

Enhancing air compressors multi fault classification using new criteria for Harris Hawks optimization algorithm in tandem with MODWPT and LSSVM classifier

Advances in Mechanical Engineering
2023, Vol. 15(12) 1–14
© The Author(s) 2023
DOI: 10.1177/16878132231216208
journals.sagepub.com/home/ade



Chemseddine Rahmoune¹ , Mohammed Amine Sahraoui² ,
Fawzi Gougam¹ , Mohamed Zair¹ and Ikhlas Meddour³

Abstract

The evolution of industrial systems toward Industry 4.0 presents the challenge of developing robust and accurate models. In this context, feature selection plays a pivotal role in refining machine learning models. This paper addresses the imperative of accurate fault diagnosis in industrial systems, focusing on air compressors. These systems, vital for efficient operations, demand early fault detection to prevent performance degradation. Conventional methods often encounter challenges due to the occurrence of similar failure patterns under comparable conditions. To address this limitation, our approach delves into a more complex scenario, where air compressors operate under diverse fault conditions. This study introduces novel feature selection criteria achieved through a fusion of the Maximal Overlap Discrete Wavelet Packet Transform (MODWPT), the Harris Hawks Optimization (HHO) algorithm, and the Least Squares Support Vector Machine (LSSVM) classifier. The synthesis of these components aims to bolster the multi-fault diagnosis accuracy and stability for each fault class. The evaluation focuses on key statistical metrics—minimum, maximum, mean, and standard deviation. Experimental outcomes underscore the method's superiority over traditional feature selection techniques. The approach excels in accuracy and stability, particularly across various fault categories, affirming the efficacy and resilience of the new criteria. The symbiotic integration of MODWPT, HHO, and LSSVM within our framework highlights its potential to elevate classification performance in the realm of industrial fault diagnosis.

Keywords

Fault diagnosis, air compressors, multi-fault classification, feature selection, Harris Hawks optimization, MODWPT, LSSVM classifier, industrial systems, industry 4.0, machine learning, accuracy, stability, signal processing, fault detection

Date received: 20 June 2023; accepted: 31 October 2023

Handling Editor: Chenhui Liang

Introduction

Air compressors are an essential tool in many industries, such as manufacturing, construction and chemical processing. They produce compressed air or gas that powers many machines, tools and equipment. However, air compressors can develop faults and breakdown, as can any mechanical system, which can lead to reduced performance and expensive downtime.¹

¹Solid Mechanics and Systems Laboratory (LMSS), University M'hamed Bougara, Boumerdes, Algeria

²Systems Engineering and Telecommunications Laboratory (LIST), University M'hamed Bougara Boumerdes, Boumerdes, Algeria

³Mechanics and Structures Laboratory, University 8 Mai 1945 Guelma, Algeria

Corresponding author:

Mohammed Amine Sahraoui, Systems Engineering and Telecommunications Laboratory (LIST), University M'hamed BOUGARA Boumerdes, Boumerdes, 35000, Algeria.

Email: ma.sahraoui@univ-boumerdes.dz



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work

without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

In order to maintain peak performance, increase the compressor's lifespan, and reduce maintenance costs, problems in the compressor must be quickly identified and corrected. Visual inspections, manual monitoring, and routine maintenance programs are examples of traditional approaches that are labor- and time-intensive. Also, they might miss some defects, especially those that are still developing.²

Modern technologies like vibration analysis,³ thermal imaging,⁴ and acoustic monitoring⁵ have been developed to get around these constraints. These methods continuously check the compressor's status and immediately detect any anomalies using a variety of sensors and equipment. These reducing technologies ensure that the compressor stays in excellent working condition for a longer period of time by identifying deficiencies early, which prevents the emergence of more serious and expensive problems.¹

The use of signal processing techniques has made it possible to quickly and precisely identify compressor faults.⁶ These methods entail examining the acoustic signal from the compressor to spot distinct patterns or features connected to various kinds of faults. First, the signal is divided into smaller, easier-to-manage signals that can be examined more quickly. Then, different algorithms are used to evaluate these signals to find traits that point to particular faults. Eventually, the features are categorized into several kinds of faults by machine learning algorithms.

In the last years, many methods for signal processing have been introduced for feature extraction, such as discrete wavelet transform (DWT),⁷ recursive Empirical Mode Decomposition (REMD),⁸ Empirical Mode Decomposition,⁹ Variational Mode Decomposition (VMD),¹⁰ Empirical Wavelet Transform (EWT)¹¹ and Wavelet Packet Transform,¹² can be used to find compressor faults, but these techniques have a certain drawback, especially in acoustic signals of different faults categories. This paper proposed, the Maximal overlap discrete wavelet Package transform (MODWPT) for feature information, the MODWPT is an improved version of the Discrete Wavelet Transform (DWT),¹³ it is a multi-resolution analysis has the ability to analyze the signals at different levels of localization by levels in the decomposition of acoustic signals in the time and frequency domains and also has the property of shift-invariance and is very highly capable of decomposing the approximate and detailed signals in the time-frequency analysis.¹⁴ Nevertheless, the obtained results from the MODWPT are complex and too large due to different faults categories in our case, thus, in order to enhance the robustness and efficiency of the feature classification, many methods can be employed for dimensionality reduction and feature selection.

Nowadays, many researchers have been task challenge to develop methods that helps to increase the accuracy of the classification by the decreased in the number of feature entries that have a direct impact on the classification results. In order to obtain these results, many optimization algorithms have been introduced to enhance feature classification through feature selection of the most pertinent and informative features,¹⁵ several optimization algorithms have been applied for feature selection such as Artificial Bee Colony (ABC),¹⁶ Genetic Algorithm (GA),¹⁷ Slime Mould Algorithm (SMA),¹⁸ Generalized Normal Distribution Optimization (GNDO),¹⁹ Manta Ray Foraging Optimization (MRFO),²⁰ has been developed, which overcomes the limitations of dimensionally reduction algorithms that suffering from higher computational complexity and helping to find the optimal features considering parameters for selection, and it gives satisfactory results in terms of global accuracy without stability calculation. By contrast, our proposed method are based on Harris Hawks algorithm (HHO)²¹ to improve the performance, accuracy and stability of each class, and generalization of a machine learning model, as well as to reduce the time and resources required for training and testing the model through the measuring the four keys of each class (Min, Max, Mean, and Standard deviation).

As an alternative to performing and achieving supervised learning of early detection and complex classification tasks, machine learning (ML) is proposed.²² This type of technique typically consists of two main steps: the first step is a key task for identification and diagnosis based on the feature extraction information and feature selection; the second step is to build a recognition and categorization model of the air compressor health condition. K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Naive Bayes (NB), Extra Tree (ET), Decision Trees (DT), Random Forests (RF), and Least Squares Support Vector Machines are a few examples of machine-learning classifiers (LSSVM).^{23,24} These techniques have been proposed to extract the best classifier in terms of accuracy and stability of each class.

This research proposes a reliable and improved multi-fault diagnosis of air-compressor. This technique is combined between the MODWPT, HHO, and LSSVM.

The paper is organized as follows: Section 2 introduces the principal diagnosis and feature extraction. Section 3 presents faults selection and performance classification. Section 4 shows the experimental benchmark description. The last section presents the obtained results with comments and conclusion. The contributions and innovations of this research is developing a fault selection and classification approach based on stability and accuracy of each class are as follows:

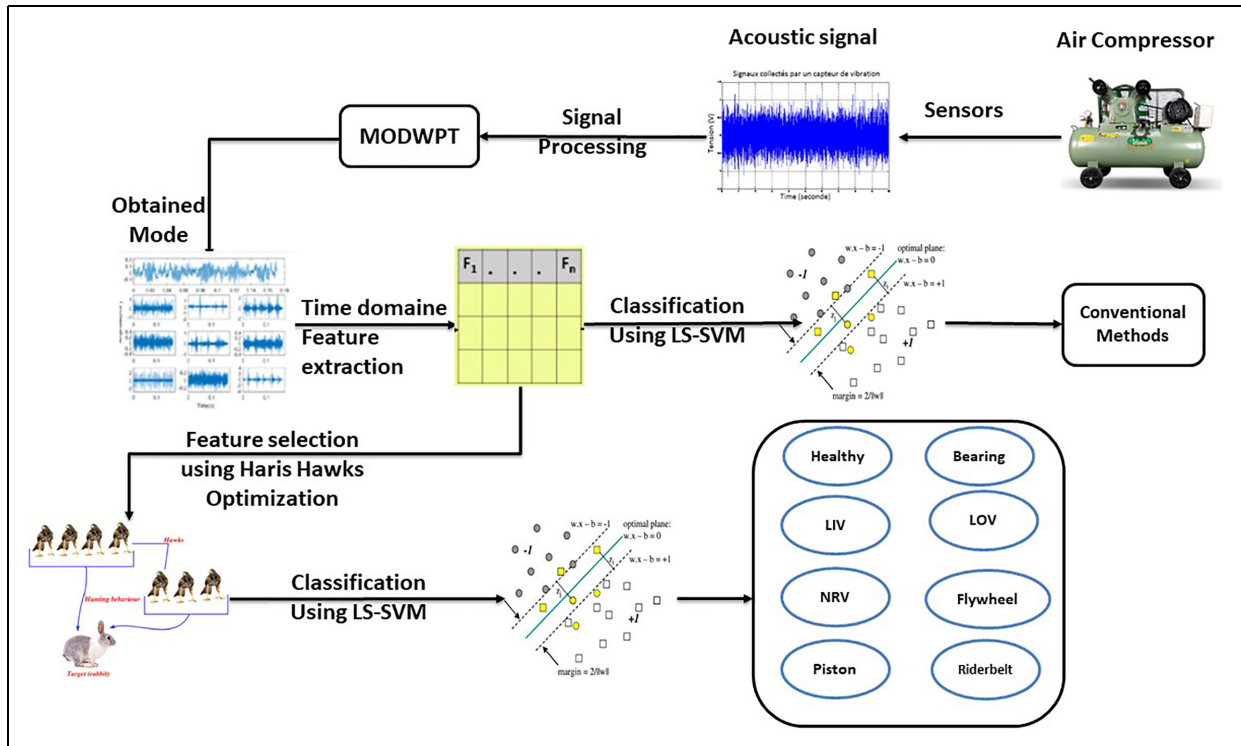


Figure 1. Flowchart of the proposed method.

- ✓ Firstly, statistical features are extracted from the Maximal overlap discrete wavelet Package transform (MODWPT) in time domain.
- ✓ During the training process, the features extraction is feed into multiple classifiers based on with and without optimization algorithms by measuring the mean, max, min, and standard deviation to obtain the higher average accuracy and best stability.
- ✓ Data processing with optimization algorithms, which can reduce the computational time by select the appropriate set of features to predict air compressor state.
- ✓ Comparison study was done between proposed approach and the existing typical methods in terms of global accuracy and stability versus accuracy and stability for each class.

Our proposed method for detecting air compressor faults can be broken down into a series of steps, as outlined in the flowchart in Figure 1. First, the acoustic signal is processed using MODWPT to extract the various AM-FM modes. Next, time domain features are extracted from these modes. To classify the faults, a LSSVM classifier is used. To further improve the feature selection process, we compare the classification results to those obtained using conventional methods and apply the Harris Hawks optimization algorithm (HHO) to eliminate unimportant parameters. Finally,

we train a model using the supervised learning method “LSSVM” to detect faults. Our method has been tested on acoustic signal data, and the results indicate that it has a high performance.

Signal processing and feature extraction

Maximal overlap discrete wavelet packet transform

Feature extraction is a technique used to reduce the dimensionality of a dataset by generating new features that can effectively summarize the existing ones. This is typically done by combining the existing features in a meaningful way, which can help to improve the accuracy of machine learning models.²⁵

One popular method for signal decomposition is the Maximal Overlap Discrete Wavelet Packet Transform (MODWPT), which is based on the intrinsic frequencies of the signal. This technique decomposes the original signal into several modes, which are then used to extract features that are combined into a global matrix for machine learning analysis.²⁶

MODWPT is an advanced version of the Discrete Wavelet Transform (DWT), which is widely used for analyzing signals in the time-frequency domain. The DWT decomposes a signal into approximation and detail coefficients using high-pass and low-pass filters, which are then sampled by a factor of 2. This process is repeated for the approximated signal at each decomposition level until the desired tolerance is reached.

MODWPT is a time-invariant transformation that is similar to DWT,²⁷ but with some important differences. In particular, the distances between peaks are made equal, and the down-sampling process is removed, resulting in coefficients of the same length as the input signal. As a result, all of the decomposed coefficients correspond to their time-series and are associated with the original signal.²⁸

To obtain the DWT of a sampled sequence of continuous time data $X = [X0, X1, \dots, XN - 1]$ is obtained by using the even-length scaling (low-pass) filter $[g_l: l = 0, L-1]$ and the wavelet (high-pass) filter $[h_l: l = 0, L-1]$, where L is a power of 2. The low-pass filters satisfy the following equation:

$$g_l = (-1)^{l+1} h_{L-l-1} \quad (1)$$

Where g_l represents the filter coefficients. The quadrature mirror filter property, which states that the product of the coefficients of the low-pass filter and high-pass filter, shifted by an integer n , is zero, links the two filters together. The following can be written:

$$\sum_{l=0}^{L-1} g_l^2 = \sum_{l=0}^{L-1} g_l g_{l+2n} = \sum_{l=-\infty}^{+\infty} g_l g_{l+2n} = 0 \quad (2)$$

Where gh represents the coefficients of the high-pass filter.

The wavelet filter coefficients can be obtained from the scaling filter coefficients using the following equation:

$$h_l = (-1)^l g_{L-l-1} \quad (3)$$

Where $h_l : l$ represents the l -th coefficient of the wavelet filter. The DWT is an effective tool for signal processing that may be applied to a variety of tasks, such as feature extraction and data compression.

Since both provide the quadrature mirror of any non-zero integer, the two filters are related to one another. The j th-level wavelet and scaling coefficients for $t = [0, N-1]$ can be derived as follows:

$$V_{j,t} = \sum_{l=0}^{L-1} g_l V_{j-1, (2t+1-l) \bmod N_{j-1}} \quad (t = 0, \dots, N_{j-1}) \quad (4)$$

$$W_{j,t} = \sum_{l=0}^{L-1} h_l V_{j-1, (2t+1-l) \bmod N_{j-1}} \quad (t = 0, \dots, N_{j-1}) \quad (5)$$

Where MOD stands for the modulus after division.

order to ensure energy conservation, the defining filters can be scaled as follows, as shown in equations (6) and (7):

$$\tilde{g}_l = \frac{g_l}{\sqrt{2}} \quad (6)$$

$$\tilde{h}_l = \frac{h_l}{\sqrt{2}} \quad (7)$$

Equation (1) can be changed as illustrated in equation (8) by using these scaling factors:

$$\sum_{l=0}^{L-1} \tilde{g}_l^2 = \frac{1}{2}, \quad \sum_{l=0}^{L-1} \tilde{g}_l \tilde{g}_{l+2n} \quad (8)$$

The expressions for the quadrature mirror filters can be updated as per equations (9) and (10):

$$\tilde{h}_l = (-1)^l \tilde{g}_{L-l-1} \quad (9)$$

$$\tilde{g}_l = (-1)^l \tilde{h}_{L-l-1} \quad (10)$$

To address the down-sampling issue, MODWT uses new filters that ensure $2^{j-1} - 1$ zeros between the elements of $[\tilde{g}_l]$ and $[\tilde{h}_l]$. The scaling coefficients $[V_{j,t}^M]$ are generated by the pyramid algorithm of MODWT, while the MODWT wavelet coefficients $[M_{j,t}^M]$ are produced as per equations (11) and (12). Here, the summation is over the range of $l = 0$ to $l-1$ and $t = 0$ to $N-1$.

$$V_{j,t} = \sum_{l=0}^{L-1} \tilde{g}_l V_{j-1, (2t+1-l) \bmod N} \quad (t = 0, \dots, N-1) \quad (11)$$

$$W_{j,t} = \sum_{l=0}^{L-1} \tilde{h}_l V_{j-1, (2t+1-l) \bmod N} \quad (t = 0, \dots, N-1) \quad (12)$$

MODWPT is a further developed method that aims to achieve perfect resolution at high frequencies. The sequence of MODWPT coefficients at level j and frequency-index n is denoted as $W_{j,n} = [W_{j,n,t}, t = 0, \dots, N-1]$, where $[W_{j,n,t}]$ is generated using equation (13).

Here, the summation is over the range of $i = 0$ to $l-1$ and $t = 0$ to N_{j-1} . The scaling coefficients $[V_{j,t}]$ are generated using equation (14), where the summation is over the range of $l = 0$ to $l-1$ and $t = 0$ to N_{j-1} .

$$W_{j,n,t} = \sum_{i=0}^{l-1} \tilde{f}_{n,l} W_{j-1, \left[\frac{n}{2} \right], (t-2^{j-1}l) \bmod N} \quad (t = 0, \dots, N_{j-1}) \quad (13)$$

$$V_{j,t} = \sum_{l=0}^{L-1} g_l V_{j-1, (2t+1-l) \bmod N_{j-1}} \quad (t = 0, \dots, N_{j-1}) \quad (14)$$

Where $\tilde{f}_{n,l} = \tilde{g}_l$ when $n \bmod 4 = 0$ or 3 , while $\tilde{f}_{n,l} = \tilde{h}_l$ when $n \bmod 4 = 1$ or 2 .¹³

Features selection and classification

Feature selection involves identifying the most relevant input features for a machine-learning task, while

ignoring those that are irrelevant or redundant. This can enhance the accuracy of a model, simplify it, and prevent over fitting.²⁹ Feature selection methods may include filter,³⁰ wrapper,³¹ or embedded³² approaches.

Using techniques like decision trees, logistic regression, support vector machines, and neural networks, classification involves grouping input data into predetermined classes.³³ The task at hand and the features of the data will choose which algorithm is being used. Several machine learning applications, including image analysis, processing of natural languages, and predictive modeling, require on classification.

The method proposed in the article employs the Harris Hawks Optimization (HHO) algorithm to optimize the objective function of the least square support vector machine (LSSVM) classification algorithm. The goal of this approach is to improve the performance of the LSSVM model through parameter optimization. The HHO algorithm is based on the hunting behavior of hawks and by integrating it with LSSVM, the proposed method aims to enhance the accuracy of classification.

Harris Hawks optimization algorithm

The HHO algorithm was first proposed in 2019²¹ as a new Meta heuristic optimization algorithm inspired by the hunting behavior of Harris Hawks in the wild. The algorithm is designed to solve a wide range of optimization problems, including continuous, discrete, and combinatorial problems.³⁴

The HHO algorithm uses a population of hawks that hunt for prey in a search space. The population of hawks is randomly initialized with each hawk's position and velocity.³⁵ The population of hawks is randomly initialized in the search space. Let N be the number of hawks in the population, and D be the dimension of the search space. Then, the position and velocity of each hawk i are initialized as follows:

- Position:

$$\begin{aligned} x_{ij} &= [x_{i1}, x_{i2}, \dots, x_{iD}], \text{ where } x_{ijk} \\ &= lb_{jk} + rand() * (ub_{jk} - lb_{jk}) \end{aligned} \quad (15)$$

Here, lb_{jk} and ub_{jk} are the lower and upper bounds of the k -th dimension of the search space, respectively. $rand()$ is a random number between 0 and 1.

- Velocity:

$$\begin{aligned} v_{ij} &= [v_{i1}, v_{i2}, \dots, v_{iD}], \text{ where } v_{ijk} \\ &= rand() * (ub_{jk} - lb_{jk}) \end{aligned} \quad (16)$$

Based on its own experience and the experience of other hawks in the population, each hawk adjusts its location and velocity during the hunting phase. Equations (17) and (18) update the location and velocity of each hawk by taking into consideration the hawk's current position, velocity, and the positions of other hawks in the population. The objective is to investigate the search space and identify a viable solution.

- Position update:

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \quad (17)$$

Here, $x_{ij}(t)$ and $x_{ij}(t+1)$ are the positions of hawk i at time steps t and $t+1$, respectively.

$v_{ij}(t+1)$ is the updated velocity of hawk i at time step $t+1$.

- Velocity update:

$$\begin{aligned} v_{ij}(t+1) &= A * v_{ij}(t) + C1 * rand() * (p_{ij} - x_{ij}(t)) \\ &+ C2 * rand() * (p_g - x_{ij}(t)) \end{aligned} \quad (18)$$

The velocities of hawk I at time steps t and $t+1$ are shown here as $v_{ij}(t)$ and $v_{ij}(t+1)$, respectively. The best hawk in the local group of hawk I is in position p_{ij} , while the best hawk overall in the population is in position p_g . Scaling factors A , $C1$, and $C2$ regulate the impact of each term in the equation. $rand()$ returns a value between 0 and 1.

Note that the update equation also includes a chaotic term to introduce randomness into the search process. The chaotic term is given by:

$$r * (rand() - 0.5) \quad (19)$$

Where r is a chaotic parameter that controls the degree of randomness, and $rand()$ is a random number between 0 and 1.

The best solutions are chosen from the current population during the updating phase, while the worst solutions are replaced with fresh hawks. The top $M\%$ of hawks in the population, where M is a predetermined number, are referred to as the elite group of hawks. The placements of a few randomly chosen hawks in the population are used in a mathematical calculation to create new hawks to replace the population's worst hawks.³⁶

In general, the HHO algorithm is a recent optimization technique that has demonstrated promising outcomes in a number of optimization problems.²¹ The algorithm is simple to implement, and its mathematical equations are easy to understand and modify and the summary of the HHO algorithm's procedures is presented as the follows pseudo-code.

Algorithm 1. Harris Hawks optimization algorithm

Initialize population of hawks with random solutions
 Calculate fitness of each hawk using objective function
 Sort hawks in descending order of fitness
 Set the global best hawk as the first (i.e., highest fitness) hawk in the population
While stopping criterion is not met:
For each hawk in the population:
If the hawk is not the global best:
 Choose a random hawk in the population as the leader
 Update the position of the hawk using the following equation:
 hawk position = hawk position + rand() * (leader position - hawk position)
If the hawk goes outside the search space:
 Bring it back inside the search space
 Calculate the fitness of the hawk's new position
If the hawk's new position is better than its previous position:
 Set the hawk's new position as its current position
If the hawk's new position is better than the global best:
 Set the hawk's new position as the global best
end If
end If
end If
end For
end While
Sort hawks in descending order of fitness

Least square support vector machine (LSSVM)

The LSSVM algorithm is a variation on the support vector machine (SVM) technique that reduces classification error by using a least squares method. Due to its ability to handle huge datasets and nonlinear classification issues with flexibility and computational efficiency, the LSSVM algorithm has grown in prominence in recent years.³⁷

The conversion of the input data into a higher-dimensional feature set via a linear combination of kernel functions is one of the key components of the LSSVM method.³⁸ The kernel function can be any function that turns the input data into a new space, such as a polynomial kernel or a radial basis function (RBF) kernel.³⁹ As a result, nonlinear classification problems that cannot be solved by a linear classifier can be handled by the LSSVM algorithm. Cross-validation can be used to discover the best kernel function based on the problem's nature.⁴⁰

The LSSVM model can be defined as follows:

$$f(x) = \text{sign}(w\hat{T} \phi(x) + b) \quad (20)$$

Where:

$f(x)$ is the output of the model for input x .

sign is a sign function that maps the output to a binary classification (e.g., -1 or 1).

w is a weight vector that determines the orientation of the hyperplane.

$\phi(x)$ is a feature vector obtained by applying a kernel function to the input x .

b is a bias term that shifts the hyper plane.

The LSSVM algorithm works by minimizing the classification error while also minimizing the complexity of the model. This is achieved by solving the following optimization problem:

$$\text{minimize } (1/2) \|w\|^2 + C \left(\sum \xi \right) \quad (21)$$

$$\text{subject to } y_i (w\hat{T} \phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0 \quad (22)$$

Where:

$\|w\|^2$ is the L2 norm of the weight vector w , which penalizes large values of w .

C is a regularization parameter that controls the trade-off between the classification error and model complexity.

ξ_i is a slack variable that allows for some misclassification of the data points.

y_i is the label of the i -th data point ($+1$ or -1).

The optimization problem can be solved using the dual formulation, which leads to a set of linear equations that can be solved efficiently. The solution of the dual problem involves only the inner products of the feature vectors, which are computed using the kernel function. This makes LSSVM computationally efficient even for large datasets.

The LSSVM algorithm has been applied to a variety of applications, including image classification, text classification, and bioinformatics. It has also been extended to handle multiclass classification and regression problems. One of the advantages of LSSVM is its ability to handle high-dimensional data with a small number of samples. However, it also has some limitations, such as the need to choose an appropriate kernel function and regularization parameter.⁴¹

Results and discussion

Dataset

The proposed methodology consists of a single stage reciprocating air compressor of collected acoustic measurements. These data sets have been obtained from a compressor that has an air pressure range of $0\text{--}35 \text{ kg/cm}^3$, driven by an induction motor with a power rating of 5 HP , 5 Am , 415 V , 50 Hz and a speed of 1440 rpm and pressure switch type PR-15 , range between 100 and 213 PSI . The data sets from this study covered eight air

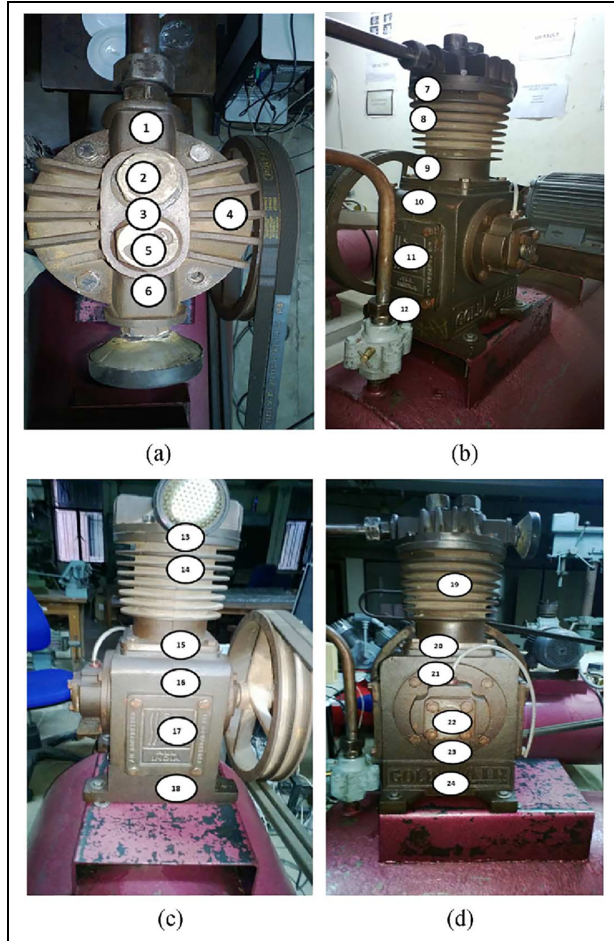


Figure 2. The positions from which acoustic signals were extracted in the air compressor were: (a) the top of the piston, (b) the side of the NRV, (c) the opposite side of the NRV, and (d) the opposite side of the flywheel.⁴²

compressor conditions which includes the healthy state and seven faulty states. These faults contain check valve (NRV) fault, leakage outlet valve (LOV) fault, leakage inlet valve (LIV) fault, driver belt fault, piston ring fault, bearing fault and flywheel fault.⁴² These data sets were acquired for period of 5 s at a sampling rate of 50 kHz using a microphone and a NIDAQ. The total number of data sets was 1800. Figure 2 shows the 24 positions of microphone placements at air compressor from which the acoustic data sets were acquired.

Nishchal K. Verma et al.⁴² recorded acoustic data from all 24 sensor positions, and after analyzing the recordings using EMD, SPA, and specific time domain features, they determined that the 8th position was the most sensitive for all compressor states. Subsequently, the researchers took 225 measurements from the chosen position (position 8) for each compressor health condition, with acoustic recordings captured at air compressor pressures ranging from 10 to 150 PSI.

Table 1. The statistical feature extraction.

Feature	Equation
Root mean square	$\sqrt{\frac{1}{n} \sum_{i=1}^n x_i ^2}$
Crest factor	$\frac{\ x\ _{\infty}}{\sqrt{\frac{1}{n} \sum_{i=1}^n x_i ^2}}$
Peak to peak	Max (x)–Min (x)
Skewness	$E\left[\left(\frac{x-\mu}{\sigma}\right)^3\right]$
Kurtosis	$E\left(\frac{x-\mu}{\sigma^4}\right)^4$
Entropy	$-\sum_i p_i \log_2(p_i)$
Mean	$\mu = \frac{1}{N} \sum_{i=1}^N A_i$
Std	$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N A_i - \mu ^2}$
Var	$\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3$
Root sum square	$X_{rss} = \sqrt{\sum_{n=1}^N x_n ^2}$
Max	Max x _i
Min	Min x _i
Mean square value	$\frac{1}{n} \sum_{i=1}^n x_i ^2$

Signal processing and feature extraction

In this subsection, we task challenge between many newest signal decomposition techniques. Therefore, the eight operations condition are decomposed the acoustic signals into set modes using Empirical Mode Decomposition (EMD), Recursive Empirical Mode Decomposition (REMD), Empirical Wavelet Transform (EWT), Variational Mode Decomposition (VMD) and Maximal Overlap Discrete Wavelet Packet Transform (MODWPT).

These methods allowed for the decomposition of the complex signals into simpler components, generating matrices that represented the extracted features from the signal. The number of matrices produced by each decomposition technique varied depending on the mode used.

As an example, firstly, the MODWPT are decomposed the acoustics signal into 16 modes (see Figure 3), each modes consists frequency and temporal information. Secondly, thirteen statistical features are then extracted from the time domain as fault signatures for each mode, resulting in a total of 208 features. The mathematical formulas for these features can be found in Table 1.

Feature optimizations and classifications:

Feature selection is the process of reducing the dimensionality of input parameters during the construction of a predictive model. It is in order to reduce the

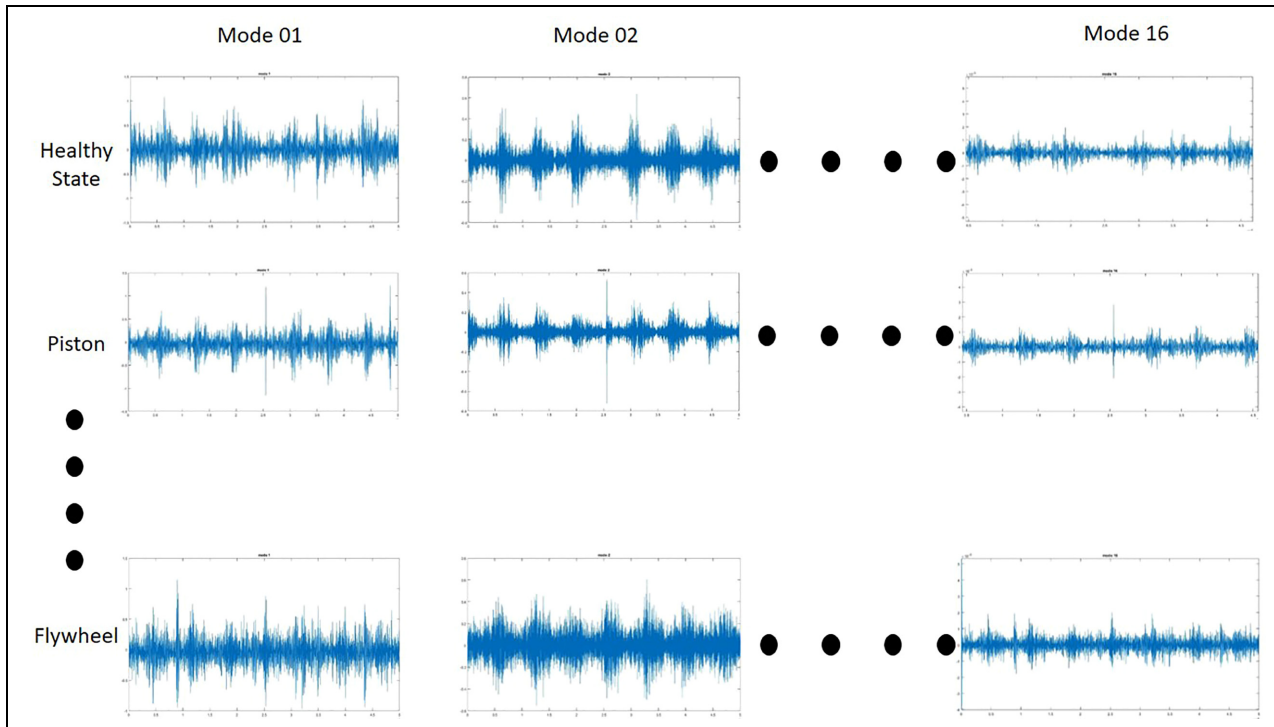


Figure 3. Acoustic signal decomposition using MODWPT for 16 modes.

number of input parameters that can confuse the results both to reduce the cost of modeling computation and, in some instances, to enhance the performance of the model, the optimization step is performed in order to select the efficient information and removing the overlapping parameters.⁴³

To demonstrate the effectiveness and robustness of this step, a simulation with and without optimization methods is performed and is presented in this subsection.

Unprocessed data without optimization methods. The statistical feature extraction from differ signal processing technique are feed into several machine learning classifiers, such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Naive Bayes (NB), Extra Tree (ET), Decision Trees (DT), Random Forests (RF), and Least Squares Support Vector Machines (LSSVM). These classifiers were trained using the extracted features to detect and identify patterns and faults state in the acoustic signal for each classes.

The classification results obtained for each decomposition technique in tandem with classifiers were recorded in a Table 2. As we can see, these comparisons illustrate the performance of different classifiers combined with decomposition methods that lead to the selection of the most effective accuracy combination for different faults occurring in the air compressor.

It can be seen from this table that, the highest level of overall accuracy are the combinations (EWT-LSSVM) and (MODWPT-LSSVM). However, to ensure the accuracy and reliability of the results, further analysis was required. To accomplish this, a detailed analysis for each combination was implemented, which allowed for a more accurate assessment of the predictive performance model. By measuring the mean, maximum, minimum, and standard deviation for each class.

To assist in the decision making process, the results for each class were tabulated for both (EWT-LSSVM) and (MODWPT-LSSVM). The Table 3 present a more comprehensive details of the performance of each class, which makes it clearer, which combination is optimal to achieve the desired result.

To validate results obtained from the previous table that LSSVM-MODWPT is the better combination for achieving high classification accuracy and stability compared to EWT-LSSVM. The detailed analysis of eight conditions states, as showing from confusion matrix that LSSVM-MODWPT consistently outperforms EWT-LSSVM for different states of air-compressor.

Moreover, as illustrated, form the Figure 4 the accuracy of each class is successfully predicted with higher percentage ranges between 96.6% and 100%. This indicates that the degree of fault differentiation in the same class mode versus the other class fault modes is the highest. On the other hand, we can see that the std

Table 2. Classification results using several classifiers with several signal decomposition techniques.

Used algorithm		KNN	DT	RF	ET	NB	SVM	LSSVM
EMD 104 features	MAX	70,7407	80,3703	87,4074	84,6296	36,6666	97,4074	99,0740
	MIN	64,8148	74,4444	83,1481	79,6296	24,4444	95,3703	96,6666
	MEAN	68,3518	76,9629	84,7962	81,9259	30,6296	96,4259	98,3703
	STD	1,7512	1,6864	1,3451	1,5815	3,8934	0,8191	0,7240
REMD 104 features	MAX	55,0793	71,9841	83,3333	76,2698	51,0317	95,7936	98,1746
	MIN	51,0317	63,2539	77,9365	72,5396	34,1269	93,3333	95,7936
	MEAN	53,3333	67,3968	79,7222	74,7936	44,3095	94,6587	97,3492
	STD	1,2515	2,5224	1,8025	1,3063	4,9539	0,8391	0,8839
VMD 130 features	MAX	59,8412	72,6190	86,5873	81,5873	64,0476	97,4603	98,6507
	MIN	56,1904	69,0476	80,6349	78,0952	59,1269	95,3174	97,1428
	MEAN	58,4285	70,7857	83,8730	80,6190	61,4206	96,2619	97,9126
	STD	1,3591	1,0166	1,8776	1,0065	1,71303	0,7542	0,4762
EWT 130 features	MAX	61,6666	70,5555	89,8148	83,5185	63,7037	99,2592	99,6296
	MIN	56,2962	64,0740	86,2962	79,2592	56,8518	97,2222	98,5185
	MEAN	58,9259	68,2407	87,9629	81,1296	60,7962	98,4444	99,0740
	STD	1,5635	2,2123	0,9876	1,0813	2,2050	0,6246	0,4000
MODWPT 208 features	MAX	88,1481	91,6666	98,1481	97,4074	81,1111	98,8888	100
	MIN	82,7777	86,4814	95,7407	95	77,4074	97,2222	99,0740
	MEAN	86,2777	89,0185	97,0185	96,4629	79,1111	97,8148	99,5555
	STD	1,5406	1,5979	0,7830	0,6843	1,2912	0,6806	0,3724

Table 3. Classification results for each class using EWT and MODWPT with LSSVM.

Used algorithm		Accuracy/class							
		Healthy	Bearing	Flywheel	LIV	LOV	NRV	Piston	Rider belt
EWT—LSSVM	MAX	100	100	100	100	100	98,5074	100	100
	MIN	94,8717	96,7213	97,4683	93,1506	96,7741	95,5882	94,7368	98,3870
	MEAN	97,3904	99,5332	99,1415	96,8456	99,3393	97,3222	98,2819	99,6998
	STD	1,3696	1,0801	0,9971	2,0789	1,1444	1,1563	1,8685	0,6350
MODWPT—LSSVM	MAX	100	100	100	100	100	100	100	100
	MIN	100	98,6301	97,1428	96	96,7213	100	98,3606	96,5517
	MEAN	100	99,8630	98,5432	99,3101	99,4108	100	99,8360	99,5102
	STD	0	0,4331	1,1359	1,4773	1,0910	0	0,5184	1,1349

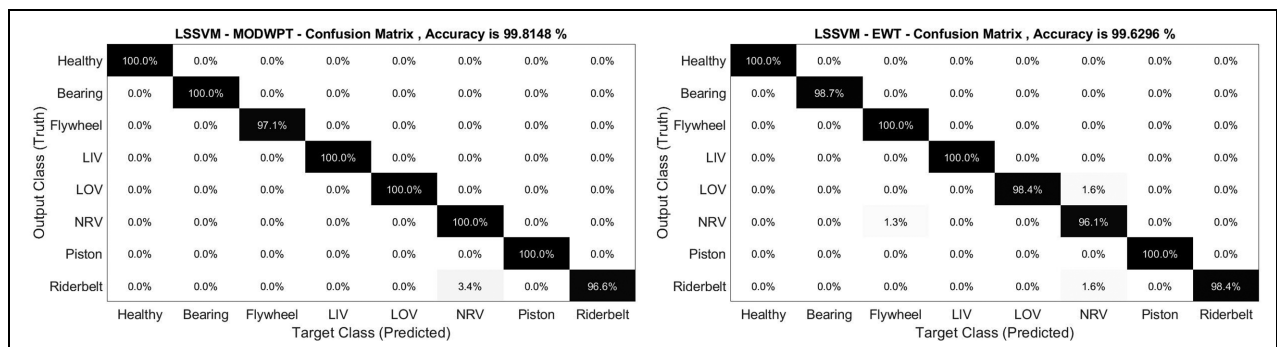


Figure 4. Classification results using MODWPT-LSSVM versus EWT-LSSVM.

Table 4. Classification results using several optimization algorithms with MODWPT signal decomposition method.

OPT algorithm	Max ACC	Min ACC	Mean ACC	Stability Std	No. of selected features
GNDO	99.6296	99.4444	99.5	0.08945	96/208
MRFO	99.8148	99.4444	99.5925	0.1171	47/208
HHO	100	99.6296	99.7593	0.125	23/208
GA	100	99.6296	99.8518	0.146	53/208
SMA	100	99.4444	99.6852	0.1962	2/208
ABC	100	99.2593	99.6481	0.2038	84/208
PSO	99.8148	99.2593	99.6296	0.2469	86/208
CS	100	99.0741	99.5	0.2767	79/208
ASO	100	99.2593	99.6296	0.2895	89/208
EO	100	99.074	99.574	0.303	21/208
MPA	99.4444	98.5185	98.9074	0.3202	71/208

stability in the damage cases (flywheel) and (Rider belt) is 1.1359, 1.1349 respectively.

We observed that, while achieved the best accuracy classification result, the stability classification results for each class can be significantly different.

For that reason, to improve rating of accuracy and stability simultaneously for each class, the feature selection techniques can be used to identify the most relevant features.

Processed data with optimization methods. To overcome the previous results, the optimization step is suggested to identify the appropriate components and remove the irrelevant alternatives. Therefore, to improve the robustness and efficiency of the proposed model, a comparison was made between several optimization algorithms (GNDO, MRFO, GA, SMA, ABC, and HHO) to identify the subset of features that provide the highest classification accuracy and also the best stability.

With the implementation of feature selection techniques with previous optimization algorithms, the LSSVM-MODWPT model can be homogeneity and improved. This Refinement lead to even higher classification accuracy and more stability, which can make the LSSVM-MODWPT combination a more powerful tool for classification tasks in various fields.

As we can see from this table that, the best results obtained after optimization step and can be considered as valuable information for future experiments to improve the performance of the classifier. The results presented in the Table 4 indicate a clear improvement in classification accuracy and stability through the successful application of optimization algorithms on the LSSVM classifier.

An important observation in this table is the variation in the standard deviation values, where lower values signify more stability in the classification. These results highlight the crucial role of optimization techniques in order to facilitate reliable and accurate classification results with the LSSVM classifier.

In order to identify the optimal results obtained among different optimization algorithms, a detailed study of four key measurements (mean, max, min, and standard deviation) was implemented for each class. By analyzing these metrics, we can obtain a thorough understanding of the performance of the LSSVM classifier and identify the most efficient optimization algorithm for each class. The inclusion of these four metrics facilitates a more thorough evaluation of the classifier's performance, allowing us to make informed decisions regarding the optimization process.

It is clear from the results presented in the Table 5 that the HHO optimization algorithm significantly outperforms the other optimization algorithms in terms of the four key measurements - mean, max, min, and standard deviation—for each class. This leads to the conclusion that the HHO algorithm is the most effective approach to optimizing the performance of the LSSVM classifier, as it consistently provides the best results on all four measurements for each class.

To carry out an evaluation of the fault recognition capability of the proposed method for eight different fault modes, a set of fault predictions on eight fault modes of the test set is performed. The confusion matrix of the prediction results is shown in Figure 5.

It can be observed from figure LSSVM-HHO that, the huge majority of samples for eight fault modes (proposed approach) are successfully predicted with higher percentage ranges between 98.5 % and 100%. This means that the degree of fault differentiation in the identical class mode from other class fault modes is the highest.

We concluded the optimization algorithm that gives the best overall performance for the LSSVM classifier, allowing us to improve its performance and reliability.

Conclusion

The comprehensive exploration of feature selection has unequivocally demonstrated its pivotal role in elevating the accuracy and stability of fault classification. This

Table 5. Classification results for each class using several optimization algorithms.

Used algorithm		Modwpt— LSSVM 208	Modwpt ABC LSSVM 84 S.F	Modwpt HHO LSSVM 23 S.F	Modwpt SMA LSSVM 2 S.F	Modwpt GA 53 S.F	Modwpt MRFO 47 S.F	Modwpt GNDO 96 S.F
<i>Accuracy/stability</i>								
Healthy	MAX	100	100	100	100	100	100	100
	MIN	100	98,5074	100	100	100	100	98,5074
	MEAN	100	98,8059	100	100	100	100	98,8059
	STD	0	0,62930	0	0	0	0	0,62930
Bearing	MAX	100	100	100	100	100	100	100
	MIN	98,6301	100	100	100	100	98,5294	98,5074
	MEAN	99,8630	100	100	100	100	99,8529	99,8507
	STD	0,4331	0	0	0	0	0,4650	0,4719
Flywheel	MAX	100	100	100	100	100	100	100
	MIN	97,1428	100	98,5075	100	98,5294	100	98,5074
	MEAN	98,5432	100	98,6567	100	99,7058	100	99,8507
	STD	1,1359	0	0,4720	0	0,6200	0	0,4719
LIV	MAX	100	100	100	100	100	100	100
	MIN	96	100	100	100	100	98,5074	98,5294
	MEAN	99,3101	100	100	100	100	99,8507	99,4117
	STD	1,4773	0	0	0	0	0,4719	0,7594
LOV	MAX	100	100	100	100	100	100	100
	MIN	96,7213	98,5075	98,5294	100	98,5074	98,5294	98,5074
	MEAN	99,4108	99,2537	99,4118	100	99,4029	99,5588	98,9552
	STD	1,0910	0,7866	0,7594	0	0,7707	0,7103	0,7209
NRV	MAX	100	100	100	100	100	100	100
	MIN	100	98,5294	100	95,5223	98,5294	97,0149	98,5294
	MEAN	100	99,8529	100	97,4626	99,7058	98,3582	99,5588
	STD	0	0,4650	0	1,5811	0,6200	1,1012	0,7103
Piston	MAX	100	100	100	100	100	100	100
	MIN	98,3606	97,0588	100	100	100	98,5294	98,5294
	MEAN	99,8360	99,2647	100	100	100	99,1176	99,5588
	STD	0,5184	1,0399	0	0	0	0,7594	0,7103
Rider belt	MAX	100	100	100	100	100	100	100
	MIN	96,5517	100	100	100	100	100	100
	MEAN	99,5102	100	100	100	100	100	100
	STD	1,1349	0	0	0	0	0	0

enhancement is particularly evident when assessing classification outcomes across distinct fault classes, further emphasizing the significance of this strategic inclusion. Notably, our investigation underscores the prominence of the MODWPT-LSSVM fusion as an exceptionally effective approach for fault classification, surpassing alternative methods. Equally compelling is the discovery of the HHO-LSSVM-MODWPT optimization algorithm as the optimal strategy for bolstering both classification stability and accuracy, reaffirming its profound impact on fault diagnosis. The implications of these findings reverberate across industries reliant on air compressors, promising tangible benefits for maintenance practices and operational efficiency. By offering a dependable means of detecting and classifying faults, our proposed approach holds the potential to revolutionize maintenance regimes, mitigate operational risks, and optimize resource utilization. Moreover, this study

transcends immediate applications, igniting the spark for future advancements in signal processing. The horizon of signal processing now beckons toward innovative optimization algorithms and signal processing techniques, poised to enhance not only air compressor systems but also broader domains of industrial automation and predictive maintenance. The groundwork laid by this research nurtures the fertile soil for burgeoning developments, propelling decision-making efficiency to new heights in diverse applications.

In summation, our investigation unfurls a new chapter in the realm of signal processing methodologies for air compressor fault detection, particularly in the context of acoustic signal diagnosis. With its multidimensional implications, this study encapsulates a transformative approach that bridges academia and industry, thus illuminating a path towards more dependable, efficient, and proactive fault diagnosis strategies.

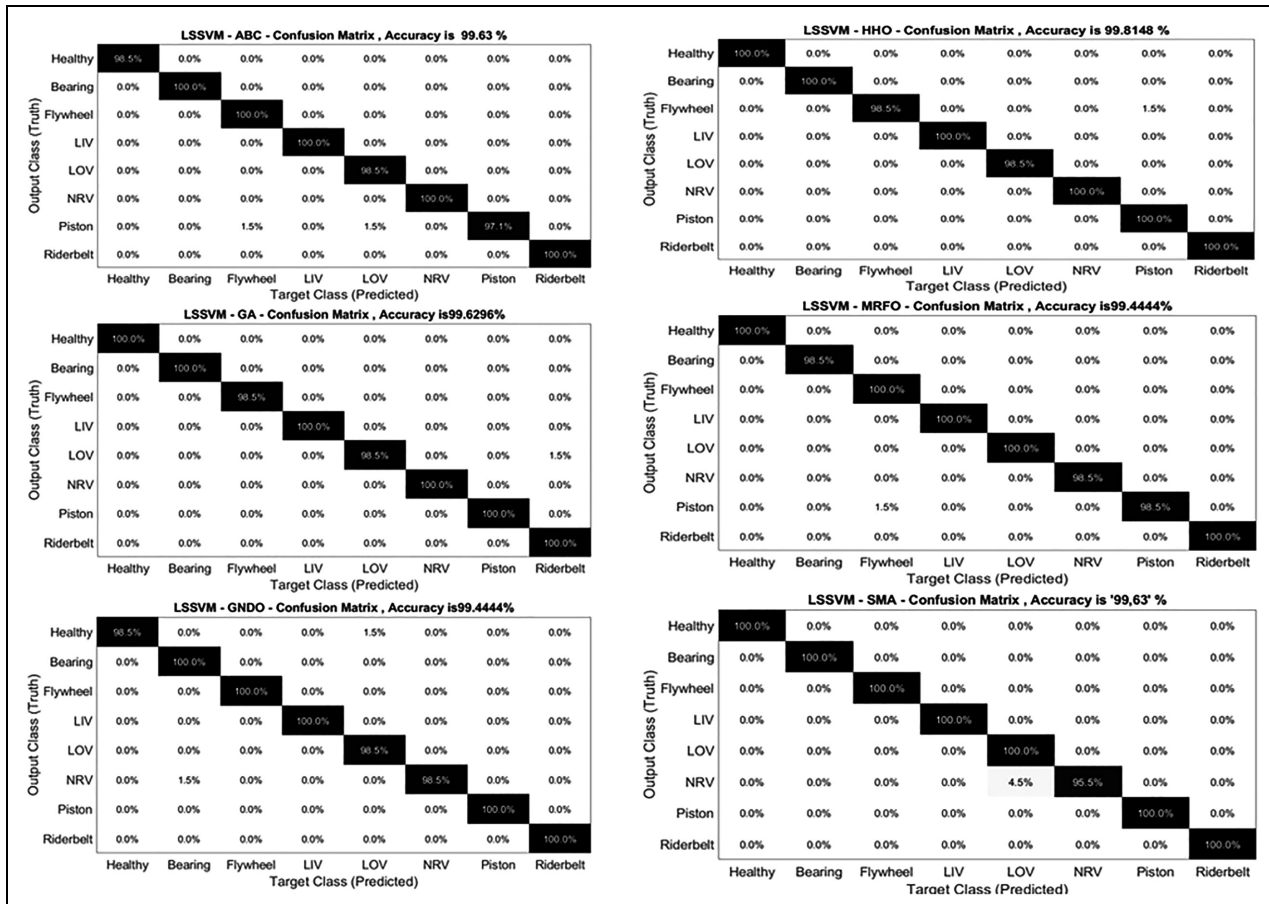


Figure 5. Confusion matrix of classification results.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Chemseddine Rahmoune  <https://orcid.org/0000-0003-0388-7455>

Mohammed Amine Sahraoui  <https://orcid.org/0009-0005-9435-1374>

Fawzi Gougam  <https://orcid.org/0000-0001-6964-1867>

References

1. Afia A, Gougam F, Rahmoune C, et al. Air compressor fault classification using MODWPT, time domain features, WSA and machine learning classifiers based on acoustic analysis. *Prog Ind Ecol Int J* 2023; 12(1/2): 192.
2. Yang W, Su Y and Chen Y. Air compressor fault diagnosis based on lifting wavelet transform and probabilistic neural network. In: *IOP conference series: materials science and engineering 2019*, Kazimierz Dolny, Poland, 21–23 November 2019, p. 012053. Bristol, UK: IOP Publishing.
3. Malla C and Panigrahi I. Review of condition monitoring of rolling element bearing using vibration analysis and other techniques. *J Vib Eng Technol* 2019; 7: 407–414.
4. Shao H, Xia M, Han G, et al. Intelligent fault diagnosis of rotor-bearing system under varying working conditions with modified transfer convolutional neural network and thermal images. *IEEE Trans Indus Inform* 2020; 17: 3488–3496.
5. Kowarski KA and Moors-Murphy H. A review of big data analysis methods for baleen whale passive acoustic monitoring. *Mar Mamm Sci* 2021; 37: 652–673.
6. Russell M and Wang P. Physics-informed deep learning for signal compression and reconstruction of big data in industrial condition monitoring. *Mech Syst Signal Proc* 2022; 168: 108709.
7. Begum M, Ferdush J and Uddin MS. A hybrid robust watermarking system based on discrete cosine transform, discrete wavelet transform, and singular value decomposition. *J King Saud Univ Comput Inf Sci* 2022; 34: 5856–5867.
8. Yao Z, Wang Z, Liu X, et al. An improved low-frequency noise reduction method in shock wave pressure

- measurement based on mode classification and recursion extraction. *ISA Trans* 2021; 109: 315–326.
9. Mousavi AA, Zhang C, Masri SF, et al. Structural damage detection method based on the complete ensemble empirical mode decomposition with adaptive noise: a model steel truss bridge case study. *Struct Health Monit* 2022; 21: 887–912.
 10. Ni Q, Ji J, Feng K, et al. A fault information-guided variational mode decomposition (FIVMD) method for rolling element bearings diagnosis. *Mech Syst Signal Proc* 2022; 164: 108216.
 11. Imane M, Rahmoune C, Zair M, et al. Bearing fault detection under time-varying speed based on empirical wavelet transform, cultural clan-based optimization algorithm, and random forest classifier. *J Vib Control* 2023; 29: 286–297.
 12. El-Hendawi M and Wang Z. An ensemble method of full wavelet packet transform and neural network for short term electrical load forecasting. *Electr Power Syst Res* 2020; 182: 106265.
 13. Ikhlef B, Rahmoune C, Toufik B, et al. Gearboxes fault detection under operation varying condition based on MODWPT, ant colony optimization algorithm and random forest classifier. *Adv Mech Eng* 2021; 13: 16878140211043004.
 14. Afia A, Rahmoune C, Benazzouz D, et al. New intelligent gear fault diagnosis method based on autogram and radial basis function neural network. *Adv Mech Eng* 2020; 12: 1687814020916593.
 15. Zebari R, Abdulazeez A, Zeebaree D, et al. A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *J Appl Sci Technol Trends* 2020; 1: 56–70.
 16. Karaboga D, Gorkemli B, Ozturk C, et al. A comprehensive survey: artificial bee colony (ABC) algorithm and applications. *Artif Intell Rev* 2014; 42: 21–57.
 17. Katoch S, Chauhan SS and Kumar V. A review on genetic algorithm: past, present, and future. *Multimed Tools Appl* 2021; 80: 8091–8126.
 18. Chen H, Li C, Mafarja M, et al. Slime mould algorithm: a comprehensive review of recent variants and applications. *Int J Syst Sci* 2023; 54: 204–235.
 19. Zhang Y, Jin Z and Mirjalili S. Generalized normal distribution optimization and its applications in parameter extraction of photovoltaic models. *Energy Convers Manag* 2020; 224: 113301.
 20. Zhao W, Zhang Z and Wang L. Manta ray foraging optimization: an effective bio-inspired optimizer for engineering applications. *Eng Appl Artif Intell* 2020; 87: 103300.
 21. Heidari AA, Mirjalili S, Faris H, et al. Harris hawks optimization: algorithm and applications. *Future Gener Comput Syst* 2019; 97: 849–872.
 22. Alzubi OA, Alzubi JA, Alweshah M, et al. An optimal pruning algorithm of classifier ensembles: dynamic programming approach. *Neural Comput Appl* 2020; 32: 16091–16107.
 23. Mahesh B. Machine learning algorithms-a review. *Inform Syst J* 2020; 9: 381–386.
 24. Shinde PP and Shah S. A review of machine learning and deep learning applications. In: *2018 Fourth international conference on computing communication control and automation (ICCUBEA)*, Pune, India, 16-18 August 2018 2018, pp. 1–6. Piscataway, NJ: IEEE.
 25. Preyanka Lakshme R and Kumar SG. A review based on machine learning for feature selection and feature extraction. In: *Advancements in smart computing and information security: first international conference, ASCIS 2022, Rajkot, India, November 24–26, 2022, revised selected papers, Part I*. Cham: Springer, 2023, pp. 144–157.
 26. Saâdaoui F, Jabeur SB and Goodell JW. Geopolitical risk and the Saudi stock market: Evidence from a new wavelet packet multiresolution cross-causality. *Finance Res Lett* 2023; 53: 103654.
 27. Bettahar T, Chemseddine R and Benazzouz D. Faults' diagnosis of time-varying rotational speed machinery based on vibration and acoustic signals features extraction, and machine learning methods. *J Vib Eng Technol* 2022; 11: 2333–2347.
 28. Afia A, Rahmoune C, Benazzouz D, et al. New gear fault diagnosis method based on modwpt and neural network for feature extraction and classification. *J Test Eval* 2019; 49: 1064–1085.
 29. Khaire UM and Dhanalakshmi R. Stability of feature selection algorithm: a review. *J King Saud Univ Comput Inf Sci* 2022; 34: 1060–1073.
 30. Remeseiro B and Bolon-Canedo V. A review of feature selection methods in medical applications. *Comput Biol Med* 2019; 112: 103375.
 31. Banerjee D, Chatterjee B, Bhowal P, et al. A new wrapper feature selection method for language-invariant offline signature verification. *Expert Syst Appl* 2021; 186: 115756.
 32. Liu H, Zhou M and Liu Q. An embedded feature selection method for imbalanced data classification. *IEEE/CAA J Autom Sin* 2019; 6: 703–715.
 33. Aggarwal S and Chugh N. Review of machine learning techniques for EEG based brain computer interface. *Arch Computat Methods Eng* 2022; 29: 3001–3020.
 34. Yıldız AR, Yıldız BS, Sait SM, et al. A new hybrid Harris hawks-Nelder-Mead optimization algorithm for solving design and manufacturing problems. *Mater Test* 2019; 61: 735–743.
 35. Alzubi OA, Alzubi JA, Al-Zoubi AM, et al. An efficient malware detection approach with feature weighting based on Harris Hawks optimization. *Cluster Comput* 2022; 25: 2369–2387.
 36. Golilarz NA, Addeh A, Gao H, et al. A new automatic method for control chart patterns recognition based on ConvNet and Harris hawks meta heuristic optimization algorithm. *IEEE Access* 2019; 7: 149398–149405.
 37. Samui P and Kothari DP. Utilization of a least square support vector machine (LSSVM) for slope stability analysis. *Sci Iran* 2011; 18: 53–58.
 38. Duan M, Liu Z, Yan D, et al. Application of LSSVM algorithm for estimating higher heating value of biomass based on ultimate analysis. *Energy Sources A Recovery Util Environ* 2018; 40: 709–715.
 39. Sharma G, Panwar A, Nasiruddin I, et al. Non-linear LS-SVM with RBF-kernel-based approach for AGC of multi-area energy systems. *IET Gener Transm Distrib* 2018; 12: 3510–3517.

40. Gao X, Wei H, Li T, et al. A rolling bearing fault diagnosis method based on LSSVM. *Adv Mech Eng* 2020; 12: 1687814019899561.
41. Wang H and Hu D. Comparison of SVM and LS-SVM for regression. In: *2005 International conference on neural networks and brain*, Beijing, China, 13–15 October 2005, pp. 279–283. Piscataway, NJ: IEEE.
42. Verma NK, Sevakula RK, Dixit S, et al. Intelligent condition based monitoring using acoustic signals for air compressors. *IEEE Trans Reliab* 2015; 65: 291–309.
43. Sahraoui MA, Rahmoune C, Meddour I, et al. New criteria for wrapper feature selection to enhance bearing fault classification. *Adv Mech Eng* 2023; 15: 16878132231183862.