

Article

Multi-Scale Feature Fusion Convolutional Neural Networks for Fault Diagnosis of Electromechanical Actuator

Yutong Song ¹ , Jinhua Du ^{1,*}, Shixiao Li ², Yun Long ¹, Deliang Liang ¹, Yifeng Liu ¹ and Yao Wang ¹

¹ The State Key Laboratory of Electrical Insulation and Power Equipment, Xi'an Jiaotong University, Xi'an 710049, China; yutongsong0204@163.com (Y.S.); jfjlgdxy2017@stu.xjtu.edu.cn (Y.L.); dliang@mail.xjtu.edu.cn (D.L.); l66351191@stu.xjtu.edu.cn (Y.L.); deborah0502@stu.xjtu.edu.cn (Y.W.)

² Langfang Power Supply Company, State Grid Jibei Electric Power Co., Ltd., Langfang 065000, China; shixiao_lee@163.com

* Correspondence: jinhuaadu@mail.xjtu.edu.cn

Abstract: Airborne electromechanical actuators (EMAs) play a key role in the flight control system, and their health condition has a considerable impact on the flight status and safety of aircraft. Considering the multi-scale feature of fault signals and the fault diagnosis reliability for EMAs under complex working conditions, a novel fault diagnosis method of multi-scale feature fusion convolutional neural network (MSFFCNN) is proposed. Leveraging the multiple different scales' learning structure and attention mechanism-based feature fusion, the fault-related information can be effectively captured and learned, thereby improving the recognition ability and diagnostic performance of the network. The proposed method was evaluated by experiments and compared with the other three fault-diagnosis algorithms. The results show that the proposed MSFFCNN approach has a better diagnostic performance compared with the state-of-the-art fault diagnosis methods, which demonstrates the effectiveness and superiority of the proposed method.

Keywords: fault diagnosis; multi scale; feature fusion; convolutional neural network; deep learning; electromechanical actuator



Citation: Song, Y.; Du, J.; Li, S.; Long, Y.; Liang, D.; Liu, Y.; Wang, Y. Multi-Scale Feature Fusion Convolutional Neural Networks for Fault Diagnosis of Electromechanical Actuator. *Appl. Sci.* **2023**, *13*, 8689. <https://doi.org/10.3390/app13158689>

Academic Editors: Adrian David Cheok and Jun Cai

Received: 3 July 2023
Revised: 21 July 2023
Accepted: 25 July 2023
Published: 27 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Due to the rapid development of semiconductor integrated circuits and digital control technology as well as the many problems existing in traditional airborne actuation systems, more electric aircraft (MEA) technology has become the main trend of modern aviation development. It emphasizes the usage of electric power instead of hydraulic power, pneumatic power, etc.; consequently, it has the advantages of light weight, high power density and efficiency, low cost, simple testing and maintenance, etc. [1]. At present, the most representative MEA of civil aircraft are Airbus A380 and Boeing 787 aircraft, and that of military aircraft is the F-35 fighter. Airbus A380 and F-35 fighter have already used electro-hydrostatic actuators to drive the primary flight control surface. Boeing 787 aircraft, on the other hand, applies electromechanical actuators (EMAs) to the secondary flight controls, like slat actuators and spoiler actuators [2]. With the advancement of power-by-wire technology, EMA, which completely abandons the hydraulic system, is bound to have better application prospects. EMA can be divided into two forms according to different driving methods: ball screw EMA with gear reduction mechanism and direct-drive ball screw EMA. The direct-drive EMA, which directly integrates the ball screw pair with the motor in its structure, eliminates the need for a gear reduction mechanism, so it has the advantages of high reliability, high efficiency, and high integration. However, as an emerging technology, achieving the same level of reliability as hydraulic servo systems is challenging for direct-drive EMA, thus limiting its large-scale application [3]. In order to ensure the safe operation and economic maintenance of aircraft, it is of great significance to accurately and timely diagnose and predict the real-time status of direct-drive EMA.

In recent years, data-driven intelligent fault diagnosis are widely used in the field of fault diagnosis. In most of the existing data-based EMA fault-diagnosis methods, feature extraction and classification are designed and perform separately rather than as a single entirety, so both cannot be optimized simultaneously. For example, dynamic response indicators derived from vibration and current signals of EMA can be utilized as classification features [4,5]. Mode decomposition [6] and information entropy [7] are also used to constitute the feature vector of EMA. Additionally, principal component analysis (PCA) is often used for the selection and fusion of EMA fault features [8,9]. After selecting the appropriate features, these features are fed into a classifier, such as BP neural network, support vector machine, and random forest, for fault classification [10,11]. The features of the aforementioned traditional data-based methods usually rely on manual design, requiring certain signal processing technology and diagnostic expertise, which can be unreliable and time-consuming. Moreover, they are mostly used in specific fields and cannot be updated online with changes in application equipment or fields. To address the defects of traditional fault-diagnosis approaches, deep learning methods such as stacked denoising autoencoder, deep belief network, and convolutional neural network (CNN) [12,13] have demonstrated great vitality in these years. Unlike traditional methods, deep learning can directly learn the effective fault features adaptively from monitoring signals and perform the fault classification at the same time, thereby achieving end-to-end fault diagnosis. Particularly, the CNN has achieved superior performance in various fault-diagnosis tasks [14–16] due to its features of weight sharing, local connection, and multiple convolution kernels. The superior performance of one-dimensional CNN (1DCNN) in the EMA fault-diagnosis problem compared to traditional data-based methods has been demonstrated in prior work [17]. Therefore, this paper aims to develop a CNN-based intelligent fault-diagnosis model for EMA.

Due to the interaction between the subsystems of fault-identification objects, the fault signals usually appear in multi-scale form. Therefore, conventional CNN methods, which only contain single-scale convolution kernels, may ignore the fault-related information. To address the challenge, several multi-scale network structures have been proposed and have achieved impressive performance. Introducing multi-scale transformation into conventional CNN can enable the network to acquire features of different receptive fields at the same level and improve the diversity and complementarity of fault-related features. For example, Jiang et al. [18] proposed a multi-scale convolutional neural network structure that can effectively extract multi-scale high-level features. The method was verified on a WT gearbox. Liu et al. [19] proposed a multiscale kernel CNN to capture the patterns of motor faults. Peng et al. [20] used a similar multi-branch structure in the feature learning process.

Despite the good performance, those multi-scale CNNs mentioned above simply combine the captured multi-scale features without taking into account differences in the importance of different branches. Therefore, the information related to the fault may not be used effectively, especially in the case of noise interference and load variation. In this paper, the fault-diagnosis object EMA often needs to operate under complex working conditions, such as the complexity and variability of work tasks, the non-linearity and non-stationarity of the signal caused by changes in speed and load, as well as strong environmental noise. To reduce the influence of the above factors, it is required that the intelligent diagnosis algorithm has better adaptability and can reduce the sensitivity to various uncertainty sources. In recent years, attention mechanisms have been widely utilized to achieve efficient resource allocation and information-capture capability of models for intelligent fault diagnosis [21], natural language processing, machine vision, and the broader fields of deep learning [22,23]. Li et al. [24] demonstrated the effectiveness of attention mechanisms in bearing intelligent fault diagnosis by locating input data segments and visualizing network-learned diagnosis knowledge. Ding et al. [25] proposed a wind turbine blade intelligent anomaly-detection method based on attention mechanism, which solves the memory occupancy problem of the input sequence and improves the accuracy of anomaly detection. Kong et al. [26] proposed an attentional recursive autoencoder hybrid model

classification algorithm for the early fault diagnosis of rotating machinery. The network is able to extract the most valuable features from the input signal.

In order to make the network fully capture the fault-related features at different scales and levels and reduce sensitivity to various sources of uncertainty, a multi-scale feature fusion CNN (MSFFCNN) is herein proposed to diagnose EMA faults. The multi-scale network structure is applied to EMA fault diagnosis for the first time. Moreover, unlike the aforementioned multi-scale methods, an attention mechanism module is used. Firstly, the attention module enables effective recalibration of feature channels, enhancing essential features and suppressing invalid ones. This improves the model's ability to focus on the most relevant information. Secondly, it further recalibrates features learned by each branch and multiscale fusion features, aiming to aggregate the optimal multiscale features. This results in improved feature representation and better model performance. To verify the performance of the proposed algorithm, an EMA system experimental platform was built, and the fault injection experiment was carried out for several common faults. The collected data were used to train and test the proposed method compared with the state-of-the-art algorithms, especially under variable load and noise conditions. This paper is organized as follows. Section 2 describes the structure of the direct-drive EMA and introduces its common faults. Section 3 details the proposed MSFFCNN for EMA fault diagnosis. Section 4 details different experiments to verify the effectiveness and superiority of the MSFFCNN. Section 5 summarizes this paper.

2. Structure and Faults Analysis of the Direct-Drive EMA

2.1. Structure of the Direct-Drive EMA

The actuator is essentially a position servo control system, which drives the load by controlling the operation of the motor to achieve the target position control. The difference between EMA and other actuators is that there is only one type of energy transmission in the electromechanical actuation system, and mechanical transmission is used instead of hydraulic transmission. The structure of the direct-drive EMA is shown in Figure 1. It is mainly composed of a controller, power conversion circuit, motor, ball screw, and load and feedback components (current sensor, resolver, and LVDT linear displacement sensor).

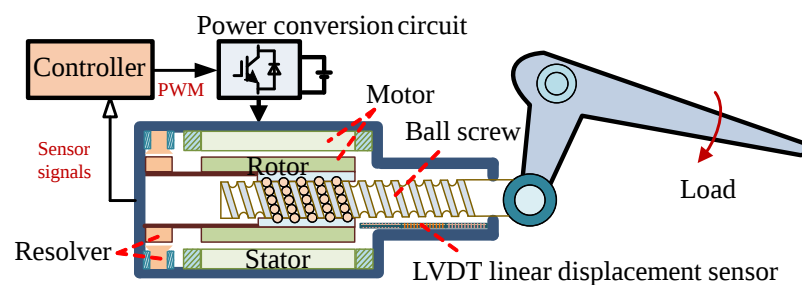


Figure 1. The structural diagram of the EMA.

Among them, the motor adopts a switched reluctance motor (SRM) with a certain fault tolerance and is matched with a direct-drive structure that cancels the gearbox, which greatly improves the reliability of EMA. When the EMA works, the controller controls the power converter by processing the flight commands as well as the feedback sensor signals, and the switching signals generated by the power converter control the motor rotation. The ball screw then converts the rotational motion of the rotor into a straight-line motion, which drives the swing of the aircraft control surface. In the whole process, the real-time measured position, velocity, and current information are fed back to the controller by the current sensor, i.e., the rotary transformer and the linear variable differential transformer (LVDT) so as to realize the closed-loop control.

This study built an dSPACE-based EMA system experimental platform, as shown in the Figure 2. The experimental platform is composed of PC, dSPACE, auxiliary power supply, IGBT drive circuit, rotary decoding circuit, power supply, EMA, load rudder,

sensors, and oscilloscope. The dSPACE hardware platform is mainly used to provide a real-time control platform for semi-real objects. In this experiment, DS1103PPC was used to output 8 PWM motor control signals through the digital I/O port of the main processor to control the opening and closing of IGBT. The SRM current analog signal output by the Hall sensor and the LVDT position signal output by the decoder board are collected by the ADC module. In addition, through the incremental encoder interface, the rotary signal of the motor can be directly input to the DS1103PPC through the rotary decoding board so as to be used for the measurement of the motor position and speed.

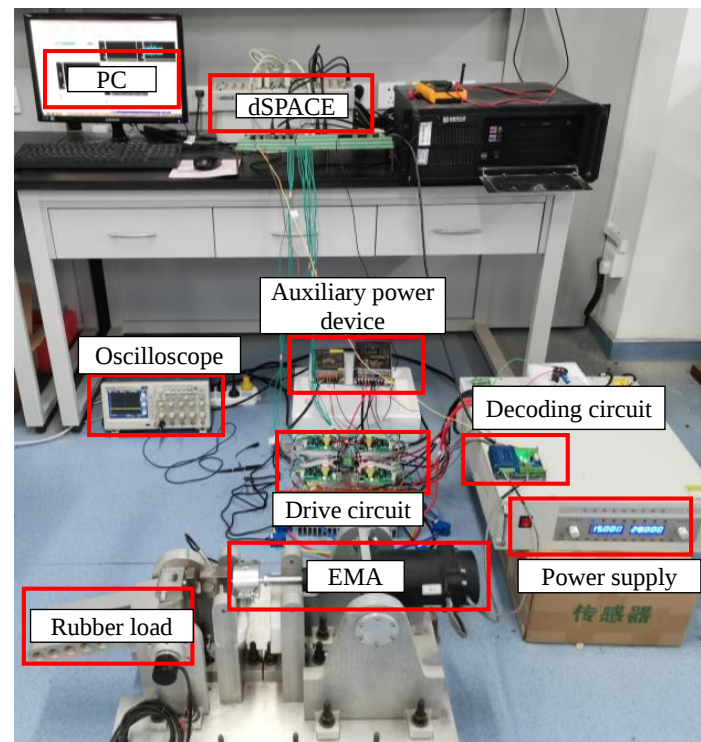


Figure 2. Experimental setup.

2.2. Fault Categories and Data Processing

From the perspective of the composition of EMA, there are four types of failures that may occur during operation: motor failure, electrical failure, mechanical failure, and sensor failure. Due to the harshness of the actual application conditions of EMA, such as overload, harsh environment, lubrication problems, and manufacturing defects, it is prone to mechanical failures. Furthermore, as a key component of the drive servo actuation system, the motor usually runs at a higher speed, accompanied by temperature rise in the housing and obvious mechanical stress, so the motor is prone to a winding short circuit and rotor-shaft eccentricity. Electrical faults mainly refer to faults in EMA's power supply and controller and sensor faults.

This paper comprehensively considers the three factors of EMA fault-occurrence frequency, degree of influence, and similarity of fault performance. The four faults of motor winding turn-to-turn short circuit, ball screw wear and jam, IGBT open circuit, and sensor deviation were selected for research. The specific experimental data and status labels are shown in Table 1. Among them, after the faults have been determined, the IGBT open circuit and the winding turns short circuit can be specific to the fault of a certain phase or a few phases according to the value of each phase's current sensor.

Table 1. Composition of experimental samples.

Label	Status	Length	Sample Type	Sample Size
C1	Normal	4096	Training	800
C2	Winding turn-to-turn short circuit			
C3	IGBT open circuit		Validation	200
C4	Ball screw wear and jam		Test	24
C5	Sensor deviation			

Taking the IGBT fault as an example, the open-circuit fault of IGBT is simulated by setting the driving signal of the IGBT to low at a certain time. Figure 3 shows the current and position signals of EMA before and after the fault. It can be seen that the fault-phase current of IGBT fault increases rapidly, and the position response is slightly deviated from that before the fault, which affects the performance of the EMA system.

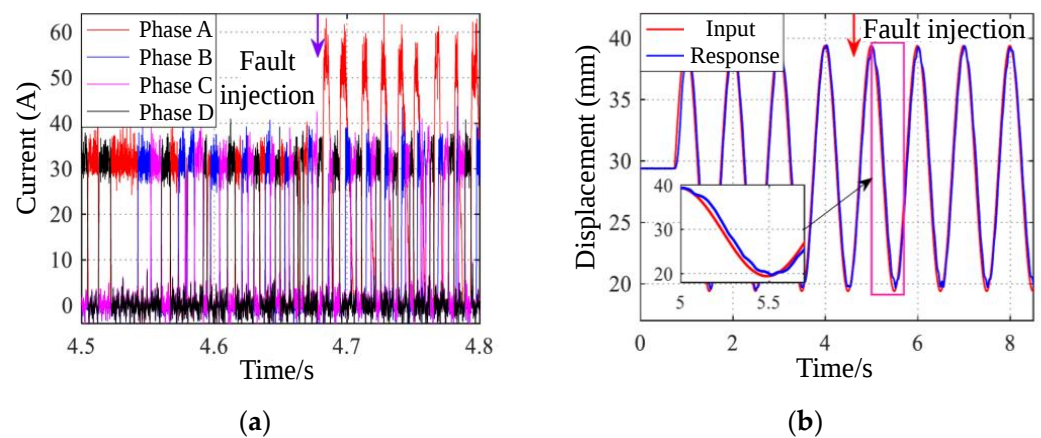


Figure 3. Experiment waveform of IGBT fault: (a) current waveform of IGBT open circuit; (b) position response of IGBT open circuit.

Under different fault conditions of EMA, the four-phase current signal output of EMA is collected at the sampling frequency of 10 kHz, and then, the sum of the four-phase current constitutes a signal. Figure 4 shows an example of the combined signals in five states, all of which have been normalized.

In order to facilitate the training of the convolutional neural network and reduce the interference of different working conditions on the model, each segment of the signal x is normalized. The size of the data is an important factor of the success of deep learning. Generally speaking, the more training samples of a network model, the better its generalization performance. Therefore, in the process of deep learning model training, data augmentation technology is often used; that is, without a substantial increase in data, limited data can generate value equivalent to more data. Taking into account that the output of the electromechanical actuator is a one-dimensional sequence signal, this paper adopts the overlapping sampling method to achieve the purpose of data amplification. The sampling method is shown in Figure 5.

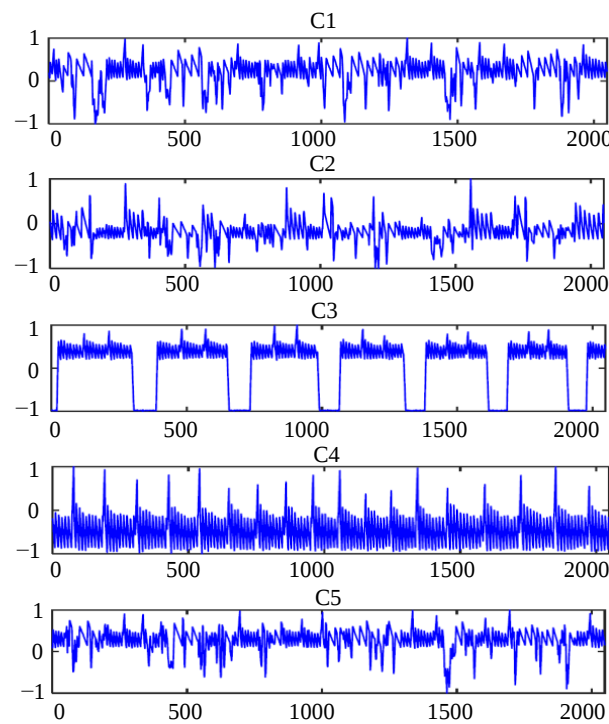


Figure 4. Signals for different states of EMA.

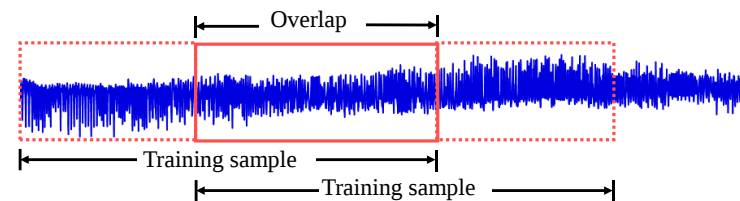


Figure 5. The data augmentation method.

The above combined signals are segmented with a certain overlap ratio so that the training set, verification set, and test set are obtained. The specific methods are as follows:

$$\begin{cases} m = \left\lfloor \frac{L-l}{l \times \lambda} \right\rfloor \\ x'_i = x'[((i-1) \times l \times (1-\lambda)) : ((i-1) \times l \times (1-\lambda) + l)]' \end{cases} \quad (1)$$

where m is the maximum number of divisible samples of each signal segment; L is the length of each signal; l is the set sample length; λ is the overlap rate; x'_i is the i -th data sample after segmentation, $i \in [1, m]$; x' is the normalized signal.

In this study, the sample length $l = 4096$ and the signal with $L = 180,000$ were divided with overlap rate $\lambda = 1/3$ to achieve sample-set amplification, resulting in a set of 6144 samples. From this set, 800 samples were randomly selected for each state as the training set, while 200 samples were reserved for validation, and 24 samples were used for test. The validation set can be utilized to monitor the occurrence of overfitting during model training. Typically, when the validation set's performance stabilizes, further training will cause the training set's performance to continue to improve, while the validation set's performance will plateau or even decline, indicating overfitting. All model accuracy tests were conducted solely on the test set to evaluate its generalization ability and ensure the reliability of the final results.

3. Proposed MSFFCNN-Based Fault-Diagnosis Method

The MSFFCNN is designed for fault diagnosis of EMA, which consists of four sequential stages: multi-scale transformation, feature learning, feature fusion, and fault classification. The overall process of the fault-diagnosis system is shown in the Figure 6.

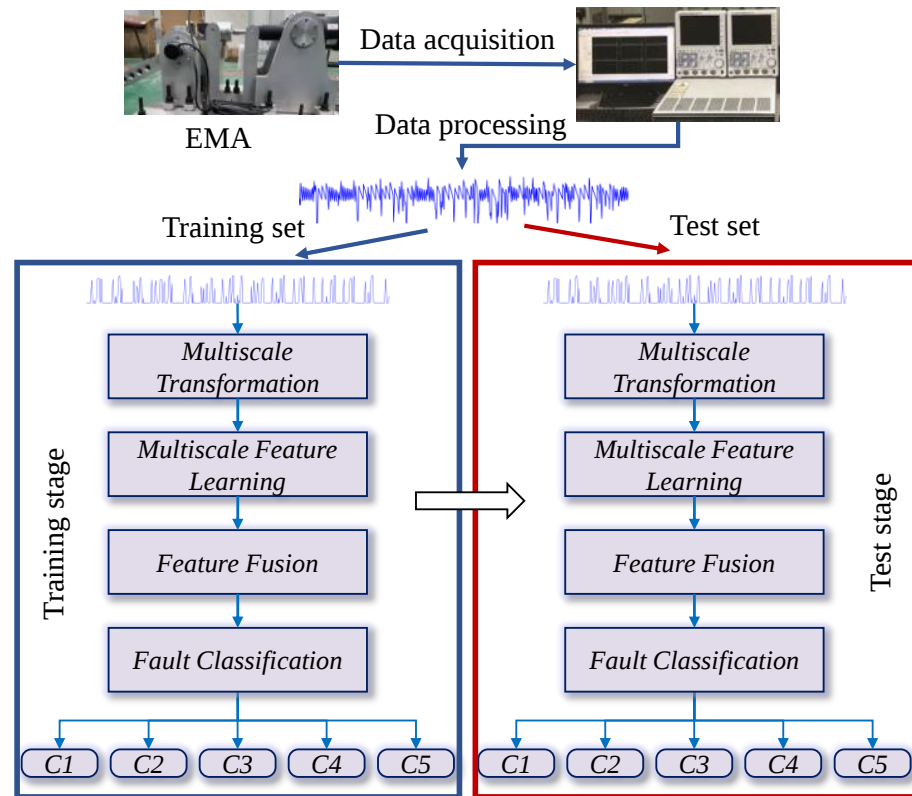


Figure 6. Flow chart of the MSFFCNN-based EMA fault-diagnosis system.

3.1. Multiscale Transformation

The so-called multi-scale is actually the sampling of different granularities of the signal. Typically, different features can be observed at different scales to complete different tasks. The multi-scale transformation stage in this paper adopts a parallel multi-branch topology. For a given 1-D signal x ($x \in \mathbb{R}^{N \times 1}$), multiple consecutive signals $\{y(k)\}$ with different granularities are constructed by a simple process of down-sampling. The representation method of multi-scale down-sampling is shown in Figure 7, and its mathematical description is as follows:

$$y_j^{(k)} = \frac{1}{k} \sum_{i=(j-1)k+1}^{jk} x_i, \quad 1 \leq j \leq \frac{N}{k}, \quad (2)$$

where k is the length of the non-overlapping window in down-sampling (also called the scale factor), and multiple filtered signals with different scales, respectively (i.e., scale1, scale2, and scale3) can be obtained. Typically, the size of k is related to the details and trends of feature learning later.

Following Figure 7, the original signal was sampled on three different scales to feed into the trunk and branch module (TBM) in the multi-scale feature learning stage.

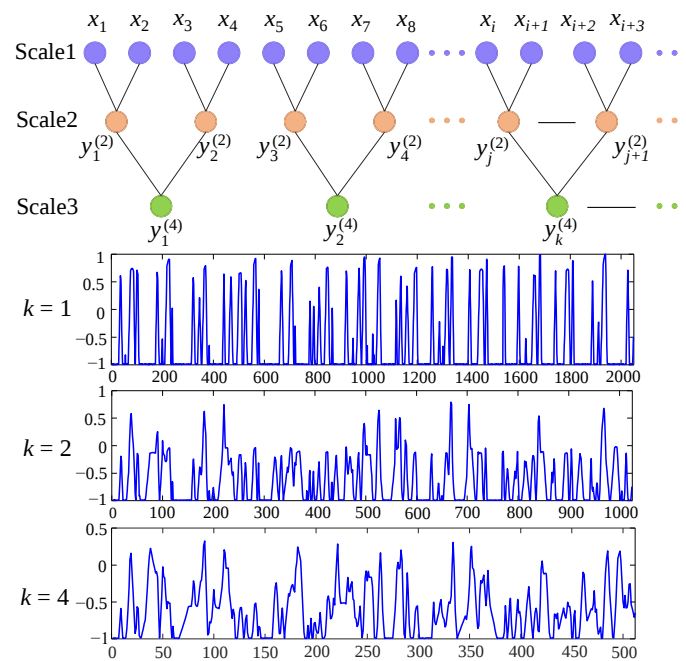


Figure 7. The representation method of multi-scale down-sampling.

3.2. Multiscale Transformation

After obtaining three granular signals with different scales, each granular signal $\{y^{(k)}\}$ ($k = 1, 2, 4$) will be fed into the TBM separately to learn useful and advanced features. As shown in Figure 8, each TBM consists of three sets of alternately stacked convolution layers and max-pooling layers.

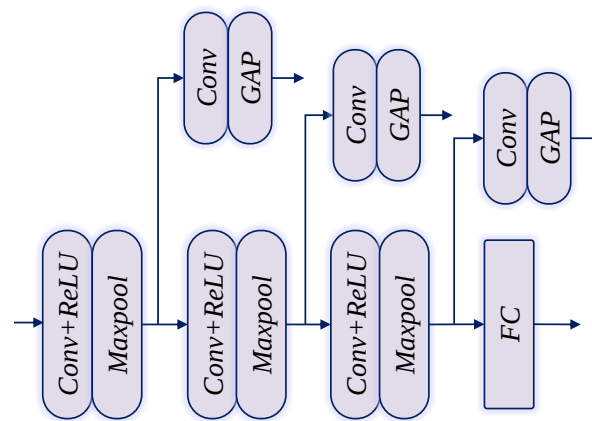


Figure 8. The trunk and branch module.

The signals pass through three pairs of stacked convolutional layers in parallel ($C1^{(k)}$, $C2^{(k)}$, and $C3^{(k)}$) and pooling layers ($P1^{(k)}$, $P2^{(k)}$, and $P3^{(k)}$), learning in an advanced and effective way from multiple granular signals of different time scales' fault characteristics. Specifically, each granular signal uses filters (convolution kernels) of different sizes so that each parallel convolutional layer at the same level can obtain the characteristics of different receptive fields and enhance its capture range of high- and low-frequency features taking into account the overall details of the input signal, thereby improving the diagnostic performance of the model.

The first convolutional layers ($C1^{(1)}$, $C1^{(2)}$, and $C1^{(4)}$) have signal lengths of N , $N/2$, and $N/4$, respectively. For each first convolution, the size m of the corresponding convolution kernel decreases as the value of k increases, which is beneficial to better extracting useful

features. Taking the i -th element a_i of the j -th output feature map $\mathbf{a}_j$ of the first convolutional layer ($C1^{(k)}$) as an example, the following is obtained:

$$a_i = \sigma(\mathbf{w}^T y_{i:i+m-1} + b), \quad (3)$$

$$\mathbf{a}_j = [a_1, a_2, \dots, a_i, \dots, a_{(L-m)/s+1}]. \quad (4)$$

Among them, $\mathbf{w}$ is the weight vector of the j -th convolution kernel; b is the bias term; $y_{i:i+m-1}$ is the m -length sub-signal of the input signal y starting from the i -th time step; $\sigma(\cdot)$ is a nonlinear activation function, namely the rectified linear unit (ReLU); s is the moving stride of the j -th convolution kernel on the signal y .

In order to increase the sparsity of the model and improve the speed of network training, maximum pooling is used to perform nonlinear down-sampling of the input feature map. Its advantage is that position-independent features can be obtained. Suppose that one input feature map is traversed with the pooling step size w , and one can be calculated for each sliding step w . For the corresponding local maximum p_j , there will be $(L-m)/ws + 1$ local maximum at the end of the traversal and will finally constitute the feature map $\mathbf{p}_t$ output by the maximum pooling layer ($P1^{(k)}$). The specific mathematical formulas are as follows:

$$p_j = \max_{(j-1)w+1 \leq i \leq jw} \{a_i\}, \quad (5)$$

$$\mathbf{p}_t = [p_1, p_2, \dots, p_j, \dots, p_{(L-m)/ws+1}]. \quad (6)$$

For each granular signal $\{y^{(k)}\}$ after $C1^{(k)}$ and $P1^{(k)}$, a certain number of new feature maps are generated. Then, these feature maps are used as the input of $C2^{(k)}$, and the same operations in Equations (3)–(6) are repeated, outputting a new feature map. Similarly, assuming that K convolution kernels are used in $C3^{(k)}$, the output of the max-pooling layer ($P3^{(k)}$) is K new feature maps, and take $\mathbf{q}^{(k)}$ as the concatenation result of the corresponding feature maps obtained after each granularity signal $\{y^{(k)}\}$ through the above process, formalized as follows:

$$\mathbf{q}^{(k)} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_K]. \quad (7)$$

Finally, the feature representation $\mathbf{q}^{(k)}$ output after each granularity signal $\{y^{(k)}\}$ undergoes continuous feature learning is flattened into a one-dimensional feature vector $\mathbf{q}$. Since the scale factor $k = 1, 2, 4$ is selected in this paper, vector $\mathbf{q}$ can be represented by the following:

$$\mathbf{q} = [\mathbf{q}^{(1)}, \mathbf{q}^{(2)}, \mathbf{q}^{(4)}]. \quad (8)$$

It can be seen from (8) that the final feature representation $\mathbf{q}$ has three different scales. Therefore, compared with the traditional single-scale representation, multi-scale feature learning has a larger feature-capture range, which is conducive to extracting rich and complementary features. The characteristics of the system provide a better distinguishing effect for the next step of fault classification.

3.3. Feature Fusion

Multi-scale feature learning realizes simple concatenations of different scale features. However, it cannot represent the difference in the importance of features. Therefore, an effective feature fusion mechanism is needed.

In this paper, an attention mechanism module is used after the feature fusion layer to distribute the weight of multi-scale feature channels. The network can selectively strengthen the useful features for fault identification and suppress invalid or even wrong information. The structure of the efficient channel attention module is shown in Figure 9.

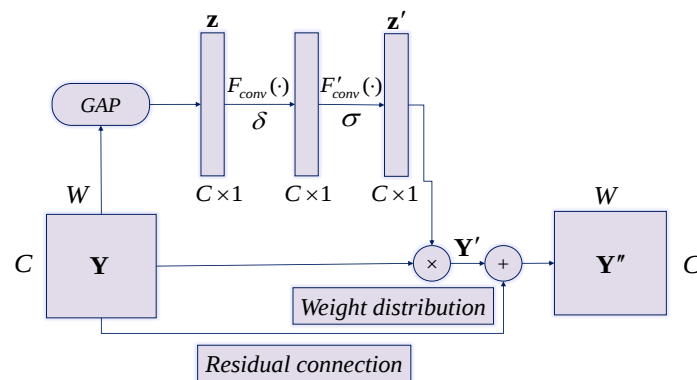


Figure 9. The attention mechanism module.

Assume that the input feature graph of the attention module is $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_C]$ ($\mathbf{y}_i \in \mathbf{R}^{W \times 1}$), where W and C are size and channel dimension of the feature graph. By using global average pooling F_{Avg} to compress information of the feature graph $\mathbf{Y}$, the channel statistical vector $\mathbf{z}$ is obtained:

$$z_i = F_{Avg}(\mathbf{y}_i) = \frac{1}{1 \times W} \sum_{j=1}^W \mathbf{y}_i(j). \quad (9)$$

After that, two fast one-dimensional convolutions are used to self-adaptively encode the channel correlation. The importance of different channels is quantified by activating function σ , thus generating the weight vector $\mathbf{z}'$ of channels. The mathematical description of this process is shown as follows:

$$\mathbf{z}' = \sigma(F'_{conv}(\delta(F_{conv}(\mathbf{z})))), \quad (10)$$

where F_{conv} is the convolution operation using the convolution kernel of $1 \times k$ and channel vector $\mathbf{z}$; F'_{conv} is the convolution operation using the convolution kernel of $1 \times k$ and vector after F_{conv} ; δ and σ are the ReLU function and Sigmoid function, respectively.

By multiplying the input feature graph $\mathbf{Y}$ and the weight vector $\mathbf{z}'$ of the channel, the calibrated feature graph $\mathbf{Y}'$ of the channel can be obtained:

$$\mathbf{Y}' = \mathbf{Y} \cdot \mathbf{z}' = [\mathbf{y}_1 z'_1, \mathbf{y}_2 z'_2, \dots, \mathbf{y}_C z'_C], \quad (11)$$

where z'_i represents the importance of the corresponding channel.

In order to prevent network degradation and improve its generalization performance, a residual connection is added after channel calibration; thus, the output of the attention mechanism module is $\mathbf{Y}'' = \mathbf{Y} + \mathbf{Y}'$.

3.4. Fault Classification

A combination of a fully connected hidden layer and a softmax layer is used to perform classification. The specific method is to first use dropout on the one-dimensional feature vector $\mathbf{q}$ obtained in the previous stage and use it as the input of the fully connected layer. The hidden layer uses ReLU as the activation function, and the softmax function is used in the output layer. In this paper, Y represents the category label of EMA's health status. Suppose it has n categories; that is, given an input sample x , the probability that sample x belongs to category c is given:

$$p(Y = c | x; \Theta) = \text{softmax}(\theta_c^T x) = \frac{\exp(\theta_c^T x)}{\sum_{j=1}^n \exp(\theta_j^T x)}, \quad (12)$$

where $\Theta = [\theta_1, \theta_2, \dots, \theta_n]$ is the parameter that needs to be learned in the model; $1/\sum_{j=1}^n \exp(\theta_j^T x)$ is the normalized function, and $\sum_{j=1}^n P_j = 1$ is the normalized function.

For any given input sample, MSFFCNN will predict a result, but it is hoped that the predicted value of the model is as consistent as possible with the true value. In order to achieve this goal, it is necessary to minimize the distance between the predicted value and the true value, which is the role of the loss function:

$$L(\theta) = -\frac{1}{m} [\sum_{i=1}^m \sum_{k=1}^K I\{y_i = k\} \log \frac{\exp(\theta_k^T x)}{\sum_{j=1}^K \exp(\theta_j^T x)}], \quad (13)$$

where m is the number of samples or the input batch size; $I\{\cdot\}$ is the index function, and when the $I\{\cdot\}$ value is true, the index function value is 1; otherwise, the index function value is 0.

In order to minimize the loss function value of the model, it is necessary to optimize and adjust the weight of the neural network, and the optimizer uses the back-propagation algorithm to complete it:

$$\theta^* = \arg \min_{\theta} L(f(x; \theta), y), \quad (14)$$

where θ^* is the optimal parameter of the model; $L(\cdot)$ is the loss function; $f(\cdot)$ and y are the output value and target value of the model, respectively.

3.5. Visualization Analysis of MSFFCNN

In order to show the classification process of the MSFFCNN, the t-SNE technology is used to visualize half of the samples in the test set. Due to the large number of network layers, only the two-dimensional distribution under one branch is shown here, as shown in Figure 10.

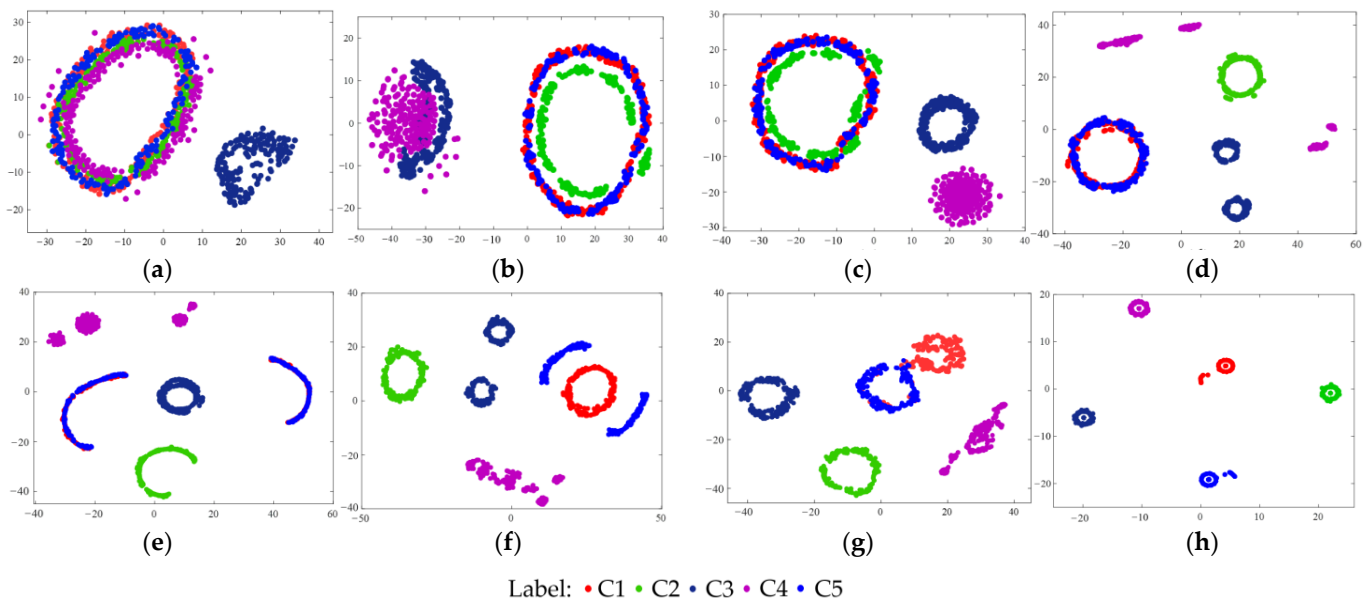


Figure 10. Two-dimensional visualization of the MSFFCNN: (a) original signal; (b) granular signal; (c–f) convolution layers; (g) feature fusion layer; (h) softmax layer.

As can be seen from Figure 10, samples of various categories of the original signal are jumbled together and completely indistinguishable. With the increase of the convolutional layers, all categories of originally linearly indivisible samples can be almost distinguished in the feature fusion layer and completely distinguishable in the softmax layer, which indicates that the nonlinear representation ability of the MSFFCNN is gradually enhanced.

In the softmax layer, all samples are kept very far apart from each other to avoid the occurrence of wrong classification, which indicates that the model has good robustness.

4. Discussion of the Fault-Diagnosis Results

In this section, the proposed MSFFCNN method is compared with conventional 1DCNN, CNN with wide first-layer kernels (WDCNN) [27], and multi-scale CNN(MSCNN) [18]. The WDCNN uses the wide kernels in the first convolutional layer to suppress high-frequency noise, but the multi-scale structure is not deployed; the MSCNN uses multi-scale transformation to combine the captured multi-scale features without taking into account differences in the importance of different branches. Therefore, comparison with the above algorithms can effectively reflect the superiority of multi-scale structure and the proposed attention mechanism-based feature fusion.

4.1. Validation Setup

The training parameters of the model are as follows: batch size is 64, training rounds are 100, the optimization algorithm is Adam, the initial learning rate is 0.001, and learning rate attenuation is set. In addition, the initialization of network weight follows the Glorot normal distribution initialization method.

The value of loss function and accuracy obtained by single training is shown in Figure 11. As can be seen, no matter the training set or the verification set, the MSFFCNN algorithm converges at about 20 epochs. In this case, the accuracy of the training set is stable at 99.9%, while the accuracy of the test set can reach about 99.8%.

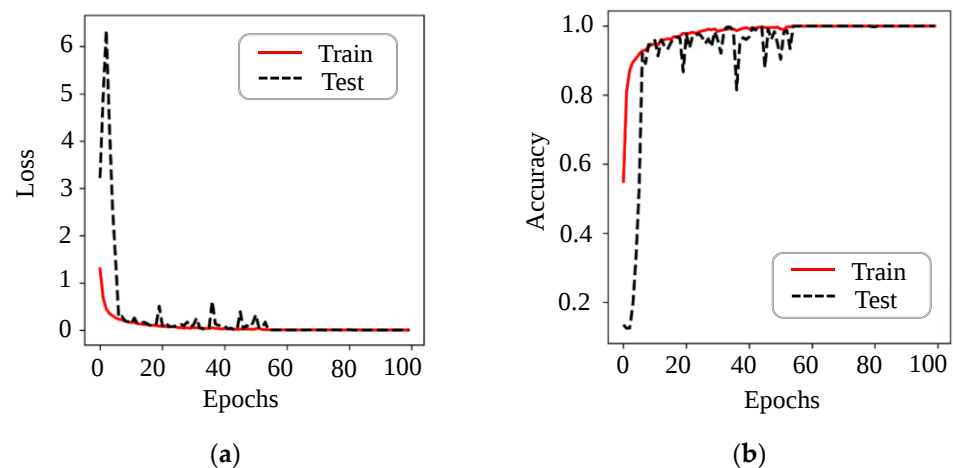


Figure 11. Training and validation of the MSFFCNN: (a) loss on training and test data; (b) accuracy on training and test data.

The constructed models were trained offline and then used for online fault diagnosis. Table 2 shows the comparison of the predicted time of a single sample between MSFFCNN and 1DCNN under five tests. As can be seen, MSFFCNN takes longer to test a single sample than 1DCNN. This makes sense because the MSFFCNN model has a deeper network structure and more parameters than 1DCNN, so it inevitably consumes more time when processing the test data. However, in practical engineering applications, a single forecast cost of 0.5 ms is totally acceptable for real-time diagnosis.

Table 2. Time of 1DCNN and MSFFCNN for single sample.

Model	Test Time of Single Sample (ms)				
	1st	2nd	3rd	4th	5th
1DCNN	0.314	0.305	0.282	0.306	0.294
MSFFCNN	0.524	0.513	0.508	0.516	0.522

4.2. Performance under Noise Environment

In order to comprehensively investigate the performance of the model, three commonly used indicators, namely accuracy, precision, and recall, are used in this section to measure the fault-classification performance of the MSFFCNN model, and the stability of the model is further evaluated through repeated tests. The definition of accuracy, precision, and recall are shown in Equations (15)–(17):

$$P_{Acc} = \frac{TP + TN}{TP + FN + FP + TN}, \quad (15)$$

$$P_{Pre} = \frac{TP}{TP + FP}, \quad (16)$$

$$P_{Rec} = \frac{TP}{TP + FN}, \quad (17)$$

where TP , FP , TN , and FN represent the number of true cases, false-positive cases, true-negative cases, and false-negative cases, respectively; P_{Acc} , P_{Pre} , and P_{Rec} represent accuracy, precision, and recall, ranging from 0 to 1; the higher the value, the better the model performance.

In practice, signals collected by EMA sensors are easily contaminated by ambient noise. Therefore, in this section, Gaussian white noise of different signal-to-noise ratios (SNR) is added to the original signal to simulate noise interference in the aviation environment. The definition of SNR is as follows:

$$SNR = 10\log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \quad (18)$$

where P_{signal} and P_{noise} represent the energy of signal and noise, respectively. The more noise contained in the signal, the smaller the SNR value.

In order to verify the anti-noise performance of the proposed MSFFCNN, the model was tested in five noise environments with SNR of -10 dB, -5 dB, 0 dB, 5 dB, and 10 dB, respectively. The fault diagnosis results are shown in Figure 12. In the environment of SNR = 5 dB with weak noise, the accuracy, precision, and recall of MSFFCNN can reach more than 98%; in the environment of SNR = -10 with strong noise, the three performance indicators can also maintain at about 77%.

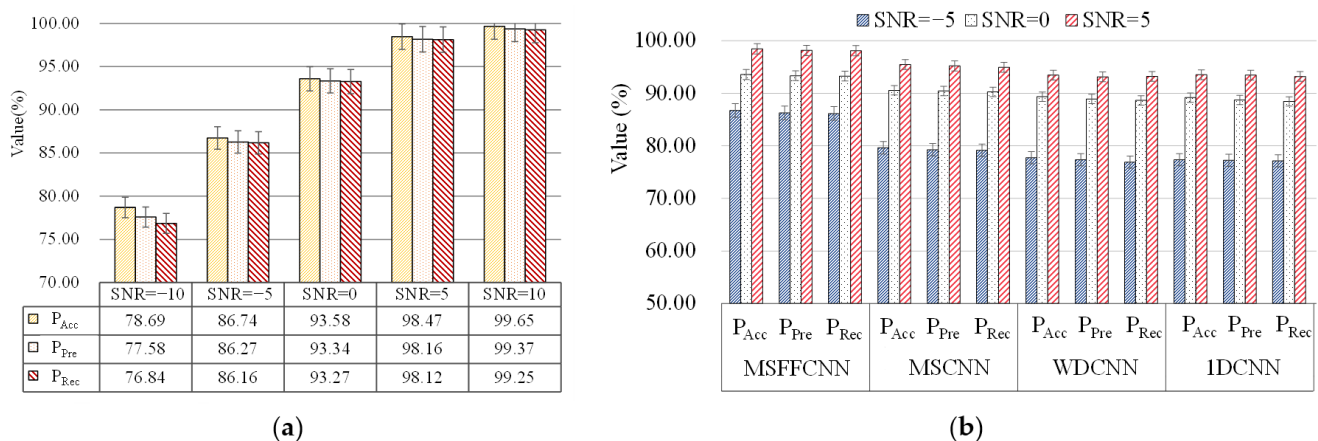


Figure 12. Performance comparison under noise environment: (a) performance of MSFFCNN under different SNR; (b) comparison with other methods under different SNR.

Furthermore, the performance difference between MSFFCNN and 1DCNN as well as the current representative MSCNN and WDCNN under the three noise environments of SNR = -5 , SNR = 0 , and SNR = 5 were compared. The results are shown in Figure 12.

As can be seen, the accuracy, precision, and recall of MSFFCNN are significantly higher than those of other reference methods under three different noise environments. MSCNN has the second-best anti-noise performance, and its three indexes can reach more than 90% when SNR = 0. WDCNN and 1DCNN can achieve a similar level of diagnostic performance, with slightly worse anti-noise performance than MSCNN.

In addition, although the diagnostic performance of all algorithms decreases to varying degrees with the increase of noise, MSFFCNN still shows excellent anti-noise ability, which means that MSFFCNN has better fault feature learning ability and recognition ability. In summary, MSFFCNN has good robustness to noise and can meet the diagnostic requirements in the actual industrial environment of environmental noise and measurement interference.

4.3. Performance under Variable Load

When the aircraft is in flight, the EMA's workload changes according to the work tasks. Accordingly, the current and angle of the motor and the position response of the EMA will change, so will the signal measured by the sensor. Figure 13 shows the normalized status signals of the EMA under different loads.

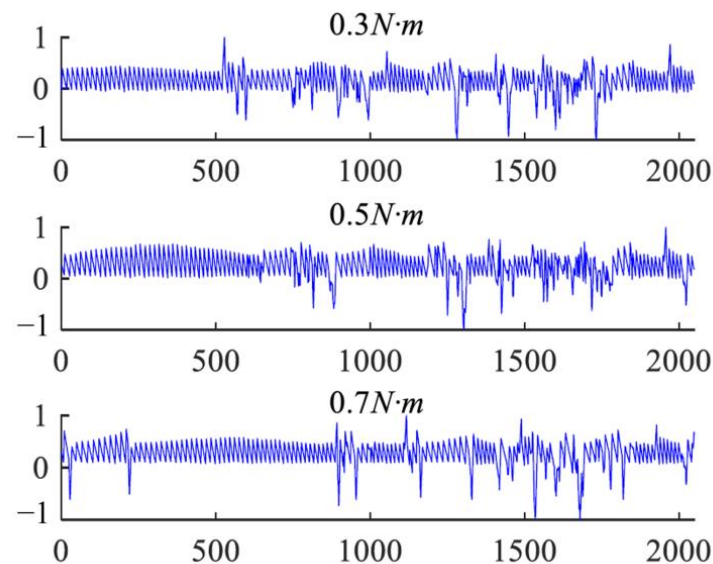


Figure 13. Normalized signals of the EMA under different loads.

As can be seen from Figure 13, there are certain differences in signal waveform and amplitude under different loads, and the greater the load changes are, the more obvious the differences will be, which will make the classifier unable to correctly classify the extracted features, thus reducing the accuracy of the intelligent diagnosis system. Therefore, it is of great practical significance to use the data under a single load to train the model and then use the trained model to fault diagnose the signal when the load changes.

In this section, the MSFFCNN is trained with data under loads of 0.3 N·m, 0.5 N·m, and 0.7 N·m, respectively, and then, the signals under the other two loads are used as the test set. Data acquired under 0.3 N·m, 0.5 N·m, and 0.7 N·m loads are defined as datasets A, B, and C respectively. After combination, there are six test conditions, namely A→B, A→C, B→A, B→C, C→A, and C→B. Moreover, in order to verify the reliability of the results, MSCNN, WDCNN, and 1DCNN are compared as the benchmark models. The variable load test results are shown in Figure 14.

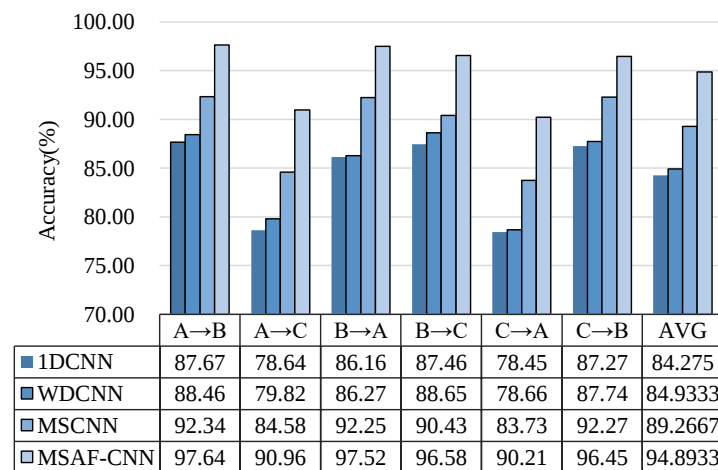


Figure 14. Performance comparison under variable load.

As shown in Figure 14, it was found that the diagnostic performance of 1DCNN and WDCNN are comparable under six variable load conditions, with an average diagnostic accuracy of about 85%. In contrast, the average diagnostic accuracy of the MSCNN model is nearly 5% higher than the previous two models, indicating that multi-scale feature learning has strong adaptability to variable load. The average diagnostic accuracy of the MSFFCNN proposed in this paper is about 95% under six variable load conditions, which is 10% higher than that of 1DCNN and WDCNN. To summarize, it can be shown that MSFFCNN has strong self-adaptive ability of variable load and can adapt to variable working environments.

5. Discussion

This paper proposes the MSFFCNN and introduces the multi-scale and feature fusion mechanism into the traditional CNN for EMA fault diagnosis.

Compared with the traditional CNN, the method in this paper utilizes a multi-scale structure to effectively and adaptively extract multi-scale, high-level features at different time scales. Moreover, the proposed attention mechanism can enhance the multi-scale features related to faults, thereby achieving a high diagnostic accuracy rate in the case of strong noise and variable loads. In order to evaluate the superiority of the proposed method in real-world industrial environment, this paper establishes a fault-diagnosis platform of the EMA system and conducts fault-injection experiments for several typical faults. The experimental results demonstrate that the proposed method has better performance than several state-of-the-art methods in scenarios with strong noise and variable loads.

Author Contributions: Conceptualization, S.L.; Methodology, Y.S.; Software, Y.S. and S.L.; Validation, S.L.; Formal analysis, Y.L. (Yun Long); Investigation, Y.L. (Yun Long); Resources, D.L.; Data curation, Y.L. (Yifeng Liu) and Y.W.; Writing—original draft, Y.S.; Writing—review & editing, J.D. and Y.L. (Yun Long); Visualization, Y.L. (Yifeng Liu) and Y.W.; Supervision, J.D. and D.L.; Project administration, J.D.; Funding acquisition, J.D. and D.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of China, grant number 52277065, and the National Key R&D Program of China, grant number 2020YFA0710500.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sayed, E.; Abdalmagid, M.; Pietrini, G.; Sa'adeh, N.-M.; Callegaro, A.D.; Goldstein, C.; Emadi, A. Review of Electric Machines in More-/Hybrid-/Turbo-Electric Aircraft. *IEEE Trans. Transp. Electrification* **2021**, *7*, 2976–3005. [\[CrossRef\]](#)
2. Garcia, A.; Cusido, I.; Rosero, J.A.; Ortega, J.A.; Romeral, L. Reliable Electro-Mechanical Actuators in Aircraft. *IEEE Aerosp. Electron. Syst. Mag.* **2008**, *23*, 19–25. [\[CrossRef\]](#)
3. Yin, Z.; Hu, N.; Chen, J.; Yang, Y.; Shen, G. A Review of Fault Diagnosis, Prognosis and Health Management for Aircraft Electromechanical Actuators. *IET Electr. Power Appl.* **2022**, *16*, 1249–1272. [\[CrossRef\]](#)
4. Ruiz-Cárcel, C.; Starr, A. Development of a Novel Condition Monitoring Tool for Linear Actuators. In Proceedings of the 12th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Oxford, UK, 9–11 June 2015.
5. Watson, M.; Smith, M.; Kloda, J.; Byington, C.; Semega, K. Prognostics and Health Management of Aircraft Engine EMA Systems. In Proceedings of the ASME 2011 Turbo Expo: Turbine Technical Conference and Exposition. American Society of Mechanical Engineers Digital Collection, Vancouver, BC, Canada, 6–10 June 2011.
6. Liu, H.; Jing, J.; Ma, J. Fault Diagnosis of Electromechanical Actuator Based on VMD Multifractal Detrended Fluctuation Analysis and PNN. *Complexity* **2018**, *2018*, e9154682. [\[CrossRef\]](#)
7. Chen, J.; Wang, L. Electromechanical Actuator Modeling and Its Application in Fault Diagnosis. In Proceedings of the 2018 International Conference on Mechanical, Electronic, Control and Automation Engineering, Qingdao, China, 30–31 March 2018.
8. Riaz, N.; Shah, S.I.A.; Rehman, F.; Khan, M.J. An Intelligent Hybrid Scheme for Identification of Faults in Industrial Ball Screw Linear Motion Systems. *IEEE Access* **2021**, *9*, 35136–35150. [\[CrossRef\]](#)
9. Chirico, A.J.; Kolodziej, J.R. A Data-Driven Methodology for Fault Detection in Electromechanical Actuators. *J. Dyn. Syst. Meas. Control* **2014**, *136*, 041025. [\[CrossRef\]](#)
10. Lou, S.; Yang, C.; Wu, P.; Kong, L.; Xu, Y. Fault Diagnosis of Blast Furnace Iron-Making Process With a Novel Deep Stationary Kernel Learning Support Vector Machine Approach. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3521913. [\[CrossRef\]](#)
11. Sun, Y.; Zhang, H.; Zhao, T.; Zou, Z.; Shen, B.; Yang, L. A New Convolutional Neural Network With Random Forest Method for Hydrogen Sensor Fault Diagnosis. *IEEE Access* **2020**, *8*, 85421–85430. [\[CrossRef\]](#)
12. Wang, J.; Zhang, Y.; Luo, C.; Miao, Q. Deep Learning Domain Adaptation for Electro-Mechanical Actuator Fault Diagnosis Under Variable Driving Waveforms. *IEEE Sens. J.* **2022**, *22*, 10783–10793. [\[CrossRef\]](#)
13. Yang, J.; Guo, Y.; Wanli, Z. An Intelligent Fault Diagnosis Method for an Electromechanical Actuator Based on Sparse Feature and Long Short-Term Network. *Meas. Sci. Technol.* **2021**, *32*, 095102. [\[CrossRef\]](#)
14. Kumar, P.; Shankar Hati, A. Convolutional Neural Network with Batch Normalisation for Fault Detection in Squirrel Cage Induction Motor. *IET Electr. Power Appl.* **2021**, *15*, 39–50. [\[CrossRef\]](#)
15. Ren, L.; Jia, Z.; Wang, T.; Ma, Y.; Wang, L. LM-CNN: A Cloud-Edge Collaborative Method for Adaptive Fault Diagnosis With Label Sampling Space Enlarging. *IEEE Trans. Ind. Inform.* **2022**, *18*, 9057–9067. [\[CrossRef\]](#)
16. Gong, R.; Tang, Z. Further Investigation of Convolutional Neural Networks Applied in Computational Electromagnetism under Physics-Informed Consideration. *IET Electr. Power Appl.* **2022**, *16*, 653–674. [\[CrossRef\]](#)
17. Li, S.S.; Du, J.H.; Long, Y. Fault diagnosis of electromechanical actuators based on one-dimensional convolutional neural network. *Trans. China Electrotech. Soc.* **2022**, *37*, 62–73. [\[CrossRef\]](#)
18. Jiang, G.; He, H.; Yan, J.; Xie, P. Multiscale Convolutional Neural Networks for Fault Diagnosis of Wind Turbine Gearbox. *IEEE Trans. Ind. Electron.* **2019**, *66*, 3196–3207. [\[CrossRef\]](#)
19. Liu, R.; Wang, F.; Yang, B.; Qin, S.J. Multiscale Kernel Based Residual Convolutional Neural Network for Motor Fault Diagnosis Under Nonstationary Conditions. *IEEE Trans. Ind. Inform.* **2020**, *16*, 3797–3806. [\[CrossRef\]](#)
20. Peng, D.; Wang, H.; Liu, Z.; Zhang, W.; Zuo, M.J.; Chen, J. Multibranch and Multiscale CNN for Fault Diagnosis of Wheelset Bearings Under Strong Noise and Variable Load Condition. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4949–4960. [\[CrossRef\]](#)
21. Lv, H.; Chen, J.; Pan, T.; Zhang, T.; Feng, Y.; Liu, S. Attention Mechanism in Intelligent Fault Diagnosis of Machinery: A Review of Technique and Application. *Measurement* **2022**, *199*, 111594. [\[CrossRef\]](#)
22. Niu, Z.; Zhong, G.; Yu, H. A Review on the Attention Mechanism of Deep Learning. *Neurocomputing* **2021**, *452*, 48–62. [\[CrossRef\]](#)
23. Guo, M.-H.; Xu, T.-X.; Liu, J.-J.; Liu, Z.-N.; Jiang, P.-T.; Mu, T.-J.; Zhang, S.-H.; Martin, R.R.; Cheng, M.-M.; Hu, S.-M. Attention Mechanisms in Computer Vision: A Survey. *Comp. Vis. Media* **2022**, *8*, 331–368. [\[CrossRef\]](#)
24. Li, X.; Zhang, W.; Ding, Q. Understanding and Improving Deep Learning-Based Rolling Bearing Fault Diagnosis with Attention Mechanism. *Signal Process.* **2019**, *161*, 136–154. [\[CrossRef\]](#)
25. Ding, J.; Lin, F.; Lv, S. Temporal Convolution Network Based on Attention for Intelligent Anomaly Detection of Wind Turbine Blades. In *Algorithms and Architectures for Parallel Processing, Proceedings of the 21st International Conference, ICA3PP 2021, Virtual Event, 3–5 December 2021*; Lai, Y., Wang, T., Jiang, M., Xu, G., Liang, W., Castiglione, A., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 193–209.

26. Kong, X.; Li, X.; Zhou, Q.; Hu, Z.; Shi, C. Attention Recurrent Autoencoder Hybrid Model for Early Fault Diagnosis of Rotating Machinery. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2505110. [[CrossRef](#)]
27. Zhang, W.; Peng, G.; Li, C.; Chen, Y.; Zhang, Z. A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals. *Sensors* **2017**, *17*, 425. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.