RESEARCH ARTICLE

RMCW: An Improved Residual Network With Multi-Channel Weighting for Machinery Fault Diagnosis

ZHENG LIU [1], HU YU [2], KUN XU[1], AND XIAODONG MIAO [1], (Member, IEEE)
[1]School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing 210000, China
[2]National Key Laboratory of Science and Technology on Helicopter Transmission, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Corresponding author: Xiaodong Miao (mxiaodong@njtech.edu.cn)

ABSTRACT Faced with increasingly complex industrial data, standard machine learning algorithms struggle to effectively extract both linear and nonlinear features. In this study, an improved residual network (ResNet) called Residual network with Independent Multi-Channel Weighting (RMCW) to tackle the nonlinear, temporally uncertain, and unevenly distributed fault. Firstly, a strategy for constructing the multi-channel vibration intrinsic mode function (IMF) images is designed to obtain the primary features by combing the empirical mode decomposition (EMD) and the gramian angular field (GAF). Secondly, a dynamic receptive field (DRF) with independent channel weighting is utilized to adaptively fuse the multi-channel features. This renders both initialization parameters for each individual channel and DRF parameters mutually independently adaptive to the fault features in the different batch. Thirdly, the RMCW model is built by inputting the fused features to the network of 9 residual building blocks. Two experimental cases verify that the propose method is effective for the machinery fault diagnosis and is superior to the comparing methods.

INDEX TERMS Feature fusion, fault diagnosis, deep learning, vibration signal.

I. INTRODUCTION

Bearings diagnosis is an important technical means of ensuring the healthy operation of rotating machinery. Due to the complex working environment and prolonged loading, bearings are prone to damage, which can have severe consequences. Therefore, it is crucial to promptly detect any faults in the bearings to prevent further damage [1], [2], [3].

In recent years, there has been extensive and multi-dimensional research conducted by scholars in the field of bearing fault diagnosis [4]. Signal processing was commonly used in relevant industries for bearing fault diagnosis, with methods such as Fourier transform, wavelet transform, variational mode decomposition (VMD) [5], [6], [7], and empirical mode decomposition (EMD) [8], [9], [10], being the most

The associate editor coordinating the review of this manuscript and approving it for publication was Haidong Shao.

prevalent. Given the complexity of vibration signals and the diverse operating conditions, relying solely on a single method for analysis is insufficient. Therefore, it is valuable to focus on researching the processing of multiple signals in order to achieve a more comprehensive analysis of vibration signals under complex operating conditions. In addition, support vector machines (SVMs) [11], [12], [13] and Artificial neural networks (ANNs) [14] were widely utilized in various applications as machine learning algorithms. For the fault diagnosis of automotive intelligent steering systems, Shi et al. [15] used a data-based SVM algorithm. They then combined a grey wolf optimizer with an undersampling procedure to optimize the original SVM and enhance the diagnostic performance. This approach proved effective in improving the accuracy and effectiveness of fault diagnosis in these systems. Movsessian et al. [16] proposed a new structural inspection framework within a traditional ANN to improve

the reliability of flaw detection and validated it with metrics such as accuracy, F1 score, and Matthews correlation coefficient with good results. Despite the remarkable advancements in fault diagnosis achieved by machine learning, the network's simplistic structure and shallow mechanism make it difficult to capture deeper hidden features in higher dimensions. This limitation often requires heavy reliance on prior feature engineering and expert experience for fault feature extraction [17].

Fault diagnosis has seen extensive research on deep learning methods [18], [19], [20], [21], including deep belief networks (DBNs) [22], [23], [24], stacked autoencoders (SAEs) [25], convolutional neural networks (CNNs) [26], [27], [28] [29], and recurrent neural networks. Deep neural networks compensate for the limitations of shallow learning machines in feature extraction and can adaptively identify hidden features, reducing the reliance of fault diagnosis algorithms on data preprocessing. For both 1D and 2D data with multiple channels, convolutional neural networks (CNNs) demonstrate outstanding performance in visual image analysis and artificial intelligence tasks. Their abilities to effectively process such data make them highly suitable for various applications [30]. The high performance of CNNs in various data forms has resulted in numerous variants with specialities such as the visual geometry group (VGG), GoogLeNet, and residual networks (ResNets). Zhao et al. [31] presented an improved ResNet by incorporating a systolic module that integrates the soft threshold denoising joint from traditional signal processing. They conducted experiments with different signal-to-noise ratios to compare the performance of the proposed model with traditional CNNs and the original ResNet. Zhao et al. [32] tackled this challenge by employing multi-mode jump fusion of fault features. They incorporated the inception module to enable seamless network connectivity without modifying the CNN module, thereby enhancing the extraction of fault feature information. The past experimental results have demonstrated that the proposed method maintained a diagnostic accuracy above 90% even under different solid noise conditions.

Using convolutional operations to adapt to neural networks is a more effective approach. To secondarily process the original data features before engaging the network model, some scholars have utilized joint processing of feature information [33], [34], [35]. Dong et al. [36] presented a fault diagnosis method for rolling bearings that combined multilayer noise reduction technology with an improved CNN. By applying singular value decomposition (SVD) to process the training samples, they obtained the intrinsic mode function (IMF). Subsequently, the IMF was superimposed on the original signal to obtain the final signal and then handed over to the improved CNN for fault diagnosis. Wu et al. [37] proposed a dual-wavelet denoising method that can retain more details of the original signal, and comparison tests showed that dual-wavelet denoising has a better performance in obtaining details of the original fault features. The experimental results showed that a small amount of feature engineering

can improve the performance of network training and bring the diagnostic effect of network models with traditional or simple structures closer to that of more advanced network structures. Signal preprocessing combined with neural network diagnosis focuses on raw sequences noise reduction and feature extraction using 1D CNNs [38], [39]. However, CNNs are primarily used for image feature extraction, and 1D CNNs cannot fully exploit the capabilities of image feature extraction [40], [41], [42].

Enlightened by feature extraction and data fusion theory, we propose an improved deep learning model RMCW, residual network with multi-channel feature weighting, for fault detection of bearings or similar objects. Specifically, a feature fusion layer with three independent channels is designed to balance the density of features from different modalities. Each independent channel deals with features via the adjustable dynamic receptive field (DRF). Using the DRF is to balance the feature density based on the adjustable parameters. With shock signal from the intrinsic mode function (IMF), the kurtosis-based feature selection is designed to select the useful features for the feature fusion layer with DRF. Furthermore, the gramian angular field (GAF) algorithm is employed to transform the data dimension for the purpose of temporally encoding the sequences while mitigating computational complexity. For performance evaluation, we apply the RMCW model to CWRU bearing dataset for a simulation experiment and further testify it with a practical bearing dataset. Compared with similar deep learning-based methods CNN and ResNet. Experimental results show that the proposed RMCW can improve fault classification performance effectively and outperform its competitors significantly.

To summarize, the RMCW model has its unique advantages over existing methods: 1) In contrast with the neural models with traditional multi-channel, the independent weighting multi-channel with DRF is utilized in RMCW can address more features from the different modalities via independent DRF parameters adjustment. 2) Comparing with extant multi-channel feature extractions, nearly all of which are focused on low abstractions such as raw signals themselves, some statistical features, and so on. Independent multi-channel with DRF can pay higher level attention to address the specific features such as numerous IMFs or others from any other modalities, leading to the acquisition of higher quality feature fusion.

Our contributions in this article are mainly listed as the following.

1) The shock degree of the IMF signals is utilized to describe the feature density in different frequency bands. Independent channel weights based on the calculated kurtosis coefficients are designed to address the problem of widely varying feature densities across channels.

2) Inspired by the color gamut adaptation of biological vision, the independent DRF in different channels is

proposed to accomplish weight adjustment of modality signals.

3) An improved deep learning-based model RCMW is built for the purpose of utilizing different independent channels with adjustable DRFs to improve model feature extraction.

The rest of this article is organized as follows. We describe related method about data processing in Section II. The RMCW network structure is expatiated in Section III. We conduct detailed discussions for simulation and practical experiment results on CWRU bearing datasets and practical slewing bearing datasets in Section IV and Section V, respectively. Finally, we conclude this article in Section VI.

II. DATA PROCESSING

Data processing is closely related to dimension reconstruction, time sequences distribution, and modalities decomposition. The most related works are described here.

A. EMPIRICAL MODE DECOMPOSITION

A complex signal can be seen as the sum of multiple different IMFs, any modal function can be linear or nonlinear, and any modes are independent. Assuming the original signal is $x(t)$, the steps of EMD are as follows:

$$h_1(t) = x(t) - m_1(t) \tag{1}$$

$$c_1(t) = h_{1k}(t) = h_{k-1}(t) - m_{1k}(t) \tag{2}$$

$$r_1(t) = x(t) - c_1(t) \tag{3}$$

$$x(t) = \sum_{}^{n} c_i(t) + r_n(t) \tag{4}$$

where $m_1(t)$ is the mean magnitude; $h_1(t)$ is the difference between $x(t)$ and $m_1(t)$; repeat (1) to obtain $h_1(t)$; $r_1(t)$ is the residual variable obtained by separation; the final residual component $r_n(t)$ is a monotonic function.

B. POLAR COORDINATES CODING

The denoised vibration signal is still a random variable in the time domain, and each sampling point corresponds to a timestamp. Assuming $x(t) = [x_1, x_2, \ldots, x_i]$, x_i is the magnitude of the corresponding timestamp. In vibration signals, the variable value at any time point xi is meaningless when detached from time. Hence, it is essential to ensure that the reconstructed variable values correspond to new valid time sequences.

To ensure the timing of information, this paper maps 1D data timing to angle values through polar coordinate coding. The information involved in the encoding is the amplitude of the signal. The specific steps are as follows:

a) To increase the timing of sampling and reduce the influence of eigenvalue variation. This paper adopts the method of piecewise aggregation approximation (PAA) for coding. In this step, take the average time sequences value of k adjacent sampling points to reduce the amplitude of the fault feature, and get new sequences $x_k(t)$.

$$x_k(t) = \left[\frac{\sum_1^k x_i}{k}, \frac{\sum_{k+1}^{2k} x_o}{2k}, \cdots \frac{\sum_{(n-1)k+1}^{nk} x_p}{nk} \right] \tag{5}$$

where x_i, x_0 and x_p are the characteristic amplitudes of the original series corresponding to the time series interval. Then the new series after PAA coding can be expressed as $x_k(t) = [x_{k1}, x_{k2}, \ldots x_{kn}]$.

b) Perform normalization processing, scale the time sequences value to the [0, 1], and standardize the importance of the feature. The normalization method in this paper is as follows.

$$x_k(t) = [e^{-|x_{k1}|}, e^{-|x_{k2}|}, \cdots, e^{-|x_{kn}|}] \tag{6}$$

Let $\delta_{kn} = e^{-|x_{kn}|}$, then the new sequences after normalization processing can be expressed as $x_k(t) = [\delta_{k1}, \delta_{k2}, \ldots \delta_{kn}]$.

c) Perform polar coordinate processing. Generate polar coordinates by taking the new time sequences kn as the polar radius and the arc cosine of the normalized value.

$$x_k(t) = [\delta_{k1}, \delta_{k2}, \cdots \delta_{kn}]$$
$$\Downarrow$$
$$x_k(t) = \{[\arccos(\delta_{k1}), k1], [\arccos(\delta_{k2}), k2],$$
$$\times \cdots, [\arccos(\delta_{kn}), kn]\} \tag{7}$$

C. GAF DIMENSION RECONSTRUCTION

The sequences $x_k(t)$ is imported into GAF to get a 2D sequences diagram. GAF methods divide into Gramian angular summation field (GASF) and Gramian angular difference field (GADF). This paper uses the GASF method to upgrade the data dimension, and the obtained GASF can be expressed as in (8) and (9), shown at the bottom of the next page.

III. RESIDUAL NEURAL METWORK WITH MULTI-CHANNEL FEATURE WEIGHTING

An enhanced ResNet model named RMCW is proposed in this paper. Unlike using multilinear CNNs, RMCW integrates the features from multiple channels of 2D images through a default convolution layer. Each channel is assigned different convolution kernels based on the proportion of the feature information, enabling comprehensive extraction of crucial features within the image.

A. DATA EXPANSION

The residual model requires data samples of sufficient scale, and the scale of IMF channel images is smaller than the original data. Data expansion is performed through overlapping sampling as shown in Fig. 1 [43].

When the total length of the signal is L, the length of the interception target is l and the interception compensation is n. Subsequently, the calculation formula for the number Z of sub-signals can be expressed as follows:

$$Z = \left\lfloor \frac{L - l}{l - n} + 1 \right\rfloor \tag{10}$$
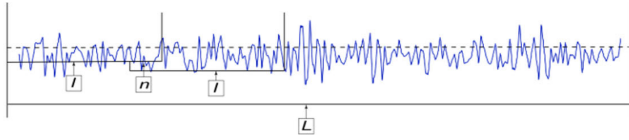
where $\lfloor \rfloor$ is the round-down operator.

FIGURE 1. The method of data expansion.

The original signal is decomposed by EMD to obtain IMFs with different shock features. Figure 3 shows that the first three IMFs have obvious impulsive characteristics, while kurtosis decreases in order. The reasons for using the first three IMFs will be explained in Section III, Part B. Then time sequences signals are saved by polar coordinate encoding, and the generated 2D time sequences diagram distributes the time sequences from the upper left to the lower right.

B. FRAMEWORK OF RMCW

The feature fusion layer in this approach utilizes kurtosis to determine the fusion proportion of each channel's information, emphasizing the crucial features while avoiding the omission of secondary information. The framework of diagnostic method is illustrated in Fig. 2. The detailed RMCW algorithm is illustrated in Algorithm 1.

C. MULTI-CHANNEL VIBRATION IMF IMAGE

We utilized EMD to decompose the raw signals to obtain IMFs that are not correlated and have prominent shock features. Besides, these shock features are distributed on various scale details. As shown in Fig. 3 [43], we select a single sample for the Hilbert analysis, and the vibration sequences is 0-0.16 s, the x-axis represents chronological information, while the y-axis represents frequency. To express the homeopathic atlas of each IMF, grayscale processing of the original atlas can represent the energy relationship of the IMFs under each detail at different scales. The IMFs in the different bands exhibit high-energy features and are independent of each other. However, an IMF signal containing the features required for fault diagnosis requires manual analysis.

Hilbert time-frequence spectrum of the IMF signal components in each band manually screened as shown in Fig. 4 [43]. The first three IMFs express the shock features of the vibration signal. In the forth IMF, the signal frequency fluctuates significantly at 0.11325(s), proving that is abnormal. By comparing time-frequency features, the first three IMF

images highlight the shock features to verify the correctness of IMFs screened for the multi-channel input.

D. NETWORK STRUCTURE

The first convolution layer in the RMCW is a multi-channel feature fusion layer, followed by 9 residual building blocks (RBB). Each RBB contains 2 convolution layers: a batch normalization (BN) layer and pooling layer. Connecting the layers and classifiers, we obtain 40 layers in total. The sheer size of the data requires deeper models that would otherwise cause gradients to disappear or make it difficult to fit complex features with diversity. Hence, the size and complexity of the dataset determine the number of residual modules.

1) DYNAMIC RECEPTIVE FIELD

Due to the variations in feature density across different image channels, the feature fusion layer employs DRFs to capture impactful features at various scale details. The primary design concept of the DRF is to allow independent parameter adjustments for each channel, with each channel receiving its own data. This way, each channel possesses its distinct field, and its parameters adapt as its internal subparameters, such as the convolution kernel size, change.

As previously described, the source data received by the three channels are derived from IMF-converted temporal images in different frequency bands. It is predictable that the information received by each channel is necessarily uneven. If the traditional RGB channel is used for feature extraction, it is inevitable that one will lose sight of the other. Similar to the reception of light by the eye in biology, humans, for example, have different sensitivities to different color bands, so the visual system adjusts its perception of different color differences and transmits them uniformly to the nerve centres to obtain visual information [44], [45], [46]. Thus the adjustment can be considered to be dynamic, and it is precisely this biological phenomenon that is analogous to the independent adjustment of the parameters of each channel in a DRF to the different bands of information.

The range in which the output of the convolution layer is mapped onto the input is defined as the receptive field. The mathematical relationship between the convolution kernel size and the receptive field can be expressed by the following function.

$$RF_i = (RF_{i+1} - 1) \times S_i + K_i \tag{11}$$

$$G_{summation} = \begin{bmatrix} \cos(\arccos(\delta_{k1}) + \arccos(\delta_{k1})) & \cdots & \cos(\arccos(\delta_{k1}) + \arccos(\delta_{kn})) \\ \vdots & \ddots & \vdots \\ \cos(\arccos(\delta_{kn}) + \arccos(\delta_{k1})) & \cdots & \cos(\arccos(\delta_{kn}) + \arccos(\delta_{kn})) \end{bmatrix} \tag{8}$$
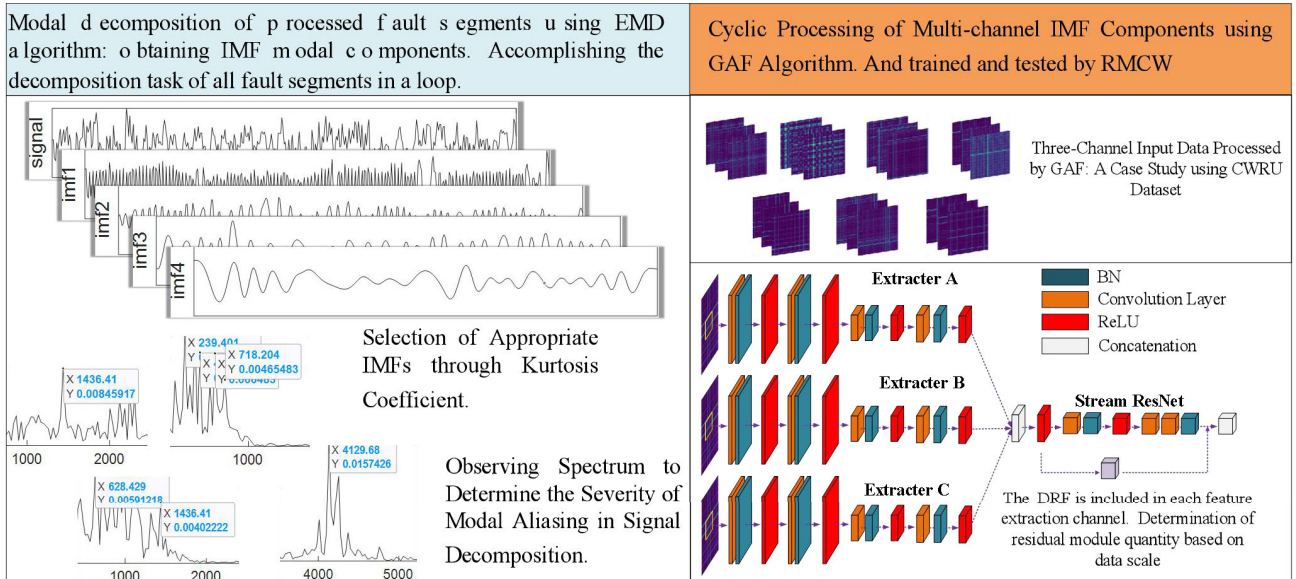
$$G_{difference} = \begin{bmatrix} \cos(\arccos(\delta_{k1}) - \arccos(\delta_{k1})) & \cdots & \cos(\arccos(\delta_{k1}) - \arccos(\delta_{kn})) \\ \vdots & \ddots & \vdots \\ \cos(\arccos(\delta_{kn}) - \arccos(\delta_{k1})) & \cdots & \cos(\arccos(\delta_{kn}) - \arccos(\delta_{kn})) \end{bmatrix} \tag{9}$$

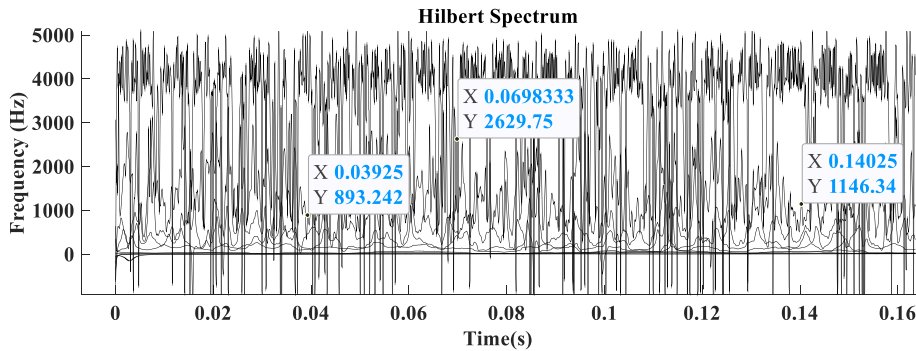**FIGURE 2.** The framework of fault diagnosis based on RMCW.

FIGURE 3. IMF Hilbert spectrum of gearbox vibration signal.

where RF_i is the size of the receptive field of the i_{th} layer, RF_{i+1} is the size of the receptive field of the $(i-1)_{th}$ layer, S_i is the convolution step size of the i_{th} layer convolution kernel, and K_i is the size of the i_{th} layer convolution kernel.

The DRF parameters are controlled by the convolution kernels in each independent channel, and they remain unaffected by other channels. The combination of channel weighting coefficients enhances the adaptation and the fault tolerance of model. As the IMF image is obtained from the original signal, the IMF at different frequency bands still belong to the same time domain. The IMF shock features can be seen as the components of the features in different frequency bands. Additionally, the GAF images quantize the original features into pixel values so that the relationship between multi-channel images can be approximated as that of the RGB channel images. The difference is that the feature density of the multi-channel images in this paper is measured by the kurtosis and requires different feature fields to balance. We utilized the same way as RGB to output fusion features,

namely element weighted summation. Output of the feature fusion layer can be expressed as follows.

$$X_j^l = \sum_{i=1}^{k} (x_{ij}^{l-1} \times w_{ij}^{l-1}) \qquad (12)$$

where X_j^l is the output of the l_{th} layer convolution of the j_{th} channel, x_{ij}^{l-1} is the input of the i_{th} feature of the $(l-1)_{th}$ layer channel, and w_{ij}^{l-1} is the weight of the i_{th} feature of the l_{th} layer of the j_{th} channel. There are k features in total.

2) MULTI-CHANNEL FEATURE WEIGHTED FUSION
The vibration signal image contains distinct impact characteristics of defects within different frequency bands. In order to fuse the channel image features of different importance, we utilized the signal kurtosis coefficient to determine the weighting coefficient. The kurtosis coefficient measures the regularity of the signal vibration. Therefore, the kurtosis
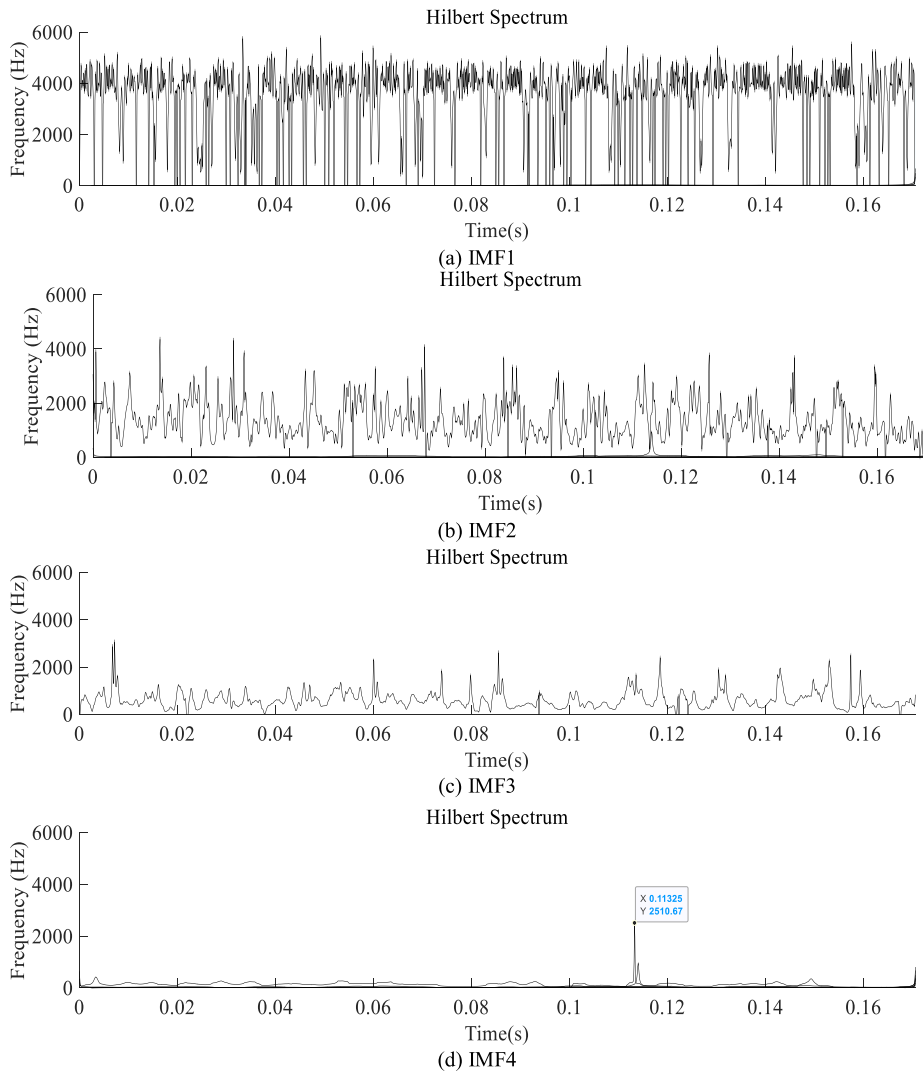
FIGURE 4. IMF 1-4 Hilbert time-frequency band.

coefficient directly affects the failure probability.

$$K_v = \frac{N \sum\limits_{i=1}^{N} (k_i - \mu)^4}{(\sum\limits_{i=1}^{N} (k_i - \mu)^2)^2} \qquad (13)$$

where K_v is the signal kurtosis of length N, k_i is the kurtosis coefficient of the i_{th} sampling point, and μ is the mean kurtosis.

The fusion weight of each channel is given by (14). After calculating the fusion weights, we perform the feature weighted fusion.

$$\lambda_j = \exp\left(\frac{mK_j}{\sum\limits_{j=1}^{m} K_j}\right) \qquad (14)$$

$$X^{last} = f\left(\sum\limits_{j=1}^{m} (\lambda_j \times X_j^{(last-1)} + b_j^{(last-1)})\right) \qquad (15)$$

TABLE 1. The design of RMCW hyperparameters.

Layer name	Structure parameters	Training parameters
Feature fusion	Kernel_size:3×3/5×5/7×7;Strides=2;M=3	
Conv2d layer	Kernel_size:3×3 Strides=2	
Ave 2dPooling	Pooling_size:1×1 Strides=2	Loss_function=Cate_crossentropy Optimizer=Adam Metrics=accuracy Batch_size=32
Fully connected	Kernel_initializer=he_normal; Kernel_regularizer=L2(0.0001)	
Softmax	/	

where X^{last} is the output of the feature fusion layer, $X_j^{(last-1)}$ is the feature output of the j_{th} fusion channel, λ_j is the weighting coefficient, m is the number of channels, and $f(x)$ is the activation function of the feature fusion layer.

3) RMCW HYPERPARAMETER SETUP
The hyperparameter settings for the RMCW are listed in Table 1 [43].

Algorithm 1 Residual Network with Independent Multi-Channel Weighting Algorithm

---

Input: Original input $\{X^E, Y^E\} \in R^{N \times (M+1)}$

1: // Assuming $\{X^E\}$ is one-dimension data; $\{X^E\}$ is label.

2: The original input is processed via EMD

3: **for** $i = 1; i < k$ **do**

4:    **for** $j = 1; j + +$ **do**

5:      **while** $r_1(t) = X^E(t) - \mathcal{H}_1(t)$ **do**

6:        // $\mathcal{H}_1(t)$ is the 1st obtained IMF and $r_1(t)$ is the 1st residual signal

7:        $r_j(t) = \mathcal{H}_j(t) - r_{j-1}(t);$

8:      **end while**

9:      Calculate $\mathcal{H}_j(t) = \frac{1}{k} \sum_{i=1}^{k} E_1(r_{j-1}(t) + \mathcal{H}_{j-1}(\delta_{j-1}(t)));$

10:      // $E(.)$ is the EMD decomposition operation

11:    **end for**

12:    Calculate $K_v(\mathcal{H}_j(t)) = N \sum_{i=1}^{N}(\mathcal{H}_j(t))^4 / (\sum_{i=1}^{N}(\mathcal{H}_j(t))^2)^2;$

13:    // $K_v(.)$ is the kurtosis calculation.

14:    Choose $\mathcal{H}_1^E, \mathcal{H}_2^E, \mathcal{H}_3^E$ by $K_1, K_2, K_3 \leftarrow K_v(\mathcal{H}_j(t));$

15:    // Return three parallel IMFs by Eq. (13)

16: **end for**

17: return $(\mathcal{H}_1^E \| \mathcal{H}_2^E \| \mathcal{H}_3^E), Y^E) \in R^{N \times (M+1)};$

18: Three groups of IMF are generated as two-dimensional images via GAF;

19: **while** $\mathcal{H}_1^E, \mathcal{H}_2^E, \mathcal{H}_3^E$ is acquired **do**

20:    **for** $L = 1; L + +$ **do**

21:      Calculate $(\mathcal{H}_L^E)_m = \frac{\mathcal{H}_L^T}{m}, m = 1, 2, \ldots, \pi$ by Eq. (5) and (10);

22:      // $(\mathcal{H}_L^E)_m$ is all sets of IMFs by screened.

23:      Generate G_{Lm}^E by Eq.(9);

24:      // G_{Lm}^E is all sets of generated GAFs.

25:    **end for**

26: **end while**

27: **return** G_{Lm}^E

28: **while** The data $(G_{Lm}^E, Y^E) \in R^{N \times (M+1)}$ is acquired **do**

29:    Initialize RMCW with residual mapping nodes n;

30:    **for** $n = 1; n \leq 3$ **do**

31:      // n is the number of channels.

32:      **for** $i = 1; i \leq k$ **do**

33:        Random $W_{ij}, \beta_{ij};$ // W_{ij}, β_{ij} are the weighting and bias, respectively.

34:        Calculate $Z_k \Leftarrow \prod_{i=1 \varphi_1 \varphi_2, \ldots \varphi_i}(G_{Lm}^E, \{W_{ij}, \beta_{ij}\}), i = 1, 2, \ldots, k;$

35:        // Z_k is the current residual feature mapping.

36:      **end for**

37:      Calculate Z_k^n

38:      // Z_k^n is all sets of residual feature mappings.

39:    **end for**

40:    if above steps finished **then**

41:      Update $Z_\xi(t)$ by Eq. (13)

42:      // Feature fusion through weighted coefficients.

43:    **end if**

44: **end while**

Output: $Z_\xi(t)$

The feature-fusion layer adopts a DRF by adjusting the receptive field range. The feature output is shown in (12). Owing to the residual structure, we added no regularization term to the RBB. The channel weights of the feature fusion layer are calculated by (14). The kurtosis and related weights for each channel are listed in Table 2 [43].

TABLE 2. The channel weighting factors of the feature fusion layer.

	Roller_fault	OR_fault	Normal	AVE	λ_j
Channel 1	3.97	5.08	2.79	4.3325	2.6754
Channel 2	4.54	7.51	1.86	4.2825	2.6452
Channel 3	3	6.81	3.36	4.5925	2.8382

IV. FAULT DIAGNOSIS MODEL APPLICATION CASE 1: CWRU BERING DATASET

This section demonstrates the efficacy of the RMCW proposed in this paper using a bearing dataset.

A. DESIGN OF SIMULATION EXPERIMENTS

In this case, both the types of CWRU deep-groove ball bearings are SKF6205-2RS and SKF6203-2RS, respectively. One cycle contains 406 sampling points. We set the sample length to 800 sampling points to ensure the feature density. Seven different fault types are labeled, each type contains 500 samples, and the training set and the test set are divided according to the ratio of 4:1. The specific distribution of bearing fault samples is shown in Table 3 [43].

TABLE 3. The describtion of datasets.

| Fault diameter/mm	0	0.18			0.36		
	health	inner race	outer race	roller	inner race	outer race	roller
Fault label	F1	F2	F3	F4	F5	F6	F7
Datasets Training	400	400	400	400	400	400	400
Test	100	100	100	100	100	100	100

B. EXPERIMENTAL RESULTS ANALYSIS

The RMCW model was trained 100 times. The accuracy and loss during the training process are shown in Fig. 5. The confusion matrix for the test results is shown in Fig. 6.

The experimental results from simulations demonstrate that the proposed method achieved an average test accuracy of 99.72% on the CWRU bearing datasets when using RMCW. The confusion matrix confirms that RMCW exhibits high diagnostic accuracy in effectively identifying various types of faults.

Experiments were conducted on a computer with an i7-10750H central processing unit and an NVIDIA GeForce GTX 1660Ti GPU. The training efficiency of the model is dependent on the hardware. The main device condition in this study is the performance of the GPU.
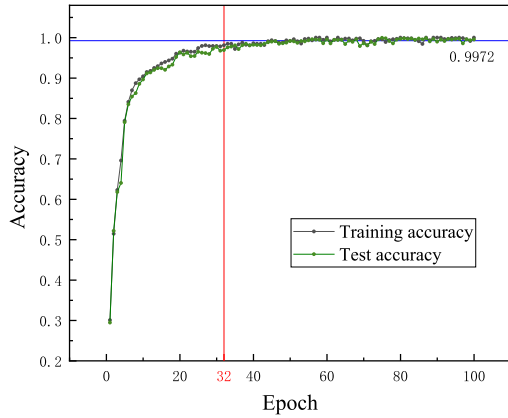
(a) Recognition accuracy on training and test sets
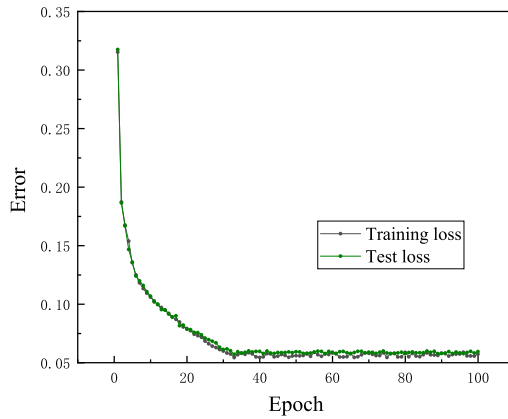
(b) The loss values on training and test sets
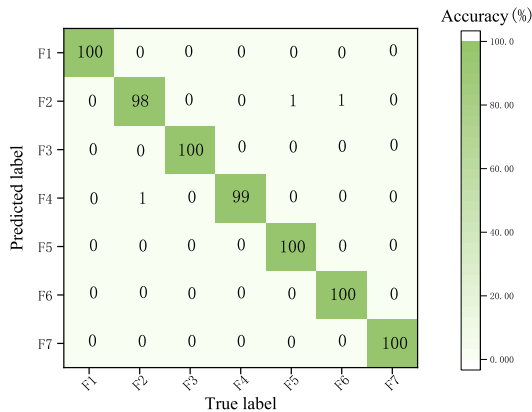
FIGURE 5. Training and testing results of RMCWIMF.

FIGURE 6. Confusion matrix of the test results (%).

C. VISUAL ANALYSIS

We utilized visualization technology to characterize the data features in the network due to the difficulty of model explanation. As shown in Fig. 7, the 2D time-sequences image of the inner raceway fault was selected into the RMCW, and the shape of the health state was observed. That can be summarized as follows: while the data dimension is continuously improving, the images become increasingly abstract and the

features become sparse. The yellow block of RBB_9, namely output_1 before the fully connected layer indicates that the feature has been activated.
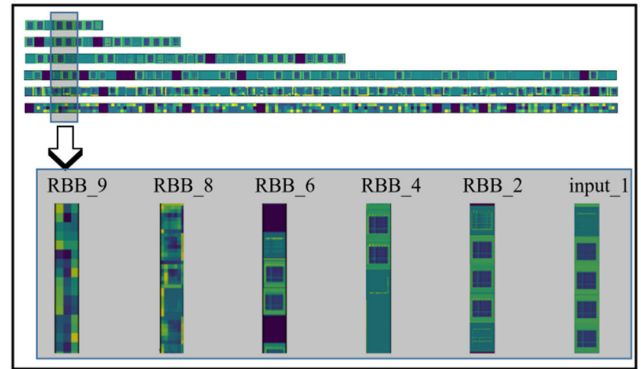
FIGURE 7. The output visualization of each residual module in RMCW for inner raceway fault.

To further explore feature extraction, the data dimensionality reduction technology is utilized to map data of different dimensions into a 3D space. Moreover, principal component analysis (PCA) combined with t-distributed stochastic neighbor embedding (T-SNE) was used to reduce data dimensionality to reduce the time cost. As shown in Fig. 8, an interval display is performed to reflect the difference in the feature distribution in each RBB. After 9 RBBs, the classification features of the data are pronounced. Softmax is the final classification layer.
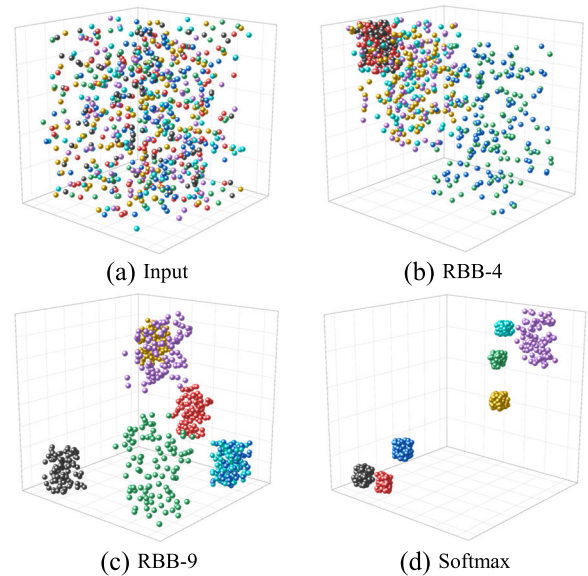

(a) Input   (b) RBB-4

(c) RBB-9   (d) Softmax

FIGURE 8. Feature visualization of RMCW by T-SNE and PCA.

V. FAULT DIAGNOSIS MODEL APPLICATION CASE 2: SLEWING BEARING DATASET

This section describes the slewing bearing fault diagnosis experiments to verify the effectiveness of RMCW.

As shown in Fig. 9, the slewing bearing test stand is loaded by a hydraulic station and the data is output by 3 sensors (vibration, temperature and torque) and collected by a data acquisition card. The datasets are provided by the Key Laboratory of Digital Manufacturing and Control Technology for Industrial Equipment in Jiangsu Province, Nanjing Tech University.

A. DESIGN OF EXPERIMENTS

The experimentation involved monitoring the vibration signals of a QNA-730-22 single-row ball slewing ring bearing using a ULT2001 low frequency capacitive accelerometer. To mitigate the interference caused by local resonance during signal acquisition, the accelerometers were positioned near the raceway surface at 90° intervals using magnetic holders, as depicted in Figure 10 [43].
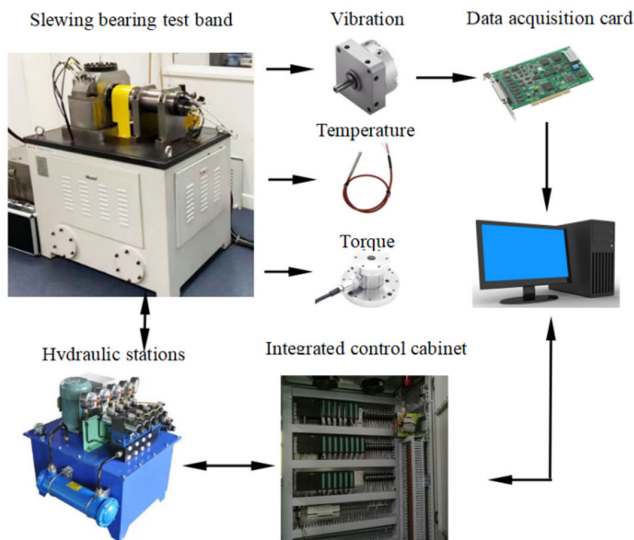

FIGURE 10. Mounting position of acceleration sensors.

TABLE 4. The experimental loading procedure.

Steps	Axial force (KN)	Overturning moment (KN · m)	Rotation speed (rpm)	Time
1	96	240	4	7 days
2	Dismantling and Inspection			
3	96	240	4	Until stuck
4	Dismantling and Inspection			

FIGURE 9. The diagram of slewing bearing test platform.

The experimental loading procedure is shown in Table 4. The slewing ring bearing was tested under full load for 10.58 days (89,520 circles) with an abnormal rattling noise and finally a stalling of the slewing bearing. The type of fault was obtained as shown in Fig. 11.

The vibration signal is obtained as shown in Fig. 12. The vibration signal data from the four measurement points further ensures the diversity and the data scale as well as prevents overfitting during the training process. The 2D multi-channel input image is acquired according to the signal processing described previously. The specific operation steps are shown in Fig. 13. The basic model parameter settings are the same as in the simulation experiment.

B. MODEL PERFORMANCE COMPARISON

We compare RMCW with the SVM, CNN and ResNet. The comparison of accuracy results for different signal-to-noise ratios (SNR) are shown in Fig. 14. The network models participating in the comparison, except for RMCW, are trained
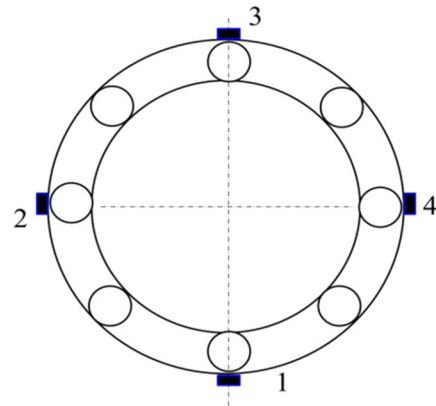

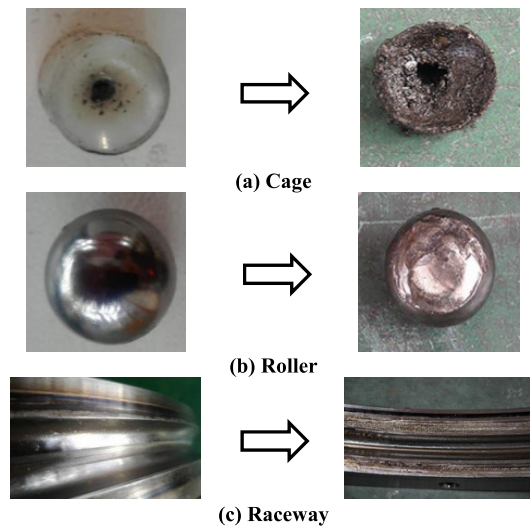
(a) Cage

(b) Roller

(c) Raceway

FIGURE 11. Fault types for accelerated bearing damage.

using the RGB channel images (reconstruction of the original signal using GAF only). The experimental results show that the classification accuracy of RMCW is better than others.

Table 5 presents the comprehensive outcomes of SVM, CNN, ResNet, and RMCW when tested with different types of noise at varying signal-to-noise ratios (SNRs). Furthermore, we utilized several evaluation metrics including precision, recall, and F1 score to assess the performance of the models. Figure 15 illustrates that RMCW exhibits exceptional performance in terms of the metrics, achieving lower

TABLE 5. Performance of experimental results under different Snrs.

Method	Training accuracy (%)	Test accuracy(%)	Precision(%)	Recall(%)	F1 score	Top-1 error(%)	MAPE
SVM	70.43	65.64	70.98	71.03	0.7100	24.63	4.833
CNN	86.21	81.31	84.21	85.62	0.8541	12.41	1.572
ResNet	95.19	90.66	96.43	93.71	0.9505	7.4	0.394
RMCW	99.74	94.11	97.89	98.12	0.9800	5.71	0.251
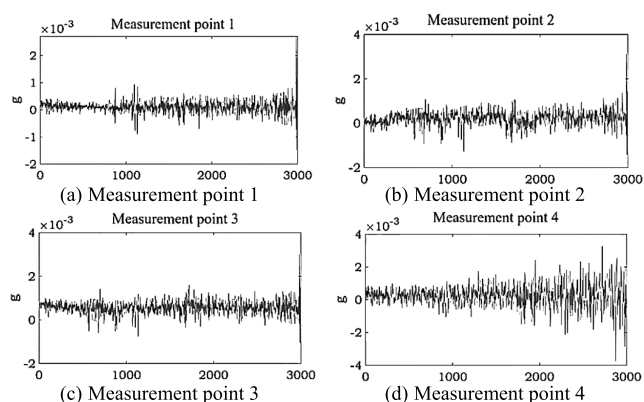
FIGURE 12. The vibration signals of slewing bearings.

FIGURE 14. Comparison of classifier performance.

FIGURE 13. The diagram of the model experiment.
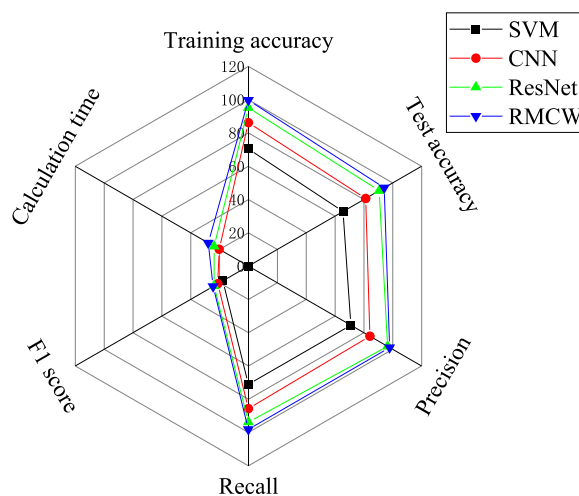
FIGURE 15. Radar plot for model performance comparison.

rates of missed detections and false alarms compared to ResNet.

The detailed results of the SVM, CNN, ResNet and RMCW with different types of noise under different SNRs are provided in Table 5. Moreover, we introduced various metrics such as precision, recall, and F1 score to evaluate the model performance, as shown in Fig. 15. The metrics of RMCW demonstrate superior performance, with lower miss and false detection rates compared to ResNet.
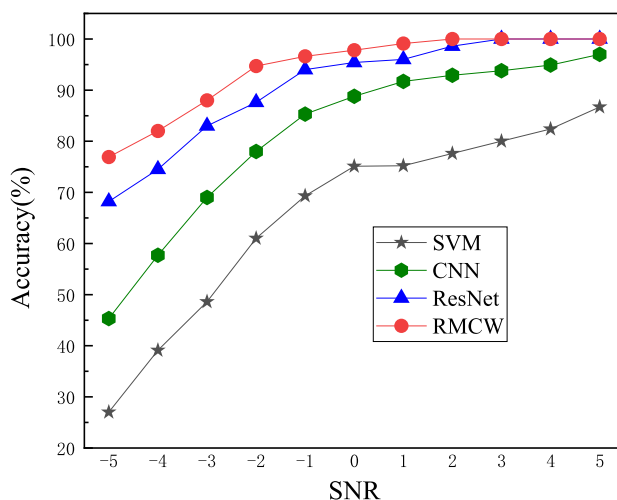
As shown in Fig. 14, RMCW has significant accuracy superiority over several other mainstream diagnostic algorithms under different SNRs, and becomes more apparent as the SNR increases. Accordingly, the average diagnostic accuracy of RMCW is improved by 3.45% over ResNet.

Average computation time of 1671.97s (27.866 min) was obtained after multiple sets of trials, based on the computer configuration described previously. It can be seen that the

difference in calculation time costs is not significant compared to other methods.

VI. CONCLUSION

In our study, a fault diagnosis framework RMCW is constructed, which is based on residual network, includes an improved method for independent multi-channel weighting optimization, and a dynamic receptive field mechanism for real bearing from a machinery system. The EMD algorithm and the GAF algorithm are utilized to obtain 2D time series fault samples with different shock degrees filtered by the kurtosis. Subsequently, CWRU bearing data are presented to illustrate the implementation procedure and validity (99.72% average accuracy achieved) of the RMCW diagnosis framework for fault classification. In addition, a real slewing bearing dataset is adopted to testify to superior fault diagnosis performance (Higher accuracy and stability at different SNRs, 3.45% higher than before improvement) of the proposed RMCW diagnosis framework.

However, it should be mentioned that there is still room for improvement in the current work. For instance, the EMD algorithm is used in our data pre-processing part, which works well but may not be the best option. Another potential improvement is to adopt the different modal information (e.g. temperature, torque) into the proposed RMCW model to get better performance.

REFERENCES

[1] Y. Jin, J. Zhang, X. Zhang, H. Xiao, B. Ai, and D. W. K. Ng, "Channel estimation for semi-passive reconfigurable intelligent surfaces with enhanced deep residual networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 11083–11088, Oct. 2021, doi: 10.1109/TVT.2021.3109937.

[2] Y. Xiao, H. Shao, M. Feng, T. Han, J. Wan, and B. Liu, "Towards trustworthy rotating machinery fault diagnosis via attention uncertainty in transformer," *J. Manuf. Syst.*, vol. 70, pp. 186–201, Oct. 2023, doi: 10.1016/j.jmsy.2023.07.012.

[3] S. Yan, H. Shao, Z. Min, J. Peng, B. Cai, and B. Liu, "FGDAE: A new machinery anomaly detection method towards complex operating conditions," *Rel. Eng. Syst. Saf.*, vol. 236, Aug. 2023, Art. no. 109319, doi: 10.1016/j.ress.2023.109319.

[4] L. Balagourouchetty, J. K. Pragatheeswaran, B. Pottakkat, and G. Ramkumar, "GoogLeNet-based ensemble FCNet classifier for focal liver lesion diagnosis," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 6, pp. 1686–1694, Jun. 2020, doi: 10.1109/JBHI.2019.2942774.

[5] J. C. Liu, H. Quan, X. Yu, K. He, and Z. H. Li, "Rolling bearing fault diagnosis based on parameter optimization VMD and sample entropy," *Acta Automatica Sinica*, vol. 48, no. 3, pp. 808–819, 2022.

[6] J. Zhang, H. Huang, T. Liu, and C. Zhang, "Multiscale fractures characterization based on ant colony optimization and two-dimensional variational mode decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 7, pp. 2562–2570, Jul. 2018, doi: 10.1109/JSTARS.2018.2837222.

[7] S. Yu and J. Ma, "Complex variational mode decomposition for slop-preserving denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 586–597, Jan. 2018, doi: 10.1109/TGRS.2017.2751642.

[8] N. Chatlani and J. J. Soraghan, "EMD-based filtering (EMDF) of low-frequency noise for speech enhancement," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 4, pp. 1158–1166, May 2012, doi: 10.1109/TASL.2011.2172428.

[9] D. Liu, L. Cui, and W. Cheng, "Flexible generalized demodulation for intelligent bearing fault diagnosis under nonstationary conditions," *IEEE Trans. Ind. Informat.*, vol. 19, no. 3, pp. 2717–2728, Mar. 2023, doi: 10.1109/TII.2022.3192597.

[10] A. Komaty, A.-O. Boudraa, B. Augier, and D. Daré-Emzivat, "EMD-based filtering using similarity measure between probability density functions of IMFs," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 1, pp. 27–34, Jan. 2014, doi: 10.1109/TIM.2013.2275243.

[11] X. Wu, W. Zuo, L. Lin, W. Jia, and D. Zhang, "F-SVM: Combination of feature transformation and SVM learning via convex relaxation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5185–5199, Nov. 2018, doi: 10.1109/TNNLS.2018.2791507.

[12] S. Dhar and V. Cherkassky, "Development and evaluation of cost-sensitive universum-SVM," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 806–818, Apr. 2015, doi: 10.1109/TCYB.2014.2336876.

[13] M. M. Adankon, M. Cheriet, and A. Biem, "Semisupervised learning using Bayesian interpretation: Application to LS-SVM," *IEEE Trans. Neural Netw.*, vol. 22, no. 4, pp. 513–524, Apr. 2011, doi: 10.1109/TNN.2011.2105888.

[14] H. H. H. Aly and M. E. El-Hawary, "A proposed ANN and FLSM hybrid model for tidal current magnitude and direction forecasting," *IEEE J. Ocean. Eng.*, vol. 39, no. 1, pp. 26–31, Jan. 2014, doi: 10.1109/JOE.2013.2241934.

[15] Q. Shi and H. Zhang, "Fault diagnosis of an autonomous vehicle with an improved SVM algorithm subject to unbalanced datasets," *IEEE Trans. Ind. Electron.*, vol. 68, no. 7, pp. 6248–6256, Jul. 2021, doi: 10.1109/TIE.2020.2994868.

[16] A. Movsessian, D. G. Cava, and D. Tcherniak, "An artificial neural network methodology for damage detection: Demonstration on an operating wind turbine blade," *Mech. Syst. Signal Process.*, vol. 159, Oct. 2021, Art. no. 107766, doi: 10.1016/j.ymssp.2021.107766.

[17] G. Li, G. Tang, G. Luo, and H. Wang, "Underdetermined blind separation of bearing faults in hyperplane space with variational mode decomposition," *Mech. Syst. Signal Process.*, vol. 120, pp. 83–97, Apr. 2019, doi: 10.1016/j.ymssp.2018.10.016.

[18] C. Han, W. Lu, H. Wang, L. Song, and L. Cui, "Multistate fault diagnosis strategy for bearings based on an improved convolutional sparse coding with priori periodic filter group," *Mech. Syst. Signal Process.*, vol. 188, Apr. 2023, Art. no. 109995, doi: 10.1016/j.ymssp.2022.109995.

[19] G. Wang, J. Qiao, J. Bi, W. Li, and M. Zhou, "TL-GDBN: Growing deep belief network with transfer learning," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 2, pp. 874–885, Apr. 2019, doi: 10.1109/TASE.2018.2865663.

[20] L. Yang, Y. Li, and Z. Wei, "Fa-Mb-ResNet for grounding fault identification and line selection in the distribution networks," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 11115–11125, Jul. 2022, doi: 10.1109/JIOT.2021.3131171.

[21] F. Cai, M. Zhan, Q. Chai, and J. Jiang, "Fault diagnosis of DAB converters based on ResNet with adaptive threshold denoising," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022, doi: 10.1109/TIM.2022.3180409.

[22] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1693–1702, Jul. 2017, doi: 10.1109/TIM.2017.2669947.

[23] K. Orphanou, A. Stassopoulou, and E. Keravnou, "DBN-extended: A dynamic Bayesian network model extended with temporal abstractions for coronary heart disease prognosis," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 3, pp. 944–952, May 2016, doi: 10.1109/JBHI.2015.2420534.

[24] A. Lee-Leon, C. Yuen, and D. Herremans, "Underwater acoustic communication receiver using deep belief network," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3698–3708, Jun. 2021, doi: 10.1109/TCOMM.2021.3063353.

[25] J. Huang, L. Cui, and J. Zhang, "Novel morphological scale difference filter with application in localization diagnosis of outer raceway defect in rolling bearings," *Mechanism Mach. Theory*, vol. 184, Jun. 2023, Art. no. 105288, doi: 10.1016/j.mechmachtheory.2023.105288.

[26] Z. Liu, Y. Lyu, L. Wang, and Z. Han, "Detection approach based on an improved faster RCNN for brace sleeve screws in high-speed railways," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4395–4403, Jul. 2020, doi: 10.1109/TIM.2019.2941292.

[27] S. Shao, R. Yan, Y. Lu, P. Wang, and R. X. Gao, "DCNN-based multi-signal induction motor fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2658–2669, Jun. 2020, doi: 10.1109/TIM.2019.2925247.

[28] B. Ma, W. Cai, Y. Han, and G. Yu, "A novel probability confidence CNN model and its application in mechanical fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021, doi: 10.1109/TIM.2021.3077965.

[29] D. T. Hoang and H. J. Kang, "A motor current signal-based bearing fault diagnosis using deep learning and information fusion," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 3325–3333, Jun. 2020, doi: 10.1109/TIM.2019.2933119.

[30] Z. Ye and J. B. Yu, "Feature extraction of gearbox vibration signals based on multi-channels weighted convolutional neural network," *J. Mech. Eng.*, vol. 57, no. 1, pp. 110–120, 2021, doi: 10.3901/JME.2021.01.110.

[31] M. Zhao, S. Zhong, X. Fu, B. Tang, and M. Pecht, "Deep residual shrinkage networks for fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4681–4690, Jul. 2020, doi: 10.1109/TII.2019.2943898.

[32] X. Q. Zhao and Y. Z. Zhang, "Improved CNN fault diagnosis method of rolling bearings under variable working conditions," *J. Xi'an Jiaotong Univ.*, vol. 56, no. 1, pp. 1–11, 2021, doi: 10.7652/xjtuxb202112013.

[33] T. Weickert, C. Benjaminsen, and U. Kiencke, "Analytic wavelet packets—Combining the dual-tree approach with wavelet packets for signal analysis and filtering," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 493–502, Feb. 2009, doi: 10.1109/TSP.2008.2007922.

[34] A. Vijayvargiya, V. Gupta, R. Kumar, N. Dey, and J. M. R. S. Tavares, "A hybrid WD-EEMD sEMG feature extraction technique for lower limb activity recognition," *IEEE Sensors J.*, vol. 21, no. 18, pp. 20431–20439, Sep. 2021, doi: 10.1109/JSEN.2021.3095594.

[35] Q. Li, B. Zheng, B. Tu, J. Wang, and C. Zhou, "Ensemble EMD-based spectral–spatial feature extraction for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5134–5148, 2020, doi: 10.1109/JSTARS.2020.3018710.

[36] S. J. Dong, X. W. Pei, W. L. Wu, B. P. Tang, and X. X. Zhao, "Rolling bearing fault diagnosis method based on multilayer noise reduction technology and improved convolutional neural network," *J. Mech. Eng.*, vol. 57, no. 1, pp. 148–156, 2021, doi: 10.3901/JME.2021.01.148.

[37] Y. Wu, G. Gao, and C. Cui, "Improved wavelet denoising by non-convex sparse regularization under double wavelet domains," *IEEE Access*, vol. 7, pp. 30659–30671, 2019, doi: 10.1109/ACCESS.2019.2903125.

[38] Z. Liu, H. Wang, J. Liu, Y. Qin, and D. Peng, "Multitask learning based on lightweight 1DCNN for fault diagnosis of wheelset bearings," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021, doi: 10.1109/TIM.2020.3017900.

[39] H. Wang, Z. Liu, D. Peng, and Y. Qin, "Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 5735–5745, Sep. 2020, doi: 10.1109/TII.2019.2955540.

[40] M. Asif, O. Nazeer, N. Javaid, E. H. Alkhammash, and M. Hadjouni, "Data augmentation using BiWGAN, feature extraction and classification by hybrid 2DCNN and BiLSTM to detect non-technical losses in smart grids," *IEEE Access*, vol. 10, pp. 27467–27483, 2022, doi: 10.1109/ACCESS.2022.3150047.

[41] D. Kollias and S. Zafeiriou, "Exploiting multi-CNN features in CNN-RNN based dimensional emotion recognition on the OMG in-the-wild dataset," *IEEE Trans. Affect. Comput.*, vol. 12, no. 3, pp. 595–606, Jul. 2021, doi: 10.1109/TAFFC.2020.3014171.

[42] S. Li, W. Dai, Z. Zheng, C. Li, J. Zou, and H. Xiong, "Reversible autoencoder: A CNN-based nonlinear lifting scheme for image reconstruction," *IEEE Trans. Signal Process.*, vol. 69, pp. 3117–3131, 2021, doi: 10.1109/TSP.2021.3082465.

[43] H. Yu, X. Miao, F. Ping, Z. Xun, and Y. Gu, "Fault feature extraction and diagnosis method based on multi-channel feature fusion residual network," in *Proc. Int. Conf. Sens., Meas. Data Anal. Era Artif. Intell. (ICSMD)*, Harbin, China, Nov. 2022, pp. 1–6, doi: 10.1109/ICSMD57530.2022.10058295.

[44] Z.-Y. Ji, X.-J. Song, H.-B. Song, H. Yang, and X.-X. Guo, "RDRF-Net: A pyramid architecture network with residual-based dynamic receptive fields for unsupervised depth estimation," *Neurocomputing*, vol. 457, pp. 1–12, Oct. 2021, doi: 10.1016/j.neucom.2021.05.089.

[45] W. Kim, A.-D. Nguyen, S. Lee, and A. C. Bovik, "DRF generation for full-reference image quality assessment," *IEEE Trans. Image Process.*, vol. 29, pp. 4219–4231, 2020, doi: 10.1109/tip.2020.2968283.

[46] Z. Ye and J. Yu, "AKSNet: A novel convolutional neural network with adaptive kernel width and sparse regularization for machinery fault diagnosis," *J. Manuf. Syst.*, vol. 59, pp. 467–480, Apr. 2021.

ZHENG LIU received the B.S. degree in vehicle engineering from Nanjing Tech University, Nanjing, China, in 2022. His research interests include fault diagnosis, target detection, and deep learning in mechanical engineering.

HU YU received the M.Eng. degree in mechanical engineering from Nanjing Tech University, Nanjing, China, in 2023. He is currently pursuing the Ph.D. degree in mechanical engineering with the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing. His research interests include mechanical vibration analysis and intelligent monitoring of equipment health status.

KUN XU received the Ph.D. degree in mechanical engineering from the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China. He is currently pursuing the joint Ph.D. degree with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. In February 2023, he joined the School of Mechanical and Power Engineering, Nanjing Tech University, as a Lecturer. His research interests include signal processing, deep learning, and fault diagnosis of rotating components.

XIAODONG MIAO (Member, IEEE) received the B.S. and Ph.D. degrees in mechanical engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2008 and 2013, respectively. He is currently an Associate Professor of mechanical engineering with Nanjing Tech University. His research interests include sensors and signal processing, machinery condition monitoring, fault diagnostics, and applications of artificial intelligence.

• • •