ACCSCIENCE PUBLISHING

Journal of Seismic Exploration

ARTICLE

Optimization of the SOM neural network model using CEEMDAN distribution entropy and ALO for seismic and blasting identification

Ailing Wang[1], Cong Pang[2,3], Guoqing Chen[1], Chawei Li[2,3], and Tianwen Zhao[4]*

[1]Mathematical Modeling Research Center, Chengdu Jincheng College, Chengdu, Sichuan, China

[2]Institute of Seismology, China Earthquake Administration, Wuhan, Hubei, China

[3]National Observation and Research Station for Wuhan Gravitation and Solid Earth Tides, Hubei Earthquake Administration, Wuhan, Hubei, China

[4]Department of Trade and Logistics, Daegu Catholic University, Gyeongsan, Daegu, Republic of Korea

Abstract

As seismic signals and artificial blasting signals exhibit high similarity in time–frequency domain features, resulting in insufficient recognition accuracy, we propose a self-organizing map (SOM) neural network classification model based on complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) multiscale distribution entropy (MDE) feature extraction and Ant Lion Optimization (ALO) algorithm improvement. The multiscale decomposition of the original seismic and blasting signals was carried out using CEEMDAN, and the distribution entropy values of the obtained multiple intrinsic mode functions were calculated to construct multidimensional feature inputs containing complexity information in the time–frequency domain. The ALO algorithm optimized the key parameters of the SOM neural network (competing layer dimensions and number of training iterations), with the root mean squared error serving as the fitness function. The optimal solution obtained by ALO optimization replaced the hyperparameter values in the original model, and multiple prediction rounds were performed on the seismic data test set to address unstable classification performance caused by random initialization in the traditional SOM network. The results revealed that the recognition performance of the CEEMDAN–MDE combined with the ALO–SOM model was significantly improved compared with machine learning models, such as linear discriminant analysis (LDA), decision tree, support vector machine, probabilistic LDA, and AdaBoost. Its recognition accuracy, recall, and F1-score were 99.3373%, 99.1479%, and 99.4557%, respectively, suggesting that this method can serve as a reliable approach for accurately differentiating between natural earthquakes and artificial blasting events, with important application value for seismic monitoring and blasting event exclusion.

Keywords: Seismic signal recognition; Complete ensemble empirical mode decomposition with adaptive noise; Self-organizing feature mapping neural network; Ant Lion Optimization algorithm; Multiscale distributional entropy

***Corresponding author:**
Tianwen Zhao
(zhaotianwen305@gmail.com)

Citation: Wang A, Pang C, Chen G, Li C, Zhao T. Optimization of the SOM neural network model using CEEMDAN distribution entropy and ALO for seismic and blasting identification. *J Seismic Explor.* doi: 10.36922/JSE025280033

Received: July 13, 2025

Revised: August 3, 2025

Accepted: August 12, 2025

Published online: August 28, 2025

1. Introduction

High-precision identification of earthquakes and explosions is one of the major challenges in seismic observation data processing. Although there are essential differences between the two in terms of earthquake mechanisms and energy release modes, they exhibit high similarity in time-domain waveforms and spectral characteristics, which is of great significance for earthquake early warning, nuclear explosion monitoring, and engineering safety assessment.[1-4] Traditional identification methods typically rely on artificial empirical features (e.g., P/S wave amplitude ratio, spectrum envelope shape, energy spectrum statistical parameters, and P-wave initial motion direction), and the recognition rate drops significantly in strong-noise environments and when the ground-motion energy difference is small.[5-7] With the increase in seismic network density and the improvement of monitoring requirements, intelligent recognition methods based on machine learning have gradually become mainstream; however, their performance is still limited by two bottlenecks: insufficient feature representation capability and weak model generalization. Therefore, integrating adaptive signal decomposition, non-linear dynamic feature extraction, and intelligent optimization models to overcome the recognition accuracy limitations of existing technologies has become a frontier research direction in seismic waveform recognition.

Techniques combining spectrum analysis and machine learning can be used for the accurate recognition of seismic waves at this stage. The integration of non-linear signal processing methods, such as wavelet transform and Hilbert–Huang transform with neural networks has been successfully applied to seismic data processing.[8-10] The development of adaptive signal decomposition methods provides a new paradigm for seismic wave feature extraction. Empirical mode decomposition (EMD) and its derivatives achieve multiscale analysis of signals by decomposing non-linear and non-stationary signals into intrinsic mode functions (IMFs).[11,12] However, traditional EMD suffers from mode aliasing. Although ensemble EMD (EEMD) alleviates this defect by adding Gaussian white noise, it introduces residual noise interference.[13,14] To address these limitations, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) has been developed. Its core innovation lies in the adaptive noise injection mechanism and residual noise isolation strategy: (i) signal contamination is avoided by adding auxiliary noise IMF components decomposed by EMD to the original signal (rather than adding raw white noise); and (ii) after each IMF extraction, integrated averaging is performed immediately to prevent residual noise transmission to low-frequency components. These improvements make CEEMDAN significantly superior to previous methods in signal completeness, modal separation, and computational efficiency, and particularly suitable for pre-processing non-stationary signals, such as seismic waves.[15,16]

At the feature quantification level, multiscale distribution entropy (MDE), an emerging representative of non-linear dynamic characteristics, effectively reveals the essential differences between earthquake and explosion signals by measuring the probability distribution complexity of IMF components at different time scales. Compared with traditional sample entropy,[17] MDE has three major advantages:

(i) Multiscale analysis capability: extracts multiscale information of time series through a coarse-graining process, avoiding the one-sidedness of single-scale analysis.
(ii) Distribution sensitivity: based on probability distribution difference (Euclidean distance or Kullback–Leibler divergence) rather than mean quantization complexity, making it more suitable for non-Gaussian distribution signals.
(iii) Noise robustness: sensitivity to random noise is significantly lower than that of sample entropy. This feature was verified in the processing of earthquake data from Maduo County, Qinghai Province, China, in 2021.

Early recognition models were mainly based on statistical classifiers (e.g., support vector machines [SVMs], random forests, and Adaboost) and shallow neural networks (e.g., backpropagation, radial basis function, and probabilistic neural network), and their performance was highly dependent on artificial feature engineering.[18-21] By extracting features, such as the number of spectral peaks and the short-time energy zero-crossing rate of seismic signals and combining them with SVM classification, recognition rates of up to 85% can be achieved in simple scenarios. However, these models face two major challenges:

(i) Feature–model decoupling: feature extraction and classification models are designed independently, resulting in information transmission loss.
(ii) Overfitting of small samples: seismic event samples are scarce and unevenly distributed, making complex models prone to local optimality.

In recent years, deep neural networks have shown great potential in seismic analysis. As an unsupervised competitive learning model, self-organizing map (SOM) networks retain the ability to reduce the dimensionality of high-dimensional data while preserving topological

structures and are adept at handling complex, non-linear, non-stationary data, making them an ideal choice for seismic signal recognition. Their core advantages are:

(i) Visual interpretability: mapping high-dimensional features to a two-dimensional grid to intuitively display the separation of earthquake and explosion clusters.
(ii) Small-sample adaptability: a robust mapping can be constructed without large-scale training data.[22]

However, the fixed network structure and sensitivity to random initialization parameters of traditional SOMs can affect the overall performance of earthquake and explosion recognition, leading to prediction results with large standard deviations (SDs).[23-25] To address this issue, intelligent biomimetic algorithms are introduced to optimize multiple network hyperparameters of the SOM model, aiming to identify the optimal hyperparameters for the training set and thereby improve the prediction accuracy and output robustness of the original model. Ant Lion Optimization (ALO) simulates the collaborative hunting mechanism of antlions, balancing global exploration and local exploitation capabilities through elite individual guidance and random walk strategies. It is particularly suitable for weight initialization and topological structure adjustment in SOM models, helping determine appropriate competition-layer parameters and initial learning rates, which are then applied to the prediction model in this study.[26,27]

To address the defects of low recognition rate, poor robustness, and imperfect feature learning of machine learning models in seismic wave recognition, the present study employs novel spectrum analysis technology and an improved unsupervised learning algorithm to develop a hybrid model for natural earthquake and artificial explosion signal recognition. The key objective is to design a series of new multiscale spectrum feature criteria based on CEEMDAN with adaptive noise and distribution entropy (DistEn),[28,29] and to propose an innovative CEEMDAN–MDE–ALO–SOM hybrid model optimized using the ALO algorithm.[30] This study makes the following important contributions to the field of natural earthquake and artificial explosion signal recognition:

(i) Novel multiscale feature extraction framework: an improved signal decomposition method based on CEEMDAN is proposed. Across the adaptive noise injection mechanism and residual noise isolation strategy, the modal aliasing problem in the traditional EMD/EEMD method is effectively solved. Combined with the 12-dimensional feature vector constructed by MDE, a comprehensive quantitative characterization of the complexity of seismic signals in the time and frequency domains is achieved for the first time.
(ii) Intelligent optimization of neural network architecture: ALO is innovatively applied to parameter optimization of the SOM neural network, and an automatic parameter adjustment mechanism is established, with the competition-layer dimension and the number of training iterations as optimization variables and root mean squared error as the fitness function. This method addresses the performance instability problem caused by the random initialization of traditional SOM and reduces the SD of classification accuracy to 1.166.
(iii) Interdisciplinary method integration: for the first time, adaptive signal processing (CEEMDAN), non-linear dynamics (MDE), bionic optimization algorithm (ALO), and unsupervised learning (SOM) are systematically integrated to construct an end-to-end intelligent recognition framework. This fusion model is theoretically innovative in the coordinated optimization of feature extraction and classification decision-making.
(iv) Large-scale empirical verification: a rigorous verification scheme, including hundreds of Monte Carlo experiments, is designed based on 414 sets of multi-source data from authoritative institutions, such as the Institute of Engineering Mechanics, China Earthquake Administration. The experimental results not only confirm the model's 99.337% recognition accuracy (F1-score = 99.456%) but also quantify its stability advantage through statistical indicators, such as the coefficient of variation (CV) (CV = 0.0117).
(v) Engineering application value: the developed feature extraction and classification module is encapsulated as a MATLAB-callable function library, supporting real-time signal processing. The model's test performance in the 2021 Qinghai Maduo earthquake aftershock sequence (recall rate = 99.148%) provides a feasible technical solution for reducing the false alarm rate of earthquake monitoring systems.

This study adopts a progressive structure of "theoretical modeling–method innovation–experimental verification–application discussion" to organize the full text. Section 2 systematically explains the mathematical principles of the ALO–SOM model, including the elite retention mechanism and random walk strategy of the ALO algorithm, the competitive learning dynamics model of the SOM network, and the complete algorithm flow chart (Figure 1). Section 3 details the CEEMDAN–MDE feature extraction method, covering the adaptive noise injection strategy of CEEMDAN, the multiscale probability distribution quantization method of MDE, and the criteria for selecting IMF components (Figures 2 and 3). Section 4 presents a rigorous controlled-variable experiment,

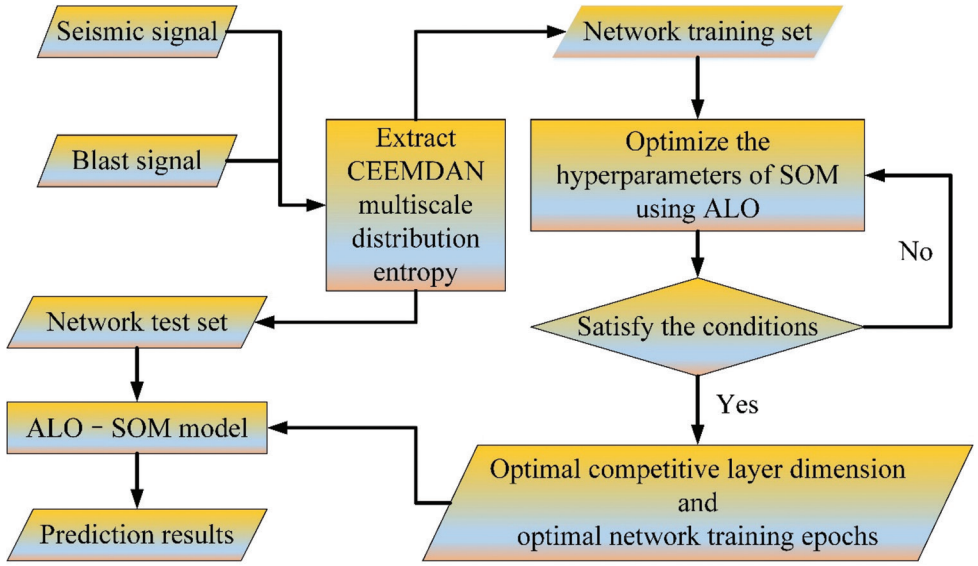

Figure 1. Flowchart of the ALO–SOM model for recognizing earthquakes and blasting.
Abbreviations: ALO: Ant Lion Optimization; CEEMDAN: Complete ensemble empirical mode decomposition with adaptive noise; SOM: Self-organizing map.

explaining the data sources and pre-processing procedures, demonstrating the ALO optimization process (Figure 4) and hyperparameter sensitivity analysis, and finally comparing six benchmark models through box plots (Figure 5) and statistical tables (Tables 1 and 2). Section 5 discusses three key issues in depth, including a comparison of spectral resolution with methods, such as EEMD–VMD, the robustness boundary in strong-noise environments, and the trade-off between computational efficiency and real-time performance. Section 6 summarizes the research results and future directions, including lightweight model deployment, cross-regional generalization testing, and the construction of a multimodal data fusion recognition framework.

2. ALO–SOM seismic wave identification model

2.1. ALO

ALO is a heuristic algorithm proposed by Mirjalili *et al.*[31] in 2015, inspired by the behavior of antlions hunting prey in nature. The algorithm imitates the habits of antlions setting traps, prey random walks, and antlions waiting to hunt. It adopts a fast convergence mechanism based on trap boundary search and elite retention, and applies roulette and random walk strategies to improve global search capabilities. It features strong robustness and simple algorithm settings. The ALO algorithm process is described in the following sections.

2.1.1. Ant random walk

The set of ant random walk steps is defined as $X_i(t)$, with the initial step number set to 0. The ant random walk is constrained within a range-limited domain, and the ant position needs to be normalized based on **Equation I**:

$$X_i^t = \frac{\left(X_i - a_i\right)\times\left(d_i^t - c_i^t\right)}{\left(b_i - a_i\right)} + c_i^t \quad \text{(I)}$$

where X_i^t is the normalized result of the ant random walk step set X_i within the feasible domain; a_i and b_i are the minimum and maximum values of the ant position vector on the i-th dimension, preset by the algorithm; and c_i^t and d_i^t are the minimum and maximum values of the ant position on the i-th dimension at the t-th iteration.

2.1.2. Antlion sets a trap

When an ant mistakenly enters the trap pit dug by the antlion, the ant's movement is restricted, and the walking formulas within the trap are given in **Equations II and III**:

$$c_i^t = A_j^t + c^t \quad \text{(II)}$$

$$d_i^t = A_j^t + d^t \quad \text{(III)}$$

Where c^t is the minimum value of all variables at the t-th iteration, d^t is the maximum value of all variables at the t-th iteration, and A_j^t is the antlion j selected at the t-th iteration.

2.1.3. Antlion trapping target

When the antlion finds an ant, it throws sand to the edge of the sand pit to prevent the ant from escaping, causing

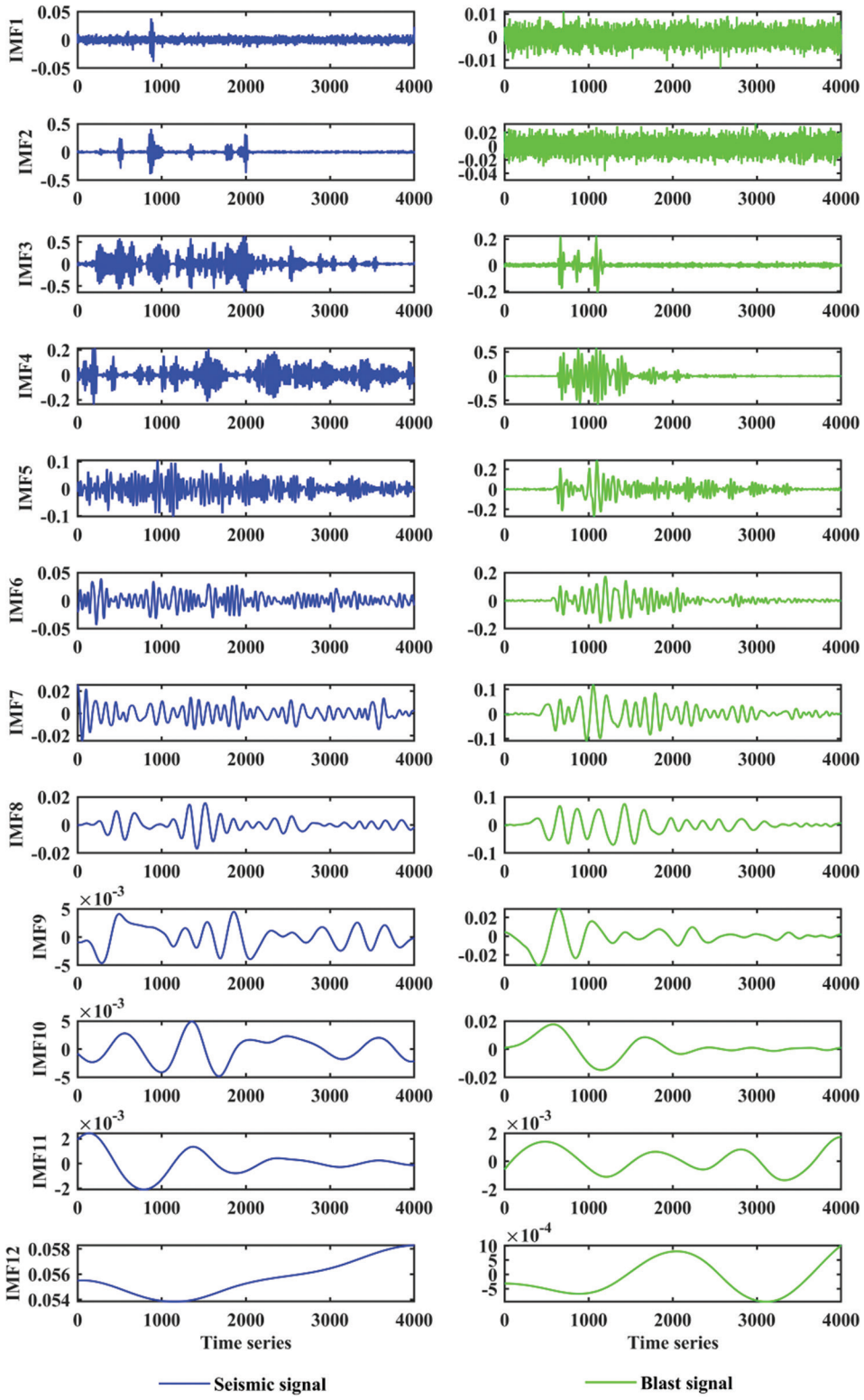

Figure 2. Comparison of complete ensemble empirical mode decomposition with adaptive noise decomposition results between single earthquake and blasting waveforms.
Abbreviation: IMF: Intrinsic mode function.

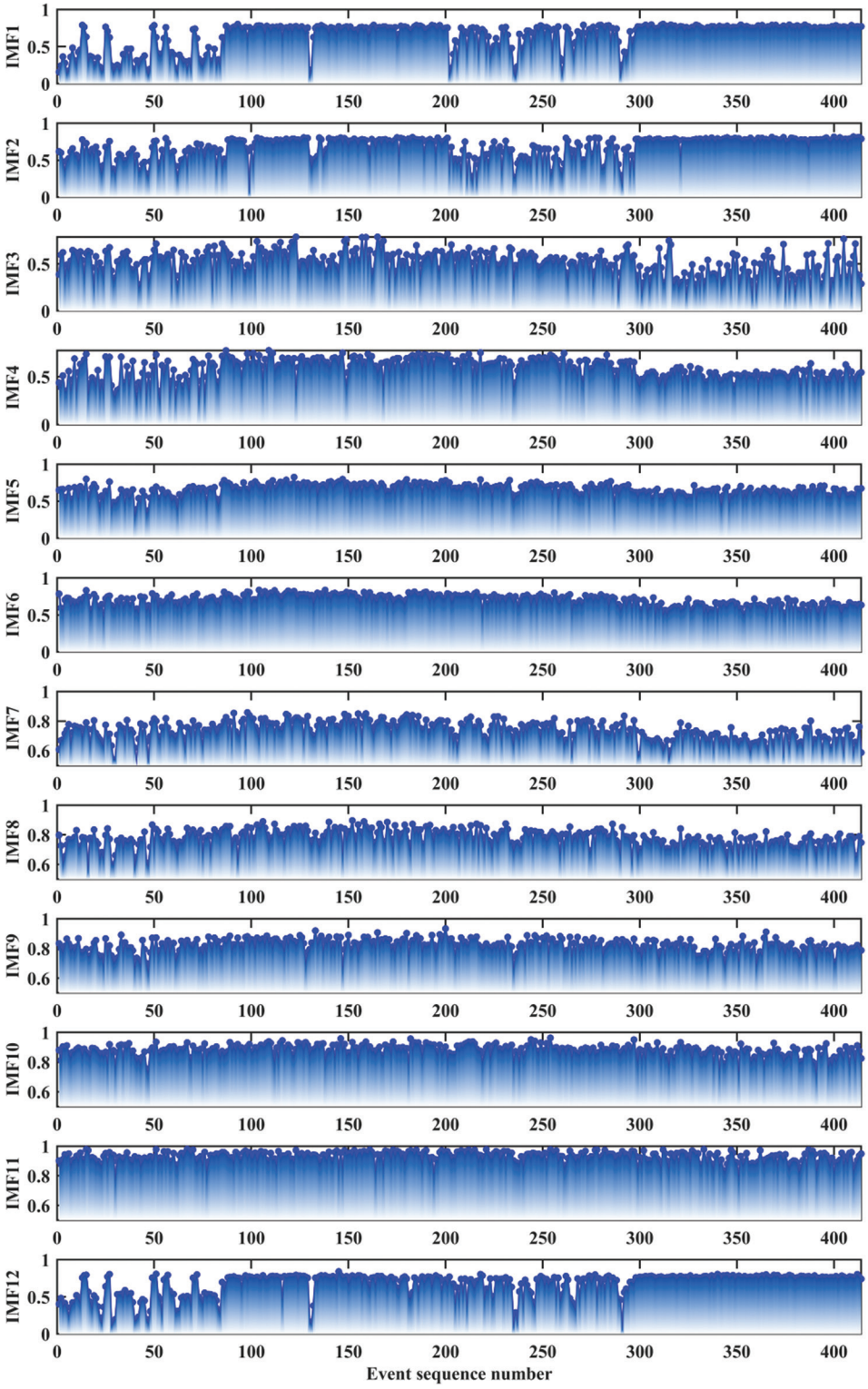

Figure 3. Distribution entropy of IMF components from complete ensemble empirical mode decomposition with adaptive noise for earthquake and explosion waveforms.
Abbreviation: IMF: Intrinsic mode function.

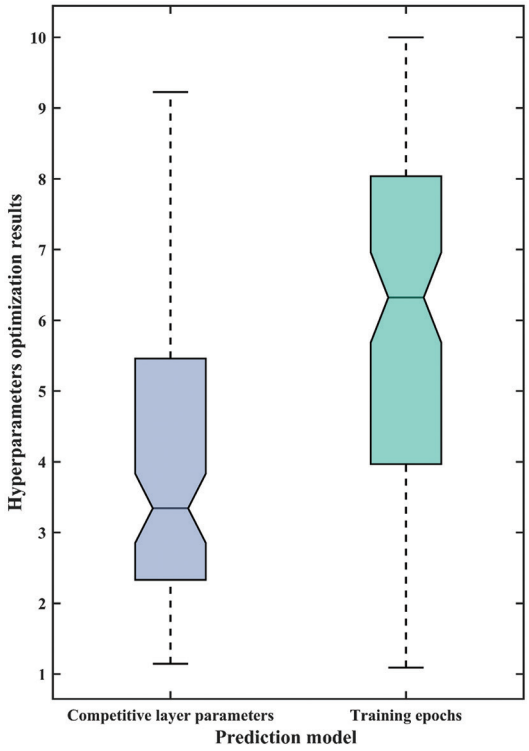

Figure 4. Hyperparameter results for the Ant Lion Optimization-optimized self-organizing map across 100 discrimination subtrials.

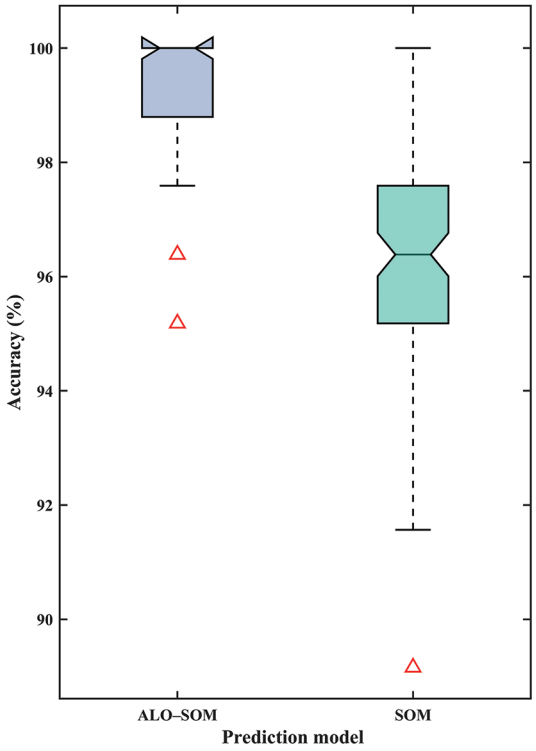

Figure 5. Box plot of 100 identification results comparing the ALO–SOM model and the SOM model. Data points marked with red triangles are identified as outliers in the box plot.
Abbreviations: ALO: Ant Lion Optimization; SOM: Self-organizing map.

the ant to slide continuously toward the antlion at the bottom of the pit. The ALO algorithm uses roulette wheel selection to identify the appropriate antlion position and dynamically narrows the trap range to speed up the capture of ants. The relevant formulas are given in **Equations IV–VI**:

$$c^t = \begin{cases} c^t, t \le 0.1T \\ T \times c^t\left(10^w \times t\right), t > 0.1T \end{cases} \quad \text{(IV)}$$

$$d^t = \begin{cases} d^t, t \le 0.1T \\ T \times d^t\left(10^w \times t\right), t > 0.1T \end{cases} \quad \text{(V)}$$

$$w = \begin{cases} 2, 0.10T < t \le 0.50T \\ 3, 0.50T < t \le 0.75T \\ 4, 0.75T < t \le 0.90T \\ 5, 0.90T < t \le 0.95T \\ 6, 0.95T < t \le T \end{cases} \quad \text{(VI)}$$

Where T is the preset upper limit of the number of algorithm iterations, and w is the dynamic weight factor related to the present number of iterations t. **Equation VI** shows a stepwise increasing trend for w.

Table 1. Statistical summary of 100 comparison tests between the ALO–SOM and SOM models

Identification model	Accuracy				
	Mean (%)	SD	Range	CV	IQR
SOM	96.5904	1.9303	10.8434	0.0200	2.4096
ALO–SOM	99.3373	1.1662	4.8193	0.0117	1.2048

Abbreviations: ALO: Ant Lion Optimization; CV: Coefficient of variation; IQR: Interquartile range; SOM: Self-organizing map; SD: Standard deviation.

2.1.4. Preying on ants

When the prey slides to the bottom of the pit, the antlion moves quickly and catches the prey. This biological phenomenon is modeled algorithmically by comparing the fitness values of the ant and the antlion. When the ant's fitness value is higher than that of the antlion, the position of the antlion is updated to the present position of the ant. The relevant formula is given in **Equation VII**:

$$Antlion_j^t = Ant_i^t, f_{obj}^{ant}\left(t\right) < f_{obj}^{antlion}\left(t\right) \quad \text{(VII)}$$

where $Antlion_j^t$ is the position of the antlion at the t-th iteration; Ant_i^t is the position of the ant at the t-th iteration;

Table 2. Statistical summary of the recognition effect of 100 rounds of six machine learning models

Identification method	Mean		
	Accuracy (%)	Recall (%)	F1-score (%)
LDA	94.4458	95.0857	96.0652
Decision tree	94.7590	96.7272	96.2864
SVM	96.7470	97.8544	97.7154
PLDA	94.7229	95.5762	96.2838
AdaBoost	94.2892	95.8346	96.0004
CNN	93.4504	94.6401	95.0182
ALO–SOM	99.3373	99.1479	99.4557

Abbreviations: ALO: Ant Lion Optimization; CNN: Convolutional neural network; LDA: Linear discriminant analysis; PLDA: Probabilistic linear discriminant analysis; SOM: Self-organizing map; SVM: Support vector machine.

and $f_{obj}^{ant}(t)$ and $f_{obj}^{antlion}(t)$ are the fitness function values of the ant and antlion, respectively, at the t-th iteration.

2.1.5. Elite antlion strategy

The movement of the ant is influenced by both the present elite antlion position and the antlion position selected by the roulette wheel. The ant position is defined by **Equation VIII**:

$$Ant_i^{t+1} = \frac{R_A^t(l) + R_E^t(l)}{2} \quad \text{(VIII)}$$

where Ant_i^{t+1} represents the position of the i-th ant in iteration $t + 1$; $R_A^t(l)$ is the latest position of the random walk steps near the antlion selected by roulette at the t-th iteration; and $R_E^t(l)$ is the position of the ant in the random walk of l steps near the elite antlion at iteration t (the best antlion position found at each iteration).

2.2. Self-organizing feature mapping neural network

SOM is an unsupervised machine learning method, also known as the Kohonen network. Its key idea is to map and compress high-dimensional data onto a two-dimensional plane while preserving the topological structure of the original data to obtain the feature similarity distribution of the output layer. It has advantages, such as effective processing of non-linear data, producing intuitive and visual results, and not requiring preset labels. The calculation steps of the SOM network are detailed in the following sections.

2.2.1. Competition process

The Euclidean distance between the input vector X and the weight vector W_j of the competition layer neuron is calculated based on **Equation IX**:

$$d_j = X - W_j = \sqrt{\sum_{i=1}^{m}\left(x_i(t) - w_{ij}(t)\right)^2} \quad \text{(IX)}$$

Where $X = (x_1, x_2, \ldots, x_m)^T$ is the input vector; W_j is the weight vector of the competition layer neuron j; w_{ij} is the weight connecting input neuron i and competition neuron j.

The best matching unit (BMU) is defined as the neuron with the smallest d_j:

$$BMU = \arg\min_j X - W_j \quad \text{(X)}$$

2.2.2. Cooperation process

During the cooperation process, the weights of the BMU and its adjacent neurons are updated. The adjacent spatial range is defined by the Gaussian function h, where σ decreases with time t, thereby dynamically adjusting the adjacent spatial range of the BMU. The formula is given in **Equation XI**:

$$h(j, BMU, t) = \exp\left(-\frac{distance(j, BMU)^2}{2\sigma(t)^2}\right) \quad \text{(XI)}$$

2.2.3. Adaptation process

To make the BMU approach the input vector, this process introduces the learning rate η and the Gaussian function h to update the BMU and the weights of its neighboring neurons. The learning rate η decreases with the number of iterations based on the initial learning rate η_0, as defined in **Equation XII**:

$$\eta(t) = \eta_0 \exp\left(-\frac{t}{\tau}\right) \quad \text{(XII)}$$

where τ is the time constant that determines the decay rate of the learning rate.

$W_j(t + 1)$ is the weight vector of the competition layer neuron j at time $t + 1$.

$$W_j(t+1) = W_j(t) + \eta(t) \cdot h(t) \cdot (x - W_j(t)) \quad \text{(XIII)}$$

2.2.4. Training process

During the training process, weights are initialized, samples are randomly selected to calculate the BMU, and the BMU, along with its neighboring weights, are updated. Simultaneously, the neighborhood radius and learning rate decrease as the number of iterations increases.

2.3. ALO–SOM model recognition principle

The classification results of SOM networks are influenced by multiple network hyperparameters and must be controlled in combination with other methods to obtain more accurate and stable unsupervised clustering results.

To achieve this, the ALO algorithm, which mimics the antlion's strategy of setting traps to hunt ants, is introduced to optimize certain SOM hyperparameters (e.g., competition layer dimension and network training iterations), thereby developing a new model capable of adaptively training for earthquake and explosion recognition on the training dataset.

Figure 1 illustrates the process of the ALO algorithm optimizing the SOM neural network hyperparameters and performing pattern recognition. The ALO algorithm uses two hyperparameters from the SOM—the dimension of the competition layer and the number of network training iterations—as optimization variables. It employs the root mean squared error between the actual prediction result vector $\vec{R}$ and the theoretical category label vector $\vec{T}$ obtained from SOM recognition using the test set as the fitness function for ALO optimization (**Equation XIV**):

$$fitness = RMSE\left(\vec{R},\vec{T}\right) \tag{XIV}$$

Based on the training set, the optimal hyperparameter values that satisfy the iteration stopping criteria are obtained, and finally, the ALO–SOM recognition model is used to identify the test set.

3. CEEMDAN MDE feature extraction

An efficient and reasonable new neural network model does not necessarily guarantee high accuracy in effectively distinguishing between earthquakes and explosions; it also requires a reliable seismic waveform feature extraction process. In this section, we provide a detailed introduction to the basic concepts and computational process of CEEMDAN, which enables the extraction of seismic waveform features across multiple frequency scales, thereby enabling a more comprehensive and precise analysis of the time–frequency differences between earthquakes and explosions.

3.1. CEEMDAN

CEEMDAN is an advanced signal decomposition method based on EMD and EEMD techniques. It enhances signal decomposition by adding complementary pairs of adaptive white noise to the original signal, performing multiple EMD decompositions, and averaging the results. This process effectively minimizes parameter interference and suppresses mode aliasing by isolating residual components. CEEMDAN overcomes key limitations of earlier methods—including modal overlap, low reconstruction accuracy, reliance on fixed parameters, and low decomposition efficiency—thus significantly improving the reconstruction purity of IMFs. It is particularly suitable for analyzing non-linear and non-stationary signals, such as complex seismic data.

(a) Step 1
m pairs of positive and negative Gaussian white noises $\omega_i(t)$(i = 1,2,…,*m*) with zero mean and constant SD are added to the original signal *x*(*t*), generating *m* synthetic noisy signals $X_i^1(t)$ based on **Equation XV**:

$$X_i^1(t) = x(t) + (-1)^q \cdot \beta \cdot \omega_i(t) \tag{XV}$$

where β is the noise coefficient related to the amplitude, and q = 1,2.

(b) Step 2
Empirical mode decomposition is applied to the noisy signals to obtain the first-order components. Each signal $X_i^1(t)$ is decomposed by EMD into an IMF component $IMF_i^1(t)$ and a residual component $r_i^1(t)$, as expressed in **Equation XVI**:

$$X_i^1(t) = IMF_i^1(t) + r_i^1(t) \tag{XVI}$$

The first-order IMFs $IMF_i^1(t)$ of all *m m* synthetic noisy signals $X_i^1(t)$ are calculated, and their arithmetic average is taken to obtain the first-order $IMF^1(t)$ $IMF^1(t)$ of the CEEMDAN algorithm. The relevant formulas are given in **Equations XVII and XVIII**:

$$IMF^1(t) = \frac{1}{m}\sum_{i=1}^{m} IMF_i^1(t) \tag{XVII}$$

$$r^1(t) = x(t) - IMF^1(t) \tag{XVIII}$$

(c) Step 3
Using a similar calculation strategy as in **Equations XV–XVIII**, the next-order component $IMF^{k-1}(t)$ is calculated step by step: for the residual $r^{k-1}(t)(k \geq 2)$ obtained in the previous step, positive and negative noise $(-1)^q \beta_{k-1} E_{k-1}(\omega_i[t])$ are added, respectively, to obtain *m* new signals $X_i^{k-1}(t)$, where β_{k-1} is the dynamically reduced noise coefficient, and $E_{k-1}(\cdot)$ is the residual after the (*k*−1)-th order EMD decomposition of the white noise $\omega_i(t)$. Performing EMD decomposition on each $X_i^{k-1}(t)$ yields *m* components $IMF_i^{k-1}(t)$, as given in **Equation XIX**:

$$X_i^{k-1}(t) = IMF_i^{k-1}(t) + r_i^{k-1}(t) \tag{XIX}$$

Taking the arithmetic average yields the (*k*−1)-th order component $IMF^{k-1}(t)$ of the CEEMDAN algorithm. The relevant formulas are given in **Equations XX and XXI**:

$$IMF^{k-1}(t) = \frac{1}{m}\sum_{i=1}^{m} IMF_i^{k-1}(t) \tag{XX}$$

$$r^{k-1}(t) = r^{k-2}(t) - IMF^{k-1}(t), k \geq 2 \tag{XXI}$$

(d) Step 4

After the iterative calculation is completed, the original signal is reconstructed. The iterative calculation from the previous step stops when the residual component becomes a monotonic function or when its extreme points are insufficient for further EMD decomposition. A total of $k-1$ CEEMDAN IMFs are obtained; the original signal $x(t)$ can be reconstructed by summing the $k-1$ IMFs and the final residual component $r^{k-1}(t)$ as follows (**Equation XXII**):

$$x(t) = r^{K-1}(t) + \sum_{k=2}^{K} IMF^{k-1}(t), \quad k = 2,...,K \tag{XXII}$$

The hyperparameters of the CEEMDAN algorithm are set as follows: the SD of white noise is 0.2, the number of noise additions is 24, and the maximum number of iterations allowed is 3,600. Figure 2 shows the CEEMDAN decomposition results of natural earthquake signals (left) and artificial blasting signals (right), where rows 1–12 correspond to the IMF_1–IMF_{12} components obtained by CEEMDAN decomposition. The waveform signal length is L = 4,000 and the components are arranged in descending order of frequency or energy size.

3.2. Calculation of MDE using CEEMDAN

DistEn is a parameter used in information theory to measure the uncertainty of data distribution or the complexity of a time series.[28,29] DistEn obtains the probability density function by directly calculating the Chebyshev distance and kernel density estimation between reconstructed vectors, thereby avoiding the problem of manually selecting the tolerance parameter r used in sample entropy. In addition, it has the advantages of being parameter-free and robust.

The original seismic signal is decomposed into several IMFs with monotonically decreasing frequencies and significant differences in energy distribution by using the adaptive noise CEEMDAN. By sequentially calculating the DistEn values of all IMF components obtained from decomposition (i.e., the DistEn of IMF_1–IMF_{12} shown in Figure 3), a one-dimensional vector that describes the different frequency distribution characteristics of the original signal is formed, referred to as MDE. This method can effectively extract multiscale pure characteristic parameters that characterize different source systems and enables robust identification of signals generated by different dynamic mechanisms. The horizontal axis of Figure 3 represents the sequence numbers of 414 earthquake and explosion waveforms, while the vertical axis shows the DistEn values of the IMF components obtained from CEEMDAN decomposition.

4. Data and experiments

In this study, we first organized the collected data and developed an experimental framework to rigorously evaluate the effectiveness and specific capabilities of the proposed method in distinguishing between earthquake and blasting events. A total of 414 sets of multi-source strong motion observation data were utilized, as follows:

(i) Earthquake case data publicly shared by the National Earthquake Data Center (data.earthquake.cn) and the Institute of Engineering Mechanics of the China Earthquake Administration: 2021 Jiangtianning MS 4.2 earthquake strong motion acceleration records (96 records); 2021 Jiangsu Dafeng MS5.0 earthquake strong motion acceleration records (117 records); and 2021 Yunnan Yangbi MS6.4 main shock aftershock sequence acceleration records (magnitude range MS2.9-3.9, 84 records).

(ii) Controllable artificial blasting test acceleration waveforms (117 records) provided by the Geotechnical Engineering Institute, China Institute of Water Resources and Hydropower Research (www.geoeng.iwhr.com).

All data pre-processing and numerical analyses were implemented on the MATLAB 2019a (The MathWorks, Inc., United States) computing platform. The formulas for accuracy, recall, and F1-score are as follows:

(i) Accuracy indicates the proportion of samples correctly predicted by the model to the total number of samples (**Equation XXIII**).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \tag{XXIII}$$

(ii) Recall refers to the proportion of positive samples correctly identified by the model to all actual positive samples (**Equation XXIV**).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \tag{XXIV}$$

(iii) The F1-score is the harmonic mean of precision and recall and is an important indicator for measuring the overall performance of the model (**Equation XXV**).

$$\text{F1} - \text{score} = \frac{2 \times \text{Accuracy} \times \text{Recall}}{\text{Accuracy} + \text{Recall}} \times 100\% \tag{XXV}$$

In these formulas, TP indicates true positive, TN represents true negative, FP is false positive, and FN signifies false negative.

Considering that the randomness of the initial parameters of the neural network may cause fluctuations

in the prediction results and lead to potential large deviations, this study conducted a 100-cycle Monte Carlo test to systematically evaluate the effectiveness and robustness of the ALO–SOM model. By introducing key hyperparameter perturbations, the stability of the model's earthquake identification performance under parameter variations was tested. The training set and test set were strictly split in an 8:2 ratio (i.e., 331:83 samples). The statistical characteristics of the repeatability test results are presented in Figures 4 and 5, Table 1.

Figure 5 presents a box plot comparing the 100 identification results of the ALO–SOM and SOM models, which simultaneously illustrates the data dispersion and statistical characteristics of the multiple rounds of prediction results from the two neural network models. The upper and lower boundaries of the box in the figure represent the upper quartile (Q3) and lower quartile (Q1) of the parameter results, respectively. The solid line inside the box indicates the median of the parameter results. Data points marked with red triangles are identified as outliers in the box plot. The horizontal solid lines above and below the box represent the maximum and minimum values, respectively. The performance difference between the ALO–SOM and SOM models was analyzed using parameters, such as the mean, SD, range, CV, and interquartile range (IQR).

Based on Figure 5, the recognition curve of the SOM model exhibits irregular oscillations, with approximately 50% of the recognition results below 97%, and a small number even below 90%, indicating a degree of instability in earthquake and explosion recognition. The SOM model optimized by ALO demonstrates stronger adaptability to data features and can effectively utilize matching competition layer dimensions and training iterations to recognize earthquakes and explosions. The worst result is above 95%, and most recognition results exceed 99%. Its performance across multiple indicators—mean, SD, range, CV, and IQR (99.3373%, 1.1662, 4.8193, 0.0117, and 1.2048, respectively)—demonstrates stronger robustness and accuracy.

Figure 4 presents a box plot of the calculation results for the ALO-optimized SOM hyperparameter values across 100 identification sub-tests. The search domain was set to [1,10,] which illustrates both the data dispersion and statistical characteristics of the 100 optimization results for the two SOM hyperparameters. The upper and lower boundaries of the box in the figure represent the upper quartile (Q3) and lower quartile (Q1) of the parameter results, respectively. The solid line within the box represents the median of the parameter results. The horizontal solid lines above and below the box indicate the maximum and minimum values, respectively. As shown in Figure 4, the SOM competition layer dimension and SOM training iterations optimized by ALO exhibit a certain degree of randomness and are not consistently stable values. This variability is related to the random division of the training set and the early termination of ALO iterations.

Table 1 presents the statistical results of the 100 comparison tests between the ALO–SOM model and the traditional SOM model. The results indicate that the ALO–SOM model outperforms the SOM model in all indicators, exhibiting higher accuracy and stability. Specifically, the average accuracy of the ALO–SOM model over 100 prediction rounds was 99.3373%, significantly higher than that of the SOM model (96.5904%), suggesting that the ALO–SOM model can provide more accurate results in earthquake identification tasks, with an accuracy improvement of 2.7469%.

In addition, the SD of the 100-round prediction accuracy for the ALO–SOM model was 1.1662, significantly lower than that of the SOM model (SD = 1.9303), indicating that its prediction results exhibit lower variability and higher stability, making it suitable for processing earthquake monitoring data that requires highly reliable results. Furthermore, the IQR of the ALO–SOM model's 100-round prediction accuracy was 4.8193, which is also significantly smaller than that of the SOM model (IQR = 10.8434), further suggesting that the ALO–SOM model's prediction results are more concentrated and less influenced by extreme values, thereby demonstrating better robustness. The CV and IQR of the ALO–SOM model highlight its superiority in prediction result consistency, with a CV of 0.0117 for ALO–SOM and 0.0200 for SOM, and an IQR of 1.2048 for ALO–SOM and 2.4096 for SOM.

To evaluate the superior performance of the ALO–SOM model and the differences in performance between machine learning algorithms based on different principles, we introduced six classic models: linear discriminant analysis (LDA), decision tree, SVM, probabilistic LDA (PLDA), convolutional neural network (CNN), and AdaBoost ensemble learning. These models have distinct strengths and weaknesses. For example, SVM is effective in handling non-linear problems but is sensitive to noisy data and outliers; LDA is suitable for large-scale linear datasets but may encounter singularity issues in the inter-class covariance matrix; AdaBoost has moderate tolerance for noise and outliers but is prone to overfitting; and decision tree models are highly interpretable but are still prone to overfitting and sensitive to outliers. The parameters for this model comparison experiment were kept constant, with 100 consecutive recognition tests conducted using a fixed training-to-test set ratio of 8:2. The comparison results are presented in Table 2 and Figure 6.

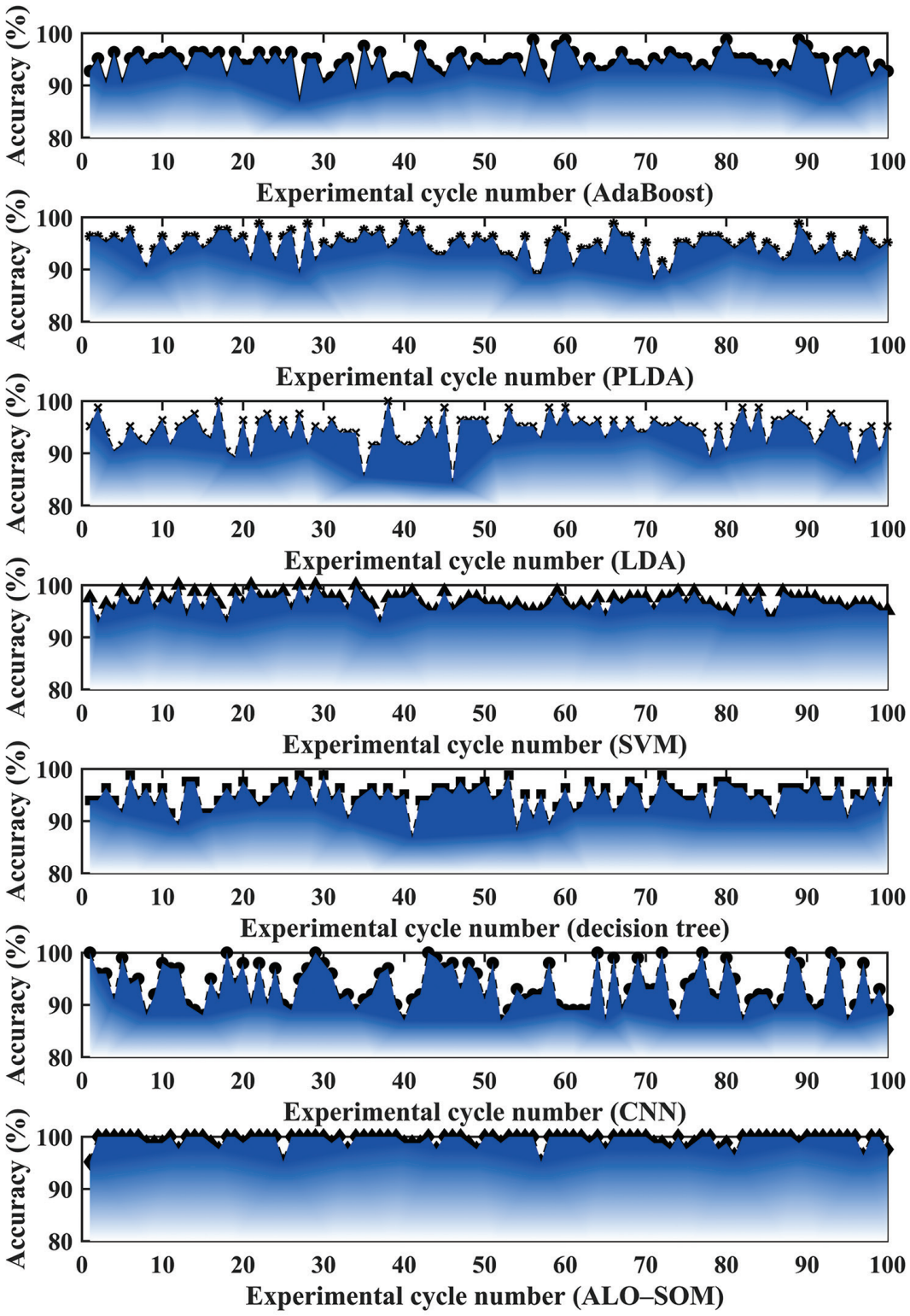

Figure 6. Comparison of seismic wave prediction performance across 100 rounds by six machine learning models.
Abbreviations: ALO: Ant Lion Optimization; CNN: Convolutional neural network; LDA: Linear discriminant analysis; PLDA: Probabilistic linear discriminant analysis; SOM: Self-organizing map; SVM: Support vector machine.

As shown in Figure 6 and Table 2, the earthquake and explosion signal recognition performance (in terms of accuracy, recall rate, and F1-score) of the ALO–SOM model was significantly better than that of classic machine learning models, including LDA, decision tree, SVM, PLDA, CNN, and AdaBoost, with all metrics exceeding 99%. This demonstrates that the model has potential applicability in small-sample earthquake event classification and offers room for further research and improvement.

To calculate the recognition accuracy of the ALO–SOM model for small-sample events, such as blast events, we employed MATLAB interpolation and resampling techniques to increase the number of blast signals from 117 to 351, while maintaining the number of natural signals at 297. The sampling rate remained at 200 Hz, resulting in a total of 648 signals used in the subsequent recognition experiments. The training-to-test set ratio was set to 8:2 (i.e., 518:130). The experimental results are illustrated in Figure 7.

Under resampling conditions, the average recognition accuracy of the ALO–SOM model over 100 prediction rounds was 99.2308%, with an SD of 0.6468, which was essentially equivalent to the recognition performance of the original experiment without resampling blast events. In addition, it was found that the mean recognition accuracy rates for both seismic and blast signals improved compared to previous results, reaching 99.5923% and 99.6385%, respectively. The SDs for accuracy rates were 0.5729 and 0.5295 for seismic and blast signals, respectively, indicating that the ALO–SOM model demonstrates good recognition performance for both seismic and blast signals in small-sample datasets.

5. Discussion

This study proposes a new hybrid model, combining CEEMDAN–MDE and ALO–SOM, for high-precision discrimination of seismic and blast signals. Experimental results showed that the model achieved excellent performance, with an accuracy of 99.337%, a recall of 99.148%, and an F1-score of 99.456%—significantly better than those of traditional machine learning methods, such as LDA, decision tree, SVM, PLDA, and AdaBoost. The combination of MDE feature extraction based on CEEMDAN and the SOM neural network optimized by ALO effectively addresses the problems of insufficient feature expression and insufficient model stability in seismic signal classification.

The practical significance of this study is that the CEEMDAN–MDE feature set provides a standardized framework for quantifying the complexity of seismic signals,

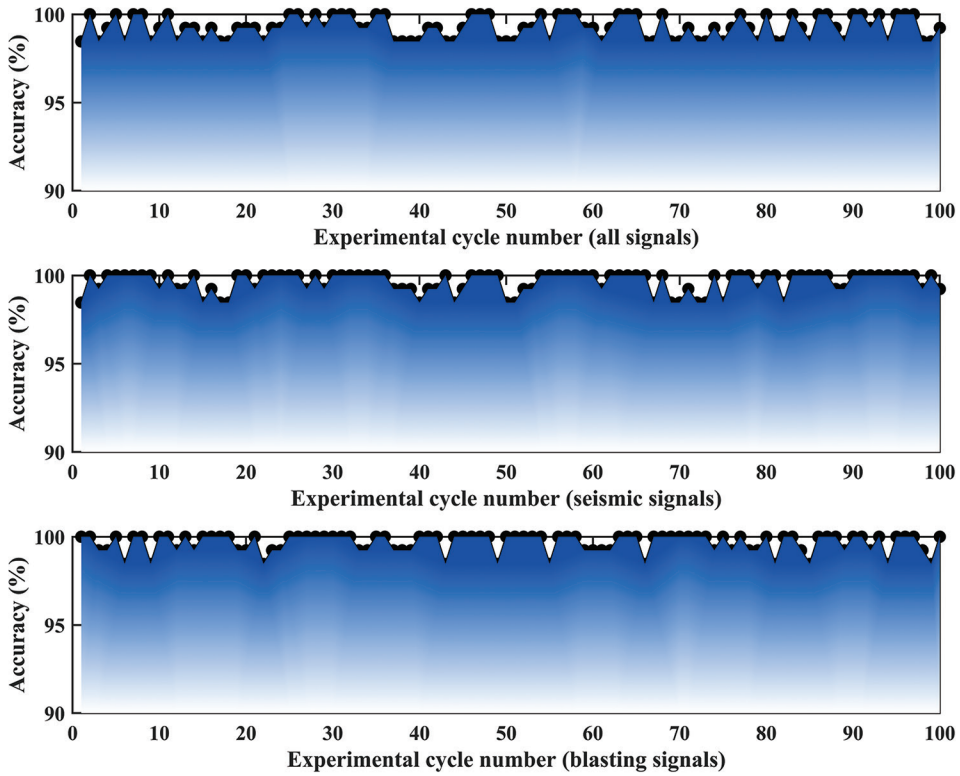

Figure 7. Seismic identification results of the Ant Lion Optimization–self-organizing map model under resampling conditions.

which can be extended to other waveform classification tasks (e.g., volcanic earthquakes, industrial vibrations). The stability and interpretability of the ALO–SOM model make it suitable for deployment in seismic monitoring systems, especially in scenarios requiring rapid discrimination between natural and man-made events (e.g., nuclear test monitoring, mine safety assessment). The high recall of the model (99.148%) is particularly important for reducing false negatives in early warning systems.

Despite the strengths of this study, several limitations remain. First, the dataset, although carefully curated, comprises 414 samples from specific regions (e.g., Jiangsu, Yunnan), which may limit the generalizability of the model in different geological environments. Second, environmental noise (e.g., wind, traffic) is not explicitly simulated, which may affect the model's performance in noisy, real-world environments. Third, the computational cost of CEEMDAN–MDE feature extraction and ALO optimization, while justified by improved accuracy, may hinder real-time applications on low-power edge devices. Finally, the model relies on all 12 initial functions (IMFs), and no feature selection was performed, which may introduce redundancy—methods, such as wrapper-based selection could simplify this process.

Future research should focus on: (i) expanding the dataset to cover global earthquake events and diverse noise conditions to enhance model robustness; (ii) exploring lightweight variants of CEEMDAN (e.g., online CEEMDAN) and entropy metrics (e.g., fuzzy entropy) to enable real-time deployment; (iii) integrating attention mechanisms or Transformer architectures to enhance feature learning; (iv) studying hybrid models that combine ALO–SOM with ensemble techniques (e.g., stacking) to address the class imbalance problem in rare event detection. In addition, applying this framework to other geophysical signal classification tasks (e.g., landslide vibration, structural health monitoring) would help verify its broader practicality.

In summary, the CEEMDAN–MDE–ALO–SOM model represents a significant advancement in the field of earthquake signal recognition, providing both theoretical innovation and practical value. Addressing its limitations through collaborative data sharing and algorithmic improvements will be key to advancing earthquake monitoring technology.

6. Conclusion

The accurate distinction between natural earthquake signals and artificial explosion signals is crucial to ensuring the reliability of earthquake early warning information release and advancing artificial intelligence-based seismology research. Based on small-sample data from multiple seismic stations across the country, this study extracted MDE features from the 12-dimensional CEEMDAN spectrum decomposition of normalized seismic wave signal, employed the SOM self-organizing feature mapping network as the basic learner, and combined the ALO algorithm to optimally tune its competition layer dimensions and training iterations, thereby improving the recognition accuracy and operational stability of the original model. Accordingly, the following conclusions are drawn:

(i) Compared with the standard SOM neural network model, LDA, decision tree, SVM, PLDA, AdaBoost ensemble learning, and other commonly used machine learning models, the ALO–SOM model achieved significantly higher earthquake and explosion recognition accuracy, with a recognition rate of 99.3373% and an SD of only 1.1662.

(ii) The multiscale spectrum feature set CEEMDAN–DisEn, which contains the IMF component DistEn values of different frequency bands, demonstrated a stronger capability for subdividing seismic wave features.

(iii) Several limitations remain in this study. First, a large amount of environmental noise may mask the effective components of seismic waves, thereby affecting the accuracy of waveform spectrum feature extraction. Second, this study used all 12-dimensional IMF components obtained by CEEMDAN decomposition for feature extraction without applying feature selection or high-dimensional data compression, which could reduce prediction efficiency and cause some loss of accuracy. If feature selection methods, such as filtering, encapsulation, and embedding, or dimension reduction methods—such as principal component analysis and low-variance filtering—were applied for feature processing, the accuracy of distinguishing between earthquakes and explosions could be further improved, while moderately reducing algorithm redundancy and iteration time. Therefore, future studies are encouraged to integrate principal component analysis to perform feature engineering compression on all IMF components from CEEMDAN, producing 2–3 new features that best represent the multiscale characteristics of seismic waveforms. This would reduce the complexity of the input matrix for the prediction model, thereby improving model training speed and prediction efficiency. Finally, due to the model's outstanding unsupervised clustering performance, it could be applied to the fault diagnosis of seismic instrument systems in the future, accurately distinguishing between data acquisition faults, power

supply faults, electromagnetic interference, and other abnormal signals, thus ensuring the effective recording and observation of environmental noise data and seismic event signals by seismic instruments.

Acknowledgments

We thank the Institute of Engineering Mechanics of China Earthquake Administration (National Earthquake Science Data Center) and the Institute of Geotechnical Engineering, China Institute of Water Resources and Hydropower Research, for providing the data support for this study.

Funding

This research was financially supported by Wuhan Gravitation and Solid Earth Tides, National Observation and Research Station Open Fund Project (No. WHYWZ202406, WHYWZ202208); the Scientific Research Fund of Institute of Seismology, CEA and National Institute of Natural Hazards, MEM (No. IS202236328, IS202436357); the Spark Program of Earthquake Technology of CEA (No. XH24025YC); the Earthquake Monitoring and Forecasting and Early Warning Tasks for 2025 (No. CEA-JCYJ-202502015); Chengdu Jincheng College Green Data Integration Intelligence Research and Innovation Project (No. 2025-2027); and the High-Quality Development Research Center Project in the Tuojiang River Basin (No. TJGZL2024-07).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Cong Pang, Tianwen Zhao, Ailing Wang

Formal analysis: Cong Pang, Guoqing Chen, Tianwen Zhao, Ailing Wang

Investigation: Tianwen Zhao, Chawei Li, Ailing Wang

Methodology: Cong Pang, Tianwen Zhao, Guoqing Chen, Chawei Li

Writing – original draft: Cong Pang, Tianwen Zhao, Guoqing Chen, Ailing Wang

Writing – review & editing: Cong Pang, Tianwen Zhao, Guoqing Chen, Ailing Wang

Availability of data

This study integrates 1,000 sets of strong earthquake observation data and explosion data from regions in and around China, primarily sourced from publicly available earthquake case data provided by the National Earthquake Data Center (data.earthquake.cn), the Institute of Engineering Mechanics, China Earthquake Administration, and the China Institute of Water Resources and Hydropower Research.

References

1. Lythgoe K, Loasby A, Hidayat D, Wei S. Seismic event detection in urban Singapore using a nodal array and frequency domain array detector: Earthquakes, blasts, and thunder quakes. *Geophys J Int*. 2021;226(3):1542-1557.

   doi: 10.1093/gji/ggab135

2. Reynen A, Audet P. Supervised machine learning on a network scale: Application to seismic event classification and detection. *Geophys J Int*. 2017;210(3):1394-1409.

   doi: 10.1093/gji/ggx238

3. Abdalzaher MS, Krichen M, Moustafa SS, Alswailim M. Using Machine Learning for Earthquakes and Quarry Blasts Discrimination. In*: 20th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)*; 2023. p. 1-6.

   doi: 10.1109/AICCSA59173.2023.10479234

4. Zheng B, Huang H, Wang T, Feng Z, Wei C. Real-Time Classification Study of Natural Earthquake and Artificial Blasting Waveforms Based on QRCN. In: 6th *International Conference on Communications, Information System and Computer Engineering* (*CISCE*); 2024. p. 449-454.

   doi: 10.1109/cisce62493.2024.10653243

5. Liu J, Zheng SH, Kang Y, Chou YQ. The focal mechanism determinations of moderate-small earthquakes using the first motion and amplitude ratio of P and S wave. *Earthquake*. 2004;24(1):19-26.

6. Wang Y, Tang F, Jiang X. Seismogenic structure of the earthquake surface rupture zone along the Maisu fault. *Acta Seismol Sin*. 2024;46(5):751-766.

   doi: 10.11939/jass.20220225

7. Zhang H, Zhang R, Zhao YG. Novel approach for energy-spectrum-based probabilistic seismic hazard analysis in regions with limited strong earthquake data. *Earthq Spectra*. 2024;40(4):2692-2711.

   doi: 10.1177/87552930241263621

8. Liu W, Qian Y, Wei X, Song W, Chen X. A wavelet-domain seismic source inversion method for submarine earthquakes based on 3-D Green's functions. *IEEE Trans Geosci Remote Sens*. 2024;62:1-16.

   doi: 10.1109/TGRS.2024.3434388

9. Lin T, Cheng J, Chen Q, Cui S. Wavelet transform-based fuzzy clustering microseismic first-arrival picking method. *IEEE Access*. 2023;11:136978-136987.

   doi: 10.1109/ACCESS.2023.3338628

10. Chen C, Wang CH, Liu JY, *et al*. Identification of earthquake

signals from groundwater level records using the HHT method. *Geophys J Int*. 2010;180(3):1231-1241.

doi: 10.1111/j.1365-246X.2009.04473.x

11. Wang YH, Liang SF, Kuo TBJ, Lin YC. Software implementation of real-time EMD-based algorithm in embedded microprocessors for wearable devices. *IEEE Trans Instrum Meas*. 2024;73:1-10.

doi: 10.1109/TIM.2024.3450102

12. Chen T, Gao S, Zheng S, *et al*. EMD and VMD empowered deep learning for radio modulation recognition. *IEEE Trans Cogn Commun Netw*. 2023;9(1):43-57.

doi: 10.1109/TCCN.2022.3218694

13. Shen W, Ding H. Observation of spheroidal normal mode multiplets below 1 mHz using ensemble empirical mode decomposition. *Geophys J Int*. 2013;196(3):1631-1642.

doi: 10.1093/gji/ggt468

14. Zhou H, Jian X,Chen S, *et al*. Leakage current de-disturbance method for distribution network type surge arrester based on EEMD-SVD and low-rank RBF neural network. *IEEE Access*. 2024;12:52097-52109.

doi: 10.1109/access.2024.3387327

15. Wu J, Wang W, Shang T, Cao J. A novel series arc fault detection method based on CEEMDAN and IFAW-1DCNN. *IEEE Trans Dielectr Electr Insul*. 2024;31(2):1020-1029.

doi: 10.1109/TDEI.2024.3360222

16. Shi LF, Zhou W, Yan X, Shi Y. Novel step detection algorithm for smartphone indoor localization based on CEEMDAN-HT. *IEEE Trans Instrum Meas*. 2024;73:1-9.

doi: 10.1109/TIM.2024.3472810

17. Pradhan SK, Chakraborty B. State of health estimation of Li-ion batteries based on sample entropy and various regression techniques. *Ionics*. 2025;31(5):4209-4225.

doi: 10.1007/s11581-025-06213-4

18. Cheraghy M, Soltanpour M, Abdalla HB, Oveis AH. SVM-based factor graph design for max-SR problem of SCMA networks. *IEEE Commun Lett*. 2024;28(4):877-881.

doi: 10.1109/lcomm.2024.3366426

19. Ren K, Zou G, Zhang S, Peng S, Gong F, Liu Y. Fault identification and reliability evaluation using an SVM model based on 3-D seismic data volume. *Geophys J Int*. 2023;234(1):755-768.

doi: 10.1093/gji/ggad095

20. Zhao G, Huang H, Lu X. Discriminating Earthquakes and Explosion Events by Seismic Signals Basing on BP-Adaboost Classifier. In: *2nd IEEE International Conference on Computer Communications (ICCC)*; 2016. p. 1965-1969.

doi: 10.1109/CompComm.2016.7925045

21. Bicego M, Cicalese F, Mensi A. RatioRF: A novel measure for random forest clustering based on the Tversky's ratio model. *IEEE Trans Knowl Data Eng*. 2023;35(1):830-841.

doi: 10.1109/tkde.2021.3086147

22. Amos AJ, Lee K, Gupta TS, Malau-Aduli BS. Validating the knowledge represented by a self-organizing map with an expert-derived knowledge structure. *BMC Med Educ*. 2024;24(416):416.

doi: 10.1186/s12909-024-05352-y

23. Liu Z, Cao J, Lu Y, Zhou P, Hu J. A hierarchical clustering method of SOM based on DTW distance for variable-length seismic waveform. *IEEE Geosci Remote Sens Lett*. 2022;19:1-5.

doi: 10.1109/LGRS.2021.3105476

24. Liu Z, Cao J, Chen S, Lu Y, Tan F. Visualization analysis of seismic facies based on deep embedded SOM. *IEEE Geosci Remote Sens Lett*. 2021;18(8):1491-1495.

doi: 10.1109/lgrs.2020.3003585

25. Chen S, Liu Z, Zhou H, Wen X, Xue Y. Seismic facies visualization analysis method of SOM corrected by uniform manifold approximation and projection. *IEEE Geosci Remote Sens Lett*. 2023;20:1-5.

doi: 10.1109/LGRS.2023.3260105

26. Dhanaraj RK, Chandraprabha M. Ant lion optimization in deep neural network for forecasting the rice crop yield based on soil nutrients. *Prog Artif Intell*. 2025;14(1):101-116.

doi: 10.1007/s13748-024-00351-y

27. Altawil IA, Mahafzah KA, Almomani A. A grid connected hybrid renewable energy system for optimal energy management based on ant-lion optimization algorithm. *J Theor Appl Inf Technol*. 2023;101(1):114-122.

28. Udhayakumar R, Karmakar C, Li P, Wang X, Palaniswami M. Modified distribution entropy as a complexity measure of heart rate variability (HRV) signal. *Entropy*. 2020;22(10):1077.

doi: 10.3390/e22101077

29. Wang Z, Gong M, Li P, Gu J,Tian W. A hypervolume distribution entropy guided computation resource allocation mechanism for the multiobjective evolutionary algorithm based on decomposition. *Appl Soft Comput*. 2021;116:108297.

doi: 10.1016/j.asoc.2021.108297

30. Saini A, Rahi OP. Optimal power flow analysis including stochastic renewable energy sources using modified ant lion optimization algorithm. *Wind Eng*. 2023;47(5):947-972.

doi: 10.1177/0309524X231169295

31. Mirjalili SZ, Saremi S, Mirjalili SM. Designing evolutionary feedforward neural networks using social spider optimization algorithm. *Neural Comput Appl*. 2015;26(8):1919-1928.

doi: 10.1007/s00521-015-1847-6