ISSN 1392-1207. MECHANIKA. Vol. 32, No. 2, 2026: 166–171

# Fault Diagnosis of Anti-Friction Bearings Based on Bi-Dimensional Extensive Empirical Wavelet Decomposition and Optimized Incremental RVM

**Yanwei LI*, Ke XU*, Xing CHEN*, Qingshou ZHENG*, Youwen CHEN**, Zihong YIN*****
**Zhejiang Railway Development Holding Group Co., Ltd, Hangzhou, 310015, China, E-mail: kxkxhz@126.com (Corresponding Author: Ke XU)*
***Zhejiang Communications Investment Group Co., Ltd, Hangzhou, 310020, China, E-mail: youwenchen_c@126.com*
****School of Civil Engineering, Southwest Jiaotong University, Chengdu, 610031, China,*
*E-mail: zihongyin2000@126.com*

https://doi.org/10.5755/j02.mech.41805

## 1. Introduction

Anti-friction bearings play a vital role in the functioning of mechanical equipment [1, 2], which is widely used in industrial manufacturing equipment, rail transportation equipment, and generator systems for bridge foundation pit construction, and so on. Faults of anti-friction bearings can lead to abnormal equipment operation, and in severe cases, cause significant economic losses and casualties [3, 4]. Consequently, the reliable operation of rotating machinery critically depends on rapid and efficient fault diagnosis. The key to achieving outstanding diagnostic results for anti-friction bearings lies in extracting the features of their vibration signals. Traditional feature extraction methods include wavelet transform and empirical mode decomposition, among others [5-8]. Empirical Wavelet Decomposition (EWD) is widely used in defect detection for adaptively extracting various modes from non-stationary signals through the construction of adaptive wavelets [9, 10]. Bi-dimensional Extensive Empirical Wavelet Decomposition (BEEWD) separates signals into fundamental and distortion components. Its key innovation lies in pre-estimating frequency thresholds based on signal amplitude modulation, thereby addressing the impact of low-frequency harmonics adjacent to the fundamental wave on signal amplitude.

Relevance Vector Machine (RVM) generally outperform Support Vector Machine in terms of generalization ability, largely due to the fact that RVMs usually require fewer support vectors and involve determining fewer training parameters [11, 12]. To further improve the generalization performance of RVM, this paper presents an incremental RVM algorithm, which adapts the original RVM formulation to create an incremental version. Incremental learning simplifies the computational process and enhances the generalization capabilities of the algorithm.

In this study, we propose a novel fault diagnosis method for anti-friction bearings that combines Bi-dimensional Extensive Empirical Wavelet Decomposition (BEEWD) and an incremental RVM algorithm optimized by Differential Evolution-enhanced Grey Wolf Optimization (DGWO). Since BEEWD produces clearer time-frequency images compared to EWD, it decomposes vibration signals from anti-friction bearings and generates corresponding time-frequency images. The structure of the enhanced incremental RVM algorithm, referred to as OIRVM, is presented. This algorithm, combined with DGWO, is employed for fault diagnosis of anti-friction bearings. The traditional Grey Wolf Optimization (GWO) algorithm is improved by incorporating differential evolution strategies, enhancing the diversity of the wolf population and thereby improving the global search capability of the grey wolves. DGWO is utilized to select the appropriate kernel parameter of the incremental RVM.

## 2. Bi-dimensional Extensive Empirical Wavelet Transform

By leveraging wavelet filter banks, EWD can detect critical frequency structures in the Fourier domain and extract individual frequency bands from the signal. In EWD,

$$\gamma_1(\omega)=\begin{cases}1 & |\omega|\le(1-\theta)\omega_1\\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\theta\omega_1}\left(|\omega|-(1-\theta)\omega_1\right)\right)\right] & (1-\theta)\omega_1\le|\omega|\le(1+\theta)\omega_1\\ 0 & others\end{cases}, \tag{1}$$

$$\kappa_n(\omega)=\begin{cases}1 & (1+\theta)\omega_n\le|\omega|\le(1+\theta)\omega_{n+1}\\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\theta\omega_{n+1}}|\omega|-(1-\theta)\omega_{n+1}\right)\right] & (1-\theta)\omega_{n+1}\le|\omega|\le(1-\theta)\omega_{n+1}\\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\theta\omega_n}\left(|\omega|-(1-\theta)\omega_n\right)\right)\right] & (1-\theta)\omega_n\le|\omega|\le(1-\theta)\omega_n\\ 0 & others\end{cases}. \tag{2}$$

the empirical scaling functions $\gamma_1(\omega)$ and empirical wavelets $\kappa_1(\omega)$ are constructed and defined accordingly to enable this adaptive decomposition framework.

Then, obtain the reconstruction signal:

$$f(t)=W_f^{\varepsilon}(0,t)\otimes\gamma_1(t)+\sum_{n=1}^{N}W_f^{\varepsilon}(n,t)\otimes\kappa_n(t), \qquad (3)$$

where $W_f^{\varepsilon}(0,t)$ is the approximation coefficient; $W_f^{\varepsilon}(n,t)$ is the detailed coefficient.

BEEWD decomposes signals into fundamental and distortion components, mitigating amplitude modulations induced by low-frequency harmonics adjacent to the fundamental frequency, enabling their detection through time-domain amplitude analysis rather than exclusive reliance on frequency-domain information. The frequency threshold (*FT*) is estimated by quantifying these amplitude variations via the number of sign changes (*NS*) in the RMS variation vector of the signal.

Evaluate *NS* in the RMS variation vector of the signal, and subsequently correlate the *NS* with *FT* using the equation provided below.

$$FT=\left(q(NS,2)+r(NS,2)-1\right)\Delta f, \qquad (4)$$

where $\Delta f$ is the frequency resolution, $q(NS, 2)$ is the quotient of *NS* and 2, and $r(NS, 2)$ is the remainder of *NS* and 2.

The innovation of BEEWD lies in its pre-estimation of the frequency threshold based on signal amplitude modulation. This estimation of FT is entirely dependent on the amplitude modulation of the signal and is not affected by variations in sampling frequency or the length of the time window.

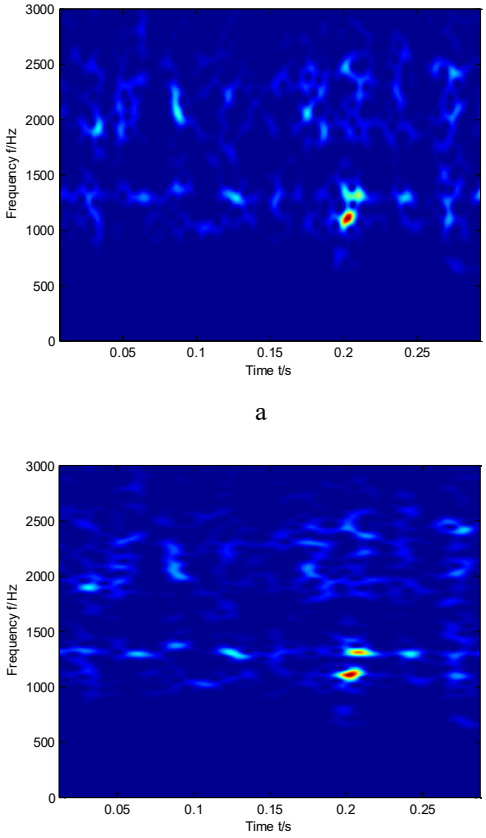
As the extracted pattern encompasses a single frequency component, the Hilbert transform can be utilized to approximate the instantaneous frequency and amplitude details. The Hilbert transform of a real-valued function determines its complex conjugate $BEEWT_H(i,t)$.

$$BEEWT_H(i,t)=\frac{\Theta}{\pi}\int_{-\infty}^{+\infty}\frac{BEEWT(i,t)}{t-\tau}d\tau, \qquad (5)$$

where $\Theta$ is the Cauchy principal value of the singular integral.

The 2D time-frequency image is constructed by mapping the instantaneous amplitudes of all BEEWD- decomposed components onto a two-dimensional grid, where the horizontal axis represents time (sample points), the vertical axis represents frequency (discretized into bins corresponding to the empirical modes), and the pixel intensity represents the normalized amplitude value, typically rescaled to a fixed resolution (128×128 pixels) for consistent input to the classification model.

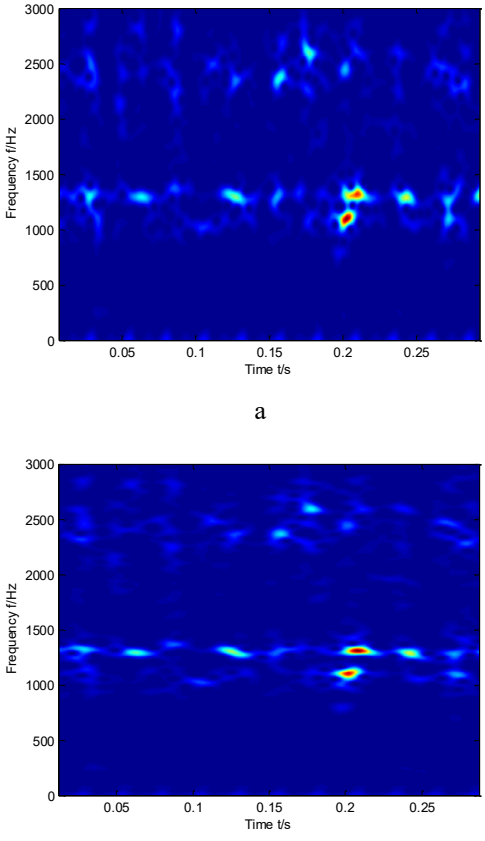
Fig. 1 presents the comparison of time-frequency images of the inner race fault sample between BEEWD and BEWD. Fig. 2 displays the comparison of time-frequency images of the outer race fault sample between BEEWD and BEWD. It is evident that the time-frequency image features generated by BEEWD are clearer compared to those from EWD. Therefore, this paper employs BEEWD to generate time-frequency images of the vibration signals of antifriction bearings.




a

b

Fig. 1 Time-frequency images of the inner race fault sample: a – BEWD, b – BEEWD




a

b

Fig. 2 Time-frequency images of the outer race fault sample: a – BEWD, b – BEEWD

## 3. Optimized Incremental RVM

The classification function of RVM is:

$$z(x)=\sum_{i=1}^{N} w_i k(x,x_i)+w_0, \tag{6}$$

where $k(x, x_i)$ is the kernel function, $w_i$ is the weight, and $w_0$ is the bias.

The likelihood of the dataset is:

$$p(t\mid w)=\prod_{i=1}^{N}\sigma\{z(x_i;w)\}^{t_i}\left(1-\sigma\{z(x_i;w)\}\right)^{1-t_i}, \tag{7}$$

where $\sigma\{z(x_i;w)\}=\frac{1}{1+e^{-z(x_i;w)}}$.

The logarithmic posterior probability is approximated via the Laplace method on two occasions. By performing two derivations, we can obtain

$$\nabla_w\nabla_w \log p(w\mid t,a)|_w=-\left(\Phi^T B\Phi+A\right), \tag{8}$$

where $A = diag(a_0, a_1, \ldots, a_n)$, $B$ is the diagonal matrix.

A fast marginal likelihood maximization method has been developed to enable simultaneous learning of kernel location and scale parameters during training. For real-time applications, processing large datasets and continuously updating the model with new data is essential. An incremental learning strategy is therefore highly valuable, prompting the modification of the original RVM formulation into an incremental variant. This incremental RVM reduces computational overhead while improving the algorithm's generalization performance.

In incremental RVM, the marginal likelihood is:

$$p(t\mid x,a)=p\left(t\mid x,\left(\Delta^T B\Delta+A\right)^{-1}\Delta^T Bt\right)p\left(\left(\Delta^T B\Delta+A\right)^{-1}\Delta^T Bt\mid a\right)(2\pi)^{N/2}\sqrt{\left|\left(\Delta^T B\Delta+A\right)^{-1}\right|}, \tag{9}$$

where $\Delta$ is the design matrix.

A differential evolution-enhanced GWO is used to determine the parameter of kernel function. The traditional GWO has limitations in the multi-class of the grey wolf population, making it prone to fast convergence and easy trapping in local optima [13, 14]. To address this, the differential evolution strategy is introduced into the traditional GWO to enhance the wolf population's diversity and boost the grey wolves' global search capability. The differential evolution-enhanced GWO algorithm is presented as follows:

$$g_i(k)=x_{r1}(k)+\mu\left[x_{r2}(k)-x_{r3}(k)\right], \tag{10}$$

where $g_i(k)$ is the mutated grey wolf, $\mu$ is the variation factor in the range of 0 and 2, $x_{r1}(k)$, $x_{r2}(k)$, and $x_{r3}(k)$ are three different individual gray wolves.

$$h_i(k)=\begin{cases} g_i(k) & rand < P \\ x_i(k) & rand \geq P \end{cases}, \tag{11}$$

where $h_i(k)$ is the crossover-generated grey wolf individual, *rand* is a random value from 0 to 1, $P$ is the crossover rate; $x_i(k)$ is the grey wolf position at iteration $k$.

$$x_i(k+1)=\begin{cases} h_i(k) & L(h_i(k))\leq L(x_i(k)) \\ x_i(k) & L(h_i(k))> L(x_i(k)) \end{cases}, \tag{12}$$

where $L(\cdot)$ is the fitness function, $x_i(k + 1)$ is the grey wolf position at iteration $k + 1$.

The process for optimizing the kernel parameter of the Incremental Relevance Vector Machine (IRVM) using the GWO algorithm enhanced by differential evolution strategy is outlined as follows:

**Initialization**: Define the range for the kernel parameter, and the DGWO algorithm is configured with a population size of 20 and a maximum iteration limit of 100, which is shown in Table 1.

**Fitness Evaluation and Role Assignment**: Calculate the fitness of each grey wolf to establish the alpha ($\alpha$)-beta ($\beta$)-delta ($\delta$) hierarchy and record their corresponding positions. In each iteration, the $\alpha$, $\beta$, and $\delta$ wolves lead the rest of the population in searching for the optimal solution.

**Population Update**: Update the positions of the grey wolves using operations like mutation, crossover, and selection. Identify the $\alpha$, $\beta$, and $\delta$ wolves from the current population based on their fitness values.

**Optimal Solution Search**: Guided by the information from the $\alpha$, $\beta$, and $\delta$ wolves, the grey wolves search for the optimal solution.

**Termination Check**: If the maximum number of iterations is reached, the process ends. Otherwise, return to Step 2 to continue the optimization process.

Table 1
Range for the kernel parameter and the parameters of DGWO

| Parameter | Value/Range |
|---|---|
| Kernel parameter search range | [0.1, 100] |
| Population size of DGWO | 20 |
| Maximum iterations of DGWO | 100 |
| Mutation factor | 0.5 |
| Crossover probability | 0.9 |

## 4. Experimental Analysis and Results

The experimental data derives from the Paderborn University Bearing Fault Dataset, with the test rig depicted in Fig. 3. The kernel parameter of the Incremental Relevance Vector Machine is optimized using the DGWO algorithm, and the kernel parameter of the optimized Incremen-

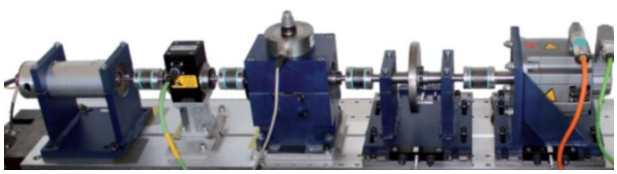


Fig. 3 Test rig of " Bearing Fault Dataset" of Paderborn

tal RVM is set to the range from 0.1 to 100. Additionally, the maximum number of iterations are set to 100.

The dataset comprising 450 samples (150 normal, 150 inner race fault, 150 outer race fault) was partitioned using stratified sampling to maintain class balance across all folds. Specifically, we employed sklearn.model_selection. Stratified K Fold with n_splits = 5, ensuring each fold contained exactly 30 samples per class (90 samples total). For each fold, 360 samples (4 folds) were used for training and 90 samples (1 fold) for testing. This stratification guarantees that the proportion of each fault category remains consistent across all training and testing subsets, preventing performance bias from class imbalance.

The remaining samples were used as the testing set. To demonstrate the enhanced performance of Bi-dimensional Extensive Empirical Wavelet Decomposition (BEEWD) over Bi-dimensional Empirical Wavelet Decomposition (BEWD), and the improved capabilities of Optimized Incremental Relevance Vector Machine (OIRVM) compared to traditional RVM, we conducted comparisons among several methods. Specifically, we tested the hybrid method combining Bi-dimensional Extensive Empirical Wavelet Decomposition and Optimized Incremental RVM (BEEWD-OIRVM) against the hybrid method of Bi-dimensional Empirical Wavelet Decomposition and traditional RVM (BEWD-RVM), as well as the traditional RVM method.

For the 1D CNN baseline, we employed a compact three-layer architecture adapted to our 128×128 time-frequency image inputs. The model consists of: 1. 64-filter Conv1D layer, kernel size 16, stride 4; 2. 128-filter Conv1D layer, kernel size 8, stride 2; 3. 256-filter Conv1D layer, kernel size 4, stride 1 – each followed by BatchNorm and max-pooling. Global average pooling reduces dimensionality before a fully-connected layer and final softmax output. Total parameters: 1.2M. For the Transformer baseline, we implemented a 4-layer encoder with 8-head self-attention (key dimension 16), feed-forward dimension 256 (GELU activation), and sinusoidal positional encoding. Input images were flattened into 256 tokens (64 pixels per token), projected to 128-dimensional embeddings. Total parameters: 0.8M.

The 1D CNN was trained using Adam optimizer ($\beta_1 = 0.9$, ($\beta_2 = 0.999$) with fixed learning rate 0.001, batch size 32, for maximum 100 epochs with early stopping (patience = 15 based on validation loss). Regularization included Dropout (0.5) and L2 weight decay ($\lambda = 0.001$). The Transformer used optimizer (weight decay ($\lambda = 0.01$) with cosine annealing schedule (initial $lr = 0.0005$, decaying to $1\times10^{-5}$), batch size 16, maximum 150 epochs, and early stopping (patience = 20).

Hyperparameters were optimized independently for each baseline using validation set performance (never

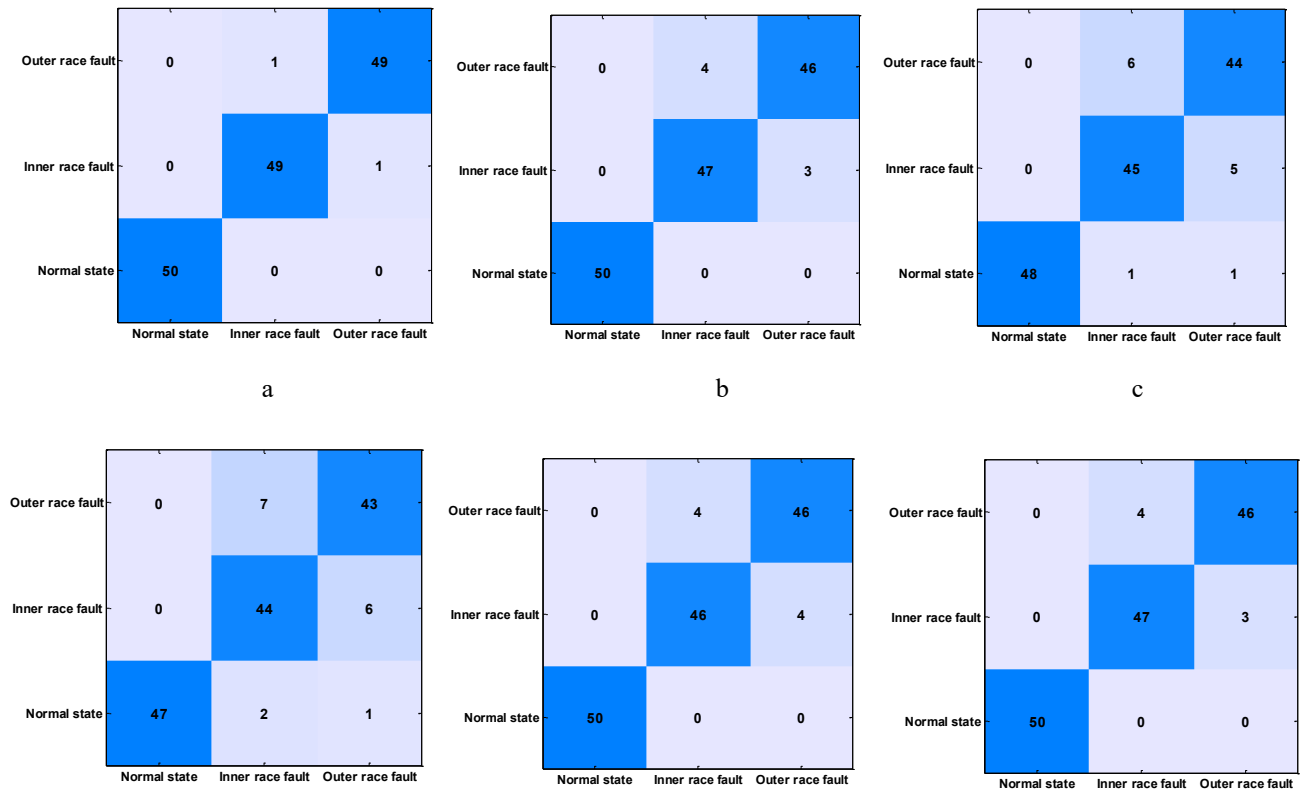



Fig. 4 Classification results of BEEWD-OIRVM, BEWD-OIRVM, BEWD-RVM, RVM,1-D CNN, Transformer: a – BEEWD-OIRVM, b – BEWD-OIRVM, c – BEWD-RVM, d – RVM, e – 1-D CNN, f – Transformer

test set). For the 1D CNN, we conducted grid search over: kernel sizes {8, 16, 32}, and filter numbers {32, 64, 128}. The reported configuration achieved highest 5-fold CV accuracy. For the Transformer, we employed Bayesian optimization via Optuna with 100 search iterations. Selection criteria prioritized highest mean accuracy with lowest standard deviation across validation folds. In Fig. 4, classification results of BEEWD-OIRVM, BEWD-OIRVM,

BEWD-RVM, RVM,1-D CNN, and Transformer are given. The 2 misclassified samples in BEEWD-OIRVM are both collected under heavy load conditions. Visual inspection of their time-frequency images reveals blurred fault characteristic frequencies, likely caused by strong background noise masking the impulsive components. This suggests the proposed method may have reduced sensitivity under extreme operating conditions. While BEEWD-OIRVM misclassified 2 samples, traditional RVM misclassified 16 samples with a more scattered error distribution across all three classes. This demonstrates that the proposed method not only reduces error quantity but also concentrates remaining errors into specific, interpretable patterns (heavy-load fault confusion), whereas weaker methods exhibit random, unexplainable failures. This approach allows clear interpretation of misclassification patterns—specifically showing that BEEWD-OIRVM's 2 misclassified samples occurred under heavy load conditions—while the quantitative headline results remain based on the full statistical aggregation across all 25 experimental runs.

All main performance numbers reported in this study originate from a stratified 5×5-fold cross-validation (CV) framework, not from a single independent hold-out test set. We adopted this approach to ensure statistical reliability and robustness of our results. The 5×5-fold CV was repeated 5 times with different random seeds, yielding 25 independent experimental runs in total. The headline accuracy of 98.4±0.8% for BEEWD-OIRVM represents the mean ± standard deviation across these 25 runs, providing a more stable estimate of true model performance than a single train-test split.

5×5-fold cross-validation is an enhanced version of cross-validation designed to improve the statistical reliability of experimental results. As indicated in Table 2, the classification accuracy of the BEEWD-OIRVM method is 98.4±0.8%, while that of BEWD-OIRVM, BEWD-RVM, traditional RVM,1-D CNN, and Transformer are 95.6±0.8%, 91.6±1.2%, 89.47±1.2%, 94.8±0.8%, 95.07±0.8%, respectively. This clearly shows that BEEWD-OIRVM outperforms the other five methods in terms of classification accuracy. Furthermore, the comparison between the classification results of BEEWD-OIRVM and BEWD-OIRVM suggests that Bi-dimensional Extensive Empirical Wavelet Decomposition (BEEWD) is more effective than Bi-dimensional Empirical Wavelet Decomposition (BEWD). Meanwhile, the comparison between BEWD-OIRVM and BEWD-RVM indicates that Optimized Incremental Relevance Vector Machine (OIRVM) is superior to traditional Relevance Vector Machine (RVM). Furthermore, BEEWD-OIRVM outperforms some modern deep learning methods, such as 1-D CNN, Transformer.

Table 2
Classification results over 5×5-fold cross-validation

| Method | Accuracy, % |
|---|---|
| BEEWD-OIRVM | 98.4±0.8 |
| BEWD-OIRVM | 95.6±0.8 |
| BEWD-RVM | 91.6±1.2 |
| RVM | 89.47±1.2 |
| 1-D CNN | 94.8±0.8 |
| Transformer | 95.07±0.8 |

## 5. Conclusions

In the process of ensemble empirical wavelet decomposition, Gaussian white noise with varying characteristics is repeatedly added to the preprocessed bearing signal prior to conducting the empirical wavelet decomposition. The Bi-dimensional extensive empirical wavelet decomposition represents an advancement from one-dimensional to two-dimensional signal processing, extending the principles of ensemble empirical wavelet decomposition. Through this method, detailed time-frequency images of the bearing signal can be generated. The classification accuracy of BEEWD- OIRVM reaches 98.4±0.8%, while that of BEWD-OIRVM is 95.6±0.8%. These results highlight that Bi-dimensional extensive empirical wavelet decomposition outperforms the conventional Bi-dimensional empirical wavelet decomposition.

Optimized incremental RVM algorithm, which adapts the original RVM formulation to create an incremental version, is used to fault diagnosis of anti-friction bearings. Incremental learning simplifies the computational process and enhances the generalization capabilities of the algorithm. Grey wolf optimization algorithm incorporating a differential evolution strategy is used to determine the parameter of kernel function. The BEWD-OIRVM achieves a classification accuracy of 95.6±0.8%, while BEWD-RVM attains 91.6±1.2%, and the traditional RVM has a classification accuracy of 89.47±1.2%. This clearly indicates that the optimized incremental RVM outperforms the traditional RVM.

The classification accuracy of the BEEWD-OIRVM method is 98.4±0.8%, while that of 1-D CNN,and Transformer are 94.8±0.8%, 95.07±0.8%,which indicates that BEEWD-OIRVM outperforms some modern deep learning methods, such as 1-D CNN, Transformer.

## Acknowledgements

This work was supported by Zhejiang Transportation Investment Group Co., Ltd. Technology Plan Project (202403).

## Data availability

The data that support the findings of this research are available from the corresponding author upon reasonable request.

## References

1. **Rodichev, A.; Novikov, A.; Gorin, A.; Tokmakova, M.** 2021. Analysis of the wear resistance of a hard anti-friction coating, applied to a plain bearing, under the conditions of boundary friction, Transportation Research Procedia 57: 573-580. https://doi.org/10.1016/j.trpro.2021.09.086.
2. **Profito, F. J.; Vladescu, S. C.; Reddyhoff, T.; Dini, D.** 2024. Numerical and experimental investigation of textured journal bearings for friction reduction, Tribology International 195: 109643. https://doi.org/10.1016/j.triboint.2024.109643.
3. **Tian, M.; An, W.; Sun, X.; Wang, L.; Chen, C.** 2025. A non-contact fault diagnosis method based on multi

information fusion networks for rolling bearings, Applied Acoustics 237: 110776. https://doi.org/10.1016/j.apacoust.2025.110776.
4. **Jiang, F.; Kuang, Y.; Li, T.; Zhang, S.; Wu, Z.; Feng, K.; Li, W.** 2025. Towards Enhanced Interpretability: A Mechanism-Driven domain adaptation model for bearing fault diagnosis across operating conditions, Mechanical Systems and Signal Processing 225: 112244. https://doi.org/10.1016/j.ymssp.2024.112244.
5. **Chen, Y.; Lei, X.; Ma, S.; Zhang, P.; Liu, X.** 2025. Fast demodulation of optical fiber Fabry-Perot pressure sensor based on the wavelet transform for blast wave measurement, Measurement 253: 117753. https://doi.org/10.1016/j.measurement.2025.117753.
6. **Baroumand, S.; Abbasi, A.R.; Mahmoudi, M.** 2025. Integrative fault diagnostic analytics in transformer windings: Leveraging logistic regression, discrete wavelet transform, and neural networks, Heliyon 11: e42872. https://doi.org/10.1016/j.heliyon.2025.e42872.
7. **Suprihadi, E.; Danila, N.; Ali, Z.** 2025. Enhancing financial product forecasting accuracy using EMD and feature selection with ensemble models, Journal of Open Innovation: Technology, Market, and Complexity 11: 100531. https://doi.org/10.1016/j.joitmc.2025.100531.
8. **Liu, P.; Song, Y.; Hou, J.; Xu, Y.** 2025. EMD-based ultraviolet radiation prediction for sport events recommendation with environmental constraint, Information Sciences 690: 121592. https://doi.org/10.1016/j.ins.2024.121592.
9. **Zhao, S.; Gao, H.; Li, X.; Li, H.; Wang, Y.; Hu, R.; Zhang, J.; Yao, W.; Li, G.** 2024. An outlier detection based two-stage EEG artifact removal method using empirical wavelet transform and canonical correlation analysis, Biomedical Signal Processing and Control 92: 106022. https://doi.org/10.1016/j.bspc.2024.106022.
10. **Nayak, A.B.; Shah, A.; Maheshwari, S.; Anand, V.; Chakraborty, S.; Kumar, T.S.** 2024. An empirical wavelet transform-based approach for motion artifact removal in electroencephalogram signals, Decision Analytics Journal 10: 100420. https://doi.org/10.1016/j.dajour.2024.100420.
11. **Pudasaini, S.; Shakya, A.; Pandey, S.P.; Paudel, P.; Ghimire, S.; Ale, P.** 2023.SMS Spam Detection using Relevance Vector Machine, Procedia Computer Science 230: 337-346. https://doi.org/10.1016/j.procs.2023.12.089.
12. **Liu, F.; Li, G.; Yang, H.** 2022. A new feature extraction method of ship radiated noise based on variational mode decomposition, weighted fluctuation-based dispersion entropy and relevance vector machine, Ocean Engineering 266: 113143. https://doi.org/10.1016/j.oceaneng.2022.113143.
13. **Ozturk, T.S.K.; Saplıoğlu, K.** 2025. Optimum design of diversion weirs with grey wolf optimization, Physics and Chemistry of the Earth, Parts A/B/C 139: 103899. https://doi.org/10.1016/j.pce.2025.103899.
14. **Ye, J.; Shi, R.; Guo, C.** 2025. Research on hierarchical emergency resource scheduling for island petrochemical enterprises based on improved multi- objective grey wolf optimization algorithm, Energy 322: 135791. https://doi.org/10.1016/j.energy.2025.135791.

Y. Li, K. Xu, X. Chen, Q. Zheng, Y. Chen, Z. Yin

FAULT DIAGNOSIS OF ANTI-FRICTION BEARINGS BASED ON BI-DIMENSIONAL EXTENSIVE EMPIRICAL WAVELET DECOMPOSITION AND OPTIMIZED INCREMENTAL RVM

S u m m a r y

The prompt and efficient diagnosis of anti-friction bearing faults is crucial for the proper functioning of rotating machinery. In this paper, a novel approach to anti-friction bearing fault diagnosis utilizing Bi-dimensional extensive empirical wavelet decomposition combined with optimized incremental RVM is presented. As an enhanced version of ensemble empirical wavelet decomposition, Bi-dimensional extensive empirical wavelet decomposition extends its application from 1D to 2D signal processing, enabling the extraction of more comprehensive anti-friction bearing features. Additionally, the optimized incremental RVM is employed for fault diagnosis of anti-friction bearings. The experimental findings indicate that Bi-dimensional extensive empirical wavelet decomposition surpasses Bi-dimensional empirical wavelet decomposition, and optimized incremental RVM outperforms traditional RVM. The proposed method incorporating Bi-dimensional extensive empirical wavelet decomposition and optimized incremental RVM proves to be highly effective for anti-friction bearing fault diagnosis.

**Keywords**: bi-dimensional extensive empirical wavelet decomposition, incremental RVM, anti-friction bearing, fault diagnosis.

Received June 5, 2025
Accepted April 24, 2026