OXFORD

Cerebral Cortex, 2024, **34**, bhae329
https://doi.org/10.1093/cercor/bhae329
Advance access publication date 15 August 2024
Original Article

A novel sand cat swarm optimization algorithm-based SVM for diagnosis imaging genomics in Alzheimer's disease

Luyun Wang[1,2,3], Jinhua Sheng[1,2,*], Qiao Zhang[4,5,6], Ze Yang[1,2], Yu Xin[1,2], Yan Song[4,5,6], Qian Zhang[1,2], Binbing Wang[1,2]

[1]School of Computer Science and Technology, Hangzhou Dianzi University, 1158 2nd Street, Hangzhou, Zhejiang 310018, China
[2]Key Laboratory of Intelligent Image Analysis for Sensory and Cognitive Health, Ministry of Industry and Information Technology of China, 215 6th Street, Hangzhou, Zhejiang 310018, China
[3]Hangzhou Vocational & Technical College, 68 Xueyuan Street, Hangzhou, Zhejiang 310018, China
[4]Beijing Hospital, 1 Dahua Road, Beijing 100730, China
[5]National Center of Gerontology, 1 Dahua Road, Beijing 100730, China
[6]Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, 1 Dahua Road, Beijing 100730, China

*Corresponding author: Jinhua Sheng, School of Computer Science and Technology, Hangzhou Dianzi University, 1158 2nd Street, Hangzhou, Zhejiang 310018, China; Key Laboratory of Intelligent Image Analysis for Sensory and Cognitive Health, Ministry of Industry and Information Technology of China, 215 6th Street, Hangzhou, Zhejiang 310018, China. Email: j.sheng@ieee.org

In recent years, brain imaging genomics has advanced significantly in revealing underlying pathological mechanisms of Alzheimer's disease (AD) and providing early diagnosis. In this paper, we present a framework for diagnosing AD that integrates magnetic resonance imaging (fMRI) genetic preprocessing, feature selection, and a support vector machine (SVM) model. In particular, a novel sand cat swarm optimization (SCSO) algorithm, named SS-SCSO, which integrates the spiral search strategy and alert mechanism from the sparrow search algorithm, is proposed to optimize the SVM parameters. The optimization efficacy of the SS-SCSO algorithm is evaluated using CEC2017 benchmark functions, with results compared with other metaheuristic algorithms (MAs). The proposed SS-SCSO-SVM framework has been effectively employed to classify different stages of cognitive impairment in Alzheimer's Disease using imaging genetic datasets from the Alzheimer's Disease Neuroimaging Initiative. It has demonstrated excellent classification accuracies for four typical cases, including AD, early mild cognitive impairment, late mild cognitive impairment, and healthy control. Furthermore, experiment results indicate that the SS-SCSO-SVM algorithm has a stronger exploration capability for diagnosing AD compared to other well-established MAs and machine learning techniques.

Key words: sand cat swarm algorithm; optimization; support vector machine; imaging genetic; Alzheimer's disease.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease characterized by an insidious onset that severely affected human behavior, especially memory decline, further impairing language ability, arithmetic ability, visual spatial perception, and executive function (Zeng et al. 2018). Current treatments for AD face enormous difficulties, and new drug development efforts have repeatedly failed. Consequently, early screening, warning, and treatment are crucial for AD prevention and intervention. For a long time, researchers have found effective information on AD pathology through the relationship between brain activity and gene expression, which is reflected in neuroimaging-genetic data (Mishra et al. 2018; Du et al. 2020). Neuroimaging data can visually elucidate the surgical states of brain using radiography technology (de Vos et al. 2018; Chandra and Dervenoulas 2019), which is also the most common diagnostic index, such as structural magnetic resonance imaging (sMRI) (Lombardi et al. 2020), functional MRI (fMRI) (Ewers et al. 2021), positron emission tomography (PET) (Ossenkoppele et al. 2021), etc. Additionally, genetic materials such as DeoxyriboNucleic Acid (DNA) (Klein et al. 2021) and Ribonucleic Acid (RNA) (Lauretti et al. 2021) significantly influence the development of brain diseases by regulating gene inheritance, expression, and neurological function (Zhang et al. 2020a). Specifically, single nucleotide polymorphisms (SNPs) offer a more precise reflection of disease-related mutations (Kamboh 2023). Therefore, brain imaging genetics is a data science field focused on comprehensive analysis of neuroimaging and genetic data that can provide effective information on AD pathology and offer insights into the early detection and treatment of AD patients.

With the development of computing technology, artificial intelligence-based computer-aided diagnosis and treatment have become essential tools for disease risk prediction and diagnosis, promoting the development of intelligent medicine. Machine learning methods, in particular, have gained popularity for extracting distinctive features that enhance the classification of AD. For instance, a novel method combining dual attention convolutional neural network (CNN) and multilayer perceptron was proposed in (Qiang et al. 2023) for computer-aided AD diagnosis by integrating multi-modality data from sMRI, genetic data, and clinical data. Another study (Meng et al. 2022) introduced a novel feature construction method and a genetic

Received: June 2, 2024. **Revised:** July 14, 2024. **Accepted:** July 25, 2024

multi-kernel SVM to identify important features distinguishing AD, early mild cognitive impairment (EMCI), and HC groups. Additionally, a hypergraph-regularized multimodal learning approach using graph diffusion for joint association learning and outcome prediction was proposed in (Wang et al. 2023b), followed by a multi-kernel SVM classifier to fuse features selected from brain imaging genetic data for Alzheimer's Disease diagnosis and prediction. These studies have demonstrated the effectiveness of various algorithms for diagnosing AD using brain imaging genetic data. However, classification performance can still be further improved by fine-tuning algorithm parameters.

Currently, support vector machine (SVM) is popular supervised learning algorithm that has been successfully used to solve data classification and regression problems due to its excellent performance and low computational cost (Cervantes et al. 2020; Dai and Zhao 2020). However, the parameter selection of kernel function in SVM can lead to poor robustness of the classifier. To address this, metaheuristic algorithms (MAs) in intelligent algorithms offer a promising approach to optimize SVM parameters and maximize the effectiveness of SVM, including swarm intelligence algorithms such as ant colony algorithm (Miao et al. 2021), sparrow search algorithm (SSA) (Xue and Shen 2020), and particle swarm optimization (Gad 2022), Gray Wolf Optimization (GWO) (Dhargupta et al. 2020), Ant Lion Optimizer (Abualigah et al. 2021), etc. These algorithms are derived by heuristic principles found in nature, simulating processes such as biological evolution, animal behavior, and natural selection to search for optimal solutions (Ma et al. 2022; Niu et al. 2024). The sand cat swarm algorithm, proposed by Amir Seyyedabbasi and Farzad Kiani in 2022 (Seyyedabbasi and Kiani 2023), appears to be a suitable metaheuristic algorithm for optimizing SVM parameters. The SCSO algorithm simulates the predatory behavior of sand cats in nature, with each sand cat exhibiting sensitivity to sound frequency. The decision to attack or search for prey is contingent upon the sound frequency of the prey. During the hunting phase, the sand cat closely pursues its prey (Wu et al. 2022). However, this can lead to weak global exploration ability and fall into a local optimum in the later stage, thereby reducing the overall optimization performance of the algorithm. This paper proposes a modified sand cat swarm optimization algorithm (SS-SCSO) for the above problem. When searching for prey, the sand cat uses a spiral search strategy, enhancing the algorithm's global exploration ability. After attacking prey, a sparrow alert mechanism is added to accelerate the convergence speed of the algorithm. This combination aims to improve the overall optimization performance and robustness of the SVM classifier.

Through these two strategies, the global search capability of the SS-SCSO algorithm has been enhanced, resulting in improved convergence. In our experiment, we employ the CEC2017 benchmark functions to validate the optimization effectiveness of the SS-SCSO algorithm. We then analyze the tables, convergence curves, box charts, and Wilcoxon rank sum tests of benchmark test functions. Finally, to verify the practicability of the SS-SCSO algorithm, we utilize it to optimize the parameters of SVM, thereby proposing a novel SS-SCSO-SVM classification algorithm applied for the diagnosis of AD.

Moreover, the key contributions of this paper are summarized below.

(1) Adding spiral search strategy and alert mechanism of SSA to the SS-SCSO algorithm enhances the global search capability and improves the convergence speed.
(2) The SS-SCSO algorithm is performed on benchmark functions to evaluate its performance and compared with other six algorithms, which proves that SS-SCSO algorithm has a better optimization effect.
(3) The paper proposes a novel comprehensive framework for early diagnosis of AD, which includes data preprocessing, feature extraction, and SS-SCSO-SVM model.
(4) The proposed SS-SCSO-SVM model is successfully applied to the AD diagnosis, and experiments results show that it outperformed several popular machine learning algorithms.

The rest of this paper is organized as follows. Section 2 introduces the related work. Section 3 explains the data and system framework. The imaging-genetic preprocessing and feature selection are presented in Section 4. Section 5 presents the SS-SCSO optimized SVM (SS-SCSO-SVM) model. Section 6 explains the experimental results of the proposed SS-SCSO on benchmark functions and imaging-genetic problems related to AD. Finally, the discussion and conclusion are provided in Section 7.

Related work

Metaheuristic algorithms are the improvement of heuristic algorithms, combing the random algorithm and the local search algorithm. These algorithms primarily solve for the optimal solution by simulating natural and human intelligence (Wu et al. 2022). The essence of metaheuristic algorithm lies in balancing the exploration and development ability of algorithms. Mas have been widely used for diagnosing different types of diseases, including Parkinson's disease (Shrivastava et al. 2017), Diabetes (Manikandan 2019), AD (Zeng et al. 2018), cancer (Meenachi and Ramakrishnan 2021), among others. Consequently, using Mas to better address the optimization problem of early diagnosis of AD has become a key research direction for many scholars. The researchers in (Zhang et al. 2023) proposed an improved Harris Hawks Optimization (HHO) algorithm named SSFSHHO, which integrated the Sobol sequence and Stochastic Fractal Search (SFS) mechanisms for optimizing the paraments of the Fuzzy k-nearest neighbor algorithm. This algorithm was employed to classify AD and MCI using the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. A new switching delayed particle swarm optimization (SDPSO) algorithm was proposed by (Zeng et al. 2018) to optimize the parameters of SVM, resulting in a novel SDPSO-SVM classifier for the diagnosis of AD and MCI. The researchers in (Shankar et al. 2019) implemented a multi-classification process by selecting ideal features from the enhancement model inspired by Group Gray Wolf Optimization (GGWO) with a CNN classifier to classify AD, MCI, and HC subjects. A hybrid algorithm was recommended in (Dhakhinamoorthy et al. 2023) that combines the Whale Optimization Algorithm and GWO, named HWGO, for segmenting brain sub regions such as the gray matter, white matter, ventricle, corpus callosum, and hippocampus (HC) to detect normal and AD brains for classification with a deep learning algorithm.

The Sand Cat Swarm optimization algorithm is one of the recent successful MAs with balanced behavior in exploration and exploitation phases. However, it does not converge quickly and may not successfully find the global optimal solution, especially for complex problems, as it may require costly operations to find the best solution since it starts late in the development stage. On the one hand, the transition to the exploration phase is delayed, which may result in not finding the best answers

to some questions. On the other hand, although the random angle used attempts to provide fast and diverse exploitation, because it is random, it is possible to encounter the same or similar regions continuously or discontinuously (Kiani et al. 2023; Adegboye et al. 2024). Therefore, some researches have been conducted to improve and address the previously mentioned limitations. Li et al. (Li and Wang 2022) proposed an SCSO based on stochastic variation and elite collaboration (SE-SCS). The SE-SCSO algorithm integrated a nonlinear periodic adjustment mechanism for balancing exploration and local exploitation ability, along with pseudo-opposition and pseudo-reflection learning mechanisms to enhance optimization-seeking efficiency and global convergence capability. Wu et al. (Wu et al. 2022) proposed the modified sand cat swarm optimization algorithm (MSCSO) to improve the mobility of the sand cat and its exploration ability. The MSCSO algorithm added the different wandering strategies, including Levy flight walking strategy for attacking prey and the triangle walking strategy for searching prey. Furthermore, the MSCSO algorithm was enhanced for the global exploration capability by using lens opposition-based learning. Niu et al. (Niu et al. 2024) developed an improve sand cat swarm optimization algorithm (ISCSO) to tackle the moving target search issue effectively. The ISCSO algorithm enhanced planning efficiency by encoding unmanned aerial vehicle (UAV) search path information into motion paths, employing elite pooling and adaptive T-distribution to escape local optima and enhance variability, and seamlessly merged search and attack methods of SCSO for balanced global and local search capabilities. Kiani et al. (Kiani et al. 2023) proposed a new SCSO algorithm (PSCSO) using the perspective of comparison within the parliamentary system, both within the party and between the parties, to explore the best search space both locally and globally. Adegboye et al. (Adegboye et al. 2024) proposed a new optimizer derived from sand cat swarm optimization (DGS-SCSO), which integrated Dynamic Pinhole Imaging and Golden Sine Algorithm to enhance the optimizer's global exploration capability and improve exploitation, facilitating convergence towards optimal solutions.

Therefore, this paper proposes a new method, the SCSO with spiral search strategy and alert mechanism of SSA (SS-SCSO), to enhance performance in the early diagnosis of AD using imaging-genetic data. In summary, the method presented in this paper significantly advances the technical capabilities in the early diagnosis of AD, contributing positively to progress in this field.

Data and system framework

Data acquisition

In this paper, the imaging-genetic data of 120 subjects are collected from the ADNI database (http://adni.loni.usc.edu) (Petersen et al. 2010), whose main propose is to promote early diagnosis of AD and the development of new treatments. The detailed information of the dataset can be listed in Table 1. Informed consent was obtained from the volunteers in accordance with the institutional review board policy. All methods are carried out in accordance with relevant guidelines and regulations. All experimental protocols are approved by the institutional review board (IRB) at Hangzhou Dianzi University (IRB-2020001). Data from 24 patients with AD (average age of 72.5±7.8, 11 females), 45 patients with EMCI (average age 71.0±6.8, 25 females), 25 patients with late mild cognitive impairment (LMCI; average age 70.5 ± 7.1, 9 females), and 26 age-matched HCs (average age 75.2 ± 6.0, 12 females) from the ADNI are analyzed in this study. Each sample contains fMRI data and SNP data. The brain MR imaging data of 120 subjects are including T1w and T2w structure data, field mapping and resting state fMRI with eyes open. The individuals' genotypes are determined using the Human 610-Quad or OmniExpress Array (Illumina, Inc., San Diego, CA) (Du et al. 2020; Bi et al. 2022a).

The SS-SCSO-SVM model

To classify different stages of AD (e.g. AD, EMCI, LMCI, and HC), a novel framework consisting of fMRI images preprocessing, genetic preprocessing, feature selection and the SS-SCSO-SVM model is proposed. The diagram of the framework is shown in Fig. 1. In Fig. 1, firstly, the imaging and genetic preprocessing techniques are used to preprocess the imaging-genetic data to deal with redundant information. Secondly, the J-HCPMMP template is used to partition each brain into 360 areas, and a functional connectivity network is generated to calculate the connectivity measures of local graph measures. And then, the general linear model (GLM) (Luo et al. 2021) and the mixed linear model (MLM) (Zhang et al. 2010) are utilized to analyze the differential SNPs significantly associated with AD. Thirdly, a feature selection algorithm is employed to identify the best features. Finally, the proposed novel SS-SCSO-SVM is applied to classify the different stages of AD, including AD, EMCI, LMCI, AD.

Imaging-genetic preprocessing and feature selection

Image preprocessing

The main steps in fMRI data preprocessing include temporal layer alignment and head motion correction, spatial normalization and smoothing, delinearization and filtering, and covariate regression (Bi et al. 2022b). We use a new J-HCPMMP (Sheng et al. 2019) method to describe the cortical architecture, function, and connectivity, which can accurately identify AD and MCI patients at different stages. The J-HCPMMP is a comprehensive mapping tool that transforms non-HCP protocol ADNI data into the HCP CIFTI gray space, utilizing T1W and fMRI data while excluding T2W data. This process involves the integration of various brain data processing toolkits such as FreeSurfer, fMRIprep, CIFTIFY, and the HCP minimal pre-processing pipeline. To begin, the T1w brain structural data undergoes preprocessing using the standard surface-based stream provided by FreeSurfer 5.0. This step entails extracting the brain cortices with smoothed, mid-thickness, pial, and inflated surfaces, which are then saved as GIFTI files. Subsequently, the fMRI data and field map undergo preprocessing with fMRIprep. Following this, the cerebral cortex is registered into the CIFTI gray space, which comprises 91,282 standard gray-ordinates. These gray-ordinates include 32,492 cortical vertices per hemisphere and 26,298 individual elements in 19 subcortical tissues. Additionally, the 32,492 vertices of the cerebral cortex are divided into 180 areas per hemisphere for further analysis and interpretation. And then, the weighted and binary connectivity matrix are then created using all the fMRI data from 360 areas.

Various network measures are computed for the network of each subject. The Brain Connectivity Toolbox (BCT toolbox) (https://sites.google.com/site/bctnet/) (Wang et al. 2023a) is employed to compute nine local graph measures for 360 graph nodes in the binary network, including matrix strength, degree (D), clustering coefficient, local efficiency, betweenness centrality, eigenvector centrality, subgraph centrality, k-coreness centrality, flow coefficient, and page rank centrality. Thus, for each subject,

Table 1. The detailed information of the dataset.

Diagnosis	HC	EMCI	LMCI	AD
Number	26	45	25	24
Age(mean±SD)	75.2 ± 6.0	71.0 ± 6.8	70.5 ± 7.1	72.5 ± 7.8
Male:Female	14:12	20:25	16:9	13:11

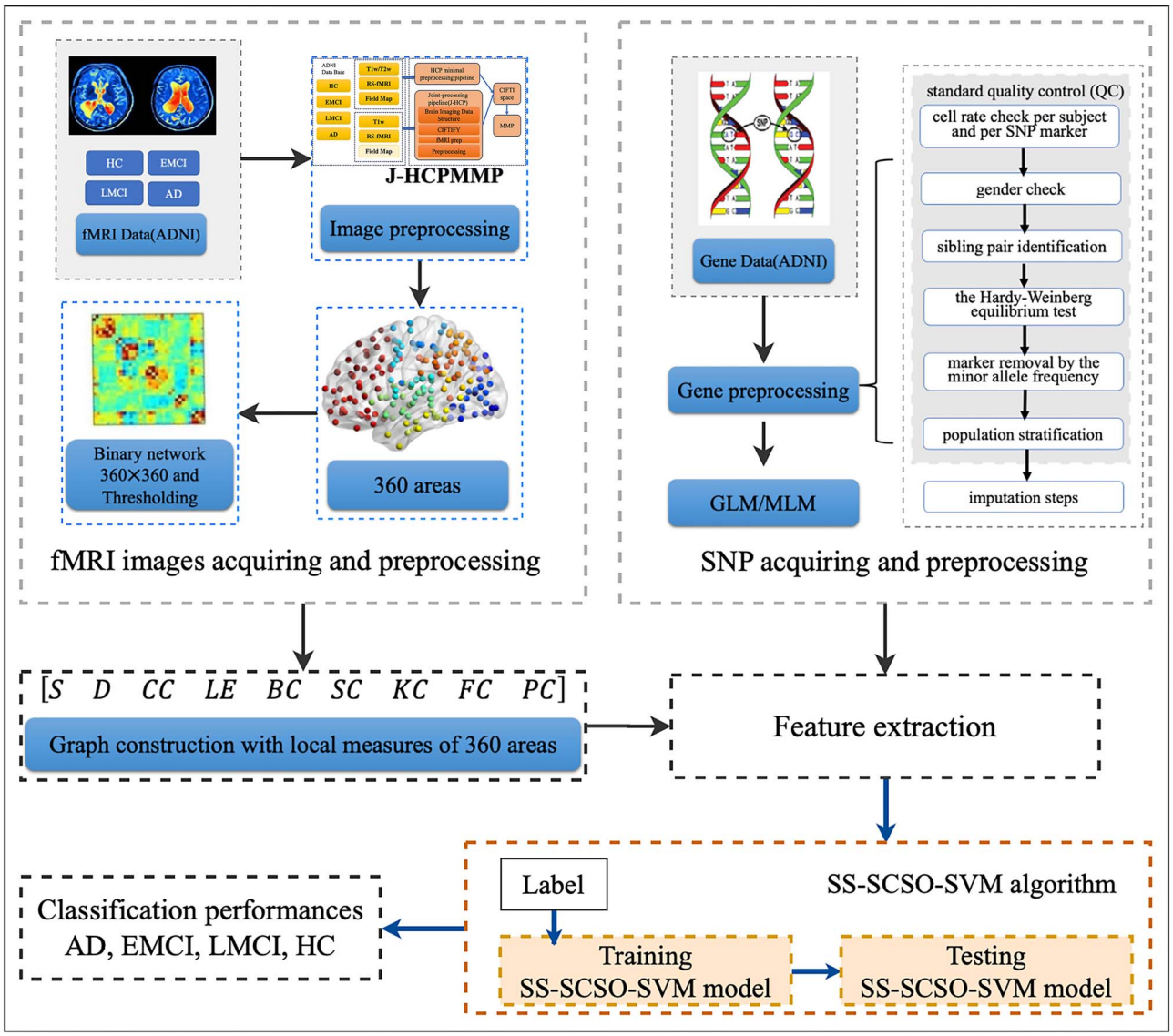

Fig. 1. The proposed framework based on the SS-SCSO-SVM algorithm.

the image features are combined to form the final feature vector comprised of 3,240 measures (360 × 9 local measures), and which are standardized to [−1, 1] prior to subsequent analysis.

Genetic preprocessing

The individuals' genotypes are determined using the Human 610-Quad or OmniExpress Array (Illumina, Inc., San Diego, CA) (Wolpin et al. 2014; Bae et al. 2023), and preprocess using the standard quality control (QC) (Anderson et al. 2010) and imputation steps. The QC criteria for the SNP data include (i) cell rate check per subject and per SNP marker, (ii) gender check, (iii) sibling pair identification, (iv) the Hardy–Weinberg equilibrium test, (v) marker removal by the minor allele frequency, and (vi) population stratification.

We use GAPIT3 software (Wang and Zhang 2021) to perform genome-wide association analysis on 120 samples in ADNI2 database. Data filtering is performed on all SNPs, the minimum allele frequency (MAF) values in the sample population detected by each SNP locus are statistically analyzed using Plink software (MAF<0.05) (Purcell et al. 2007). Subsequently, the GLM and MLM are used to analyze the differential SNPs that are significantly associated with the AD (SNPs associated with the AD) (p-value < 0.05). SNPs in GLM and MLM are intersected to acquire common SNPs. There were 2,418 SNPs significantly associated with the AD by the GLM, and a total of 2,396 SNPs significantly associated with

the AD are achieved using the MLM. Then, 2,318 common SNPs were obtained using the intersection progressing.

Feature selection

Feature selection algorithm is an important part of machine learning (Sha et al. 2023). It helps to strengthen the understanding of the relationships between features and feature values, reduce the number of features, and improve classification accuracy. As part of the feature selection process, we use the minimal redundancy maximal relevance (MRMR) algorithm (Wang et al. 2022b; Wang et al. 2023a). The MRMR algorithm (Kumar and Reddy 2021) aims to maximize the correlation between features and the target variable while minimizing the correlation between the features themselves. It evaluates the relevance of features by calculating the mutual information between each feature and the target variable. Features with high mutual information with the target variable and low mutual information with other features are selected. This approach not only ensures that the selected features are highly relevant to the classification task but also reduces redundancy among them, resulting in a more effective and discriminative subset of features. MRMR is defined as follows:

$$MRMR = \underset{s}{\mathbf{MAX}}\left\{\frac{1}{|S|}\sum_{x_i\in s} I(x_i;c) - \frac{1}{|S|^2}\sum_{x_i,x_j\in s} I\left(x_i,x_j\right)\right\} \quad (1)$$

The correlation between the features set S and the class Cis defined by the average value of all mutual information values between each feature x_i and C. Finally, we demand the feature set S with the maximum correlation-minimum redundancy.

The SS-SCSO optimized SVM model

Support vector machine

SVM is a common supervised learning algorithm mainly used for classification and regression problems (Suk and Shen 2013; Chauhan et al. 2019). Its basic idea is to map data to a high-dimensional space and establish a hyperplane. Specifically, for binary classification problems, the SVM algorithm treats data points as n-dimensional vectors and finds a hyperplane to maximize the interval between two types of data samples. The LIBSVM toolbox (http://www.csie.ntu.edu.tw/cjlin/libsvm/) (Chang and Lin 2011) is used to apply the SVM based on radial basis function (RBF-SVM) algorithm (Anyanwu et al. 2022) to classification in MATLAB. The hyperplane as given in Eq. (2).

$$g(x) = w^T \bullet x + b = 0 \quad (2)$$

where w is the normal vector, which determines the direction of the hyperplane, b determines the distance between the hyperplane and the origin. In order to find parameters w and b to maximize the interval $\gamma = \frac{2}{\|w\|}$, the objective function is converted into a dual optimization problem by introducing Lagrangian coefficients,

$$\underset{w,b}{\min}\underset{\alpha_i\geq 0}{\max} L(w,b,\alpha) = \frac{1}{2}\|w\|_2^2 - \sum_{i=1}^{n}\alpha_i\left[y_i\left(w^T x_i + b\right) - 1\right],$$
$$s.t.\ (0 \leq \alpha_i \leq c, i = 1, 2, \cdots, n) \quad (3)$$

where $L(w,b,\alpha)$ represents the Lagrangian function, α_i represents the Lagrangian coefficient, and c is a penalty factor which is the upper bound of α_i. Note that $y_i = 1$ if x_i belongs to the one class, otherwise $y_i = -1$ when x_i belongs to the other class. The final classification decision function is:

$$f(x) = \mathbf{sign}\left(g(x)\right) = \mathbf{sign}\left(w^* \bullet x + b^*\right) = \mathbf{sign}\left(\sum_{i=1}^{n}\alpha_i^* y_i K(x, x_i) + b^*\right) \quad (4)$$

where α_i^* is the optimal Lagrangian coefficient, w^* and b^* are the optimal w and b, **sign** $(\bullet)$ represents the symbolic function, $K(x, x_i)$ is the kernel function.

In this paper, SVM is used to solve two and multi classification problems, and the "one-to-rest" method (Yuan and Chu 2006) is used to transform multiple binary classification problems into a single problem. This algorithm takes any one class m of these four classes as one category, and the remaining classed as another category, to construct a two-category SVM classifier. Thus, for the case of four classes, a total of four two-class SVM classifications can be constructed device.

$$f^m(x) = \mathbf{sign}\left(g(x)\right) = \mathbf{sign}\left(w^{m*} \bullet x + b^{m*}\right)$$
$$= \mathbf{sign}\left(\sum_{i=1}^{n}\alpha_i^{m*} y_i^m K(x, x_i) + b^{m*}\right) \quad (5)$$

This paper uses the RBF kernel function (Anyanwu et al. 2022) method to convert the dot product operation in the high-dimensional space into the operation of the kernel function in the low-dimensional input space, thereby solving the problem of difficult operations in high-dimensional space and unsatisfactory results.

$$K_{RBF}(x, x_i) = e^{\left(\frac{-\|x-x_i\|^2}{2\sigma^2}\right)} = e^{-\gamma\|x-x_i\|^2} \quad (6)$$

$$\gamma = \frac{1}{2\sigma^2} \quad (7)$$

where σ is the variance and hyperparameter, γ defines the D of influence of training sample and testing sample. xand x_i represent the Euclidean distance $\|x - x_i\|^2$ between and points xand x_i. The kernel function of the RBF can be optimized by specifying the kernel parameter σ in Eq. (6). The proper selection of parameter values is crucial to the performance of SVM. The two parameters associated with the RBF kernel are the γ for the RBF kernel and the penalty parameter c for the SVM model. The parameter γ of the RBF kernel is inversely proportional to σ (Eq. (7)) and signifies the influence of a single training example. The penalty parameter c specifies the influence of the kernel on the model and guides the SVM optimization process in determining the extent to which misclassification should be minimized for each training example. This parameter is universal across all SVM kernels and balances the trade-off between misclassification of training instances and the simplicity of the decision surface (Wainer and Fonseca 2021). A low value of c serves to regularize the decision surface, while a high c aims to accurately classify all training samples. The novel of SS-SCSO-SVM algorithm is selected to optimize the parameter using the parameter optimization principle, in order to achieve the best bias variance balance and improve the best generalization performance of the RBF kernel.

The SS-SCSO algorithm

Original SCSO

Motivated by the survival behavior of sand cats in nature, the SCSO algorithm is a metaheuristic optimization algorithm development by Amir Seyyedabbasi and Farzad Kiani in 2022 (Seyyedabbasi and Kiani 2023). According to the behavior of sand cat, the algorithm target can be divided into two stages

by searching the prey operation and attacking prey operation of the sand cat. Based on this, a mechanism is proposed to achieve the balance in exploration and exploitation stages.

(1) **Initialize Population**

Like all metaheuristic optimization algorithms, the SCSO algorithm first generates a uniformly distributed random population within the range, where l_b and u_b are the upper and lower bounds of each population in each dimension, and the initialization formula is,

$$X_i = l_b + \mathbf{rand}\,(0,1) \times \left(u_b - l_b\right). \tag{8}$$

In the problem space, a single sand cat is modeled as a $1 \times \mathbf{dim}$ array, which encodes the search space, where **dim** represents the dimension. To initialize the SCSO algorithm, create an initialization matrix of size $N \times \mathbf{dim}$ using the sand cat population, where N represents population of cats, in proportional to the dimension of the problem.

$$Sand\ Cat_i = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N\times \mathbf{dim}} = \begin{bmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1\mathbf{dim}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1} & \cdots & x_{ij} & \cdots & x_{idim} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{Nj} & \cdots & x_{Ndim} \end{bmatrix}_{N\times \mathbf{dim}} \tag{9}$$

Among them, $Sand\ Cat_i$ is the population matrix of the sand cat swarm, X_i is the $i-th$ sand cat swarm, and x_{ij} is the dimension of the $i-th$ population.

The fitness of each sand cat is obtained through the fitness function of the problem to be solved. The fitness function $Fintness$ of the sand cat swarm is:

$$Fintness = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N\times 1} = \begin{bmatrix} f(x_{11}, x_{12}, \cdots, x_{1d}) \\ \vdots \\ f\left(x_{i1}, x_{i2}, \cdots, x_{id}\right) \\ \vdots \\ f\left(x_{N1}, x_{N2}, \cdots, x_{Nd}\right) \end{bmatrix}_{N\times 1} \tag{10}$$

(2) Search for Prey

The prey search mechanism of sand cats relies on low-frequency noise emission, and the expression of each sand cat is $X_i = (x_{i1}, x_{i2}, x_{i3}, \cdots, x_{idim})$. Sand cats can perceive low frequencies below 2 kHz, so in order to search for prey, it is assumed that the sand cat's sensitivity range r_G is from 0 to 2 kHz to achieve the effect of gradually approaching the prey without losing or skipping it. S_M is inspired by the auditory characteristics of sand cats and is assumed to have a value of 2. The specific formula is shown in (11):

$$r_G = S_M - \left(\frac{S_M \times t}{T}\right). \tag{11}$$

where, t is the current number of iterations, T is the maximum number of iterations. During the search step, each current search agent's position update is based on a random position. The search agent is then able to explore new spaces in the search space. In order to avoid falling into local optimum, the sensitivity range of each sand cat is different. **rand** $(0, 1)$ represents a random number from 0 to 1, that is,

$$r = \mathbf{rand}\,(0,1) \times r_G. \tag{12}$$

Each sand cat updates its position based on the optimal candidate position Pos_{bc}, current position Pos_c, and its sensitivity range r. Therefore, the sand cat can find other best prey positions, with new positions located between the current position and the prey position. At the same time, the randomness ensures the low running cost and low complexity of the algorithm.

$$Pos\,(t+1) = r \bullet \left(Pos_{bc}(t) - \mathbf{rand}\,(0,1) \bullet Pos_c(t)\right). \tag{13}$$

(3) Attack Prey

When the sand cat starts to attack its prey, it first uses the optimal position Pos_b and the current position$Pos_c$to generate a random position. The distance (Pos_{rnd}) between the sand cat and prey is shown by Eq. (14) to simulate the process of the sand cat attacking prey, that is:

$$Pos_{rnd} = \left|\mathbf{rand}\,(0,1) \bullet Pos_b(t) - Pos_c(t)\right|, \tag{14}$$

$$Pos\,(t+1) = Pos_b(t) - r \bullet Pos_{rnd} \bullet cos\theta. \tag{15}$$

At the same time, it is assumed that the sensitivity range of the sand cat is a circle, and the direction of movement can be determined by a random angle θ on the circle. Since the chosen random angle is between 0° and 360°, its value is between -1 and 1. The SCSO algorithm uses the roulette selection algorithm to select a random angle for each sand cat. The sand cat can approach the hunting position and finally capture the prey through Eq. (15) to avoid falling into the local optimum.

(4) Balancing Mechanism

The main parameter controlling the constant transition between the search phase and the predation phase is R, which allows the SCSO to switch between the two phases. When $|R| \le 1$, the sand cat's next position is any position between the current position and the hunting position, otherwise the sand cat's task is to find a new solution globally.

$$R = 2 \times \mathbf{rand}\,(0,1) - r_G \tag{16}$$

$$X\,(t+1) = \begin{cases} Pos_b(t) - r \bullet Pos_{rnd} \bullet cos\theta, |R| \le 1; \\ r \bullet \left(Pos_{bc}(t) - \mathbf{rand}\,(0,1) \bullet Pos_c(t)\right), |R| > 1. \end{cases} \tag{17}$$

The pseudo-code is shown in Algorithm 1.

The proposed method of SS-SCSO

In the SCSO algorithm, each sand cat progressively closes in on its prey, endowing the algorithm with a strong exploitation capability. However, during the later stages of the SCSO algorithm, each sand cat is susceptible to getting trapped in local optimum, hindering its ability to discover an optimal position. To enhance the mobility of the sand cat and the algorithm's exploration capability, in this paper, three operations are integrated into the SCSO algorithm to improve search capability of the optimizer, including ICMIC chaotic mapping, spiral search strategy and alert mechanism of SSA. The specific operations are presented in detail as follows.

(1) Add ICMIC Chaos Map

In the swarm intelligence optimization algorithm, the distribution state of the initial population is crucial to the convergence speed and optimization accuracy of the algorithm. In order to increase the randomness and diversity of the sand cat group,

Algorithm 1. The framework of SCSO.

Input:
T: the maximum iterations;
Initialize a population of N Sand cat;
Initialize Sand cat swarm: $Pos_i\ (i = 1, 2, \cdots, N)$;
Calculate the fitness function based on the objective function;
Initialize the r, r_G, R;
Output:
Pos_{bc}.
1. **while** $t \le T$ **do**
2. Update r, r_G, R;
3. Check if any sand cats exceed the search space;
4. Calculate the fitness value of each sand cat and find the best candidate location Pos_{bc};
5. **for** $i = 1$ **to** N **do**
6. Obtain a random angle based on the Roulette Wheel Selection $(0^\circ \le \theta \le 360^\circ)$;
7. **if** $|R| \le 1$ **then**
8. Update the sand cat's search position based on Eq. (15);
9. **else**
10. Update the sand cat's search position based on Eq. (13);
11. **end if**
12. **end for**
13. $t = t + 1$;
14. **end while**
15. return Pos_{bc}.

ICMIC chaos mapping is added, and the formula is as follows:

$$\begin{cases} z_{n+1} = \sin\left(\frac{2}{z_n}\right) \\ -1 \le z_n \le 1, z_n \ne 0 \end{cases} \tag{18}$$

(2) Search for Prey

The formula of the search prey phase of the original sand cat swarm algorithm is:

$$Pos\ (t+1) = r \bullet \left(Pos_{bc}(t) - \mathbf{rand}\ (0,1) \bullet Pos_c(t)\right) \tag{13}$$

This algorithm adds a spiral search strategy and a spiral search strategy factor in the prey search stage, that is,

$$z = e^{k \times \cos\left(\pi\left(1 - \frac{t}{t_{max}}\right)\right)} \tag{19}$$

where z is the spiral search factor. The spiral search strategy is added in the prey search stage so that the sand cat has multiple search paths to better adjust its position, thus improving the global search performance of the algorithm.

$$Pos\ (t+1) = e^{z*l} \bullet \cos\left(2\pi l\right) \bullet r \bullet \left(Pos_{bc}(t) - \mathbf{rand}\ (0,1) \bullet Pos_c(t)\right) \tag{20}$$

After incorporating spiral search, the sand cat swarm will search in the search space in a spiral form, expanding the ability to explore unknown areas, allowing the algorithm to jump out of the local optimum, and effectively improving the algorithm's global search capability.

(3) Integrated sparrow alert mechanism

After the second stage of the sand cat swarm, the sparrow alert mechanism is added, and the sand cat swarm is integrated with the sparrow alert mechanism, so that the sand cat swarm algorithm has a faster convergence speed. When a sand cat swarm realizes danger, sand cats on the edge of the swarm will quickly move to a safe area to get a better position, while sand cats in the middle of the group will move randomly to get closer to other sand cats, that is,

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta \bullet \left| X_{i,j}^t - X_{best}^t \right|, if f_i > f_g; (a) \\ X_{i,j}^t + K \bullet \left(\frac{\left| X_{i,j}^t - X_{worst}^t \right|}{(f_i - f_w) + \varepsilon} \right), if\ f_i = f_g; (b) \end{cases} \tag{21}$$

where X_{best} is the current global optimal position. β represents the step control parameter, which is a random number obeying the normal distribution with a mean of 0 and a variance of 1. $K \in [-1, 1]$ is a random number, and f_i is the fitness value of the current sand cat individual. f_g and f_w are the current global best and worst fitness values respectively. ε is a constant to avoid zero in the denominator.

If $f_i > f_g$, it indicates that this sand cat is on the edge of the population and is vulnerable to attacks from natural enemies. When $f_i = f_g$, it indicates that the sand cat located in the middle of the population is aware of danger and needs to approach other sand cats to avoid predation. K represents the direction of sparrow movement and is also a step size control parameter.

In order to address the shortcomings of traditional SCSO algorithms and improve their optimization search performance, this paper proposes an improved SS-SCSO algorithm by incorporating a spiral search strategy during the prey search phase and a fusion sparrow alert mechanism. The pseudocode and flowchart of proposed algorithm (SS-SCSO) are presented in Algorithm 2 and Fig. 2, respectively.

Computational complexity analysis

Time complexity is an important index for evaluating algorithm efficiency, reflecting the algorithm's ability to optimally solve problems. In the SCSO algorithm and SS-SCSO algorithm, it is assumed that N is the population scale, the maximum number of iterations is T, Dim refers to the dimension of the detection space. The time complexity of the SCSO algorithm is $O\left(N \times Dim \times T\right)$. In SS-SCSO, the specific definitions of each complexity are:

1) The initialization parameter time is $O(1)$.
2) Initialization population individual phase $O\left(N \times Dim\right)$.
3) Time complexity for sand cats to search and prey $O(N \times Dim \times T)$.
4) Time complexity for sand cats to combine the sparrow alert mechanism $O\left(N \times Dim \times T\right)$.

Therefore, the overall complexity of the SS-SCSO is $O(N \times Dim \times T)$. The performance of SS-SCSO is significantly improved without consuming the higher complexity.

The SS-SCSO-SVM model

A novel hybrid classifier, namely the SS-SCSO-SVM model, is proposed, which uses SS-SCSO algorithm to optimize the penalty factor c and kernel parameter γ to improve the classification performance. The classification accuracy of SVM is selected as the fitness function of SS-SCSO, thus, its formula is as follows,

$$fitness = \frac{TP + TN}{TP + FP + FN + TN}. \tag{22}$$

where the result of the model correctly predicting the positive class is true positive (TP). The result of the model correctly predicting the negative category is true negative (TN). False positive (FP) is the result of the misclassified positive class. False negative (FN) represents the prediction of the negative class that does not meet the standards (Subashini et al. 2009).

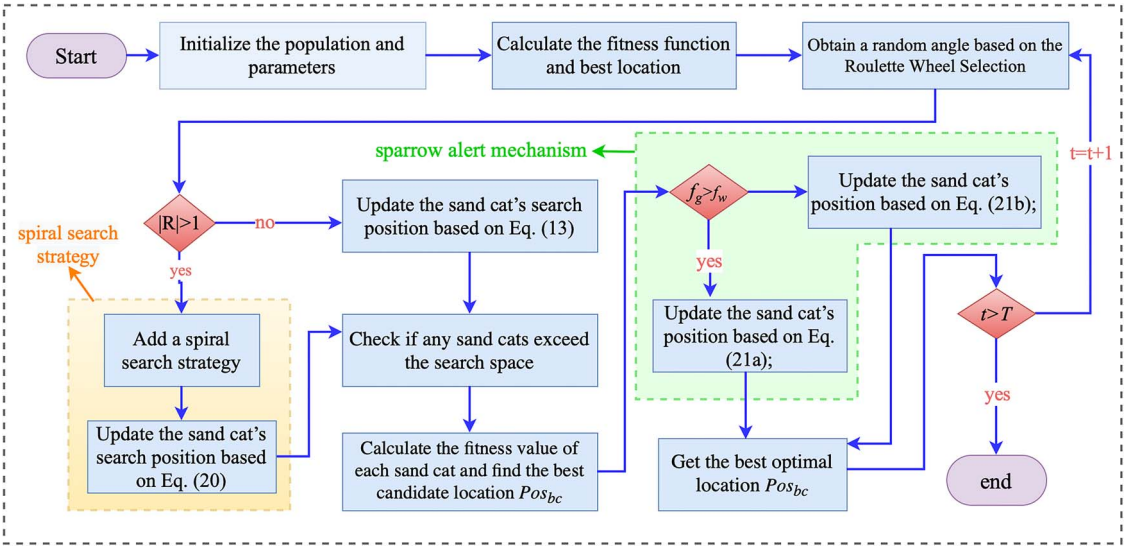

Fig. 2. The proposed framework based on the SS-SCSO-SVM algorithm.

Algorithm 2. The framework of SS-SCSO.

Input:
T: the maximum iterations;
SD: the number of the sand cat who perceive the danger;
Initialize a population of N Sand cat;
Initialize Sand cat swarm by ICMIC using Eq. (18);
Calculate the fitness function based on the objective function and best location;
Initialize the r, r_G, R;
Output:
Pos_{bc};
1. **while t ≤ T do**
2. Update r, r_G, R;
3. **for i = 1 to N do**
4. Obtain a random angle based on the Roulette Wheel Selection ($0^{\circ} \le \theta \le 360^{\circ}$);
5. **if** $|R| \le 1$ **then**
6. Update the sand cat's search position based on Eq. (13);
7. **else**
8. Add a spiral search strategy,
9. Update the sand cat's search position based on Eq. (20);
10. **end if**
11. **end for**
12. Check if any sand cats exceed the search space;
13. Calculate the fitness value of each sand cat and find the best candidate location Pos_{bc};
14. **for i = 1 to SD do**
15. **if** $f_i > f_g$ **then**
16. Update the sand cat's position based on Eq. (21a);
17. **else**
18. Update the sand cat's position based on Eq. (21b);
19. **end if**
20. **end for**
21. Get the current new location;
22. if the new location is better than before, update it;
23. $t = t + 1$;
24. **end while**
25. return Pos_{bc}.

In this study, we utilize imaging-genetic data as input features for classify different stages of AD by binary classifiers, which are evaluated in 5-fold cross-validation. We calculate six metrics for performance evaluation, including accuracy (ACC), sensitivity (SN), specificity (SP), Matthew Correlation Coefficient (MCC), precision, and F1-measure (Zhang et al. 2020b; Sheng et al. 2024). For the four-classes task, we focus on three metrics in confusion matrix analysis for performance evaluation, including accuracy, recall, and precision. The computation of these metrics for four-classes task is illustrated in Fig. 3. Analysis of the confusion matrix provides a comprehensive performance metrics including counts of true positives, true negatives, false positives, and false negatives.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}. \tag{23}$$

$$Sensitivity = Recall = \frac{TP}{TP + FN}. \tag{24}$$

$$Specificity = \frac{TN}{FP + TN}. \tag{25}$$

$$Matthew\ Correlation\ Coefficient = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \tag{26}$$

$$Precision = \frac{TP}{TP + FP}. \tag{27}$$

$$F1 - measure = \frac{2TP}{2TP + FP + FN}. \tag{28}$$

where TP/TN is the number of positive and negative samples correctly predicted, and FP/FN is the number of positive and negative samples incorrectly predicted (Zhang et al. 2020b). We choose MCC as the primary measure because it is a balanced measure even if class sizes vary widely.

Experiment results

All the experiments in this paper are completed on the computer with M2 Max processor with 32GB memory and an operating system of macOS 13.5 using matlab2022a.

The experimental results of the proposed SS-SCSO on benchmark functions

To thoroughly assess the performance of the SS-SCSO algorithm, this paper uses the CEC2017 benchmark functions to verify

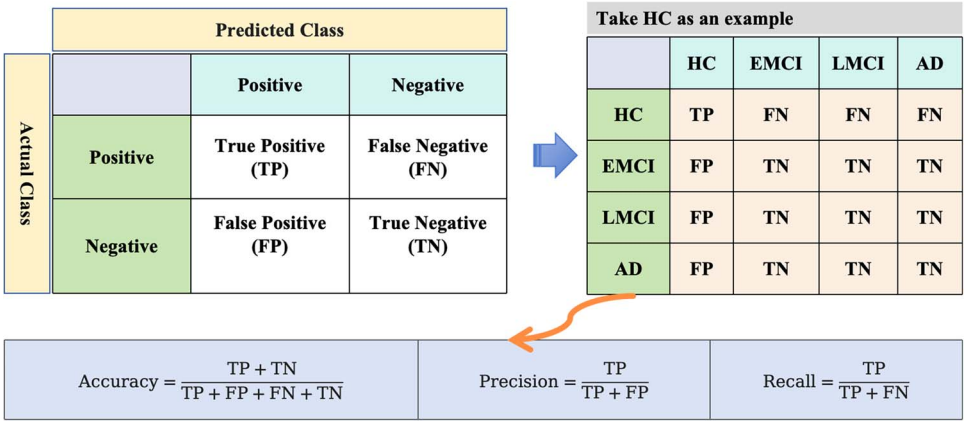

Fig. 3. Calculation of evaluation metrics for four-classes task.

Table 2. The parameter settings for the comparative algorithms.

Algorithms	Parameter values
DBO	$k \in (0, 0.02]$; $b \in (0, 1)$; $\alpha \in [-1, 1]$;
SSA	$v_0 = 0$; $\alpha \in (0, 1]$; $R_2 \in [0, 1]$; $ST \in [0.5, 1.0]$;
HHO	$r \in (0, 1)$; $u \in [0, 1]$; $v \in [0, 1]$; $\beta \in 1.5$;
SABO	$r_{i,d} \in [0, 1]$;
SCSO	$SM = 2$; Roulette Wheel selection $\in [0, 360]$;
MSCSO	$C = 0.35$; $SM = 2$; $\beta \in [0, 2\pi]$; Roulette Wheel selection $\in [0, 360]$;
SS-SCSO	$SM = 2$; Roulette Wheel selection $\in [0, 360]$;

the performance of the SS-SCSO algorithm. To better show the optimization effect, SS-SCSO is also compared with other up-to-the-minute optimizers, such as dung beetle optimizer algorithm (DBO) (Xue and Shen 2023), SSA (Xue and Shen 2020), SCSO (Seyyedabbasi and Kiani 2023), HHO algorithm (Heidari et al. 2019), subtraction average based optimizer algorithm (SABO) (Trojovský and Dehghani 2023), and MSCSO (Wu et al. 2022). The crucial parameter settings of these algorithms are given in Table 2.

To comprehensively evaluate the optimization effectiveness of the SS-SCSO algorithm, we employ the CEC2017 benchmark function for testing in this section. Table S1 in Supplementary material shows the introduction of the CEC2017 benchmark functions. To better describe the optimization process and convergence speed of each algorithm, set the number of populations of each algorithm $N = 30$, the dimension $\mathbf{dim} = 30$. The swarm scale for all algorithms is 30, and the total number of iterations $T = 1000$. And then, each algorithm runs 30 times independently and calculates three indexes for the results of each algorithm on the function, namely the best, the mean, and the standard deviation. And then, convergence curve, Wilcoxon rank sum test, and box graph are analyzed in the experiment. Each index can clearly reflect the optimization ability of each algorithm.

The statistical results and six comparison algorithms for the SS-SCSO algorithm in CEC2017 benchmark functions are shown in Table S2 in Supplementary material. Based on the data in Table S2, it can be concluded that the SS-SCSO algorithm has demonstrated good performance. In F1, F4, F10, F11, F12, F14-F15, F18, F19, F25, F26, and F28-F30, the SS-SCSO algorithm can be obtained a better fitness value compared with other comparison algorithms. However, in F4, F10, F11, F19, and F26, the standard deviation is lower than that of SSA or HHO. In F17, F22, and F24, the average fitness value and standard deviation are insufficient. The average fitness value and standard deviation of F23 and F27 are superior to other algorithms. In F5, F6, F7, F9, F13, F16, F20, and F21, the DBO algorithm or SSA algorithm can obtain a better fitness value. In F3, the MSCSO algorithm can obtain a better fitness value. Generally, according to the analysis in Table S2, the addition of a spiral search strategy and a sparrow alert mechanism has improved the exploration ability of the algorithm, making the SS-SCSO algorithm have stronger optimization ability. To visualize the ranking distribution of each algorithm, we present the stacked bar charts, including best, mean and stand deviations in Fig. 4(a). It is evident that our proposed SS-SCSO consistently ranks in the top three, demonstrating performance comparable to the CEC winners. Similarly, in Fig. 4(b), we illustrate the Sankey ranking of the seven algorithms, showing that our proposed optimizer mostly maintains the top-three position across different functions. Additionally, both Fig. 4(a) and Fig. 4(b) indicate that SSA consistently ranks in the top three, demonstrating strong performance. Therefore, these results clearly establish that the developed SS-SCSO algorithm, combined with the alert mechanism of SSA, is the most competitive algorithm.

The convergence curves of the SS-SCSO algorithm and other six comparison algorithms in F1, F4, F10, F11, F12, F14, F15, F18, F19, F26, F28, and F30 are shown in Fig. 5. The horizontal direction indicates the number of iterations, with 1,000 iterations. The vertical line represents the best fitness value corresponding to the number of iterations. As can be shown, the SS-SCSO algorithm has better convergence ability in all test functions in terms of convergence value, outperforming the other comparison algorithms in general. They can jump out of the local optimum and obtain better fitness values. However, some other algorithms fall into local optima, resulting in the algorithm not converging better. Additionally, from the convergence curve, it can be seen that the SS-SCSO algorithm can find a better position and has a fast convergence speed.

Figure S1 shows a box plot of 29 independent operations for seven algorithms, which describe the data using five statistical measures: minimum, upper quartile, median, lower quartile, and maximum. It can be seen that the SS-SCSO algorithm mostly maintains its lowest point, indicating that the algorithm has achieved good results. Some box plots have very small differences because it is easy to find a good value in the function, resulting in a small variance. Based on the above analysis, the SS-SCSO algorithm has achieved good results in the CEC2017 benchmark functions. From the p-value in the Table S3 in Supplementary material,

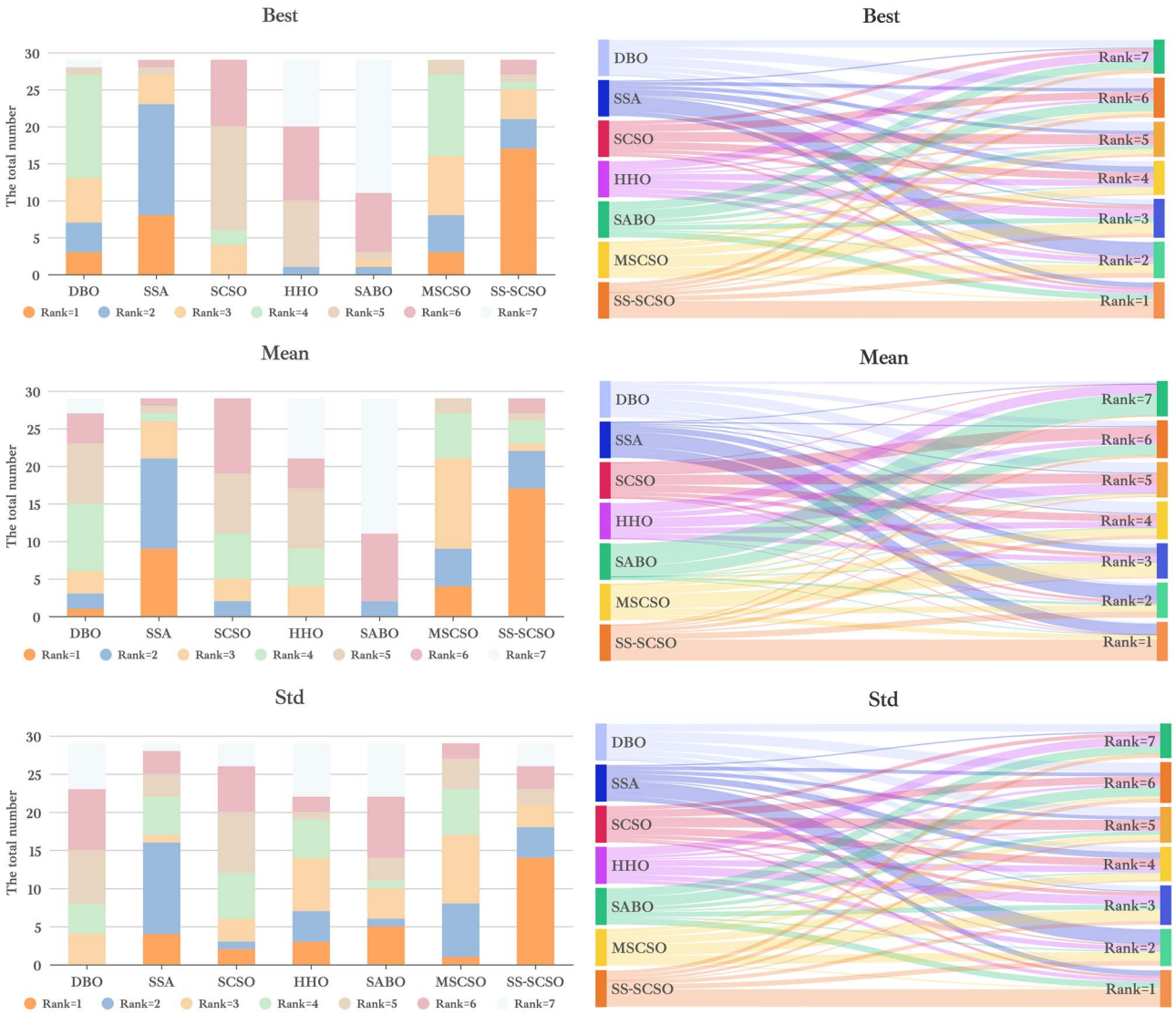

Fig. 4. (a) The ranking stacked bar chart of different algorithms on CEC2017; (b) the ranking Sankey of different algorithms on CEC2017.

it can be seen that most of them are less than 5% by the Wilcoxon rank sum test results, indicating a lower similarity between the SS-SCSO algorithm and the six comparison algorithms. However, compared to other algorithms, SSA algorithm and MSCSO algorithm have more functions with partial values greater than 5%, this means that fitness values obtained in these functions of these two algorithms are not significantly different from those of the SS-SCSO algorithm. Additionally, some other functions are greater than 5%, indicating that the difference between the values obtained by the SS-SCSO algorithm and the comparison algorithm is not significant, and the difference between the fitness values obtained by the SS-SCSO algorithm and the comparison algorithm is very small. Due to the majority being less than 5%, this indicates that the SS-SCSO algorithm has significant differences in most functions compared to other algorithms.

By selecting typical test functions in the CEC2017 benchmark functions, the graphical representation of SS-SCSO, SCSO, and MSCSO is comparatively analyzed, as well as the trajectory, average fitness, search history and convergence value of SS-SCSO in the first dimension (Wang et al. 2022a). Figure 6(c) shows the iterative trajectory of SS-SCSO in the first dimension, providing insights into the fluctuation state in this specific dimension. Figure 6(b) and (d) show the average fitness value and convergence value of SS-SCSO in multiple iterations respectively. Obviously, SS-SCSO exhibits initial volatility, but converges to an enhanced stability state as iterations proceed. This trend shows that the population gradually narrows its search scope and eventually disperses around the optimal solution. In Fig. 6(e), the 2D distribution of search history positions across these features is shown. As can be seen from the figure, the search positions basically surround the optimal solution. Furthermore, this phenomenon emphasizes that enhanced population diversity during the search phase facilitates a more efficient and comprehensive exploration of the search space. The results show that the overall volatility of the SS-SCSO algorithm tends to decrease, and the convergence speed of the SS-SCSO algorithm is faster than the SCSO algorithm and MSCSO algorithm, indicating that the SS-SCSO algorithm has better convergence performance than the other two algorithms and can find higher quality solutions faster.

In short, analysis of the experimental results above indicates that the introduction of ICMIC chaos mapping, spiral search strategy, and sparrow alert mechanism significantly improved the permeance of the algorithm. The ICMIC chaos mapping improves the diversity of the population and the travers ability of the

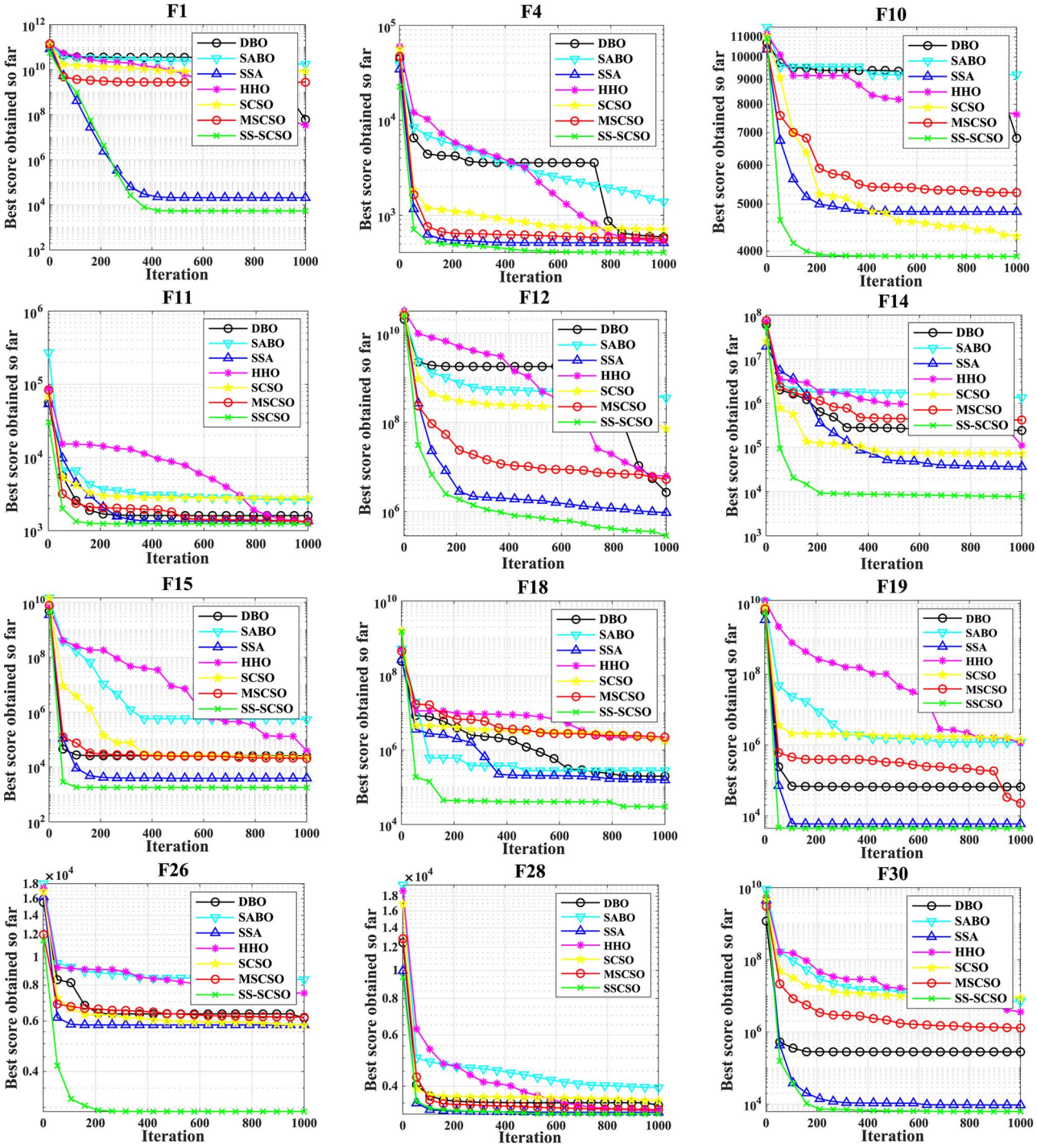

Fig. 5. Convergence curve of the benchmark function optimization algorithm on CEC 2017.

population traversal space. The spiral search strategy increases the possibility of the algorithm to escape local optimum and effectively improves the global search performance of the algorithm. Integrating the sand cat swarm with the sparrow's alert mechanism can make the sand cat swarm algorithm improve the convergence speed faster. Therefore, the proposed method has been more successful in solving various benchmark problems than other related MAs.

The experiment results of SS-SCSO-SVM for AD

In this paper, we consider a total of six binary and one four classification tasks, namely HC vs. AD, EMCI vs. AD, LMCI vs. AD, EMCI vs. LMCI, HC vs. EMCI, HC vs. LMCI, and HC vs. EMCI vs. LMCI vs. AD. The feature selection algorithm based on three different algorithms (MRMR, sparse linear regression feature selection algorithms based on stationary selection (SS-LR) (Zhang et al. 2021), and Fisher Score (FS) (Liao et al. 2023)) identifies the top 60 features, comprising the top 30 fMRI features and the top 30 SNP features. These selected features are then used with the SS-SCSO-SVM classifier to achieve the optimal classification accuracy. The K-fold (K = 5) cross validation (KCV) method (Elmaz et al. 2020) is employed to evaluate the performance of the SS-SCSO-SVM model for all classification tasks in unbiased manner. In each fold of KCV, 80% of data are selected for training the

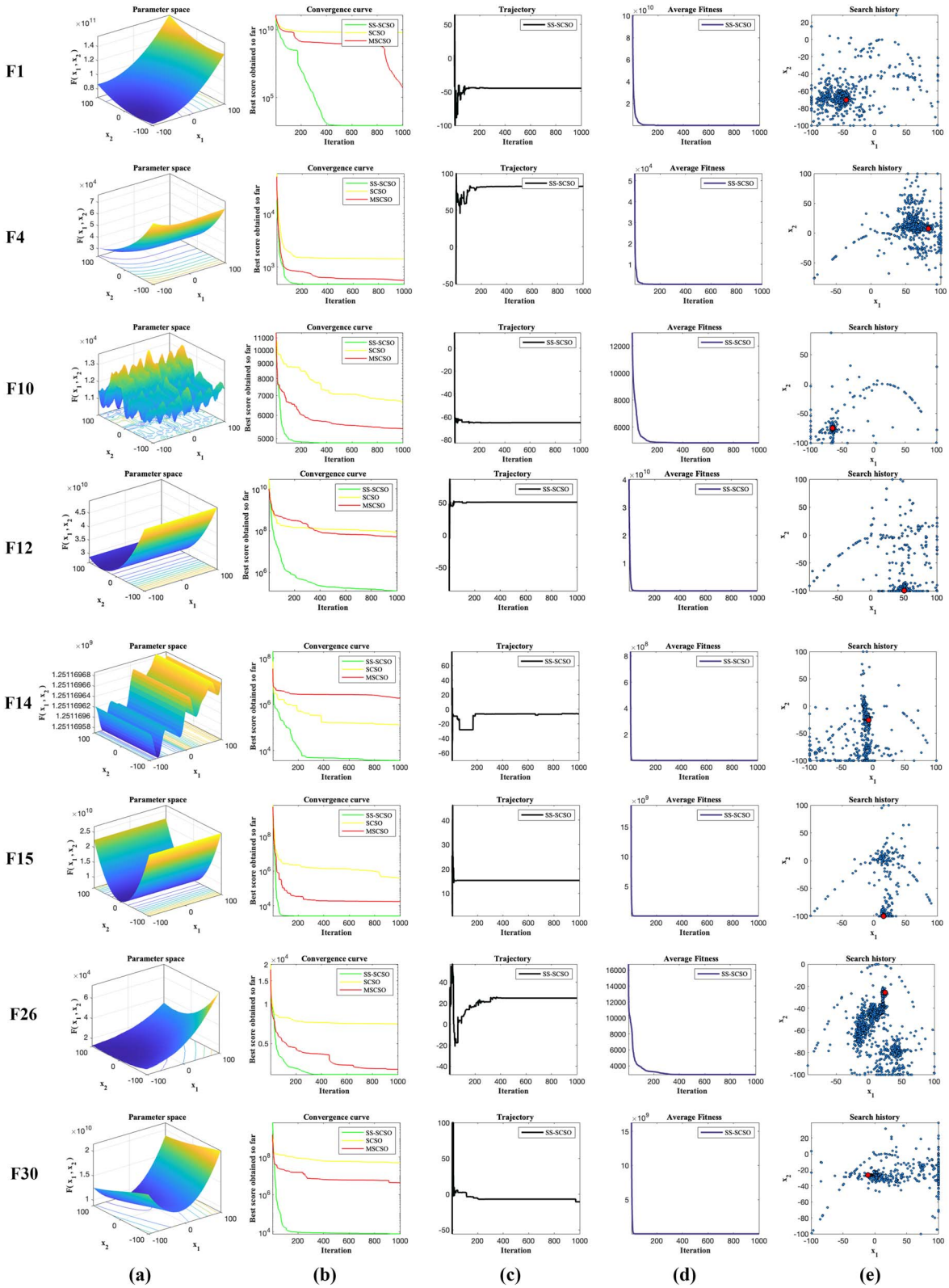

Fig. 6. (a) Visual representations of benchmark mathematical functions, (b) comparison convergence curves among SS-SCSO, SCSO and MSCSO, (c) SS-SCSO's trajectory, (d) SS-SCSO's average fitness, (e) SS-SCSO's search history.

Table 3. Performance of SS-SCSO-SVM for classify with six different classification tasks.

Tasks	ACC	SN	SP	MCC	precision	F1-measure
HC vs. AD	0.984	0.977	0.970	0.972	0.970	0.965
EMCI vs. AD	0.957	0.914	0.920	0.903	0.951	0.929
LEMCI vs. AD	0.927	0.901	0.882	0.875	0.899	0.899
EMCI vs. LMCI	0.910	0.865	0.909	0.861	0.889	0.865
HC vs. EMCI	0.918	0.875	0.903	0.904	0.916	0.902
HC vs. LMCI	0.945	0.929	0.935	0.933	0.932	0.935
HC vs. EMCI vs. LMCI vs. AD	0.720	0.702	0.687	0.614	0.720	0.643

Table 4. Classification of results and performance of different feature selection algorithms using SS-SCSO-SVM.

Feature selection	HC vs. AD	EMCI vs. AD	LEMCI vs. AD	EMCI vs. LMCI	HC vs. EMCI	HC vs. LMCI	HC vs. EMCI vs. LMCI vs. AD
MRMR	**0.984**	**0.957**	**0.927**	**0.910**	**0.918**	**0.945**	**0.720**
SS-LR	0.963	0.938	0.896	0.887	0.908	0.927	0.678
FS	0.946	0.933	0.913	0.876	0.901	0.897	0.562

Table 5. The classification results of SS-SCSO-SVM compared with SCSO-SVM and MSCSO-SVM.

Model	HC vs. AD	EMCI vs. AD	LEMCI vs. AD	EMCI vs. LMCI	HC vs. EMCI	HC vs. LMCI
SS-SCSO-SVM	**0.984**	**0.957**	**0.927**	**0.910**	**0.918**	**0.945**
MSCSO-SVM	0.958	0.919	0.877	0.856	0.882	0.930
SCSO-SVM	0.934	0.862	0.802	0.839	0.828	0.911

model, while remaining 20% are selected for calculating accuracy. Each experiment is repeated 10 times to mitigate random effects.

Six evaluation metrics, including ACC, sensitivity, specificity, MCC, precision and F1-measure, of the SS-SCSO-SVM is illustrated in Table 3. The results show that the proposed SS-SCSO-SVM framework achieves classification accuracies of 0.984, 0.957, 0.927, 0.910, 0.918, 0.945, and 0.720 for HC vs. AD, EMCI vs. AD, LMCI vs. AD, EMCI vs. AD, HC vs. EMCI, HC vs. LMCI, and HC vs. EMCI vs. LMCI vs. AD, respectively. As shown in Table 4, the MRMR algorithm outperforms the SS-LR and FS algorithms in terms of the classification performance.

In this study, the SS-SCSO-SVM model is adopted as the classifier. To validate the performance of SS-SCSO-SVM, we also tested the MSCSO and SCSO algorithms for optimizing the SVM model. The same procedures used for SS-SCSO-SVM are applied to the MSCSO-SVM and SCSO-SVM model with six different classification tasks, with six metrics for performance evaluation. Six evaluation metrics are used to make comparisons among SS-SCSO-SVM, MSCSO-SVM, and SCSO-SVM. The purpose of this study is to determine whether SS-SCSO-SVM exhibits superior classification performance compared to other SCSO swarm intelligence optimization algorithms. The comparison results for six different classification tasks are presented in Table 5 and Fig. 7. As shown, SS-SCSO-SVM outperforms other SCSO swarm intelligence optimization algorithms in all six different classification tasks for AD diagnosis, demonstrating significant advantages in terms of the six-evaluation metrics. Therefore, the effective of the SCSO and MSCSO algorithms for the SVM parameter optimization is compared, and the optimization effect of SS-SCSO on predicting model parameters is verified. Additionally, the classification performance of SS-SCSO-SVM is found to be relatively more stable.

To conduct a thorough and comprehensive analysis to verify the predictive performance of the proposed model, we compare six machine leaning models with the SS-SCSO-SVM, including SVM (Chauhan et al. 2019), Linear Discriminant Analysis (LDA) (Bandos et al. 2009), CNN (Li et al. 2018), AdaBoost (Wang and Sun 2021) and Decision tree (Mishra et al. 2020) algorithms. The results obtained in Table 6 indicate that the SS-SCSO-SVM outperforms all other comparative algorithms in all binary and four-class classification tasks.

Next, the predictive capabilities of SS-SCSO-SVM are compared with other algorithms using confusion matrix analysis. Confusion matrices visually display misclassifications across different classes, highlighting the algorithm's performance for each class. Figure 8 presents confusion matrixes for SS-SCSO-SVM algorithm and five other algorithms, which are obtained by training and testing with test data. The classification results across all four classes underscore the superior performance of the SS-SCSO-SVM model compare to the others.

By analyzing the imaging genomic data obtained from the ADNI database, the proposed SS-SCSO-based SVM algorithm demonstrates superior performance, outperforming several SCSO-based SVM algorithms and other machine learning techniques in diagnosing AD. As shown Table 5 and Table 6, the classification accuracy of the SVM algorithm optimized by various SCSO methods is better than the traditional machine learning algorithms. This suggests that ensemble swarm intelligent optimization algorithm can further improve the classification accuracy of traditional classifiers. Therefore, combing the proposed SS-SCSO model with other machine learning techniques may yield promising results. For example, the proposed method or other advanced swarm intelligence optimization algorithms could be used to optimize the classifiers within deep learning frameworks, potentially improving their performance.

Discussion and conclusion

The SCSO algorithm is a recently proposed swarm intelligence optimization algorithm that mimics the hunting behavior of

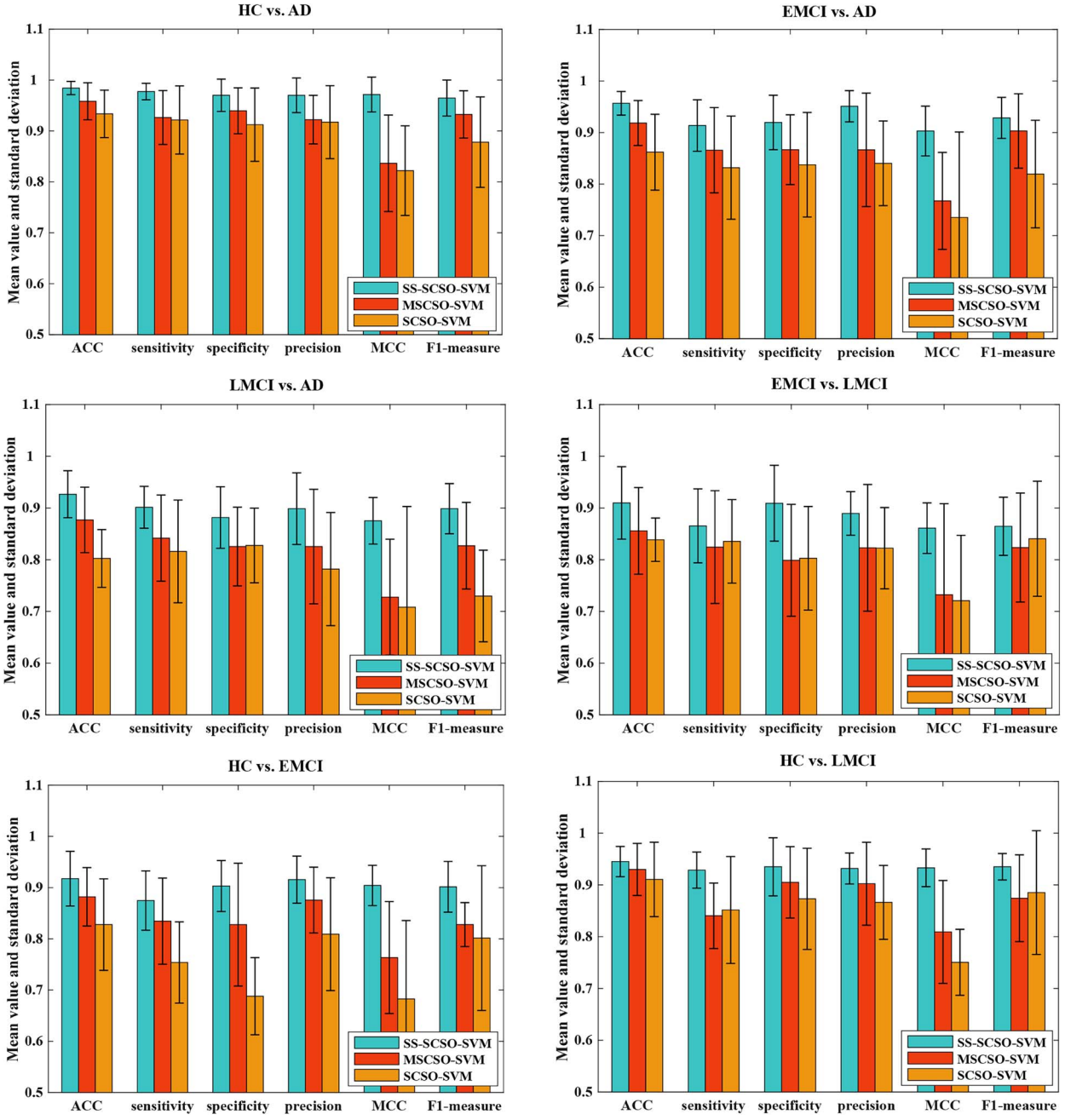

Fig. 7. Results of SS-SCSO-SVM compared with MSCSO-SVM and SCSO-SVM models for six different classification tasks in six metrics for performance evaluation.

Table 6. The classification results of SS-SCSO-SVM compared with other algorithms for two classify of six different classification tasks and four classify.

	Task	SS-SCSO-SVM	SVM	LDA	CNN	AdaBoost	Decision tree
Two classify	HC vs. AD	**0.984**	0.935	0.895	0.924	0.867	0.924
	EMCI vs. AD	**0.957**	0.824	0.764	0.835	0.793	0.809
	LEMCI vs. AD	**0.927**	0.809	0.782	0.726	0.762	0.745
	EMCI vs. LMCI	**0.910**	0.823	0.703	0.780	0.736	0.794
	HC vs. EMCI	**0.918**	0.738	0.828	0.713	0.705	0.773
	HC vs. LMCI	**0.945**	0.823	0.805	0.852	0.748	0.832
Four classify	HC vs. EMCI vs. LMCI vs. AD	**0.720**	0.667	0.608	0.625	0.583	0.615

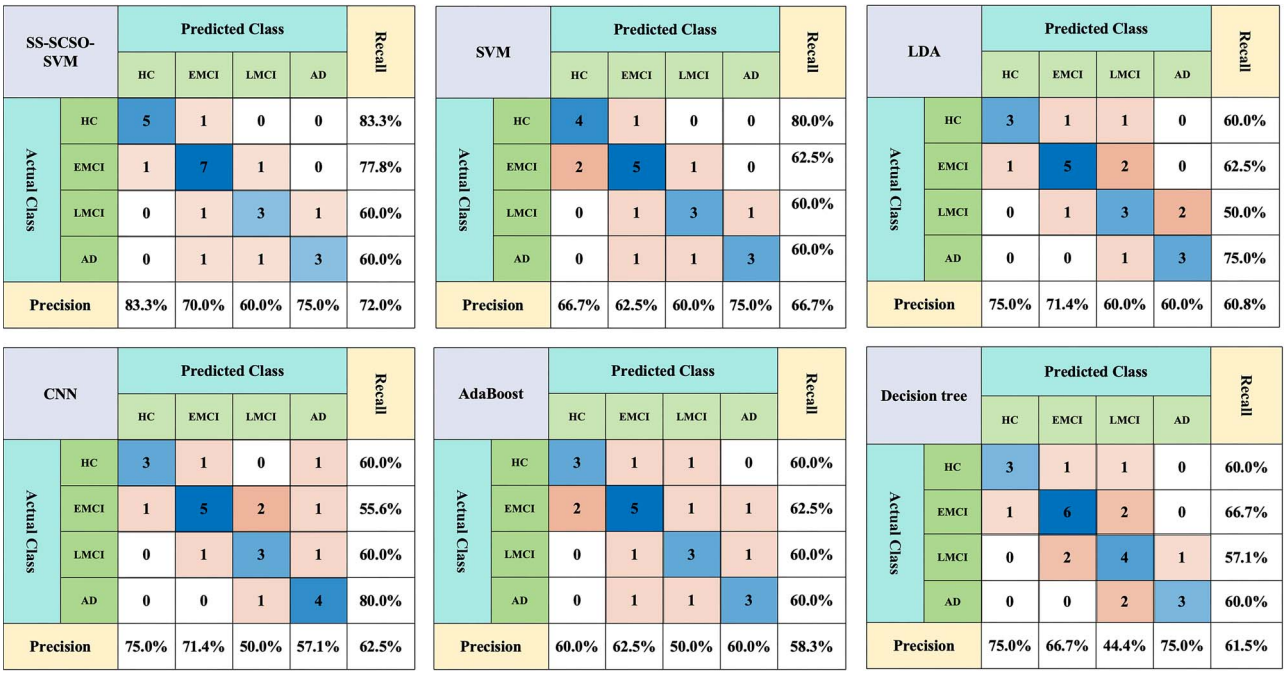

Fig. 8. Confusion matrix results of different models for four-classes task.

sand cats. While effective, the SCSO algorithm struggles with exploration capabilities in its later stages, often falling into local optima and facing difficulties in convergence. To address these problems, an improved SS-SCSO is proposed, which incorporates the spiral search mechanism and the sparrow alert mechanism, enhancing the algorithm's global search ability and accelerating convergence speed. The performance of SS-SCSO is evaluated on the CEC2017 benchmark functions, and it achieve optimal solutions for most functions compared to other classis MAs. The imaging-genetic data used in our study is relatively small, and SVM has been shown to perform well on small-scale datasets, achieving good classification performance on limited datasets. And then SVM performs superiorly when dealing with high-dimensional data, especially through the use of the RBF kernels functions, which can handle complex nonlinear relationships. Furthermore, this paper proposes a new framework based on the SS-SCSO-SVM model for AD diagnosis using imaging-genetic data, including data preprocessing, feature selection and classification.

Analyzing imaging-genetic data from the ADNI database, the SS-SCSO-based SVM algorithm shows superior performance, surpassing the MSCSO-SVM, SCSO-SVM models, and several advanced machine learning algorithms in diagnosing AD. The proposed SS-SCSO-SVM model achieves excellent classification accuracies of 98.4% for HC vs. AD, 95.7% for EMCI vs. AD, 92.7% for LMCI vs. AD, 91.0% for EMCI vs. LMCI, 91.8% for HC vs. EMCI, 94.5% for HC vs. LMCI, and 72.0% for HC vs. EMCI vs. LMCI vs. AD. It is worth mentioning that compared with other studies, the AD diagnostic framework proposed in this paper uses brain imaging-genetic data (fMRI and SNP data) as inputs, and each subject's category labels include AD, EMCI, LMCI, and HC, indicating that a comprehensive range of information is considered.

In the future, we aim to expand the sample size and include different imaging modalities to encompass more patients and healthy controls, thereby better capturing the impact of AD on imaging genetics. we aim to expand the sample size and different imaging modalities to include more patients and healthy controls to more accurately capture the impact of AD imaging genes. Additionally, we plan to explore the combination of the SS-SCSO algorithm and deep learning techniques for imaging-genetic data to further enhance AD diagnosis.

Acknowledgments

This work was supported by the National Natural Science Foundations of China (No. 62271177), Key Program of the Natural Science Foundation of Zhejiang Province (No. LZ24F010007), and the Natural Science Foundation of Zhejiang Province (No. LTGY23F020004).

Author contributions

Luyun Wang (Conceptualization, Data curation, Formal analysis, Methodology, Resources, Writing—original draft), Jinhua Sheng (Funding acquisition, Investigation, Supervision, Writing—review & editing), Qiao Zhang (Conceptualization, Validation), Ze Yang (Data curation, Formal analysis), Yu Xin (Data curation, Software), Yan Song (Validation), Qian Zhang (Data curation, Formal analysis), Binbin Wang (Data curation, Formal analysis).

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

Jinhua Sheng is supported by the National Natural Science Foundations of China (No. 62271177), and the Key Program of the Natural Science Foundation of Zhejiang Province (No. LZ24F010007). Luyun Wang is supported by the Natural Science Foundation of Zhejiang Province (No. LTGY23F020004).

Conflict of interest statement: None declared.

Data availability

Data used in preparation of this article are obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). MATLAB codes are available from GitHub (https://github.com/Soft1319/Algorithm.git).

Ethical statement

This study was approved by the institutional review board (IRB) at Hangzhou Dianzi University (IRB-2020001) and the ethics committee at Beijing Hospital (2022BJYYEC-375-01).

References

Abualigah L, Shehab M, Alshinwan M, Mirjalili S, Elaziz MA. Ant lion optimizer: a comprehensive survey of its variants and applications. *Archives of computational methods in engineering*. 2021:28(3):1397–1416. ISSN 1134-3060. https://doi.org/10.1007/s11831-020-09420-6.

Adegboye OR, Feda AK, Ojekemi OR, Agyekum EB, Khan B, Kamel S. DGS-SCSO: enhancing sand cat swarm optimization with dynamic pinhole imaging and golden sine algorithm for improved numerical optimization performance. *Sci Rep*. 2024:14(1):1491. ISSN 2045-2322. https://doi.org/10.1038/s41598-023-50910-x.

Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. *Nat Protoc*. 2010:5(9):1564–1573. ISSN 1754-2189. https://doi.org/10.1038/nprot.2010.116.

Anyanwu GO, Nwakanma CI, Lee J-M, Kim D-S. Optimization of RBF-SVM kernel using grid search algorithm for DDoS attack detection in SDN-based VANET. *IEEE Internet Things J*. 2022:10(10):8477–8490. ISSN 2327-4662. https://doi.org/10.1109/JIOT.2022.3199712.

Bae J, Logan PE, Acri DJ, Bharthur A, Nho K, Saykin AJ, Risacher SL, Nudelman K, Polsinelli AJ, Pentchev V, et al. A simulative deep learning model of SNP interactions on chromosome 19 for predicting Alzheimer's disease risk and rates of disease progression. *Alzheimers Dement*. 2023:19(12):5690–5699. ISSN 1552-5260. https://doi.org/10.1002/alz.13319.

Bandos TV, Bruzzone L, Camps-Valls G. Classification of hyperspectral images with regularized linear discriminant analysis. *IEEE Trans Geosci Remote Sens*. 2009:47(3):862–873. ISSN 0196-2892. https://doi.org/10.1109/TGRS.2008.2005729.

Bi X-a, Mao Y, Luo S, Hao W, Zhang L, Luo X, Luyun X. A novel generation adversarial network framework with characteristics aggregation and diffusion for brain disease classification and feature selection. *Brief Bioinform*. 2022a:23(6):bbac454. ISSN 1467-5463. https://doi.org/10.1093/bib/bbac454.

Bi X-a, Zhou W, Luo S, Mao Y, Xi H, Zeng B, Luyun X. Feature aggregation graph convolutional network based on imaging genetic data for diagnosis and pathogeny identification of Alzheimer's disease. *Brief Bioinform*. 2022b:23(3):bbac137. ISSN 1467-5463. https://doi.org/10.1093/bib/bbac137.

Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A. A comprehensive survey on support vector machine classification: applications, challenges and trends. *Neurocomputing*. 2020:408:189–215. ISSN 0925-2312. https://doi.org/10.1016/j.neucom.2019.10.118.

Chandra A, Dervenoulas G. Marios Politis, and Alzheimer's Disease Neuroimaging Initiative. Magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment. *J Neurol*. 2019:266(6):1293–1302. ISSN 0340-5354. https://doi.org/10.1007/s00415-018-9016-3.

Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*. 2011:2(3):1–27. ISSN 2157-6904. https://doi.org/10.1145/1961189.1961199.

Chauhan VK, Dahiya K, Sharma A. Problem formulations and solvers in linear SVM: a review. *Artif Intell Rev*. 2019:52(2):803–855. ISSN 0269-2821. https://doi.org/10.1007/s10462-018-9614-6.

Dai Y, Zhao P. A hybrid load forecasting model based on support vector machine with intelligent methods for feature selection and parameter optimization. *Appl Energy*. 2020:279:115332. ISSN 0306-2619. https://doi.org/10.1016/j.apenergy.2020.115332.

de Vos F, Koini M, Schouten TM, Seiler S, van der Grond, Lechner A, Schmidt R, de Rooij, Rombouts SARB. A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer's disease. *NeuroImage*. 2018:167:62–72. ISSN 1053-8119. https://doi.org/10.1016/j.neuroimage.2017.11.025.

Dhakhinamoorthy C, Mani SK, Mathivanan SK, Mohan S, Jayagopal P, Mallik S, Qin H. Hybrid whale and Gray wolf deep learning optimization algorithm for prediction of Alzheimer's disease. *Mathematics*. 2023:11(5):1136. ISSN 2227-7390. https://doi.org/10.3390/math11051136.

Dhargupta S, Ghosh M, Mirjalili S, Sarkar R. Selective opposition based grey wolf optimization. *Expert Syst Appl*. 2020:151:113389. https://doi.org/10.1016/j.eswa.2020.113389.

Du L, Liu K, Yao X, Risacher SL, Han J, Saykin AJ, Guo L, Shen L. Detecting genetic associations with brain imaging phenotypes in Alzheimer's disease via a novel structured SCCA approach. *Med Image Anal*. 2020:61:101656. ISSN 1361-8415. https://doi.org/10.1016/j.media.2020.101656.

Elmaz F, Büyükçakır B, Yücel Ö, Mutlu AY. Classification of solid fuels with machine learning. *Fuel*. 2020:266:117066. ISSN 0016-2361. https://doi.org/10.1016/j.fuel.2020.117066.

Ewers M, Luan Y, Frontzkowski L, Neitzel J, Rubinski A, Dichgans M, Hassenstab J, Gordon BA, Chhatwal JP, Levin J, et al. Segregation of functional networks is associated with cognitive resilience in Alzheimer's disease. *Brain*. 2021:144(7):2176–2185. ISSN 0006-8950. https://doi.org/10.1093/brain/awab112.

Gad AG. Particle swarm optimization algorithm and its applications: a systematic review. *Archives of computational methods in engineering*. 2022:29(5):2531–2561. ISSN 1134-3060. https://doi.org/10.1007/s11831-021-09694-4.

Heidari AA, Mirjalili S, Faris H, Aljarah I, Mafarja M, Chen H. Harris hawks optimization: algorithm and applications. *Futur Gener Comput Syst*. 2019:97:849–872. ISSN 0167-739X. https://doi.org/10.1016/j.future.2019.02.028.

Kamboh MI. Genomics and functional genomics of Alzheimer's disease. *Neurotherapeutics*. 2023:19(1):152–172. ISSN 1933-7213. https://doi.org/10.1007/s13311-021-01152-0.

Kiani F, Anka FA, Erenel F. PSCSO: enhanced sand cat swarm optimization inspired by the political system to solve complex problems. *Adv Eng Softw*. 2023:178:103423. ISSN 0965-9978. https://doi.org/10.1016/j.advengsoft.2023.103423.

Klein H-U, Trumpff C, Yang H-S, Lee AJ, Picard M, Bennett DA, De Jager. Characterization of mitochondrial DNA quantity and quality in the human aged and Alzheimer's disease brain. *Mol Neurodegener*. 2021:16(1):1–17. ISSN 1750-1326. https://doi.org/10.1186/s13024-021-00495-8.

Kumar MRRBS, Reddy MR. A C4. 5 decision tree algorithm with MRMR features selection based recommendation system for tourists. *Psychology and Education*. 2021:58(1):3640–3643. ISSN 0033-3077. https://doi.org/10.17762/pae.v58i1.1352.

Lauretti E, Dabrowski K, Praticò D. The neurobiology of non-coding RNAs and Alzheimer's disease pathogenesis: pathways, mechanisms and translational opportunities. *Ageing Res Rev*. 2021:71:101425. ISSN 1568-1637. https://doi.org/10.1016/j.arr.2021.101425.

Li F, Liu M, Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks. *Comput Med Imaging Graph*. 2018:70:101–110. ISSN 0895-6111. https://doi.org/10.1016/j.compmedimag.2018.09.009.

Li Y, Wang G. Sand cat swarm optimization based on stochastic variation with elite collaboration. *IEEE Access*. 2022:10:89989–90003. ISSN 2169-3536. https://doi.org/10.1109/ACCESS.2022.3201147.

Liao D, Zhang Z-Q, Guo Z-P, Tang L-R, Yang M-H, Wang R-P, Liu X-F, Liu C-H. Disrupted topological organization of functional brain networks is associated with cognitive impairment in hypertension patients: a resting-state fMRI study. *Neuroradiology*. 2023:65(2):323–336. ISSN 0028-3940. https://doi.org/10.1007/s00234-022-03061-1.

Lombardi G, Crescioli G, Cavedo E, Lucenteforte E, Casazza G, Bellatorre A-G, Lista C, Costantino G, Frisoni G, Virgili G, et al. Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment. *Cochrane Database Syst Rev*. 2020:3(3):CD009628. ISSN 1469-493X. https://doi.org/10.1002/14651858.CD009628.pub2.

Luo L, Tang Z-z, Schoville SD, Zhu J. A comprehensive analysis comparing linear and generalized linear models in detecting adaptive SNPs. *Mol Ecol Resour*. 2021:21(3):733–744. ISSN 1755-098X. https://doi.org/10.1111/1755-0998.13298.

Ma J, Hao Z, Sun W. Enhancing sparrow search algorithm via multi-strategies for continuous optimization problems. *Inf Process Manag*. 2022:59(2):102854. ISSN 0306-4573. https://doi.org/10.1016/j.ipm.2021.102854.

Manikandan K. Diagnosis of diabetes diseases using optimized fuzzy rule set by grey wolf optimization. *Pattern Recogn Lett*. 2019:125:432–438. ISSN 0167-8655. https://doi.org/10.1016/j.patrec.2019.06.005.

Meenachi L, Ramakrishnan S. Metaheuristic search based feature selection methods for classification of cancer. *Pattern Recogn*. 2021:119:108079. ISSN 0031-3203. https://doi.org/10.1016/j.patcog.2021.108079.

Meng X, Wei Q, Meng L, Liu J, Yue W, Liu W. Feature fusion and detection in Alzheimer's disease using a novel genetic multi-kernel SVM based on MRI imaging and gene data. *Genes*. 2022:13(5):837. ISSN 2073-4425. https://doi.org/10.3390/genes13050837.

Miao C, Chen G, Yan C, Yuanyuan W. Path planning optimization of indoor mobile robot based on adaptive ant colony algorithm. *Comput Ind Eng*. 2021:156:107230. ISSN 0360-8352. https://doi.org/10.1016/j.cie.2021.107230.

Mishra S, Blazey TM, Holtzman DM, Cruchaga C, Yi S, Morris JC, Benzinger TLS, Gordon BA. Longitudinal brain imaging in preclinical Alzheimer disease: impact of APOE ε4 genotype. *Brain*. 2018:141(6):1828–1839. ISSN 0006-8950. https://doi.org/10.1093/brain/awy103.

Mishra S, Mallick PK, Tripathy HK, Bhoi AK, González-Briones A. Performance evaluation of a proposed machine learning model for chronic disease datasets using an integrated attribute evaluator and an improved decision tree classifier. *Appl Sci*. 2020:10(22):8137. ISSN 1454-5101. https://doi.org/10.3390/app10228137.

Niu Y, Yan X, Wang Y, Niu Y. An improved sand cat swarm optimization for moving target search by UAV. *Expert Syst Appl*. 2024:238:122189. ISSN 0957-4174. https://doi.org/10.1016/j.eswa.2023.122189.

Ossenkoppele R, Smith R, Mattsson-Carlgren N, Groot C, Leuzy A, Strandberg O, Palmqvist S, Olsson T, Jögi J, Stormrud E, et al. Accuracy of tau positron emission tomography as a prognostic marker in preclinical and prodromal Alzheimer disease: a head-to-head comparison against amyloid positron emission tomography and magnetic resonance imaging. *JAMA neurology*. 2021:78(8):961–971. ISSN 2168-6149. https://doi.org/10.1001/jamaneurol.2021.1858.

Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, Jack CR, Jagust WJ, Shaw LM, Toga AW, et al. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology*. 2010:74(3):201–209. ISSN 0028-3878. https://doi.org/10.1212/WNL.0b013e3181cb3e25.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker, Daly MJ. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007:81(3):559–575. ISSN 0002-9297. https://doi.org/10.1086/519795.

Qiang Y-R, Zhang S-W, Li J-N, Li Y, Zhou QY, Alzheimer's Disease Neuroimaging Initiative. Diagnosis of Alzheimer's disease by joining dual attention CNN and MLP based on structural MRIs, clinical and genetic data. *Artif Intell Med*. 2023:145:102678. ISSN 0933-3657. https://doi.org/10.1016/j.artmed.2023.102678.

Seyyedabbasi A, Kiani F. Sand cat swarm optimization: a nature-inspired algorithm to solve global optimization problems. *Eng Comput*. 2023:39(4):2627–2651. ISSN 0177-0667. https://doi.org/10.1007/s00366-022-01604-x.

Sha T, Zhang Y, Peng Y, Kong W. Semi-supervised regression with adaptive graph learning for EEG-based emotion recognition. *Math Biosci Eng*. 2023:20(6):11379–11402. ISSN 1547-1063. https://doi.org/10.3934/mbe.2023505.

Shankar K, Lakshmanaprabu SK, Khanna A, Tanwar S, Rodrigues JJPC, Roy NR. Alzheimer detection using group Grey wolf optimization based features with convolutional classifier. *Computers & Electrical Engineering*. 2019:77:230–243. ISSN 0045-7906. https://doi.org/10.1016/j.compeleceng.2019.06.001.

Sheng J, Wang B, Zhang Q, Liu Q, Ma Y, Liu W, Shao M, Chen B. A novel joint HCPMMP method for automatically classifying Alzheimer's and different stage MCI patients. *Behav Brain Res*. 2019:365:210–221. ISSN 0166-4328. https://doi.org/10.1016/j.bbr.2019.03.004.

Sheng J, Zhang Q, Zhang Q, Wang L, Ze Yang Y, Xin, and Binbing Wang. A hybrid multimodal machine learning model for detecting Alzheimer's disease. *Comput Biol Med*. 2024:170:108035. ISSN 0010-4825. https://doi.org/10.1016/j.compbiomed.2024.108035.

Shrivastava P, Shukla A, Vepakomma P, Bhansali N, Verma K. A survey of nature-inspired algorithms for feature selection to identify Parkinson's disease. *Comput Methods Prog Biomed*. 2017:139:171–179. ISSN 0169-2607. https://doi.org/10.1016/j.cmpb.2016.07.029.

Subashini TS, Ramalingam V, Palanivel S. Breast mass classification based on cytological patterns using RBFNN and SVM. *Expert Syst Appl*. 2009:36(3):5284–5290. ISSN 0957-4174. https://doi.org/10.1016/j.eswa.2008.06.127.

Suk H-I, Shen D. Deep learning-based feature representation for AD/MCI classification. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013: 16th International Conference*. Nagoya, Japan, September 22–26, 2013, Proceedings, Part II 16, 583–590. https://doi.org/10.1007/978-3-642-40763-5_72.

Trojovský P, Dehghani M. Subtraction-average-based optimizer: a new swarm-inspired metaheuristic algorithm for solving

optimization problems. *Biomimetics*. 2023:8(2):149. ISSN 2313-7673. https://doi.org/10.3390/biomimetics8020149.

Wainer J, Fonseca P. How to tune the RBF SVM hyperparameters? An empirical evaluation of 18 search algorithms. *Artif Intell Rev*. 2021:54(6):4771–4797. ISSN 0269-2821. https://doi.org/10.1007/s10462-021-10011-5.

Wang W, Sun D. The improved AdaBoost algorithms for imbalanced data classification. *Inf Sci*. 2021:563:358–374. ISSN 0020-0255. https://doi.org/10.1016/j.ins.2021.03.042.

Wang J, Zhang Z. GAPIT version 3: boosting power and accuracy for genomic association and prediction. *Genomics proteomics & bioinformatics*. 2021:19(4):629–640. ISSN 1672-0229. https://doi.org/10.1016/j.gpb.2021.08.005.

Wang M, Wang J-S, Li X-D, Zhang M, Hao W-K. Harris hawk optimization algorithm based on Cauchy distribution inverse cumulative function and tangent flight operator. *Appl Intell*. 2022a:52(10):10999–11026. ISSN 0924-669X. https://doi.org/10.1007/s10489-021-03080-0.

Wang Y, Li X, Ruiz R. Feature selection with maximal relevance and minimal supervised redundancy. *IEEE transactions on Cybernetics*. 2022b:53(2):707–717. ISSN 2168-2267. https://doi.org/10.1109/TCYB.2021.3139898.

Wang L, Sheng J, Zhang Q, Zhou R, Zhongjin Li Y, Xin, and Qian Zhang. Functional brain network measures for Alzheimer's disease classification. *IEEE. Access*. 2023a:11:111832–111845. ISSN 2169-3536. https://doi.org/10.1109/ACCESS.2023.3323250.

Wang M, Shao W, Huang S, Zhang D. Hypergraph-regularized multimodal learning by graph diffusion for imaging genetics based Alzheimer's disease diagnosis. *Med Image Anal*. 2023b:89:102883. ISSN 1361-8415. https://doi.org/10.1016/j.media.2023.102883.

Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, Arslan AA, Beane-Freeman L, Bracci PM, Buring J, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat Genet*. 2014:46(9):994–1000. ISSN 1061-4036. https://doi.org/10.1038/ng.3052.

Wu D, Rao H, Wen C, Jia H, Liu Q, Abualigah L. Modified sand cat swarm optimization algorithm for solving constrained engineering optimization problems. *Mathematics*. 2022:10(22):4350. ISSN 2227-7390. https://doi.org/10.3390/math10224350.

Xue J, Shen B. A novel swarm intelligence optimization approach: sparrow search algorithm. *Systems science & control engineering*. 2020:8(1):22–34. ISSN 2164-2583. https://doi.org/10.1080/21642583.2019.1708830.

Xue J, Shen B. Dung beetle optimizer: a new meta-heuristic algorithm for global optimization. *J Supercomput*. 2023:79(7):7305–7336. ISSN 0920-8542. https://doi.org/10.1007/s11227-022-04959-6.

Yuan S-F, Chu F-L. Support vector machines-based fault diagnosis for turbo-pump rotor. *Mech Syst Signal Process*. 2006:20(4):939–952. ISSN 0888-3270. https://doi.org/10.1016/j.ymssp.2005.09.006.

Zeng N, Qiu H, Wang Z, Liu W, Zhang H, Li Y. A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease. *Neurocomputing*. 2018:320:195–202. ISSN 0925-2312. https://doi.org/10.1016/j.neucom.2018.09.001.

Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Jianming Y, Arnett DK, Ordovas JM. Mixed linear model approach adapted for genome-wide association studies. *Nat Genet*. 2010:42(4):355–360. ISSN 1061-4036. https://doi.org/10.1038/ng.546.

Zhang Q, Sidorenko J, Couvy-Duchesne B, Marioni RE, Wright MJ, Goate AM, Marcora E, Huang K-L, Porter T, Laws SM, et al. Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture. *Nat Commun*. 2020a:11(1):4799. ISSN 2041-1723. https://doi.org/10.1038/s41467-020-18534-1.

Zhang Y-H, Jin M, Li JR, Kong XY. Identifying circulating miRNA biomarkers for early diagnosis and monitoring of lung cancer Biochimica et Biophysica Acta (BBA)-molecular basis of disease. 2020b:1866:165847 ISSN 0925-4439.

Zhang T, Liao Q, Zhang D, Zhang C, Yan J, Ngetich R, Zhang J, Jin Z, Li L. Predicting MCI to AD conversation using integrated sMRI and rs-fMRI: machine learning and graph theory approach. *Front Aging Neurosci*. 2021:13:688926. ISSN 1663-4365. https://doi.org/10.3389/fnagi.2021.688926.

Zhang Q, Sheng J, Zhang Q, Wang L, Yang Z, Xin Y. Enhanced Harris hawks optimization-based fuzzy k-nearest neighbor algorithm for diagnosis of Alzheimer's disease. *Comput Biol Med*. 2023:165:107392. ISSN 0010-4825. https://doi.org/10.1016/j.compbiomed.2023.107392.