis.
(2) The 1-D signals are transformed into 2-D images through signal-to-image conversion. They can show the invisible information of the original signal, such as texture characteristics and energy distribution. In this way, the marine centrifugal fan's failure can be reflected.
(3) Lightweight CNN is used for fault adaptive extraction, which reduces the complexity of the mode and improves the operation efficiency. It can realize end-to-end fault diagnosis. The overall diagnosis accuracy can reach 99.3%. Compared with other models, this model has a high diagnostic accuracy while maintaining light weight. In addition, the model has good noise immunity.

As exploratory basic research, this method can provide a reference for the fault diagnosis of other equipment in the cabin and also for the research of intelligent fault diagnosis systems of ships. It should be noted that multiple compound faults may exist in marine centrifugal fans. In future research work, the diversity of the data set should be further improved.

REFERENCES

1. Y. H. Tan, J. D. Zhang, H. Tian, D. Y. Jiang, L. Guo, G. M. Wang, Y. J. Lin, "Multi-label classification for simultaneous fault diagnosis of marine machinery: A comparative study," Ocean Engineering, vol. 239, p. 109723, 2021, doi: 10.1016/j.oceaneng.2021.109723.

2. G. H. Yan, Y. H. Hu, J. W. Jiang, "A Novel Fault Diagnosis Method for Marine Blower with Vibration Signals," Polish Maritime Research, vol. 29, no. 2, pp. 77-86, 2022, doi:10.2478/POMR-2022-0019.

3. Y. Xie and T. Zhang, "Fault Diagnosis for Rotating Machinery Based on Convolutional Neural Network and Empirical Mode Decomposition," Shock and Vibration, vol. 2017, pp. 11-12, 2017, doi: 10.1155/2017/3084197.

4. Z. Guan, Z. Liao, K. Li, and P. Chen, "A Precise Diagnosis Method of Structural Faults of Rotating Machinery based on Combination of Empirical Mode Decomposition, Sample Entropy, and Deep Belief Network," Sensors (Basel), vol. 19, no. 3, p. 591, 2019, doi: 10.3390/s19030591.

5. M. Kuai, G. Cheng, Y. Pang, and Y. Li, "Research of Planetary Gear Fault Diagnosis Based on Permutation Entropy of CEEMDAN and ANFIS," Sensors (Basel), vol. 18, no. 3, p. 782, 2018, doi: 10.3390/s18030782.

6. R. Nishat Toma, C.-H. Kim, and J.-M. Kim, "Bearing Fault Classification Using Ensemble Empirical Mode Decomposition and Convolutional Neural Network," Electronics, vol. 10, no. 11, p. 1248, 2021, doi: 10.3390/ELECTRONICS10111248.

7. W. Jiang, Y. H. Xu, Z. Chen, N. Zhang, and J. Z. Zhou, "Fault diagnosis for rolling bearing using a hybrid hierarchical method based on scale-variable dispersion entropy and parametric t-SNE algorithm," Measurement, vol. 191, p. 110843, 2022, doi: 10.1016/j.measurement.2022.110843.

8. S. Zhou, M. H. Xiao, P. Bartos, M. Filip, and G. S. Geng, "Remaining Useful Life Prediction and Fault Diagnosis of Rolling Bearings Based on Short-Time Fourier Transform and Convolutional Neural Network," Shock and Vibration, vol. 2020, p. 8857307, 2020, doi: 10.1155/2020/8857307.

9. X. C. Zhang, H. W. Li, W. Y. Meng, Y. F. Liu, P. Zhou, C. He, Q. B. Zhao, "Research on fault diagnosis of rolling bearing based on lightweight convolutional neural network," Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 44, no. 10, p. 462, 2022, doi:10.1007/s40430-022-03759-6.

10. A. G. Howard et al., "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint, 2017, doi: 10.48550/arXiv.1704.04861.

11. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520, doi: 10.48550/arXiv.1801.04381.

12. X. Y. Zhang, X. Y. Zhou, M. X. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6848-6856, doi: 10.48550/arXiv.1707.01083.

13. N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient CNN architecture design," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 116-131, doi: 10.48550/arXiv.1807.11164.

14. S. Z. Hou, W. Guo, Z. Q. Wang, and Y. T. Liu, "Deep-Learning-Based Fault Type Identification Using Modified CEEMDAN and Image Augmentation in Distribution Power Grid," IEEE Sensors Journal, vol. 22, no. 2, pp. 1583-1596, 2022, doi: 10.1109/Jsen.2021.3133352.

15. M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in 2011 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2011, pp. 4144-4147, doi: 10.1109/ICASSP.2011.5947265.

16. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Communications of the ACM, vol. 60, no. 6, pp. 84-90, 2017, doi: 10.1145/3065386.

17. S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3-19, doi: 10.48550/arXiv.1807.06521.

18. J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132-7141, doi: 10.48550/arXiv.1709.01507.

19. L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," Journal of Machine Learning Research, vol. 9, pp. 2579-2605, 2008.

20. K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.