# Research on Bearing Fault Diagnosis Method based on Attention Entropy-COA-SVM

Bing Liu[a*], Yanshuai Yang[b], Zhentao Mi[c], Ruifeng Zhang[d], Yuxi Song[e], Guoqing Zhu[f]

China Academy of Machinery Science & Technology Oingdao Branch Co.,Ltd., Qingdao, 266300, China

[a*]2573725015@qq.com, [b]yang_yanshuai@163.com
[c]mizhentao2016@163.com, [d]1456351736@qq.com
[e]2257172526@qq.com, [f]1026600545@qq.com

**Abstract.** Aiming at the problem of bearing fault diagnosis, a fault diagnosis method based on Attention Entropy (AE) and Support Vector Machines (SVM) optimized by Coati Optimization Algorithm (COA) is proposed. Firstly, the attention entropy feature of bearing vibration signal is extracted, which can extract the key information in the signal more effectively, so as to enhance the identifiability of fault features and improve the accuracy of fault identification. Secondly, the attention entropy feature is input into the COA-SVM model for bearing fault classification. Through the verification on the fault data set of Case Western Reserve University, the results show that the method can effectively identify different fault types. Through comparative experiments, the method based on attention entropy has obvious advantages in fault detection and classification performance, and the accuracy of this method can reach 94.64 %.

**Keywords:** Attention entropy, Support vector machine, Long-nose raccoon optimization algorithm, Fault diagnosis.

## 1 Introduction

The concept of entropy is increasingly used to reveal the characteristics of nonlinear signals, especially in the analysis of rolling bearing faults, which has become an important tool to measure the complexity of vibration signals[1]. In the process of bearing fault diagnosis, selecting the appropriate entropy algorithm has a significant impact on the accuracy of the diagnosis results.

In order to improve the application effect of entropy theory in the field of bearing fault diagnosis, many researchers have deepened and optimized it. Li Weimin[2] proposed a fault diagnosis method of asynchronous motor based on approximate entropy and support vector machine. The vibration signal samples were obtained by fault simulation test, and the approximate entropy fault feature vector was calculated. The classification model constructed shows higher diagnostic accuracy in fault diagnosis


A. Mallik et al. (eds.), *Proceedings of the 2025 6th International Conference on Management Science and Engineering Management (ICMSEM 2025)*, Atlantis Highlights in Economics, Business and Management 4,
https://doi.org/10.2991/978-94-6463-845-5_104

accuracy. However, Chang Lin et al.[3] based on sample entropy and SVM rolling bearing fault diagnosis method, and compared with EEMD and CEEMDAN algorithm. The results show that the sample entropy features extracted from the signals decomposed by the ICEEMDAN algorithm can better cluster and distinguish different rolling bearing operating states. The SVM classification method has higher classification accuracy and provides an effective way for multi-fault diagnosis of rolling bearings. Bian Dongxue[4] proposed to use LMD to decompose the bearing vibration signal into multiple product components, and calculate the approximate entropy of each component. Then, the first three components are selected according to the correlation coefficient, and their approximate entropy values are calculated as feature vectors. Finally, the feature vectors are input into the improved PSO-ELM model to identify and classify different fault types of rolling bearings.

On the basis of previous studies, the author of this paper proposes a new bearing fault diagnosis method, which combines the advantages of attention (AE) entropy and support vector machine (SVM). Firstly, the vibration signals of the bearing in different states are collected, and the feature vectors are extracted by the attention entropy algorithm. Then, the feature vector is input into the improved PO-SVM classification model to realize the identification and classification of bearing fault state. This method can effectively extract bearing fault features and has high classification accuracy.

## 2 Attention Entropy

Yang et al[5] proposed a new time series complexity measurement tool-attention entropy, which can focus on the changes of key points in time series. Compared with the traditional entropy, the attention entropy has stronger robustness, is not affected by the timing length, and does not need to set hyperparameters. The implementation process is shown in Fig.1.

(1)It is assumed that each data point in the vibration signal represents an independent subsystem, and the change of its state can be understood as the adaptation process of the subsystem to the environment. In this case, the peak points of the signal can effectively reflect the upper and lower boundary changes of the local state of the subsystem, so these peak points can be regarded as the key points of the system.

(2)In order to identify the key points in the time series, the difference between the minimum value $\{\min-\min\}$, the difference between the minimum value and the maximum value $\{\min-\max\}$, the difference between the maximum value and the minimum value $\{\max-\min\}$, the difference between the maximum value $\{\max-\max\}$ can be used. By calculating these differences, the interval between the adjacent key points can be determined.

(3)In order to quantify the complexity of the interval between adjacent core points, Shannon entropy can be used for calculation. The formula is as follows:

$$H(x) = -\sum_{x=1}^{b} p(x)\log_2 p(x) \tag{1}$$

In the formula, $H(x)$ represents the Shannon entropy of the interval between adjacent core points; $x$ represents the set of adjacent core point intervals; $p(x)$ represents the probability that the interval is $x$.

(4)The mean value of Shannon entropy calculated by different strategies in (2) is defined as the attention entropy of the vibration signal.

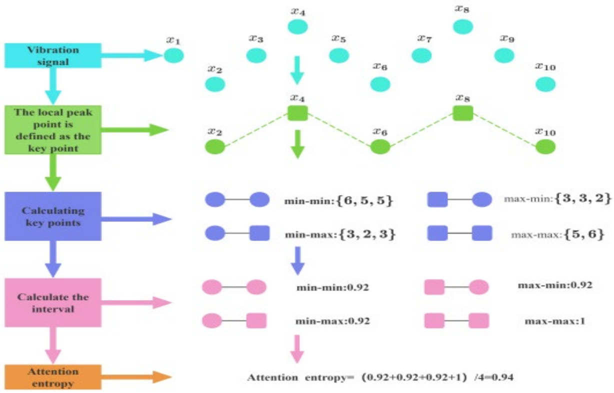



**Fig. 1.** Attention Entropy schematic diagram

# 3 COA-SVM Fault Diagnosis Model

## 3.1 COA Algorithm

(1) Population initialization

Randomly arrange the initial position of the raccoon group in the search space:

$$x_{i,j} = l_j + r \cdot (u_j - l_j) \tag{2}$$

In the formula, $x_{i,j}$ is the position of the $i$-th long-nosed raccoon in the $j$-dimension; $l_j$ is the lower bound of the $j$-dimensional variable; $u_j$ is the upper bound of the $j$-dimensional variable; $R$ is a random number between $[0,1]$.

(2) Exploration stage

In the first stage of the COA algorithm, the behavior strategy of the raccoon group besieging the lizard is simulated to update the population position in the search space. In this process, some raccoons will go up the tree to block and intimidate lizards, while other raccoons wait under the tree. Once the lizard falls from the tree in fright, the raccoons under the tree attack and chase their prey[6]. This strategy of imitating the behavior of raccoons in nature enables the algorithm to make effective position changes in the search space, thus reflecting the global search and exploration ability of COA in dealing with optimization problems.

The position expression of half of the raccoons is:

$$x_{i,j}^{P1} = x_{i,j} + R \cdot \left(I_j^G - I \cdot x_{i,j}\right) \tag{3}$$

In the formula, $x_{i,j}^{P1}$ is the new position of the $i$-th long-nosed raccoon in the $j$-th dimension; b is a set $\{1,2\}$ random integer; $I_j^G$ is the original location of the lizard.

The rest of the long-nosed raccoons begin to move in space:

$$x_{i,j}^{P1} = \begin{cases} x_{i,j} + R \cdot \left(I_j^{G,new} - I \cdot x_{i,j}\right), F_I^G < F_i \\ x_{i,j} + R \cdot \left(x_{i,j} - I_j^{G,new}\right), else \end{cases} \tag{4}$$

In the formula, $I_j^{G,\ new}$ is the new position of lizard; $F_i$ is the objective function before updating the position; $F_I^G$ is the objective function of the lizard after changing its position.

(3) Development phase

In the second stage of the COA algorithm, the algorithm draws on the escape mechanism of the raccoon when it encounters a predator to adjust its position in the search space. When a raccoon is attacked by predators, it will quickly escape from its current location and move towards a more secure area[7]. This strategy makes the raccoon tend to a better position in the search process, thus demonstrating the development potential of COA in local search.

$$x_{i,j}^{P2} = x_{i,j} + (1-2R) \cdot \left(l_j^{local} + R \cdot \left(u_j^{local} - l_j^{local}\right)\right) \tag{5}$$

In the formula, $x_{i,j}^{P2}$ is the new position of the $i$-th long-nosed raccoon in the $j$-th dimension; $l_j^{local}$ is the lower boundary of the current position; $u_j^{local}$ is the upper boundary of the current position.

After two stages of updating the position of all long-nosed raccoons in the search space, the algorithm completes the current iteration. Subsequently, the COA algorithm will continue to perform the next iteration until the final number of iterations is reached. When all iterations are completed, the algorithm outputs the optimal solution as the result.

### 3.2 COA-SVM Model Diagnosis Process

SVM is a pattern recognition technology that relies on kernel learning. The core idea is to upgrade the original low-dimensional data to high-dimensional space through nonlinear mapping, and to find a classification boundary plane that can maximize the interval based on the criterion of minimizing structural risk. SVM has efficient computing power in classification tasks. However, its key parameters depend on manual adjustment, which may lead to overfitting or underfitting of the algorithm, thus affecting its overall performance. Therefore, the author adopts the method of COA optimization SVM to realize the effective recognition of features.

The COA-SVM fault diagnosis process is as follows:

(1) The original data is standardized, and the data set is randomly assigned to divide the training set and the test set for subsequent model training and verification.

(2) Set the parameters of the COA algorithm, including population size, maximum number of iterations, etc. The penalty parameter C and kernel function parameters of SVM are taken as the optimization objectives, and the classification error rate is taken as the fitness function. The purpose is to find the parameter combination that can minimize the error rate in the iterative process.

(3) Initialize the long-nosed raccoon population and calculate the fitness function value of each individual. In the iterative process, the parameters in the population are updated according to the foraging behavior of the long-nosed raccoon.

(4) In each iteration, the fitness of the long-nosed raccoon population is evaluated. If the termination condition is satisfied ( such as the maximum number of iterations or the fitness value is lower than the preset threshold ), the current optimal parameter is output ; otherwise, continue to update the parameters iteratively.

(5) The optimal parameter combination obtained by the COA algorithm iteration is applied to the SVM model, and the model is trained using the training set.

(6) Use the test set to test the trained SVM model, evaluate the classification performance of the model, and output the fault identification results of the test samples.

## 4 Example Analysis

In this experiment, the bearing data set of Case Western Reserve University in the United States was used for experimental verification and analysis. The test bench and its schematic diagram are shown in Fig.2.The detailed information of the bearing fault samples is shown in Table 1.

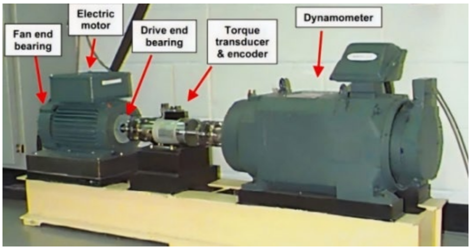



**Fig. 2.** Test bench

**Table 1.** Details of bearing fault samples

| Label | Failure mode | Fault Size/mm | Number of samples |
|---|---|---|---|
| 1 | Normal | 0 | 90 |
| | Inner Ring Fault | 0.178 | |
| 2 | Inner Ring Fault | 0.356 | 90 |
| | Inner Ring Fault | 0.534 | |
| | Rolling Element Fault | 0.178 | |
| 3 | Rolling Element Fault | 0.356 | 90 |
| | Rolling Element Fault | 0.534 | |

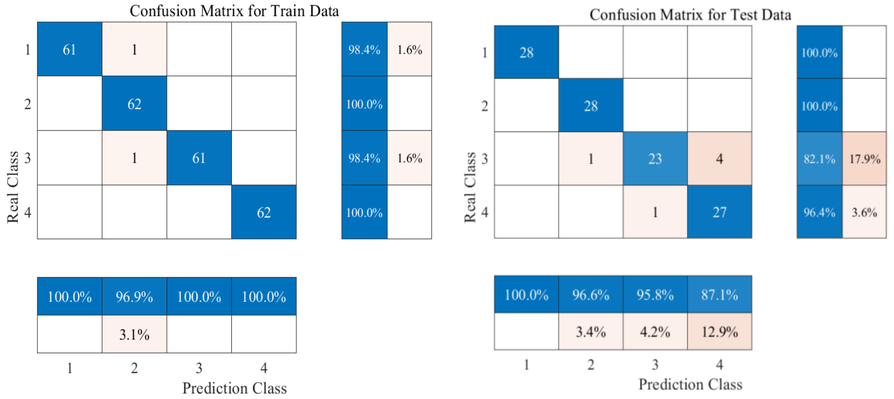
| | | | |
|---|---|---|---|
| | Outer Race Defect | 0.178 | |
| 4 | Outer Race Defect | 0.356 | 90 |
| | Outer Race Defect | 0.534 | |

According to the steps of improving COA to optimize SVM, the population size of COA algorithm is set to 30. In the network parameter configuration of COA and SVM, the population size is also set to 30, and the maximum number of iterations is set to 100. For each working condition, 248 sets of data were randomly selected to form a training set, and the remaining 112 sets of data were used for the test set.

By comparing the confusion matrix in Fig.3 with the diagnosis output results of the model in Table 2, according to the diagnosis results, the AE-COA-SVM model is 2.68 % higher than the test set of the AE-SVM model, and the accuracy of the training set is improved by 0.8 %. It can effectively identify different fault types of rolling bearings and has certain application potential.Compared with other models such as AEAE-PSO-SVM and AE-GA-SVM, the accuracy of AE-COA-SVM model is the best, but its calculation time is slightly lower than other models.

**Table 2.** Output results of model diagnosis

| Model | Accuracy of test set | Accuracy of training set | Time |
|---|---|---|---|
| AE-SVM | 91.96% | 98.39% | 45.5s |
| AE-COA-SVM | 94.64% | 99.19% | 38.7s |
| AE-PSO-SVM | 96.2% | 97.2% | 35.2s |
| AE-GA-SVM | 94.4% | 96.4% | 42.6s |




(a) Confusion matrix of AE-COA-SVM training set (b) Confusion matrix of AE-COA-SVM test set

**Fig. 3.** AE-COA-SVM confusion matrix

# 5 Conclusion

Aiming at the phenomenon that the bearing working environment is changeable and prone to failure, this paper proposes a fault identification method based on support vector machine ( SVM ) optimized by attention entropy and long nose raccoon optimization algorithm ( COA ). This method combines modern computer technology with advanced algorithms and has been successfully applied to the actual fault type identification process.

The COA algorithm is used to optimize the SVM model, which can find the best combination of penalty parameter c and kernel function parameter g more efficiently and accurately, thus significantly improving the classification accuracy of SVM. The optimized SVM is applied to bearing fault diagnosis. Compared with the unoptimized SVM, the experimental results confirm that the SVM method optimized by attention entropy combined with COA algorithm has better performance in fault diagnosis accuracy.

# References


1. Bian Jun, Jing Laixing, Liu Yanqiu. Application of entropy theory in bearing fault diagnosis[J]. Journal of Dalian Polytechnic University, 2024, 43(04): 300-306.
2. Li Weimin, Ma Jizhao, Lei Xiaozhu. Research on Fault Diagnosis for Asynchronous Motor Based onApproximate Entropy and Support Vector Machine[J]. Machine Tool & Hydraulics, 2021, 49(05): 173-176+155.
3. Dan Changlin, Li Sanyan, Zhang Bin. Research on fault diagnosis method for rolling element bearings based onsample entropy and SVM[J]. China Measurement & Test, 2020, 46(11): 37-42.
4. Bian Dongxue, Zhang Jinping. Bearing Fault Diagnosis Based on LMD Approximate Entropyand Improved PSO-ELM[J]. Machine Tool & Hydraulics, 2023, 51(14): 227-232.
5. YANG J W, CHOUDH ARY G I, RAHARDJA S, et al. Classification of in terbeat interval time-series using Attention Entropy[J]. IEEE Transactions on Affective Computing, 2020.
6. Liu Jianfeng, Li Zhiyuan, Zhou Yaru. Transformer windings based on leakage field and ICOA-ResNet early fault diagnosis[J]. Power System Protection and Control, 2024, 52(09): 99-110.
7. Li Bingwei, Ye Shuxia, Qi Liang, et al. Data Correction Method for Low-cost Gas Sensors Based onCOA-GRU Algorithm[J]. Instrument Technique and Sensor, 2024, (03):120-126.