A SIGNAL PROCESSING AND OPTIMIZATION-BASED APPROACH FOR BEARING FAULT DIAGNOSIS USING SVMD, SVD, AND WO-SVM

AO HUNG LINH
Faculty of Mechanical Engineering, Industrial University of Ho Chi Minh City, Viet Nam,
Corresponding author: aohunglinh@iuh.edu.vn
DOIs: https://www.doi.org/10.46242/jstiuh.v80i2.5907

Abstract.
Roller bearing acceleration vibration signals are inherently non-stationary and noisy. However, benchmark datasets commonly used in pattern recognition often contain minimal noise, whereas real-world bearing vibration signals typically include both the bearing's inherent vibration and additional noise from related machine components. To enhance the accuracy of fault diagnosis methods when processing real-world signals, Gaussian noise was intentionally added to the original signals. Subsequently, the Successive Variational Mode Decomposition (SVMD) method was employed to decompose these signals into multiple modes. The Singular Value Decomposition (SVD) technique was then applied to construct a feature matrix from these modes, which served as input to the Support Vector Machine (SVM) classifier. The SVM hyperparameters (C and γ) were optimized using the Walrus Optimizer (WO) algorithm. Experimental analysis on signals under both normal and faulty bearing conditions demonstrates that the proposed fault diagnosis method achieves higher classification accuracy and faster convergence compared to other methods.
Keywords: Fault diagnosis, Roller Bearing, Successive Variational Mode Decomposition, Singular Value Decomposition, Walrus Optimizer, Support Vector Machine.

1 INTRODUCTION

The sound and secure functioning of industrial setups relies heavily on identifying malfunctions in rotating components. Rolling element bearings are susceptible to various types of faults, which may lead to severe operational failures if not detected in a timely manner. Among various monitoring techniques, vibration signal analysis has proven to be one of the most effective approaches. However, vibration signals acquired from faulty bearings are often non-stationary, non-linear, and heavily contaminated with environmental noise.To effectively decompose these complex signals, several methods have been developed, chẳng hạn như Empirical Mode Decomposition (EMD) [1] và Variational Mode Decomposition (VMD) [2]. While VMD addresses some limitations of EMD , it still extracts all modes simultaneously and can be sensitive to parameter initialization. Recently, Successive Variational Mode Decomposition (SVMD) [3] has been proposed as a superior alternative, extracting intrinsic mode components (IMCs) sequentially with better adaptability and lower computational complexity. For classification tasks, Support Vector Machine (SVM) has become a widely used intelligent tool due to its excellent performance in high-dimensional spaces [4]. The classification accuracy of SVM is highly dependent on two key hyperparameters: the penalty parameter (C) and the kernel exponent (γ). Traditionally, these parameters are selected based on user experience or heuristic trials, which often fail to find the global optimum. Metaheuristic optimization algorithms, such as Particle Swarm Optimization (PSO) [5], Genetic Algorithm (GA) [6], and Whale Optimization Algorithm (WOA) [7], are commonly employed to automate this tuning process. Despite these advancements, current literature reveals a significant gap. Studies combining VMD with SVD and SVM [8, 9] or VMD-SVD with Adaboost [10] have shown high accuracy, but they are typically validated on "clean" benchmark datasets with minimal noise. In real-world industrial environments, fault signatures are frequently buried under significant background noise. Unlike most previous studies [8-10] that rely on pristine laboratory data, this research intentionally introduces Additive White Gaussian Noise (AWGN) into the original signals to simulate harsh machinery operations and test the robustness of the diagnostic framework. Furthermore, while researchers have explored SVMD-based indices [11] or combined SVMD with Salp Swarm Optimization [12], these methods may still struggle with convergence speed and local optima in high-noise conditions. There is currently no comprehensive framework that integrates the sequential precision of SVMD, the feature extraction strength of SVD, and an SVM classifier optimized by the state-of-the-art Walrus Optimizer (WO). To fill this gap, this paper introduces the AWGN-SVMD-SVD-WO-SVM model.

© 2026 Trường Đại học Công nghiệp thành phố Hồ Chí Minh

By intentionally adding noise to the training and testing phases, we demonstrate that our hybrid approach achieves superior stability and classification accuracy compared to traditional methods using PSO, GA, or WOA.

The paper is structured as follows: Section 2 introduces the AWGN, SVMD, and SVD methods. Section 3 introduces the WO optimization algorithm. Section 4 presents the parameter optimization of SVM based on the WO algorithm. The application of the WO-SVM method to diagnose bearing damage is presented in Section 5. The experimental process and results are detailed in Section 6, and the conclusion is presented in Section 7 of this paper.

2 SUCCESSIVE VARIATIONAL MODE DECOMPOSITION- SINGULAR VALUE DECOMPOSITION

2.1 Successive Variational Mode Decomposition

It is known that VMD's simultaneous IMF extraction leads to significantly increased computation time with a high number of signal modes (K). In contrast, SVMD is a sequential mode extraction algorithm, which accelerates convergence and avoids extracting unwanted modes [3]. Mathematically, the method suggests that the input signal $f(t)$ breaks down into two parts: the Lth mode ($u_L(t)$) and the residual signal ($f_r(t)$), shown as:

$$f(t) = u_L(t) + f_r(t) \tag{1}$$

where $f_r(t)$ comprises two components: the sum of previously obtained modes and the unprocessed signal part $f_u(t)$.

$$f_r(t) = \sum_{i=1}^{L-1} u_i\,(t) + f_u(t) \tag{2}$$

The iterative expression for SVMD is as follow:

$$\hat{u}_L^{n+1}(\omega) = \frac{\hat{f}(\omega)+\alpha^2(\omega-\omega_L^n)^4\hat{u}_L^n(\omega)+\frac{\hat{\lambda}(\omega)}{2}}{[1+\alpha^2(\omega-\omega_L^n)^4]\left[1+2\alpha(\omega-\omega_L^n)^2+\sum_{i=1}^{L-1}\frac{1}{\alpha^2(\omega-\omega_i)^4}\right]} \tag{3}$$

In this context, α serves as the weight controlling the data fidelity term. Since its magnitude is typically quite high, the update formula for ω_L can be simplified as shown below:

$$\omega_L^{n+1} = \frac{\int_0^\infty \omega\left|\hat{u}_L^{n+1}(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_L^{n+1}(\omega)\right|^2 d\omega} \tag{4}$$

The Lagrange multipliers λ are expresed as follows:

$$\hat{\lambda}^{n+1} = \hat{\lambda}^n + \tau\left[\hat{f}(\omega) - \left(\hat{u}_L^{n+1}(\omega) - f_u^{n+1}(t) + \sum_{i=1}^{L-1} u_i^{n+1}(\omega)\right)\right] \tag{5}$$

Ultimately, the revised expression is derived by applying the double ascent method as shown below:

$$\begin{cases} y(\omega) = \dfrac{\alpha^2(\omega-\omega_L^n)^4\left[\hat{f}(\omega)-\hat{u}_L^{n+1}(\omega)-\sum_{i=1}^{L-1}\hat{u}_i(\omega)+\frac{\hat{\lambda}(\omega)}{2}\right]-\sum_{i=1}^{L-1}\hat{u}_i(\omega)}{1+\alpha^2\left(\omega-\omega_L^{n+1}\right)^4} \\ \hat{\lambda}^{n+1} = \hat{\lambda}^n + \tau\left[\hat{f}(\omega) - \left(\hat{u}_L^{n+1}(\omega) - f_u^{n+1}(t) + \sum_{i=1}^{L-1} u_i^{n+1}(\omega)\right)\right] \end{cases} \tag{6}$$

where τ is the update parameter and ω_L is the center frequency.

In this research, the number of decomposition modes for the SVMD algorithm is established at $K = 3$. This configuration is strategically chosen based on the physical characteristics of bearing vibration signals, which typically comprise the fundamental rotation frequency, fault-induced impulse components, and background noise. By setting $K = 3$, the SVMD process can effectively isolate these primary constitutive components without leading to over-decomposition. This ensures the extraction of a concise yet highly discriminative feature matrix during the subsequent SVD stage, while simultaneously reducing the computational complexity for the metaheuristic optimization process.

2.2 Singular Value Decomposition

Singular value decomposition (SVD) has many advantages such as its noise filtering ability, noise sensitivity, and high-resolution spectrum analysis. This technique aims to analyze a matrix to produce singular values, singular vectors.

Let Σ be an $M \times N$ matrix, defined by the equation $\Sigma = E\Delta V^T$

(7)

Here, $E = [e_1, e_2, e_3,\ldots, e_n] \in R^{N\times N}$ and $V = [v_1, v_2, v_3,\ldots,v_n] \in R^{M\times M}$ are orthonormal matrices (i.e., $V^TV = I$ and $E^TE = I$). The matrix $\Delta \in R^{N\times M}$ is a diagonal matrix $\Delta = [\mathrm{diag}\{\sigma_1, \sigma_2,\ldots,\sigma_p\}]$, where $p = \min(N, M)$ and the diagonal elements are ordered singular values such that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$. Within this context, e_i and v_i correspond to the ith left and right singular vectors of Σ, respectively, and σ_i are the singular values of Σ.

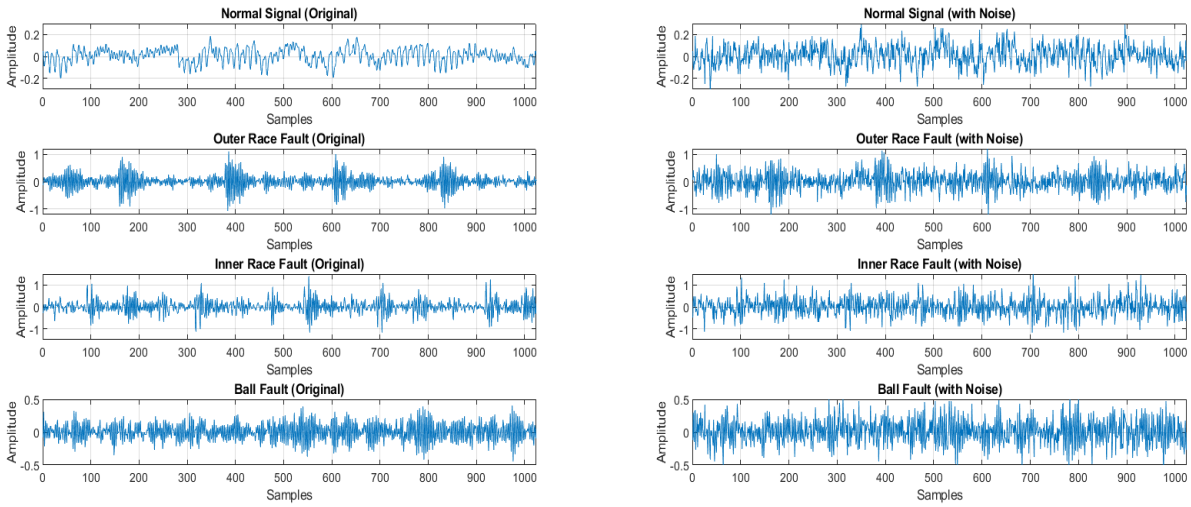

Figure 1: Roller bearing acceleration vibration signals with noise (left figure) and without noise (right figure) under four conditions.

2.3 Additive White Gaussian Noise (AWGN)

The addition of Gaussian white noise serves to simulate realistic operating environments, where sensor signals are often affected by various sources of noise such as mechanical vibrations, electromagnetic interference, and background disturbances. By introducing noise into the vibration signals, the robustness and generalization capability of the diagnostic algorithm can be more effectively evaluated. This ensures that the model is not only accurate under ideal conditions but also reliable in practical applications where noise is inevitable.

The generated signals are scaled relative to the original ones according to the signal-to-noise ratio (*SNR*). This metric, defined as the ratio of signal power to noise power, is expressed in decibels as:

$$SNR_{dB} = 10log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) \tag{8}$$

where P_{signal} is the power of origin signal and P_{noise} is the power of noise. In this study, $SNR_{dB} = 0$ dB was added to the original signal to verify the effectiveness of the proposed method.

Figure 1 presents the acceleration signals of a bearing under four different health conditions: normal, outer race fault, ball fault, and inner race fault. The raw signals are shown in the left column, while the right column depicts the corresponding signals after the addition of Gaussian white noise. This simulates practical scenarios where sensor measurements are often contaminated by environmental or electronic disturbances.

2.4 Feature Matrix Based on SVMD And SVD

In this research, we first augment the raw roller bearing signals $x(t)$ by adding Gaussian noise to enhance the robustness of the decomposition process. The noise-added signal, denoted as $f(t)$, is then processed using the SVMD method to decompose it into several Modes (*MDs*). The resulting *MDs* are subsequently separated into two initial feature vector matrices, M_1 and M_2.

$$M_1 = \begin{bmatrix} MD_1 \\ MD_2 \\ \vdots \\ MD_J \end{bmatrix}, \; M_2 = \begin{bmatrix} MD_{J+1} \\ MD_{J+2} \\ \vdots \\ MD_n \end{bmatrix} \tag{9}$$

where $J = n/2$ if n is an even number, and $J = (n+1)/2$ if n is an odd number. From these initial feature vector matrices, we can extract the characteristics of the roller bearing vibration signal $f(t)$. Furthermore, fault feature vectors are identified as the singular values, which inherently reflect the fundamental

characteristics of the vector matrices M_1 and M_2, and consequently, the roller bearing vibration signal. Table 1 presents the singular values extracted via the SVMD-SVD method from four rolling bearing samples, with Gaussian noise added to the original signals.

Table 1: Acceleration vibration signals of four rolling bearing samples after being added with Gaussian noise and extracted by SVMD-SVD method.

Condition of roller bearing	Singular value of fault feature derived from Equation (7) σ_i			
	σ_1	σ_2	σ_3	σ_4
Normal (NOR)	0.8533	0.1573	0.7129	0.0710
Outer-race fault (ORF)	1.3530	0.4883	0.9730	0.1828
Inner-race fault(IRF)	1.8247	1.2743	1.0712	0.2389
Ball fault (BF)	0.7061	0.4828	0.4973	0.0936

3 WALRUS OPTIMEZ ALGORITHM

The central idea of the Walrus Optimizer (WO) algorithm is to create a new metaheuristic optimization algorithm by mathematically modeling the intelligent natural behaviors of walruses [13]. The algorithm mimics the life cycle and survival strategies of these animals, particularly their methods of foraging for food, migrating, and defending against predators, to effectively solve complex optimization problems.

3.1 Initialization

The optimization process begins with initialization, where random candidate solutions X are generated within the solution space. This step is performed as follows:

$$X = LB + rand(UB - LB) \tag{10}$$

In this optimization, the walruses serve as the agents. Their search for solutions is bounded by lower (LB) and upper (UB) limits, with their movements influenced by *rand*, a uniform random vector between 0 and 1. With each iteration their positions matrix X are continuously updated and can be expressed as follows:

$$X = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,d} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1} & X_{n,12} & \cdots & X_{n,d} \end{bmatrix}_{n\times d} \tag{11}$$

where d is the dimension of design variables, n is the population size,.

The fitness values for every search agent are presented in Equation (11)

$$F = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,d} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ f_{n,1} & f_{n,12} & \cdots & f_{n,d} \end{bmatrix}_{n\times d} \tag{12}$$

3.2 Danger signals and safety signals

Walruses assign 1 or 2 animals to patrol the area while the group is foraging or sleeping. When there is a sign of trouble, the patrolling walruses will send out a distress signal. The danger signal and safety signal in WO can be expressed as follows:

$$Danger_signal = A^{*}R \tag{13}$$

Danger factors A and R can be presented as follows:

$$A = 2 \times \alpha \tag{14}$$

$$R = 2 \times r_1 - 1 \tag{15}$$

in which $\alpha = 1 - \frac{t}{T}$, the value of α drops from 1 to 0 over the iterations, t, where T represents the total number of iterations.

The safety signal in WO is established in relation to the danger signal, and can be expressed as follows:

$$Safety_signal = r_2 \tag{16}$$

where, r_1 and r_2 are random numbers in the range of (0, 1).

3.3 Migration Phase (Exploration):

Inspired by the migration of walrus herds to better feeding grounds or safer locations, this phase guides the search agents (solutions) towards the best solution found so far. This allows the algorithm to converge on promising regions of the search space that were identified during exploration. The walrus position is updated as follows:

$$X_{i,j}^{t+1} = X_{i,j}^{t} + \text{Migration_step} \tag{17}$$

$$Migration_step = (X_m^t - X_n^t) \times \beta \times r_3^2 \tag{18}$$

$$\beta = 1 - \frac{1}{1+\exp(-\frac{t-T/2}{T}\times 10)} \tag{19}$$

Let $X_{i,j}^t$ denote the current coordinate of the *i*th walrus in the *j*th dimension, while $X_{i,j}^{t+1}$indicates its new position after the update. The parameter *Migration_step* determines how far the walrus moves. X_m^t and X_n^t refer to two randomly selected vigilant walruses from the group at iteration *t*. The migration step's control factor, β, evolves smoothly as the iterations proceed. The random variable r_3 is drawn from a uniform distribution in the interval (0, 1).

3.4 Exploitation Phase (Defensive Behavior):

During reproduction, two distinct behaviors occur: onshore roosting and underwater foraging. The mathematical model for these can be presented as follows:

3.4.1 Roosting behavior

Step 1: Redistribution of male walruses

The male walrus's position is updated using a Halton sequence distribution. To ensure both randomness and uniformity, we divide the search area into several sections and then select a random point from each section.

Step 2: Position update of female walruses

The male walrus ($Male_{i,j}^t$) and the lead walrus (X_{best}^t) both influence the female walrus ($Female_{i,j}^t$).

$$Female_{i,j}^{t+1} = Female_{i,j}^t + \alpha \times \left(Male_{i,j}^t - Female_{i,j}^t\right) + (1-\alpha) \times (X_{best}^t - Female_{i,j}^t) \tag{20}$$

where, $Female_{i,j}^{t+1}$ is the new position for the *i*th female walrus on the *j*th dimension, $Male_{i,j}^t$ and $Female_{i,j}^t$are the positions of the *i*th male and female walruses on the *j*th dimension.

Step 3: Position update of juvenile walruses

Young walruses on the outskirts of the group are frequently hunted by predators such as killer whales and polar bears. As a result, they must constantly adjust their location to reduce the risk of being caught.

$$Juvenile_{i,j}^{t+1} = (O - Juvenile_{i,j}^t) \times P \tag{21}$$

$$O = X_{best}^t + Juvenile_{i,j}^t \times LF \tag{22}$$

In this context, $Juvenile_{i,j}^{t+1}$denotes the updated location of the ith juvenile walrus along the *j*th dimension, while $Juvenile_{i,j}^t$ represents its current coordinate on that axis. The parameter *P* serves as the distress factor, randomly selected within the range (0, 1). *O* indicates the designated safe position, and *LF* is a vector generated from a Lévy distribution, modeling the Lévy flight behavior.

$$Levy(\alpha) = 0.05 \times \frac{x}{|y|^{\frac{1}{\alpha}}} \tag{23}$$

where *x* and *y* are two normally distributed variables, *x* $N(0, \sigma_x^2)$, *y* $N(0, \sigma_y^2)$.

$$\sigma_x = \left[\frac{\Gamma(1+\alpha)\sin(\pi\frac{\alpha}{2})}{\Gamma\left(\frac{1+\alpha}{2}\right)\alpha 2^{\frac{(\alpha-1)}{2}}}\right]^{\frac{1}{\alpha}}, \sigma_y = 1, \alpha = 1.5 \tag{24}$$

where, σ_x and σ_y are the standard deviation, $\Gamma(x) = (x+1)!$.

3.4.2 Foraging behavior

a) Fleeing behavior:

$$X_{i,j}^{t+1} = X_{i,j}^t \cdot R - \left|X_{best}^t - X_{i,j}^t\right| \cdot \mathrm{r}_4^2 \tag{25}$$

The term $\|X_{best}^t - X_{i,j}^t\|$ indicates the Euclidean distance between the current walrus and the best-known solution at iteration *t*, while r_4 is a uniformly distributed random number in the range (0, 1).

b) Gathering behavior:

Walruses exhibit cooperative foraging behavior by adjusting their movements based on the spatial positions of other individuals within the population. The exchange of locational information among group members facilitates the collective identification of marine regions with higher food availability.

$$X_{i,j}^{t+1} = (X_1 + X_2)/2 \tag{26}$$

$$\begin{cases} X_1 = X_{best}^t - a_1 \times b_1 \times \left|X_{best}^t - X_{i,j}^t\right| \\ X_2 = X_{second}^t - a_2 \times b_2 \times \left|X_{second}^t - X_{i,j}^t\right| \end{cases} \tag{27}$$

$$a = \beta \times r_5 - \beta \tag{28}$$

$$b = \tan(\theta) \tag{29}$$

In these formulations, X_1 and X_2 represent weighting factors that influence the gathering behavior of walruses, while X_{second} denotes the position of the second walrus at iteration t. The term $\left|X_{second}^t - X_{i,j}^t\right|$ quantifies the distance between the current walrus and the second individual. Parameters a and b serve as gathering coefficients, r_5 is a uniformly distributed random variable within the interval (0, 1), and θ is an angular parameter ranging from 0 to π.

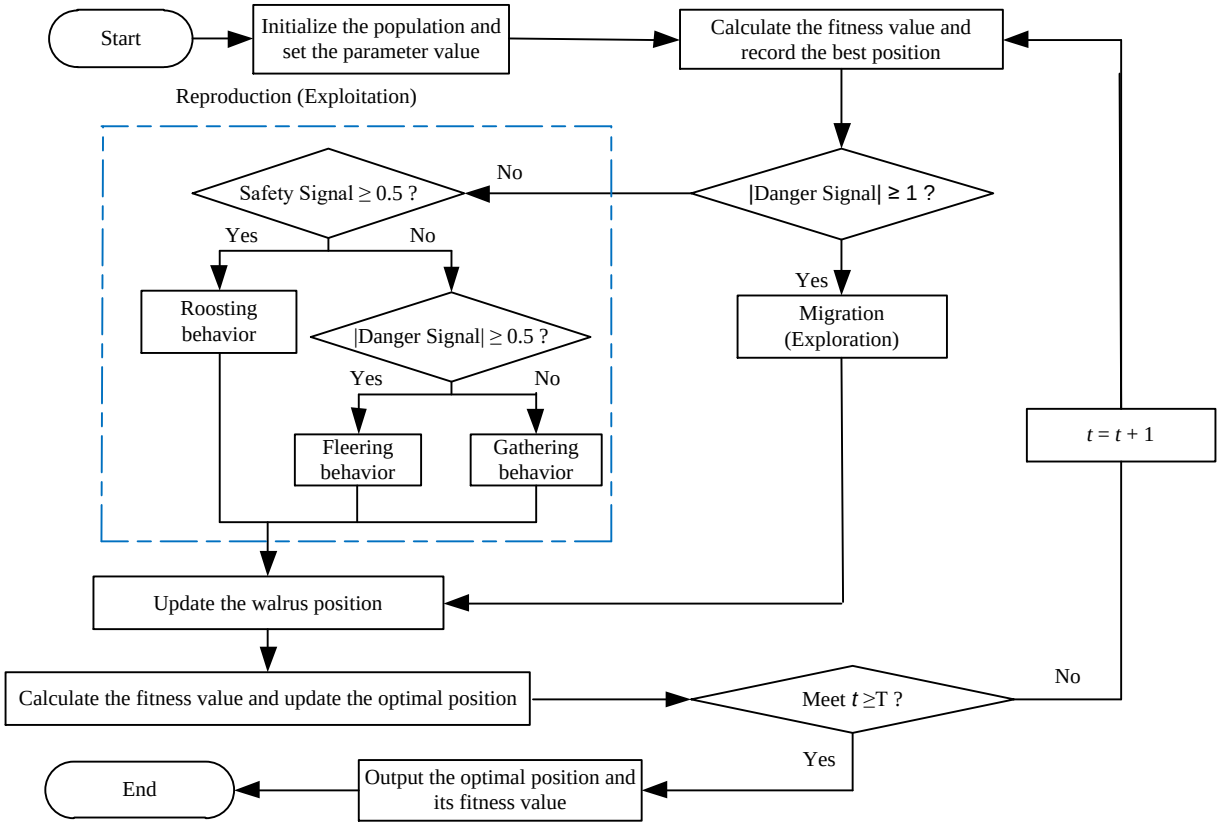

Figure 2: Flowchart of Walrus Optimizer algorithm.

4 WO-BASED SVM PARAMETERS OPTIMIZATION

4.1 Support Vector Machine

The standard binary SVM classifies input vectors into two distinct categories using a sign function. To enhance class separability, SVM transforms training samples from the input space into a higher-dimensional feature space through the application of a mapping function Φ.

Given a training dataset $G = \{(x_i, y_i); i = 1, 2, ..., l\}$, where each sample $x_i \in R_d$ is labeled by its class $y \in \{+1, -1\}$, the objective function for cases where the training data isn't linearly separable in the input feature space is formulated as follows [4]:

$$Min\ \phi(\omega) = \frac{1}{2}\langle\omega|\omega\rangle + C\sum_{i=1}^{l}\xi_i \tag{30}$$

$$subject\ to\ y_i(\langle\omega.\phi(x_i)\rangle + b) \geq 1 - \xi_i,\ \ \xi_i \geq 0, i = \{1,2,\dots,l\}$$

where ω denotes the normal vector of the hyperplane, C represents the penalty parameter, b is the bias term, ξ_i are slack variables constrained to be non-negative, and $\Phi(x)$ denotes the mapping function.

By introducing the non-negative Lagrange multipliers $\alpha_i \geq 0$, the optimization problem can be reformulated as follows:

$$Max\ L(\omega, b, \alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i\, \alpha_j y_i y_j K(x_i, x_j) \tag{31}$$

$$subject\ to\ 0 \le \alpha_i \le C, \sum_{i=1}^{l} \alpha_i y_i = 0$$

The decision function is expressed as:

$$f(x) = sgn\left[\sum_{i=1}^{l} \alpha_i y_i K(x_i . x) + b\right] \tag{32}$$

The radial basis function (RBF) kernel is employed in Equation (30) to transform the initial problem domain into a Gaussian one, as shown by the following formula:

$$K(x_i, x_j) = exp\left(-\gamma \|x_i - x_j\|^2\right) \tag{33}$$

where γ is the parameter that determines the influence or reach of each individual training sample.

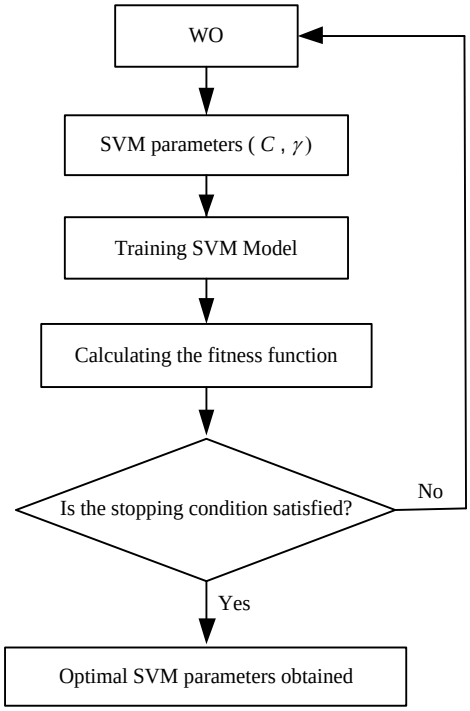

Figure 3: The parameter optimization flowchart of SVM based on WO

4.2 SVM Parameter Optimization Based on the Walrus Optimizer Algorithm

As highlighted in previous studies, the performance of Support Vector Machines (SVM) is highly sensitive to its parameter settings. Typically, the selection of SVM parameters relies heavily on user experience. In this paper, the authors propose employing the Walrus Optimizez (WO) algorithm to determine the optimal SVM parameters.

The parameter pair (C, γ) are the optimization variables, and the SVM validation error acts as the fitness function, defined as follows:

$$G(C, \gamma) = \mathrm{Test}_{\mathrm{Error}_{\mathrm{SVM}}}(C, \gamma) \tag{34}$$

where $G(C, \gamma)$ is a fitness function and $Test_{Error_{SVM}}(C, \gamma)$ is define:

$$Test_{Error_{SVM}} = \frac{Number\ of\ incorrect\ classification\ in\ validation\ samples}{Total\ number\ of\ samples\ in\ validation\ set} \tag{35}$$

The Walrus Optimization (WO) algorithm is used to optimise the SVM training process. It finds the best SVM parameters to reduce validation error and improve the SVM's ability to generalize new data. Each individual in the WO population is randomized at the start, after which the SVM algorithm computes its corresponding output weight matrix. After this, WO evaluates the "fitness" of each individual in the population. This entire process repeats until a predefined stopping criterion is met. Figure 3 illustrates the complete WO-SVM algorithm.

4.2.1 Parameter Setting For Algorithms

Parameters of the WO algorithm were configured as follows: *SearchAgents*: 20; *Max_iter* = 30; K = 3; *alpha* = 2000; N = 500; *num_runs* = 20. Parameters of the PSO algorithm were fixed with the values given in the Ref. [14, 15]; that is, W = 0.75, $c_1 = c_2$ = 1.5, the numbers of particles was 20, and the iteration count was 30. Parameters of the GA algorithm were configured as follows: Population Size (*NP*): 20; Maximum Number of Generations (*Max_Gen*): 30; Termination Criteria: 10^{-6}.

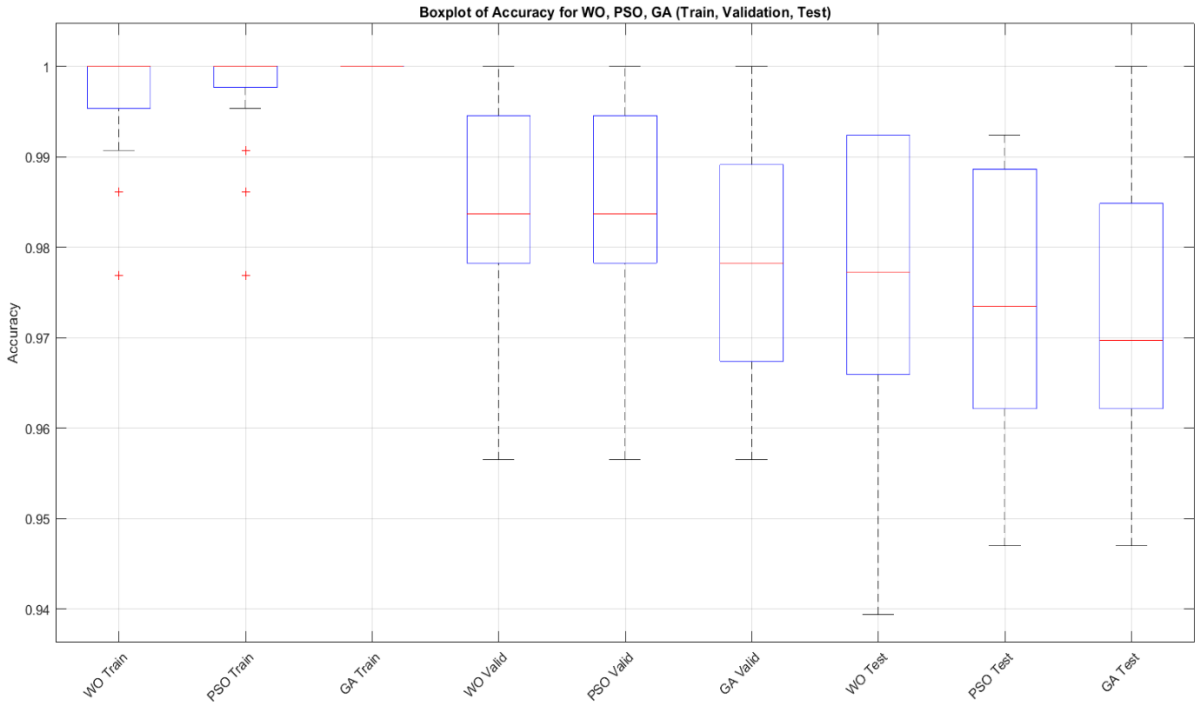

Figure 4: Boxplot accuracy of WO, PSO, and GA over 20 runs

4.2.2 Algorithm Performance Analysis For Algorithms

a) Analysis of Boxplot Accuracy for WO, PSO, and GA

The "Boxplot of Accuracy for WO, PSO, GA (Train, Validation, Test)" illustrates the performance consistency of the three optimization methods across different datasets (Figure 4).

- *Training Accuracy*: All three methods (WO, PSO, and GA) exhibit exceptionally high and stable accuracy on the training set, with median values at or very close to 100%. The minimal spread and short whiskers indicate robust performance and low variability across 20 runs.
- *Validation Accuracy*: WO demonstrates the highest validation accuracy, with its median residing between 98% and 99% and showing relatively low dispersion. PSO's validation accuracy is slightly lower than WO's, with a median also in the 98-99% range but a marginally lower central tendency. GA records the lowest validation accuracy, with a median below 98% and wider interquartile range, signifying greater variability.
- *Test Accuracy*: WO consistently achieves the highest test accuracy, with a median between 97% and 98% and a narrow boxplot, indicating high consistency across runs. PSO's test accuracy is slightly inferior to WO's, with its median below 97.5% and a wider spread. GA shows the lowest test accuracy among the three, with its median between 96% and 97% and the broadest boxplot, reflecting the highest variability in performance on the test set.

From the above analysis, it can be seen that WO consistently outperforms PSO and GA in terms of accuracy on both validation and test sets, exhibiting higher median accuracies and superior stability (narrower interquartile ranges). This suggests that WO possesses better generalization capabilities for the given problem.

b) Analysis of Optimization Algorithm Convergence Curves

Based on the "Average Convergence Curves over 20 runs" plot, the convergence performance of three optimization algorithms WO, PSO, and GA are analyzed as follows (Figure 5):

- *Convergence Speed*: WO demonstrates superior convergence speed, exhibiting a significant reduction in "Mean Best Score" within the initial iterations and achieving stability the earliest. PSO also shows rapid

convergence but is slower than WO. In contrast, GA exhibits the slowest convergence speed among the three algorithms.

- *Achieved Optimal Value*: WO attains the lowest "Mean Best Score", maintaining this level from around the 8th-9th iteration until the end. PSO converges to a value which is higher than WO. GA yields the least optimal result with the highest convergence value.
- *Stability*: All three algorithms exhibit stability after convergence; however, WO and PSO achieve stability at significantly better score levels compared to GA.

From the above analysis it can be seen that WO demonstrates superior optimization performance in this problem, characterized by faster convergence and the ability to find a more optimal solution compared to both PSO and GA.

5 APPLICATION OF WO-SVM TO ROLLER BEARING FAULT DIANOSIS

5.1 Data Acquisition

The dataset utilized in this study was obtained from the Case Western Reserve University Bearing Data Center Website (CWRUBDCW), with authorization from Professor Loparo [16] (see Figure 6). The experimental setup consisted of a 2-horsepower Reliance Electric motor, a torque transducer/encoder, a dynamometer, and associated control electronics. Data were acquired at a sampling frequency of 485063 Hz, with the motor operating at 1797 rpm. The bearings used in the experiment were SKF deep groove ball bearings, specifically the 6205-2RS JEM model. To create the dataset, artificial faults measuring 0.007 inches in diameter were introduced into the test bearings using electro-discharge machining. The dataset then comprises 440 samples, representing four distinct bearing conditions: normal (NOR), inner race fault (IRF), outer race fault (ORF), and ball fault (BF). These samples were divided into three subsets: one for testing and the other two for training and validation. Details of the vibration signal samples are summarized in Table 2.

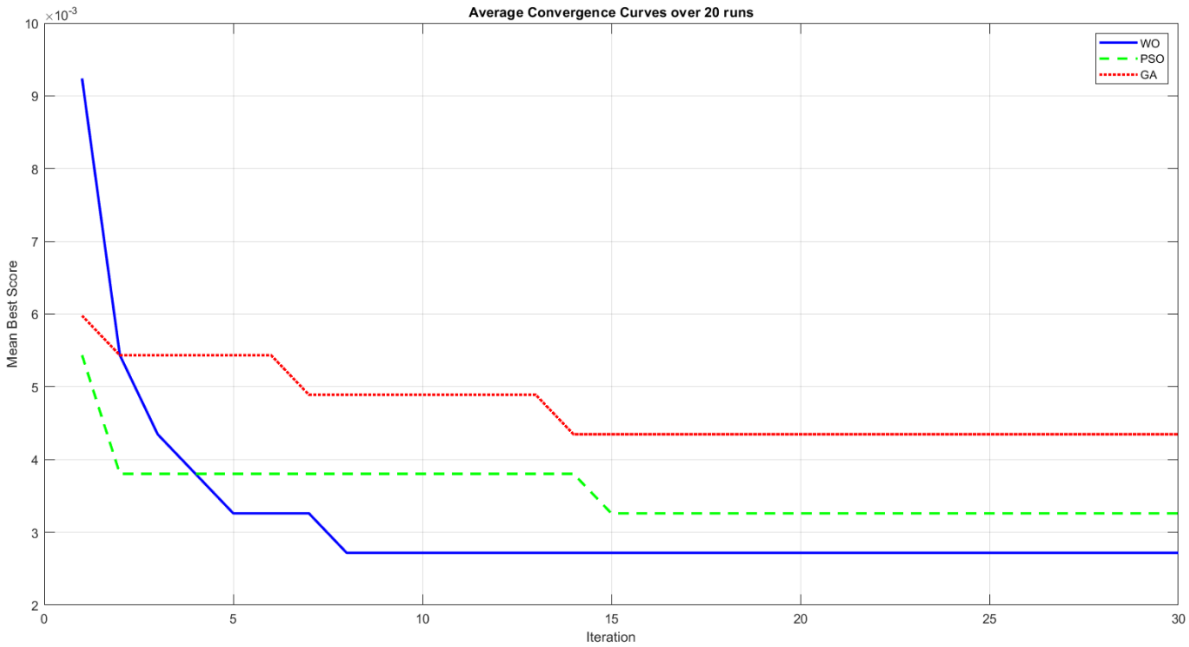

Figure 5: Average Convergence Curve of WO, PSO, and GA over 20 runs

Table 2: Collection of vibration signal samples

Roller bearing conditions	Class	Collection samples	Training samples	Validation samples	Test samples
Inner-race fault (IRF)	1	110	66	22	22
Outer-race fault (ORF)	2	110	66	22	22
Ball fault (BF)	3	110	66	22	22
Normal (NOR)	4	110	66	22	22

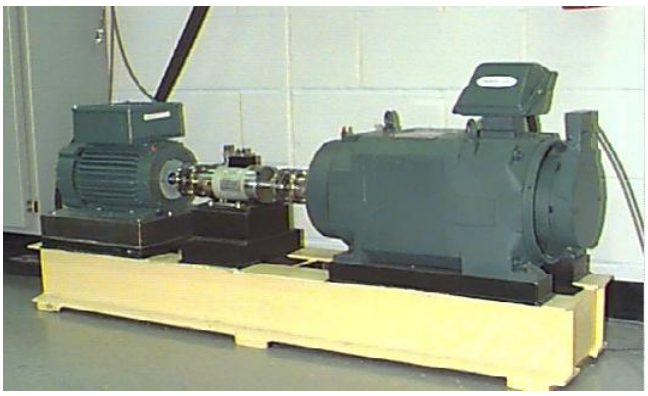

Figure 6: The roller bearing vibration data collection model was extracted from [16].

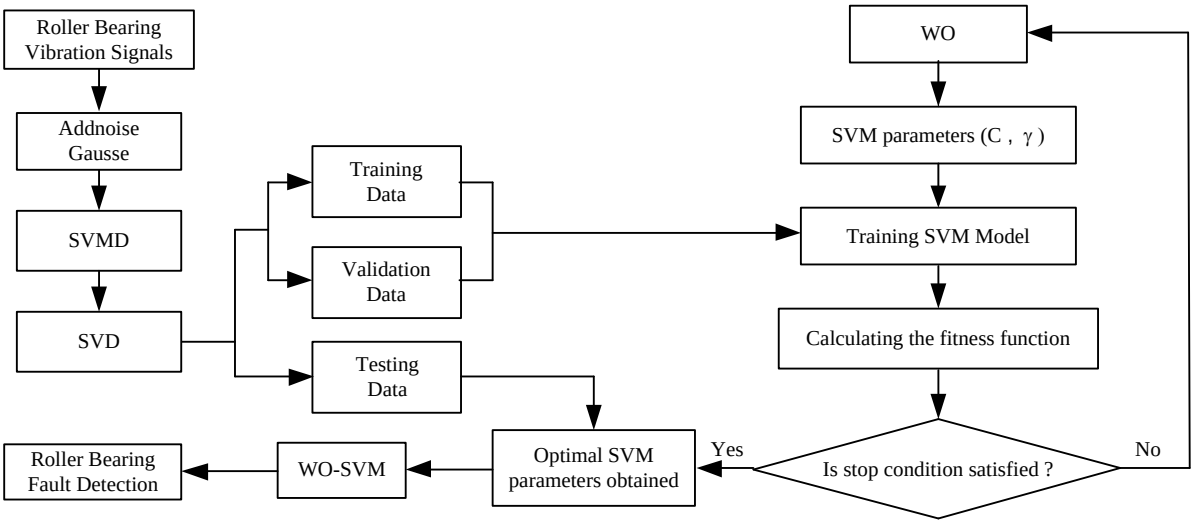

Figure 7: Roller bearing fault detection method based on SVMD-SVD and WO-SVM.

5.2 Application

Based on the above analysis, it is evident that the SVMD-SVD representations of roller bearing vibration signals under various operating conditions and fault types exhibit clear distinctions. In this study, the fault characteristics extracted from each σ_i are utilized as input vectors for the WO-SVM classifier. This approach effectively identifies the operational state and fault category of the bearing. The flowchart illustrating the proposed roller bearing fault diagnosis framework, which integrates SVMD-SVD and WO-SVM, is presented in Fig. 7. The experimental procedure is outlined as follows:

(1) Add Gaussian white noise to the bearing acceleration vibration signal at a signal-to-noise ratio (SNR) of 0 dB to simulate real-world noisy conditions.

(2) Apply SVMD to the vibration signals, resulting in a set of mode functions MD_k .

(3) Perform SVD on each selected mode function MD_k using Equation (9).

(4) Form the final feature vector Z by combining the singular values extracted from the SVD. This vector, referred to as the SVMD-SVD vector, is denoted as (Eq. 36):

$$Z = [\sigma_1, \sigma_2, \ldots, \sigma_m] \tag{36}$$

(5) Divide the feature dataset into three subsets with a ratio of 60% for training, 20% for validation, and 20% for testing.

(6) Evaluate the similarity and distribution of the data by computing the maximum and minimum values across the dataset.

(7) Use WO-SVM for classification:

- Train and validate the Walrus Optimization-based SVM (WO-SVM) classifier using the training and validation sets.
- After optimization, extract the optimal hyperparameters C and γ, which are then used for final testing.

In the four-class classification problem, we implement a One-vs-Rest (OvR) strategy. This approach involves constructing four independent binary classifiers, each optimized using the WO algorithm and employing Support Vector Machines (WO-SVMi, i = 1, 2, 3, 4). Each classifier is trained to recognize a specific target class, labeling samples belonging to that class as +1 and all other classes as −1. Here are the details for each classifier (Table 3):

Table 3: Description of Four WO-SVM Classifiers in One-vs-Rest Strategy

Classifier	Target Class (+1)	Other Classes (−1)	Purpose	Final Decision Rule
WO-SVM$_1$	IRF (Inner Race Fault)	ORF, BF, NOR	Detects inner race damage	If decision score from WO-SVM$_1$ is the highest → classify as **IRF**
WO-SVM$_2$	ORF (Outer Race Fault)	IRF, BF, NOR	Detects outer race damage	If decision score from WO-SVM$_2$ is the highest → classify as **ORF**
WO-SVM$_3$	BF (Ball Fault)	IRF, ORF, NOR	Detects ball damage	If decision score from WO-SVM$_3$ is the highest → classify as **BF**
WO-SVM$_4$	NOR (Normal Condition)	IRF, ORF, BF	Identifies normal condition (no fault)	If decision score from WO-SVM$_4$ is the highest → classify as **NOR**

Table 4: The comparing result means of the classification of bearing failure of the classifiers SVMD-SVD-WO-SVM with SVMD-SVD-PSO-SVM, and SVMD-SVD-GA-SVM methods.

Method	C and γ optimal of test		Accuracy TRAIN	Accuracy VALID	Accuracy TEST	Cost time total (s)
	C	γ				
SVMD-SVD-WO-SVM	191.5339	0.9219	100.00%	99.41%	97.12%	30.26
SVMD-SVD-PSO-SVM	277.0014	1.0992	100.00%	98.41%	96.97%	32.13
SVMD-SVD-GA-SVM	209.2350	1.4308	100.00%	97.82%	95.45%	35.56

6 RESULTS AND DISCUSION

Table 3 indicates that the dataset exhibits similar values for the training, validation, and test sets, with $\boldsymbol{x}_i$ representing the dataset. Consequently, strong training and validation outcomes can be expected to yield favorable test results.

Fig. 5 illustrates WO's faster convergence speed compared to PSO or GA. WO reached a validation error value of 2.8×10^{-3} in the 7th generation, whereas PSO attained the error value of 2.8×10^{-3} in the 15th generation. GA, on the other hand, reached a validation error value of 4.3×10^{-3} in the 14th generation.

Table 4 presents a comparative evaluation of three hybrid fault diagnosis frameworks: SVMD-SVD-WO-SVM, SVMD-SVD-PSO-SVM, and SVMD-SVD-GA-SVM. All models follow a consistent diagnostic pipeline in which Gaussian white noise is first added to the original bearing vibration signals to simulate real-world disturbances and assess the robustness of each approach under noisy conditions. Subsequently, the signals are decomposed using SVMD, and key features are extracted via SVD. The resulting features are then classified using a SVM, whose parameters are optimized through different metaheuristic algorithms: WO, PSO, and GA, respectively.

Among the evaluated approaches, the SVMD-SVD-WO-SVM model achieves the highest performance with accuracies of 100%, 99.41%, and 97.12% for the training, validation, and testing sets, respectively. While the PSO-SVM (100%, 98.41%, 96.97%) and GA-SVM (100%, 97.82%, 95.45%) also demonstrate excellent learning abilities, they show a more noticeable decline in performance on the testing set. This indicates that while all models converge well during training, the SVMD-SVD-WO-SVM framework possesses superior generalization and robustness—especially under noisy conditions—whereas the PSO and GA-based models are relatively more prone to overfitting. The superior performance of the proposed SVMD-SVD-WO-SVM framework is also closely linked to the optimal selection of the decomposition level, K = 3. This specific configuration allows the SVMD process to effectively isolate the most discriminative components of the bearing vibration-namely the fundamental frequency, fault-induced impulses, and the added Gaussian noise. By capturing these three essential modes, the subsequent SVD stage is provided with a high-quality feature matrix that is free from the interference of over-decomposition or insufficient signal separation. This precise feature representation significantly simplifies the search space for the Walrus Optimizer, enabling the SVM to achieve a near-perfect classification accuracy while maintaining robust generalization across all datasets

The superior optimization performance of WO in tuning SVM hyperparameters (C, γ) can be attributed to its unique biological modeling of walrus behaviors. Specifically, the use of Halton Sequence in the redistribution of male walruses ensures a more uniform coverage of the search space compared to the pure random initialization used in PSO and GA. This low-discrepancy distribution prevents the population from clustering in limited areas, thereby enhancing the global exploration of potential SVM parameter pairs. Furthermore, the integration of Levy Flight in the update process for juvenile walruses provides a mechanism for escaping local optima. By modeling movement with a Lévy distribution—characterized by many small steps and occasional long jumps—WO can effectively "jump" out of stagnated regions in the complex SVM error landscape where PSO and GA might prematurely converge.

From the perspective of computational efficiency, the WO-SVM model is also the most time-efficient, requiring only 30.26 seconds, which is faster than the PSO-SVM (32.13s) and GA-SVM (35.56s) approaches. This reflects the potential of the WO algorithm to simultaneously optimize classification performance and reduce computational overhead.

In summary, the inclusion of Gaussian noise provides a realistic benchmark for robustness, and the experimental results clearly show that the WO-based optimization framework offers the most reliable and efficient performance for fault diagnosis under practical, noisy conditions. The C and γ of each classifier are presented in Table 4, in which the classification results of the proposed method give an error equal to PSO-SVM and lower than GA-SVM with the validation data set. However, as shown in Fig.5, the convergence speed of WO-SVM is faster than PSO, so it gives better classification results in the test data set. The classification results in the test set are also presented in Table 4. At the same time, the program running time of the SVMD-SVD-DE-SVM method is shorter than the SVMD-SVD-PSO-SVM and SVMD-SVD methods -GA-SVM.

7 CONCLUSION

This paper proposed a robust hybrid fault diagnosis framework, SVMD-SVD-WO-SVM, designed to handle bearing vibration signals under realistic noisy conditions. The experimental results demonstrate that the proposed method significantly outperforms PSO-SVM and GA-SVM in terms of both classification accuracy and convergence efficiency.The core strength of the Walrus Optimizer (WO) in this framework lies in its superior ability to tune the SVM hyperparameters (C, γ). By integrating the Halton Sequence, the algorithm ensures a more uniform and comprehensive exploration of the search space, preventing the population from becoming trapped in localized clusters. Additionally, the incorporation of Lévy Flight provides a powerful stochastic mechanism for the model to "jump" out of local optima, a common limitation observed in PSO and GA. These mechanisms collectively ensure that the SVMD-SVD-WO-SVM model achieves an optimal balance between global exploration and local exploitation, resulting in a high test accuracy of 97.12% and a faster convergence rate.In conclusion, the combination of advanced signal decomposition (SVMD), robust feature extraction (SVD), and an enhanced metaheuristic optimizer (WO) provides a reliable and efficient solution for bearing fault diagnosis. Future work could explore the scalability of this WO-optimized framework for multi-fault classification in even more complex industrial environments.

ACKNOWLEDGMENT

The author expresses gratitude to Professor K. A. Loparo of Case Western Reserve University for permitting the use of data from the Bearing Data Center in this study.

REFERENCES

[1] N. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences,* vol. 454, pp. 903-995, 1998.

[2] K. Dragomiretskiy and D. Zosso, "Variational Mode Decomposition," *IEEE Transactions on Signal Processing,* vol. 62, no. 3, pp. 531-544, 2014.

[3] M. Nazari and S. M. Sakhaei, "Successive variational mode decomposition," *Signal Processing,* vol. 174, p. 107610, 2020.

[4] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Springer: New York. , 1995.

[5] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, pp. 1942-1948, 1995.
[6] J. H. Holland, Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence: The MIT Press, 1992.
[7] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," *Advances in Engineering Software*, vol. 95, pp. 51-67, 2016.
[8] Y. ZHANG and W. HU, "Fault diagnosis of rotating machinery based on VMD-SVD and SVM," *Journal of Mechanical and Electrical Engineering*, vol. 39, no. 3, pp. 324-329, 2022.
[9] K. D. Chen Jian, Sun Taihua, Zhang Lei, "Rolling bearing fault diagnosis method based on SVD-VMD and SVM," *Journal of Electronic Measurement and Instrumentation*, vol. 36, no. 1, pp. 220-226, 2022.
[10] T. Z. C. Lu, X. Liu, R. Qin, and J. Ma, "Cylindrical roller bearing fault diagnosis based on VMD-SVD and Adaboost classifier method," *Vibroengineering PROCEDIA*, vol. 17, pp. 19–24, 2018.
[11] Y. Guo, Y. Yang, S. Jiang, X. Jin, and Y. Wei, "Rolling Bearing Fault Diagnosis Based on Successive Variational Mode Decomposition and the EP Index," (in eng), *Sensors (Basel)*, vol. 22, no. 10, 2022.
[12] T. Thelaidjia, N. Chetih, A. Moussaoui, and S. Chenikher, "Successive variational mode decomposition and blind source separation based on salp swarm optimization for bearing fault diagnosis," *The International Journal of Advanced Manufacturing Technology*, vol. 125, no. 11, pp. 5541-5556, 2023.
[13] M. Han, Z. Du, K. F. Yuen, H. Zhu, Y. Li, and Q. Yuan, "Walrus optimizer: A novel nature-inspired metaheuristic algorithm," *Expert Systems with Applications*, vol. 239, p. 122413, 2024.
[14] F. Ardjani, K. Sadouni, and M. Benyettou, "Optimization of SVM MultiClass by Particle Swarm (PSO-SVM)," in *2010 2nd International Workshop on Database Technology and Applications*, pp. 1-4, 2010.
[15] K. Chandramouli and E. Izquierdo, *Image Classification using Chaotic Particle Swarm Optimization*., pp. 3001-3004, 2006.
[16] Loparo. K. A, *Bearings vibration dataset, Case Western Reserve University*. Available: https://engineering.case.edu/bearingdatacenter/download-data-file (Accessed on September 10, 2025)

PHƯƠNG PHÁP TIẾP CẬN DỰA TRÊN XỬ LÝ TÍN HIỆU VÀ TỐI ƯU HÓA ĐỂ CHẨN ĐOÁN HƯ HỎNG Ổ LĂN BẰNG PHƯƠNG PHÁP SVMD, SVD VÀ WO-SVM

AO HÙNG LINH
Khoa Công nghệ Cơ khí, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
Tác giả liên hệ: aohunglinh@iuh.edu.vn

Tóm tắt. Tín hiệu dao động gia tốc của ổ lăn vốn có tính chất không dừng và chứa nhiều nhiễu. Tuy nhiên, các tập dữ liệu chuẩn thường được sử dụng trong nhận dạng mẫu lại chứa rất ít nhiễu, trong khi tín hiệu rung ổ lăn trong thực tế thường bao gồm cả dao động vốn có của ổ lăn và nhiễu phát sinh từ các thành phần máy liên quan. Để nâng cao độ chính xác của các phương pháp chẩn đoán lỗi khi xử lý tín hiệu thực, nhiễu Gaussian đã được chủ động thêm vào các tín hiệu gốc. Tiếp theo, phương pháp phân tích chế độ biến thiên liên tiếp (SVMD) được áp dụng để phân tách các tín hiệu này thành nhiều chế độ. Sau đó, kỹ thuật phân tích giá trị kỳ dị (SVD) được sử dụng để xây dựng ma trận đặc trưng từ các chế độ này, đóng vai trò làm đầu vào cho bộ phân loại máy vector hỗ trợ (SVM). Các siêu tham số của SVM (C và γ) được tối ưu hóa bằng thuật toán Walrus Optimizer (WO). Kết quả thực nghiệm trên các tín hiệu ở cả điều kiện ổ lăn bình thường và hư hỏng cho thấy phương pháp chẩn đoán được đề xuất đạt độ chính xác phân loại cao hơn và tốc độ hội tụ nhanh hơn so với các phương pháp khác.

Từ khóa. Chẩn đoán lỗi, Ổ lăn, Phân tích chế độ biến thiên liên tiếp, Phân tích giá trị kỳ dị, Bộ tối ưu hóa Walrus, Máy vectơ hỗ trợ.

Received on September 23– 2025
Revised on January 16 – 2025