

# 多策略改进哈里斯鹰优化算法及其在高维特征选择中的应用

邱从波

温州大学计算机与人工智能学院, 浙江 温州

收稿日期: 2026年3月26日; 录用日期: 2026年4月23日; 发布日期: 2026年4月30日

## 摘要

为解决原始哈里斯鹰优化算法(HHO)在迭代过程中易陷入局部最优、种群多样性下降、后期收敛精度不足等问题, 提出一种融合多策略的改进哈里斯鹰优化算法(IHHO)。该算法首先引入经验互益策略, 整合历史种群的有效搜索信息, 增强种群个体间的信息交互能力, 提升算法跳出局部最优的潜能; 其次设计能量自适应扰动策略, 依据猎物逃逸能量动态调整扰动比例, 实现全局广域探索与局部精准开发的动态平衡; 最后提出精英方差引导开发策略, 通过计算精英个体的逐维方差确定局部搜索步长, 解决算法后期搜索的局部振荡问题, 提升局部优化精度。通过时间复杂度分析明确了IHHO的计算成本, 同时给出了算法的完整实现步骤。为验证算法的性能, 选取CEC2020标准测试函数集, 将其与原始HHO、白鲸优化算法(BWO)、算术优化算法(AOA)等5种经典群智能优化算法进行对比实验, 设置相同的实验参数与迭代条件, 从最优值、平均值、标准差等指标进行分析, 并通过Friedman统计检验与收敛曲线验证算法的综合性能。实验结果表明, 所提出的算法在CEC2020测试函数集上的Friedman平均排名为1.43, 显著优于对比算法, 在复杂高维非线性优化问题中展现出更优的全局寻优能力、收敛速度与鲁棒性。将IHHO算法应用于高维特征选择问题, 以 $k$ -最近邻分类错误率为适应度函数, 选取8个高维数据集开展实验, 从适应度值、分类准确率与特征压缩效果等维度进行验证。结果显示, IHHO在Breast、CNS、Leukemia等多数高维数据集中取得了更低的平均适应度值与更高的分类准确率, 部分数据集分类准确率达到100%, 所选特征子集兼具良好的判别性与泛化性, 有效实现了高维特征的精准筛选与冗余特征消除。研究表明, 所提多策略改进方法能显著提升哈里斯鹰优化算法的综合性能, 并在连续函数优化与高维特征选择领域均具有良好的应用效果与推广价值。

## 关键词

哈里斯鹰优化, 群智能算法, 特征选择, 高维优化

## Multi-Strategy Improved Harris Hawks Optimization Algorithm and Its Application in High-Dimensional Feature Selection

Congbo Qiu

## Abstract

To address the problems that the original Harris Hawks Optimization (HHO) algorithm is prone to falling into a local optimum, suffers from the decline of population diversity, and has insufficient convergence accuracy in the later iteration stage, an improved Harris Hawks Optimization algorithm fused with multiple strategies (IHHO) is proposed. Firstly, an experience mutual benefit strategy is introduced into the algorithm to integrate the effective search information of historical populations, enhance the information interaction ability among population individuals, and improve the algorithm's potential to jump out of the local optimum. Secondly, an energy-adaptive perturbation strategy is designed to dynamically adjust the perturbation ratio based on the prey's escape energy, thereby achieving a dynamic balance between global wide-area exploration and local precise development. Finally, an elite variance-guided exploitation strategy is proposed, which determines the local search step size by calculating the dimension-wise variance of elite individuals, thus solving the problem of local oscillation in the later search stage of the algorithm and improving the local optimization accuracy. The computational cost of IHHO is clarified through time complexity analysis, and the complete implementation steps of the algorithm are given simultaneously. To verify the optimization performance of IHHO, the CEC2020 standard test function set is selected for comparative experiments with five classic swarm intelligence optimization algorithms including the original HHO, Beluga Whale Optimization (BWO) and Arithmetic Optimization Algorithm (AOA). The same experimental parameters and iteration conditions are set, and the performance is analyzed from the indicators such as the optimal value, mean value and standard deviation. In addition, the comprehensive performance of the algorithm is verified by the Friedman statistical test and convergence curves. The experimental results show that the Friedman average ranking of IHHO on the CEC2020 test function set is 1.43, which is significantly better than that of the comparison algorithms. IHHO exhibits superior global optimization ability, convergence speed and robustness in solving complex high-dimensional nonlinear optimization problems. The IHHO algorithm is applied to the high-dimensional feature selection problem, with the  $k$ -nearest neighbor classification error rate as the fitness function. Eight high-dimensional datasets are selected for experiments, and the algorithm is verified from the dimensions of fitness value, classification accuracy and feature compression effect. The results show that IHHO achieves lower average fitness values and higher classification accuracy in most high-dimensional datasets such as Breast, CNS and Leukemia, with the classification accuracy reaching 100% in some datasets. The selected feature subsets have both good discriminability and generalization ability, which effectively realize the accurate screening of high-dimensional features and the elimination of redundant features. The research shows that the proposed multi-strategy improvement method can significantly improve the comprehensive performance of the Harris Hawks Optimization algorithm, and IHHO has good application effects and popularization value in the fields of continuous function optimization and high-dimensional feature selection.

## Keywords

Harris Hawks Optimization, Swarm Intelligence Algorithm, Feature Selection, High-Dimensional Optimization

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

群智能优化算法作为模拟自然生物群体行为的随机搜索方法, 凭借结构简单、鲁棒性强、无需梯度信息依赖等优势, 已成为复杂非线性优化问题的重要求解工具, 广泛应用于函数优化、特征选择、路径规划等领域[1]。哈里斯鹰优化算法(HHO)通过模拟哈里斯鹰协同捕猎机制, 融合探索与开发动态切换策略, 在各类优化问题中展现出良好基础性能[2]。已有研究通过终端替换、自适应策略改进[3][4], 或结合机器学习框架[5]、Transformer 预测模型[6], 将 HHO 拓展至约束优化、轨迹规划、疾病预测等场景并取得显著成效。

然而在特征选择领域, 原始 HHO 仍存在明显局限: 种群个体信息交流单一导致多样性下降, 易陷入局部最优; 猎物逃逸能量驱动的扰动模式固定, 难以适配全局探索与局部开发的动态平衡; 后期搜索阶段优势个体易产生局部振荡, 降低收敛精度与优化效率, 这些缺陷限制了其在高维、多模态复杂场景中的应用。特征选择作为数据挖掘与机器学习的关键前置步骤, 核心是从高维原始特征中筛选强判别性子集, 以消除冗余、降低计算复杂度、提升模型泛化能力。高维特征选择本质为组合优化问题, 传统方法存在搜索效率低、解质量差等问题, 而群智能优化算法为其提供了高效求解思路, 因此设计高性能改进算法具有重要理论与实际意义。

针对原始 HHO 缺陷与高维特征选择需求, 本文提出融合多策略的改进哈里斯鹰优化算法(IHHO), 从三方面优化原始算法: 引入经验互益策略整合历史经验, 增强种群全局搜索能力; 设计能量自适应混合扰动策略, 动态调整扰动比例; 提出精英方差引导开发策略, 基于精英个体统计特征优化局部搜索步长, 解决后期振荡问题。为验证性能, 首先在 CEC2020 标准测试函数集上, 与经典群智能算法从最优值、平均值等指标进行对比, 结合 Friedman 检验与收敛曲线分析综合优势; 随后将 IHHO 应用于高维特征选择, 以  $k$ -最近邻分类错误率为适应度函数, 在 8 个生物医学数据集上验证, 从适应度值、分类准确率等维度评估实际应用效果。本文通过多策略改进提升 HHO 综合性能, 为复杂优化问题提供高效求解方法, 同时为高维特征选择提供新技术思路。

## 2. 哈里斯鹰优化算法

哈里斯鹰优化算法(Harris Hawks Optimization, HHO)是 Heidari 等人提出的一种新的群智能优化方法, 它灵感来自于哈里斯鹰群体捕猎行为的复杂过程[2]。实际捕猎中, 它会表现出侦察、围捕、追踪、攻击等各种层次的合作模式, 依靠不断改变的捕食行为一步步逼近目标。HHO 把这一自然现象用数学建模的方式表达出来。

### 2.1. 探索阶段

HHO 引入了逃逸能的概念。该算法使用该逃逸能量的值来确定进入哪个搜索阶段。其表达式如下:

$$E = 2E_0 \left( 1 - \frac{Iter}{MaxIter} \right) \quad (1)$$

其中  $E_0$  表示猎物的初始能量, 取值范围为  $[-1, 1]$ , 该值在每次迭代中自动更新。 $Iter$  表示当前迭代次数,  $MaxIter$  表示算法设计时确定的最大迭代次数。HHO 算法的更新过程可以简化为两个阶段: 探索 and 开发。当  $|E| > 1$  时, 算法在目标空间进行广泛搜索。相反, 当  $|E| < 1$  时, 算法倾向于关注局部开发。

当  $|E| \geq 1$  时, 算法为探索阶段。此时, 猎物具有很强的逃避性, 鹰群要在一个广阔的空间范围去寻找目标。为了提高种群的多样性以及全局搜索的能力, 个体位置更新方式要更加重视增强种群的探索性。该阶段的位置更新公式可表述为:

$$X_i(t+1) = \begin{cases} X_{rand}(t) - r_1 \times |X_{rand}(t) - 2r_2 \times X_i(t)|, & q < 0.5 \\ (X_{rabbit}(t) - X_m(t)) - r_3 \times (lb + r_4 \times (ub - lb)), & q \geq 0.5 \end{cases} \quad (2)$$

其中,  $X_{rand}(t)$  为当前种群中随机选取的个体,  $X_m(t)$  为种群平均位置,  $lb$  和  $ub$  分别表示搜索空间的下界和上界。  $q, r_1, r_2, r_3, r_4$  为  $[0, 1]$  的随机数。平均位置计算公式为:

$$X_m(t) = \frac{1}{N} \sum_{i=1}^N X_i(t) \quad (3)$$

其中,  $X_i(t)$  表示迭代  $t$  中第  $i$  个鹰的位置,  $N$  表示鹰的总数。

## 2.2. 过渡阶段

HHO 算法种群中每一个个体都对应着一个候选解, 而目前全局最优个体就是猎物, 用  $X_{rabbit}(t)$  来表示猎物的位置。设第  $i$  个个体在第  $t$  次迭代时的位置为  $X_i(t)$ , 算法利用猎物逃逸能量来评价当前搜索状态, 具体表达式为以下所示的公式:

$$E = 2E_0 \times \left(1 - \frac{t}{T}\right) \quad (4)$$

$$E_0 = 2r - 1 \quad (5)$$

式中,  $T$  为最大迭代次数,  $r$  为随机数。

## 2.3. 开发阶段围捕策略

当  $|E| < 1$  时, 算法进入开发阶段。此时, 猎物逃逸能力减弱, 鹰群把重点放在当前最佳解上进行细致的搜寻, 从而加快收敛速度并加强局部搜索的效果。HHO 算法根据猎物的状态和围捕方式的不同, 在开发阶段提出了不同的位置更新方法。

### 2.3.1. 软围捕

当  $|E| \geq 0.5$  且  $r \geq 0.5$  时, 猎物仍具有一定逃逸能力。这时采用软包围策略, 其位置更新公式为:

$$X_i(t+1) = \Delta X(t) - E |J \times X_{rabbit}(t) - X_i(t)| \quad (6)$$

$$\Delta X(t) = X_{rabbit}(t) - X_i(t) \quad (7)$$

$$J = 2 \times (1 - r_5) \quad (8)$$

式中,  $\Delta X(t)$  表示当前个体与猎物之间的位置差,  $J$  为猎物的随机跳跃强度,  $r_5$  为  $[0, 1]$  的随机数。

### 2.3.2. 硬围捕

当  $|E| < 0.5$  且  $r \geq 0.5$  时, 猎物体力下降、逃逸能力较弱时。这时采用硬包围策略, 其位置更新公式为:

$$X_i(t+1) = X_{rabbit}(t) - E \times |X_{rabbit}(t) - X_i(t)| \quad (9)$$

### 2.3.3. 渐进式软围捕并快速俯冲

当  $|E| \geq 0.5$  且  $r < 0.5$  时, 算法结合 Levy 飞行实现快速俯冲。如下公式所示:

$$X_i(t+1) = \begin{cases} Y, & \text{if } f(Y) < X_i(t) \\ Z, & \text{if } f(Z) < X_i(t) \end{cases} \quad (10)$$

$$Y = X_{rabbit}(t) - E \times |J \times X_{rabbit}(t) - X_i(t)| \quad (11)$$

$$Z = Y + S \times LF(d) \quad (12)$$

其中,  $S$  为随机向量,  $LF(d)$  表示  $d$  维 Levy 飞行。

### 2.3.4. 渐进式硬围捕并快速俯冲

当  $|E| < 0.5$  且  $r < 0.5$  时, 采用更强的俯冲式更新:

$$X_i(t+1) = \begin{cases} Y, & \text{if } f(Y) < X_i(t) \\ Z, & \text{if } f(Z) < X_i(t) \end{cases} \quad (13)$$

$$Y = X_{rabbit}(t) - E \times |J \times X_{rabbit}(t) - X_m(t)| \quad (14)$$

$$Z = Y + S \times LF(d) \quad (15)$$

## 3. 改进哈里斯鹰优化算法

### 3.1. 经验互益策略

本文引入经验互益策略[7], 它加强了种群对于过去搜索过程中的重要经验, 提高了个体之间的合作水平。该策略的关键之处就在于把历史种群所拥有的经验数据转化成引导向量加入到当下的个体状态更新当中, 从而提升个体跳出局部最优的潜能以及拓展解空间的探寻幅度。由于算法早期需要具有较高的随机性才能找到潜在的全局最优解, 在探索和开发阶段的过渡区间里才使用该种经验互益策略。

为了描述历史经验对当前个体的作用, 首先保存上一代种群的位置矩阵  $H$ 。对于当前个体  $X_i$ , 随机选取 3 个历史个体索引  $r_1$ 、 $r_2$  和  $r_3$ , 并逐维构造 3 个经验向量  $E_1^h$ 、 $E_2^h$  和  $E_3^h$ 。其第  $d$  维可表示为:

$$E_{1,d}^h = \begin{cases} H_{r_1,d}, & \text{rand} < p_h \\ X_{i,d}, & \text{其他} \end{cases} \quad (16)$$

$$E_{2,d}^h = \begin{cases} H_{r_2,d}, & \text{rand} < p_h \\ X_{i,d}, & \text{其他} \end{cases} \quad (17)$$

$$E_{3,d}^h = \begin{cases} H_{r_3,d}, & \text{rand} < p_h \\ X_{i,d}, & \text{其他} \end{cases} \quad (18)$$

其中,  $p_h$  为历史经验继承概率。

根据上述公式可知, 个体按照一定概率继承历史种群的的经验信息, 既可以使个体不偏离当前搜索路径, 又可以很好地将以前的经验加以整合。为了保证经验共享策略在优化中起到主导作用, 采用权重因子  $w_t$  的方式进行计算, 其计算公式为:

$$w_t = 1 - \frac{\left| 2 \times \frac{t}{T} - 1 \right|}{0.4} \quad (19)$$

其中,  $t$  为当前评估次数,  $T$  为最大评估次数。

当  $w_t < 0$  时, 令  $w_t = 0$ 。权重值在搜索初期慢慢提高, 在后期却慢慢地降低, 以此来达到经验协同策略在某些时期的有效控制目的。在此基础上, 构造经验互益候选解如下公式所示。

$$X_{emb} = X_i + w_t \times (\alpha \times R \times (E_1^h - X_i) + \beta \times (1 - R) \times (E_2^h - E_3^h)) \quad (20)$$

其中,  $R$  为  $[0, 1]$  上的随机向量,  $\alpha$  和  $\beta$  分别为经验受益项与经验互补项的权重系数。

为了使经验互益策略与传统的 HHO 算法的位置更新机制互相配合工作，需要把经验互益策略产生的候选解同 HHO 算法得到的初始候选解  $X_H$  进行融合，从而得到改进后的经验互益候选解。融合过程可以使用如下数学式子来表示。

$$X_{embs}^{final} = X_H + w_i (X_{embs} - X_i) \tag{21}$$

### 3.2. 能量自适应 Cauchy-Gaussian 扰动策略

本文提出一种能量自适应调节的 Cauchy-Gaussian 混合扰动策略。此方法用 Cauchy 分布的重尾性提高远距离随机跳跃的能力，用 Gaussian 分布的正态化平稳性达到高精度局部搜索的效果。利用实时观测到的逃逸能量变化情况，自主地对两种扰动分量所占比例予以改变，进而改善算法在全局搜索和局部改进两者之间所取得的结果表现状况。如下图 1 所示。

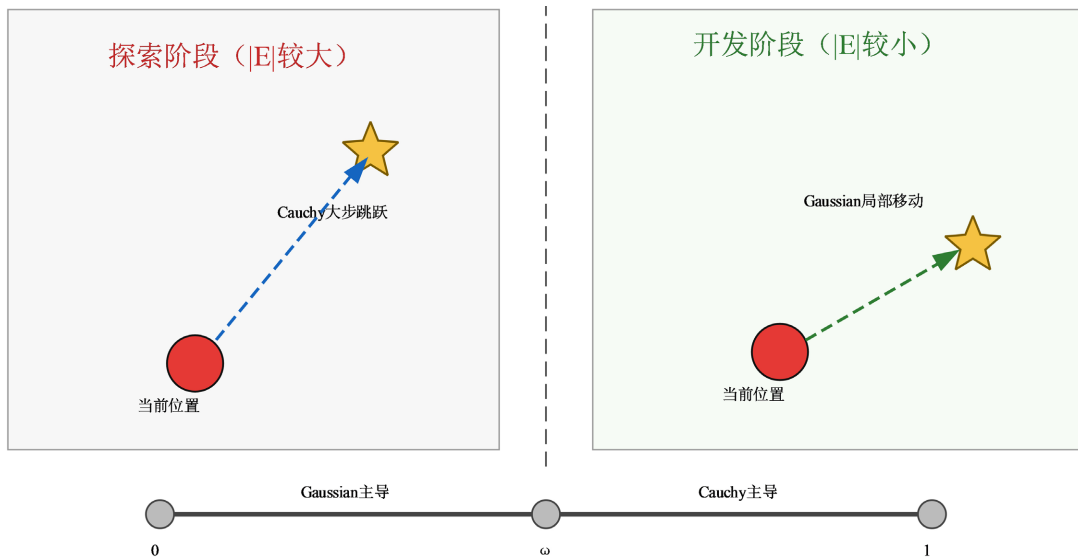


Figure 1. Energy-adaptive Cauchy-Gaussian perturbation strategy schematic diagram  
图 1. 能量自适应 Cauchy-Gaussian 扰动策略示意图

首先，根据逃逸能量构造归一化权重如下公式所示：

$$\omega = \min \left( 1, \frac{|E|}{2} \right) \tag{22}$$

其中， $E$  为当前个体对应的逃逸能量。

当  $|E|$  较大时， $\omega$  较大，说明当前个体更倾向于执行全局探索；当  $|E|$  较小时， $\omega$  较小，说明当前个体更倾向于局部开发。随后，从当前种群中随机选取一个个体  $X_r$ ，并分别生成 Cauchy 随机向量  $C$  与 Gaussian 随机向量  $G$ 。为了使 Gaussian 扰动在搜索后期逐渐减弱，其扰动尺度设置如下公式所示。

$$\sigma_t = \sigma_0 \times \left( 1 - \frac{t}{T} \right) \tag{23}$$

其中， $\sigma$  为 Gaussian 扰动的初始尺度参数。

之后，使用公式(23)构造能量自适应 Cauchy-Gaussian 扰动候选解。

$$X_{cg} = X_i + \omega \times C \times (X_i - X_r) + (1 - \omega) \times G \times (X_{rabbit} - X_i) \tag{24}$$

其中,  $X_i$  表示当前个体,  $X_{rabbit}$  表示全局最优个体。

### 3.3. 精英方差引导开发策略

本文提出了一个用逐维方差来引导的精英方差, 用部分取代传统的完整协方差矩阵的方式, 保留了精英群体的统计性质, 并减少高维空间导致的计算量。该策略示意图如图 2 所示。

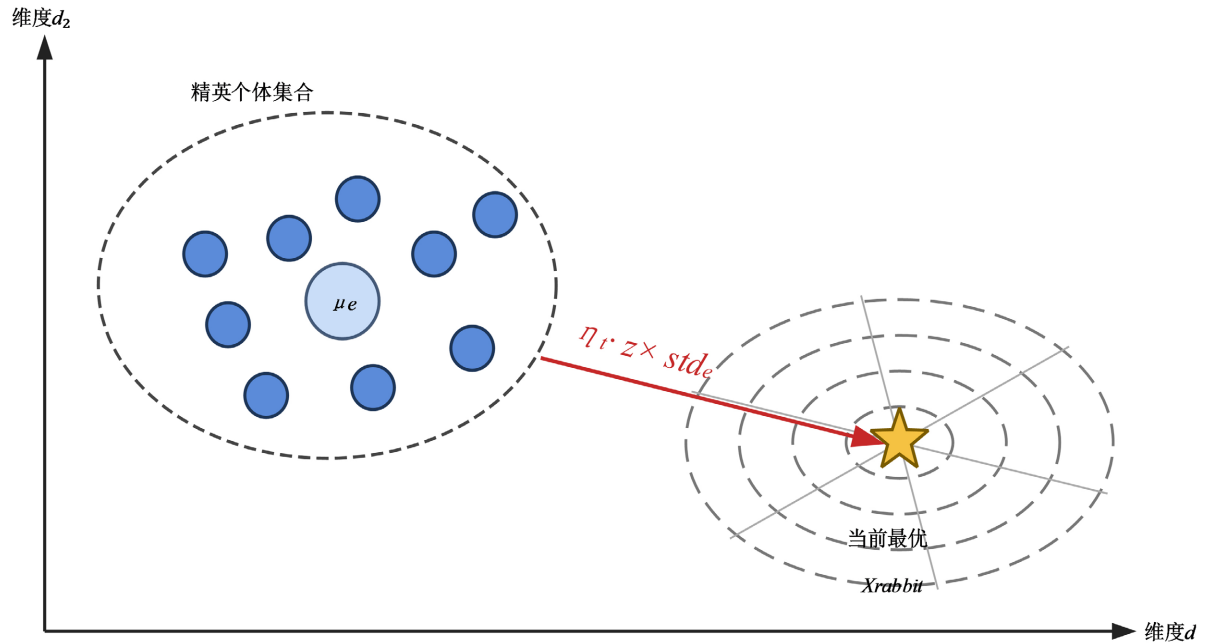


Figure 2. Elite variance-guided exploitation strategy schematic diagram  
图 2. 精英方差引导开发策略示意图

首先, 根据当前种群适应度值进行升序排序, 选取前  $M$  个个体构成精英子集, 如下公式所示。

$$M = \max(2, \lceil \rho N \rceil) \quad (25)$$

其中,  $\rho$  为精英比例,  $N$  为种群规模。设精英子集为  $\varepsilon = \{X_1, X_2, \dots, X_M\}$ , 之后, 进一步计算统计特征。对第  $d$  维而言, 精英个体的样本方差定义为:

$$\text{var}_d = \text{Var}(X_{1,d}, X_{2,d}, \dots, X_{M,d}), \quad d = 1, 2, \dots, D \quad (26)$$

之后, 得到精英子集的逐维标准差向量如下公式所示。

$$\text{std}_e = \sqrt{\text{var} + \varepsilon} \quad (27)$$

其中,  $\varepsilon$  为防止数值不稳定的微小常数。由此可见,  $\text{std}_e$  能够反映精英个体在各维度上的离散程度, 并可作为局部搜索步长的重要依据。

该策略主要作用于后期开发过程, 当搜索进程进入后期, 连续  $L$  次更新未获得更优解时, 启用精英方差引导开发策略。此时, 定义随迭代逐渐减小的步长因子为:

$$\eta_t = \eta_0 \times \left(1 - \frac{t}{T}\right) \quad (28)$$

其中,  $\eta_0$  为初始步长参数。

最后，以当前全局最优个体  $X_{rabbit}$  为中心，结合标准高斯随机向量  $z$  与精英逐维标准差向量  $std_e$ ，构造局部开发候选解。如下公式所示：

$$X_{cov} = X_{rabbit} + \eta_i \times (z \times std_e) \quad (28)$$

### 3.4. 时间复杂度和算法实现

#### 3.4.1. 时间复杂度

设种群规模为  $N$ ，问题维度为  $D$ ，最大迭代次数为  $T$ 。本文改进的哈里斯鹰优化算法(IHHO)在保留原始 HHO 搜索框架的基础上，引入经验互益策略、能量自适应扰动策略与精英方差引导开发策略，同时新增种群排序与精英子集筛选操作，其时间复杂度可逐模块细化分析。初始化阶段与 HHO 保持一致，复杂度为  $O(ND)$ ；迭代阶段中，原始位置更新复杂度仍为  $O(TD)$ 。IHHO 在每轮迭代需对种群适应度进行排序，累计为  $O(TM\log N)$ 。经验互益策略通过历史种群信息构造引导向量，累计开销均为  $O(TD)$ 。精英方差引导开发策略需对精英个体执行逐维方差与标准差计算，全迭代累计为  $O(TND)$ 。IHHO 的总时间复杂度为  $O(IHHO) = O(ND + TD + TM\log N + TND)$ 。在实际场景中，种群规模  $N$  通常取固定值，此时  $O(M\log N)$  为常数项，算法整体复杂度可近似为  $O(TD + TND)$ ，与原始 HHO 同阶。在典型实验参数( $N=30$ 、 $D=10\sim 1000$ )下，IHHO 相较于原始 HHO 的额外计算量占比约为 20%~45%，在可接受范围内，并未显著提升算法复杂度量级，却有效改善了原始 HHO 易陷入局部最优、后期收敛精度不足等缺陷，在计算开销与优化性能间取得良好平衡。

#### 3.4.2. 算法实现

本文所提出的 IHHO 算法的操作流程可以总结为如下几步：

**步骤 1：**初始化种群，计算种群适应度值，确定全局最优个体，同时保存初始历史经验信息并初始化个体停滞次数。**步骤 2：**用原始 HHO 的位置更新规则产生初始解集。**步骤 3：**当搜索进程处于探索与开发衔接区间时，采用经验互益策略对历史有益经验进行继承与重组，生成经验互益候选解。**步骤 4：**根据实时得到的逃逸能量值，用动态调整的扰动机制产生候选解集来达到全局搜索和局部优化。**步骤 5：**当算法进入后期阶段或个体出现连续停滞时，采用精英方差引导开发策略生成局部开发候选解。**步骤 6：**对各候选解执行竞争选择并更新种群位置，随后更新历史经验库与停滞计数。若达到最大评价次数，则输出全局最优解，否则继续迭代。

## 4. 实验结果与分析

为了验证 IHHO 的优化效果，本小节选取了 CEC2020 测试函数用于验证实验。哈里斯鹰优化算法(Harris Hawks optimization, HHO)、白鲸优化算法(Beluga Whale Optimization, BWO)、算术优化算法(arithmetic optimization algorithm, AOA)、正余弦优化算法(Sine Cosine Algorithm, SCA)和爬行动物搜索算法(Reptile Search Algorithm, RSA)作为对比实验。每个算法的种群个数  $N=30$ ，最大评价次数  $T = dim \times 10,000$ 。这些算法的参数如表 1 所示。

### 4.1. Friedman 统计性检验分析

为进一步评估算法的综合表现，采用 Friedman 检验进行多指标排名比较，结果见表 2。由表可知，IHHO 平均得分为 1.43，位居首位，彰显其强劲的局部搜索能力、稳定的全局寻优特性及良好普适性。从各函数表现来看，IHHO 在 F1~F8 中均名列前茅，在 F9~F10 中亦具明显优势；F4 函数对各算法综合性能区分度低，虽各算法平均值接近，但 IHHO 仍保持较高排名，对整体结果影响较小。这一优异表现得益于其合理的设计与优化机制。Friedman 检验结果充分表明，IHHO 综合性能优于所有对比算法。

**Table 1.** Parameter setting of each algorithm  
**表 1.** 各算法参数设置

算法	参数	数值	算法	参数	数值
HHO	$S$	0.01	RSA	$a$	0.1
	$\beta$	1.5		$\beta$	0.005
BWO	$W_f$	[0.1, 0.05]	IHHO	$S$	0.01
AOA	$C1$	0.2		$\beta$	1.5
	$C3$	3		$ph$	0.85
	$\mu$	25		$L$	5
	$\sigma$	3		$elite\_ratio$	0.2
SCA	$C1$	[0, 2]		$\sigma_0$	0.1
	$C2$	[0, 2]			
	$C3$	[0, 2]			

**Table 2.** The Friedman average ranking of each algorithm in CEC2020  
**表 2.** 各算法在 CEC2020 的 Friedman 平均排名

函数	IHHO	HHO	BWO	AOA	SCA	RSA
CEC1	1.00	2.00	5.00	4.00	3.00	6.00
CEC2	1.00	2.00	4.90	3.20	4.00	5.90
CEC3	1.00	2.60	4.30	4.60	2.50	6.00
CEC4	3.45	3.45	3.45	3.45	3.75	3.45
CEC5	1.00	2.40	5.00	3.90	2.70	6.00
CEC6	1.00	2.50	4.40	4.60	2.50	6.00
CEC7	1.00	2.50	5.00	2.5	4.00	6.00
CEC8	1.00	2.00	4.00	5.00	3.00	6.00
CEC9	2.00	2.60	2.50	4.60	3.50	5.80
CEC10	1.90	1.60	4.90	3.70	2.90	6.00
平均值	1.43	2.36	4.34	3.95	3.18	5.71

## 4.2. 收敛曲线分析

图 3 显示, IHHO 在 CEC2020 函数集上优化效率优异: 迭代初期即展现快速收敛特性, 后续持续保持稳定性能提升, 其收敛曲线整体水平低于其他对比算法, 体现出强劲的全局寻优能力与良好的局部优化性能。

这一优势源于经验互惠策略提升了种群中后期信息交互效率, 有效避免过早收敛; 精英方差引导开

发策略强化了优质解区域的精细搜索，保障了算法的收敛性与稳定性；能量自适应扰动策略虽初期有助于跳出局部最优。

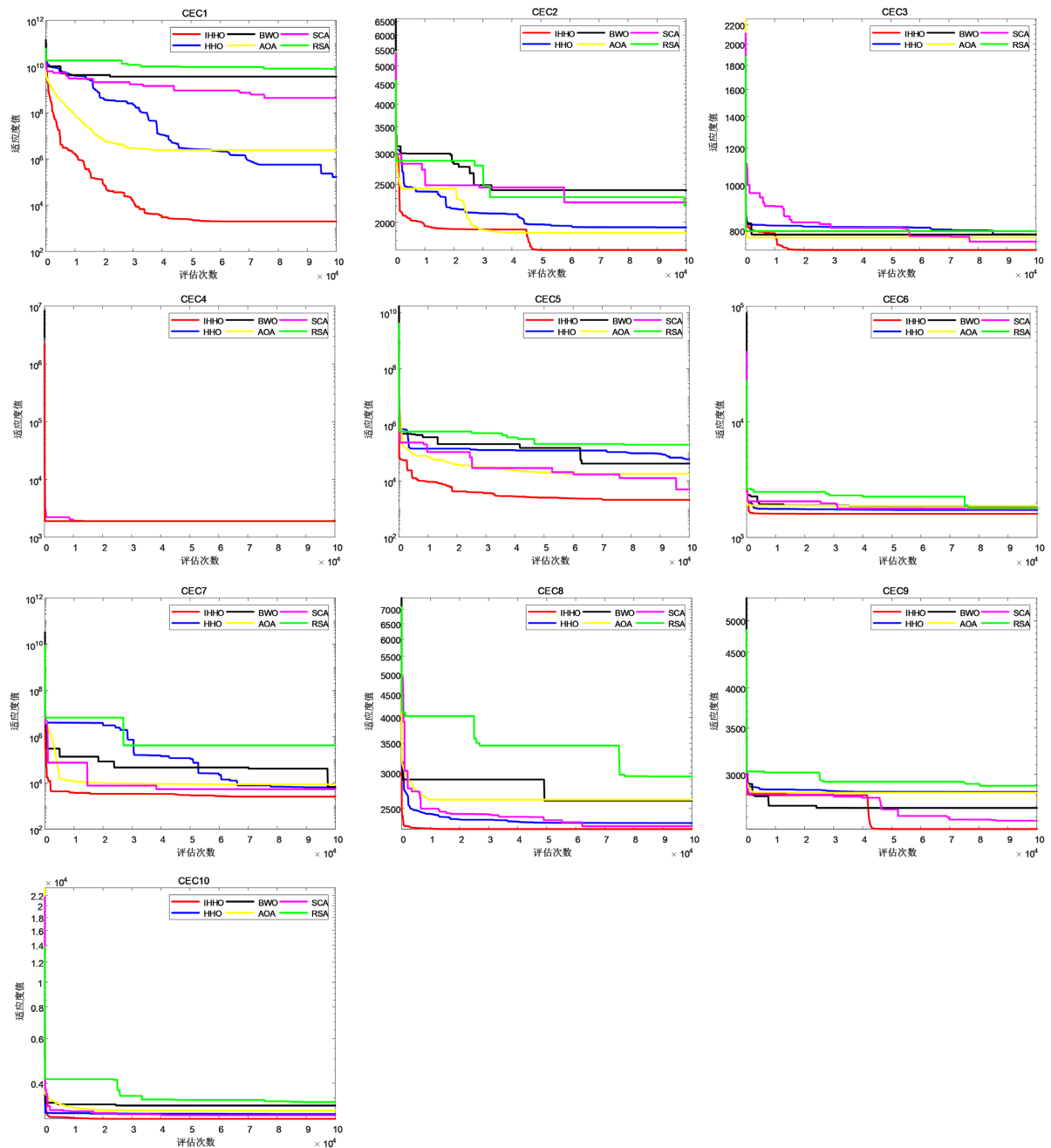


Figure 3. CEC2020 convergence curve of test function  
图 3. CEC2020 测试函数的收敛曲线

### 4.3. 消融实验

为验证三种改进策略，本文在 CEC2020 测试集上开展消融实验，实验结果见附录中的表 A1。其中 IHHO 为完整算法，IHHO1 (去除经验互益)、IHHO2 (去除能量自适应扰动)、IHHO3 (去除精英方差引导

开发)。

各消融版本表现显示：IHHO2 在 F1、F7 上平均值显著偏高，表明能量自适应扰动策略是提升全局搜索与跳出局部最优的关键；IHHO1 在 F5、F6 上更优，IHHO3 在 F2、F3、F8 上表现较好，说明经验互益策略增强中期协同搜索能力，精英方差引导策略提升后期开发精度。综上，三策略互补性强，使完整 IHHO 获得更均衡的优化性能。整体来看，完整 IHHO 综合竞争力最优。平均值指标下，在 F1、F4、F7、F10 上取得最优结果，F3、F8 保持强竞争力；最优值指标下，在 7 个函数上获最优或并列最优，验证了三策略协同实现了全局探索、局部开发与结果稳定性的良好平衡。

#### 4.4. 参数敏感性分析

在参数敏感性方面，本文设计一套科学的正交方案表，以测试 IHHO 所引入的新参数的敏感性。本文将  $ph$ 、 $L$ 、 $elite\_ratio$  和  $\sigma_0$  作为待分析的关键参数，并采用正交设计构造 9 组实验方案。在表 3 中列出了正交实验中各算法编号对应的参数组合。在附录中的表 A2 中是种群大小  $N = 30$ ， $dim = 30$ ，每个算法在 IEEE CEC2020 独立运行 30 次获得的统计结果。

试验结果表明： $ph = 0.85$  可平衡全局探索与局部开发需求，适配不同类型优化问题。 $L=5$  能兼顾搜索覆盖性与局部精细寻优能力，降低收敛波动； $elite\_ratio=0.2$  实现精英引导与种群多样性的最优折中，提升算法稳定性； $\sigma_0 = 0.10$  可保证初始种群多样性，同时避免过度干扰收敛效率。

综合分析， $ph=0.85$ 、 $L=5$ 、 $elite\_ratio=0.2$ 、 $\sigma_0=0.10$  组合具备最优泛化性能，能使 IHHO 算法在单峰、多峰及复杂混合问题中均保持均衡稳定的寻优效果，实现探索能力、开发能力与收敛效率的协同最优。本文后续实验均采用该参数配置。

**Table 3.** Parameter combinations of each algorithm in the orthogonal experiment  
**表 3.** 正交实验中各算法的参数组合

算法	$ph$	$L$	$elite\_ratio$	$\sigma_0$
IHHO1	0.75	3	0.1	0.05
IHHO2	0.75	5	0.2	0.10
IHHO3	0.75	7	0.3	0.15
IHHO4	0.85	3	0.2	0.15
IHHO5	0.85	5	0.3	0.05
IHHO6	0.85	7	0.1	0.10
IHHO7	0.95	3	0.3	0.10
IHHO8	0.95	5	0.1	0.15
IHHO9	0.95	7	0.2	0.05

## 5. 高维特征选择问题的应用

### 5.1. 选择特征解

对于每个解，维数等于特征数，每个维度表示对应特征的指数。举个例子，当数据中有 100 个特征时，一个解有 100 个维度。

在每个解中，维度被限制在  $[0, 1]$  的范围内。来确定一个特征是否会被选中，并使用 0.5 的静态阈值，

如下式所示：

$$\begin{cases} X_i^d > 0.5 & \text{选择} \\ X_i^d \leq 0.5 & \text{不选择} \end{cases} \quad (29)$$

其中  $X_i^d$  是第  $i$  个个体的  $d$  维度。

## 5.2. 适合度函数

为验证本文提出的算法在特征选择过程中的优化效果，本文使用分类错误率作为目标函数来评价解所选择的特征的性能。目标函数的计算方法如下：

$$Fitness = CEE = \frac{\text{错误的分类数}}{\text{样本总数}} \quad (30)$$

其中 CEE 是指  $k$ -最近邻( $k$ -Nearest Neighbor, KNN)算法的分类错误率， $k$  设为 5。KNN 中几个邻居之间的距离使用欧几里德距离计算，定义如下：

$$DE(Y, X) = \sqrt{\sum_{j=1}^{dim} (X^j - Y^j)^2} \quad (31)$$

其中  $X$  和  $Y$  是指给定实例中的特定特征， $dim$  是使用的特征总数。对每个数据集，实例被分为训练样本和测试样本(70%, 30%)。训练样本用于 FS 评估，而测试样本在测试阶段被隐藏。为了减少过度拟合问题，我们在文本中采用了  $K = 10$  的分层交叉验证。交叉验证方法首先将训练样本划分为十倍相同大小。然后用  $9(K-1)$  个数据集来训练 KNN 分类器，而剩余的 1 个折叠作为验证集。通过替换验证集和训练集，重复 10 次评估过程。

## 5.3. 实验结果分析

### 5.3.1. 实验结果统计与分析

为验证 IHHO 在高维特征选择中的有效性，选取 8 个高维数据集开展实验[8]，以适应度值( $k$ -最近邻分类错误率)和分类准确率为评价指标，各算法独立运行 1000 次。由图 4 可知，IHHO 在 Breast、CNS、Leukemia 等数据集上取得最优或接近最优的平均适应度值，在 Breast、CNS 数据集上达成最高分类准确率，Leukemia 系列数据集准确率达 1，实现了低适应度与高准确率的统一，验证其对关键特征的高效识别能力。IHHO 的核心优势在于兼顾特征筛选效率与分类稳定性：所选特征子集重复验证性强、鲁棒性优，在多数数据集上展现出更均衡的综合性能，适用于实际高维特征选择场景。对比算法均存在明显局限：原始 HHO 优化精度不足，易陷入局部最优；BWO 性能波动大，对数据集适应性不均衡；AOA 对高维空间探索粗放，难获最优特征子集；SCA 受周期性搜索影响，稳定性欠佳；RSA 易受随机性干扰，筛选与分类性能较差。

尽管 IHHO 在 Colon、Ovarian 等少数数据集上未达最优，但整体上仍以低适应度、高准确率的统一表现，彰显出强鲁棒性与泛化性，在高维特征选择领域具备显著应用潜力。

### 5.3.2. 各算法收敛曲线及直方图分析

在图 5 中展示了各个算法的精确度和特征数量结果，分析结果可知，IHHO 在 8 个高维特征选择数据集上性能优异，相较原始 HHO 收敛速度更快、迭代次数更少、适应度值更低，优化路径平滑且后期波动小，在 Breast、CNS 等高维数据集上可快速寻得优质解，全局搜索与局部优化能力优势突出。对比算法均存在性能短板：HHO 因缺乏历史经验继承与自适应扰动机制，易出现收敛停滞与局部波动；BWO 搜索稳定性差、曲线波动显著；AOA 收敛速度偏慢；SCA 受周期性更新规则约束，振荡明显；RSA 面对高维任务鲁棒性不佳、波动较大。

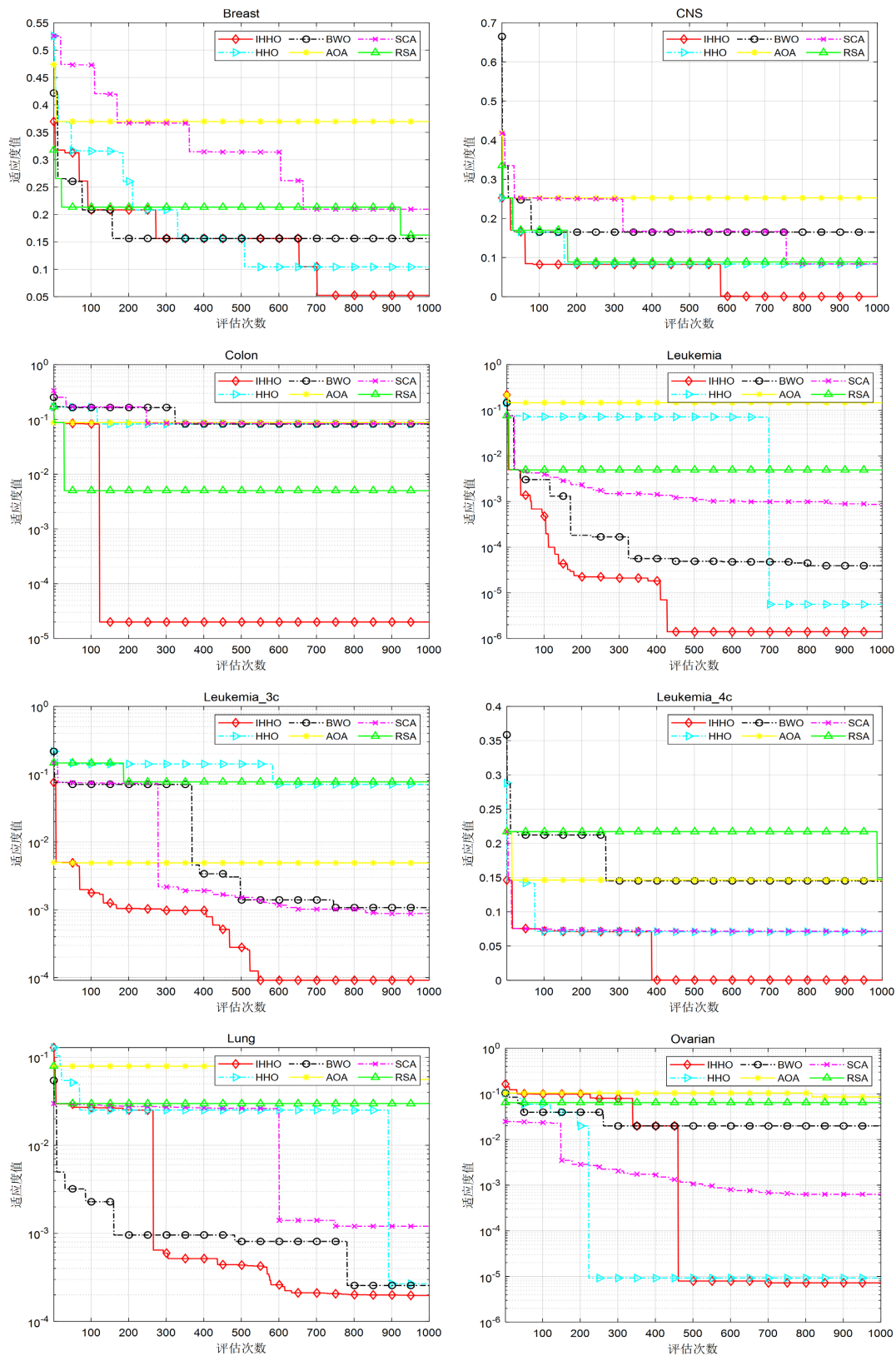
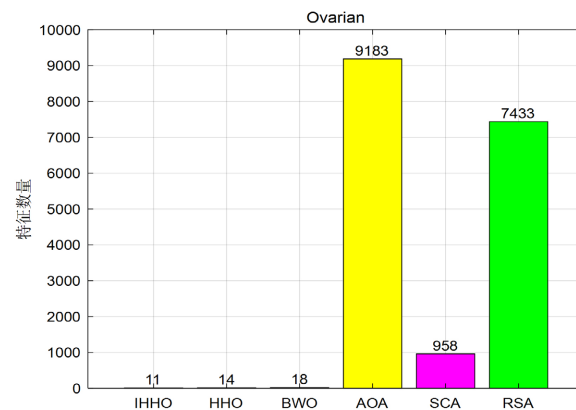
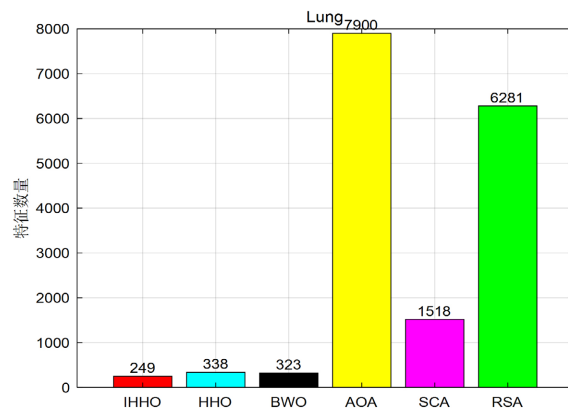
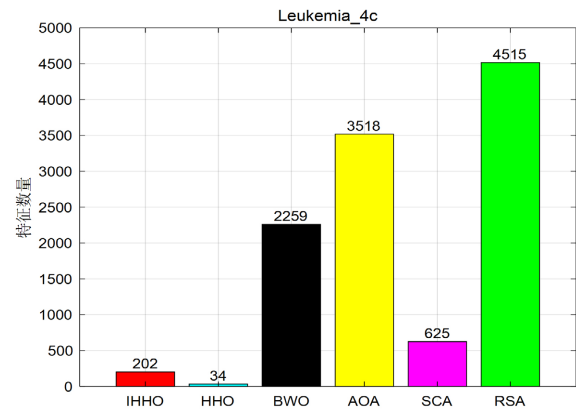
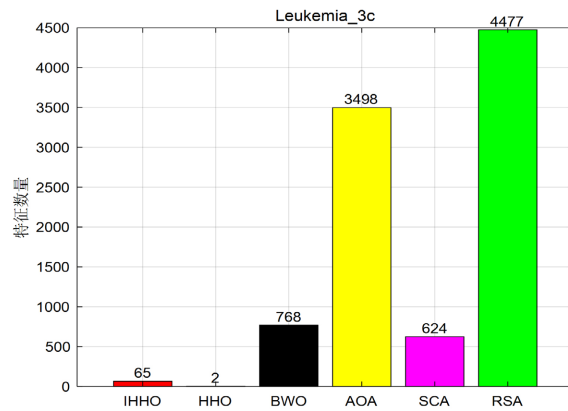
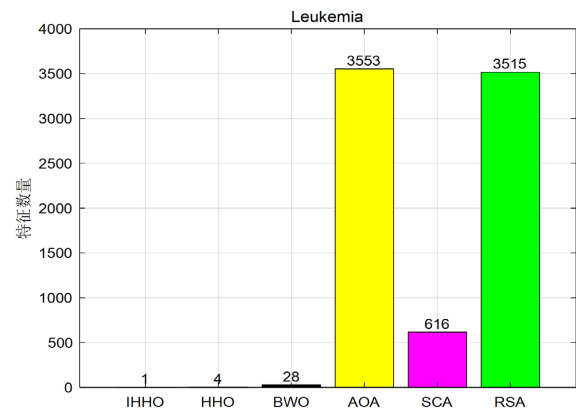
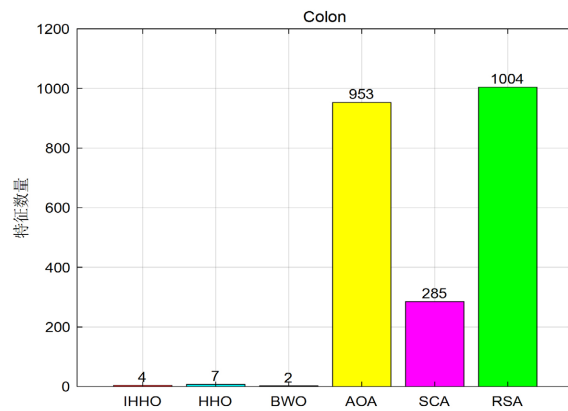
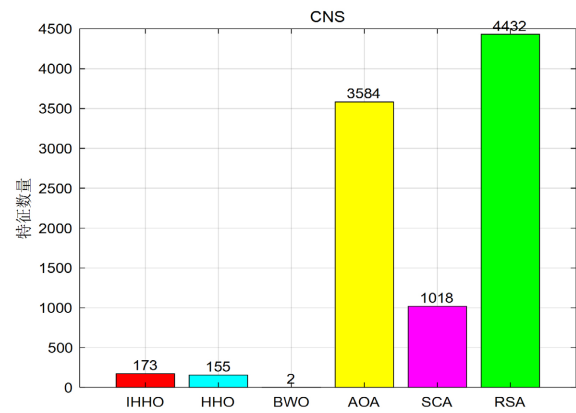
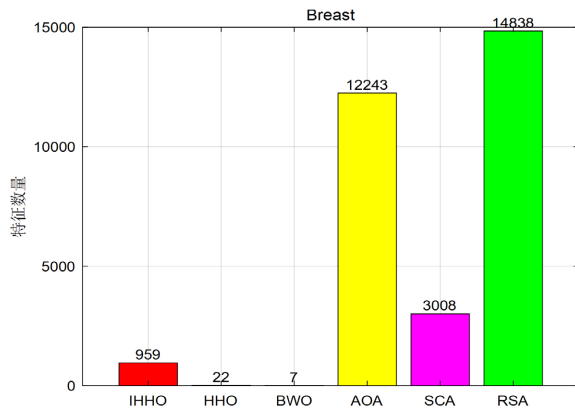


Figure 4. Convergence curves of each algorithm  
图 4. 各个算法的收敛曲线



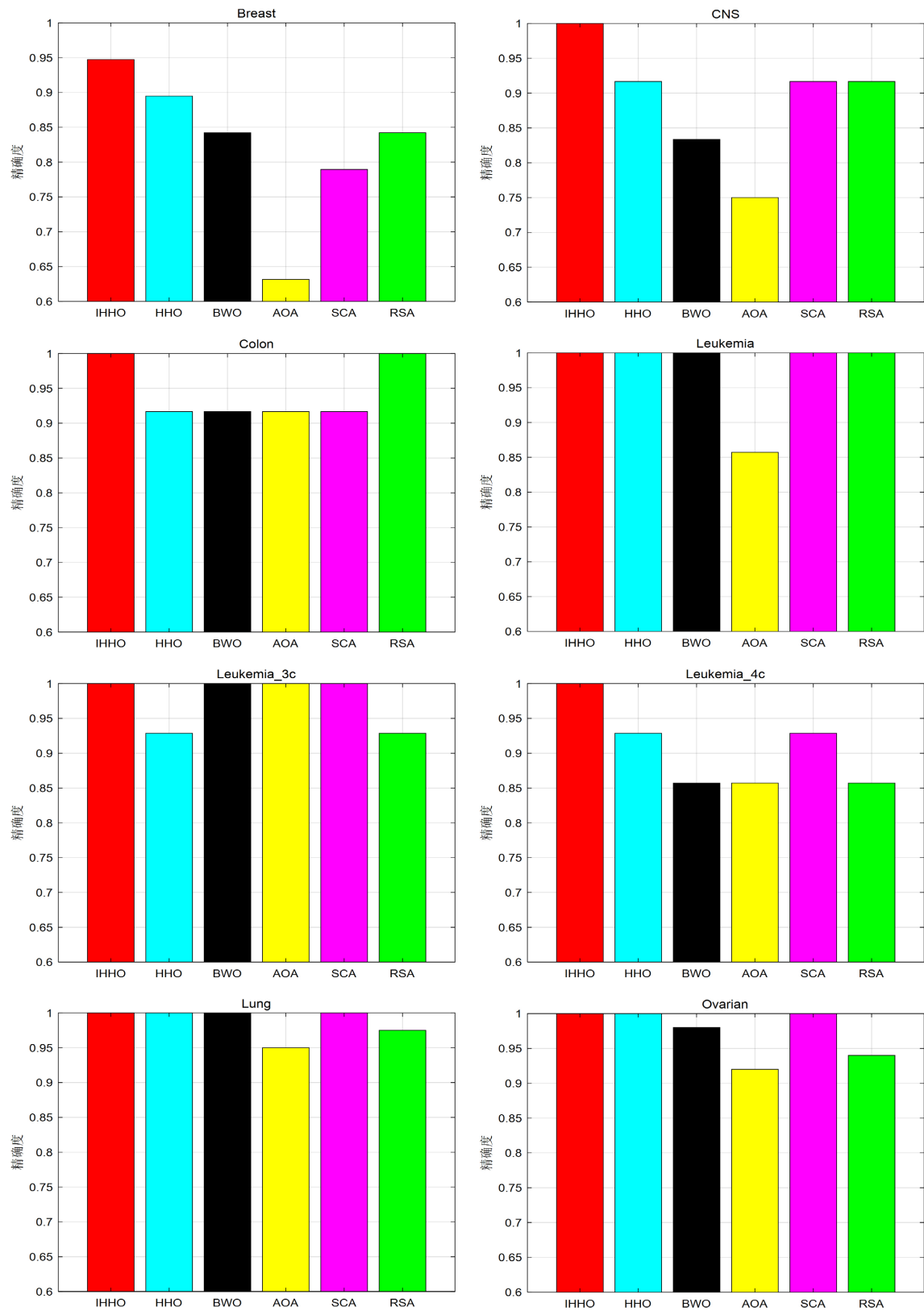


Figure 5. Accuracy and feature size of each algorithm  
 图 5. 各个算法的精确度和特征数量对比图

综上所述, IHHO 在收敛速度、寻优稳定性、分类精度及特征压缩效率等多维度实现协同提升, 并在多数数据集上分类准确率达到或优于最优水平, 所选特征子集判别性强, 可有效保留数据核心信息, 能在剔除冗余特征的同时保持高分类精度, 泛化性能与实用价值突出。

## 6. 总结与展望

本文针对 HHO 易陷入局部最优、收敛精度不足等问题, 提出一种多策略改进的哈里斯鹰优化算法。算法从三个层面进行增强: 引入经验互益策略, 利用历史信息提升跳出局部最优的能力; 设计能量自适应扰动, 动态平衡全局探索与局部开发; 提出精英方差引导开发, 依据精英方差调整搜索步长, 提高后期收敛精度。与其他算法比较, IHHO 在 Friedman 平均排名 1.43、收敛速度与鲁棒性上均显著优于对比算法。进一步, 我们将 IHHO 应用于特征选择问题。实验结果显示, 其在多数数据集上取得了更低的分类错误率与更高的准确率, 证明了其在特征筛选中的有效性。IHHO 有效克服了原算法的缺陷, 在连续优化与特征选择中均表现出优越性, 为优化问题提供了更高效的解决方案。同时, 本研究工作仍可从以下两方面进行深入探索: (1) 拓展应用场景: 将算法扩展至多目标、动态优化问题, 并进一步验证其在工业调度、故障诊断、等实际场景中的适用性。(2) 推进算法融合与工程化: 探索 IHHO 与深度学习模型在结构优化、同时研究算法的并行化与轻量化, 以降低计算成本, 推动工程应用。

## 参考文献

- [1] Xie, Y., Wang, X., Shen, Z., Sheng, Y. and Wu, G. (2023) A Two-Stage Estimation of Distribution Algorithm with Heuristics for Energy-Aware Cloud Workflow Scheduling. *IEEE Transactions on Services Computing*, **16**, 4183-4197. <https://doi.org/10.1109/tsc.2023.3311785>
- [2] Heidari, A.A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M. and Chen, H. (2019) Harris Hawks Optimization: Algorithm and Applications. *Future Generation Computer Systems*, **97**, 849-872. <https://doi.org/10.1016/j.future.2019.02.028>
- [3] Zhong, C. and Li, G. (2022) Comprehensive Learning Harris Hawks-Equilibrium Optimization with Terminal Replacement Mechanism for Constrained Optimization Problems. *Expert Systems with Applications*, **192**, Article 116432. <https://doi.org/10.1016/j.eswa.2021.116432>
- [4] Xu, J., Chang, X., Liu, Y., Wang, X. and Ren, C. (2025) Time and Energy Optimal Trajectory Planning for Freight Train Cleaning Robot Based on Seventh-Degree Polynomial Interpolation and an Improved HHO. *Alexandria Engineering Journal*, **122**, 165-179. <https://doi.org/10.1016/j.aej.2025.02.083>
- [5] Arafat, K.M.Y., Hossain, A., Ikfat, M., Amin, M.A., Tanvir, K., Gomes, D., et al. (2026) FRF-HHO: Early Ovarian Cancer Prediction Using Explainable Fuzzy Random Forest Optimized by Harris Hawks Algorithm. *Advances in Biomarker Sciences and Technology*, **8**, 219-235. <https://doi.org/10.1016/j.abst.2026.01.003>
- [6] Chen, W., Wu, Y., Pei, T., Lin, J. and Yuan, G. (2026) Predictive Maintenance Strategy for Photovoltaic Power Systems: Collaborative Optimization of Transformer-Based Lifetime Prediction and Opposition-Based Learning HHO Algorithm. *Energy Engineering*, **123**, 1-10. <https://doi.org/10.32604/ee.2025.070905>
- [7] Jia, H. and Rao, H. (2025) Experience Exchange Strategy: An Evolutionary Strategy for Meta-Heuristic Optimization Algorithms. *Swarm and Evolutionary Computation*, **98**, Article 102082. <https://doi.org/10.1016/j.swevo.2025.102082>
- [8] Zhu, Z.X., Ong, Y.S. and Dash, M. (2007) Markov Blanket-Embedded Genetic Algorithm for Gene Selection. *Pattern Recognition*, **40**, 3236-3248. <https://doi.org/10.1016/j.patcog.2007.02.007>

## 附录

**Table A1.** Comparative results of ablation experiments  
**表 A1.** 消融实验对比结果

函数	类型	IHHO	IHHO1	IHHO2	IHHO3
CEC1	最优值	100.1816219	179.0194529	920.1838076	124.152673
	平均值	3025.429576	3121.56455	17568.08117	3989.328378
	标准差	2272.055252	3324.836303	19181.38624	2505.967331
CEC2	最优值	1261.008086	1222.103113	1728.764034	1141.818861
	平均值	1673.075604	1689.887727	1993.044106	1651.224203
	标准差	220.2060058	296.4154123	197.2350708	289.3657035
CEC3	最优值	723.4136924	731.8087324	759.4327764	726.5158163
	平均值	745.9574164	742.7554031	774.1630011	741.3516437
	标准差	16.99125132	8.768171304	10.96718582	9.357888836
CEC4	最优值	1900.000000	1900.000000	1900.000000	1900.000000
	平均值	1900.000000	1900.000000	1900.000000	1900.000000
	标准差	0.000000	0.000000	0.000000	0.000000
CEC5	最优值	2619.450635	1841.020607	2747.560625	2614.257818
	平均值	7479.62155	6381.951784	6510.147278	6637.129544
	标准差	3071.777641	4406.265886	3304.945299	2790.255031
CEC6	最优值	1601.419527	1601.091103	1601.523054	1601.665554
	平均值	1704.577392	1664.229098	1725.62388	1684.435674
	标准差	78.50064181	84.39894744	79.62210307	66.73611863
CEC7	最优值	2181.674821	2481.643986	2621.40265	2223.368947
	平均值	3471.71939	4201.380054	6230.323705	3714.268399
	标准差	2009.32852	1336.672175	4627.002787	2366.747345
CEC8	最优值	2237.896725	2302.840496	2302.668528	2242.75927
	平均值	2298.527043	2305.199127	2312.285743	2292.109033
	标准差	21.47673143	2.239054405	5.875642117	25.79094703
CEC9	最优值	2500.000000	2500.000000	2500.402097	2746.591886
	平均值	2706.820254	2736.391751	2774.968071	2756.484517
	标准差	109.4617393	83.59318359	103.6367406	5.662602746
CEC10	最优值	2897.742869	2945.207404	2898.539065	2897.937917
	平均值	2929.925277	2947.169537	2931.151535	2933.068818
	标准差	27.69836705	2.842998168	28.38833792	23.55594983

**Table A2.** Results of parameter sensitivity analysis  
**表 A2.** 参数敏感性分析结果

函数	类型	IHHO	IHHO1	IHHO2	IHHO3	IHHO4	IHHO5	IHHO6	IHHO7	IHHO8	IHHO9
CEC1	最优值	106.5	104.1	100.0	128.2	111.5	241.9	106.0	104.4	100.6	102.0
	平均值	4085.5	3383.0	3575.1	4162.3	4649.5	4001.8	3126.7	3785.6	3560.2	3936.1
	标准差	3044.9	3010.3	2806.9	3531.7	3846.8	3463.9	2852.7	3538.2	3757.6	3877.2
CEC2	最优值	1270.1	1241.4	1241.1	1230.3	1245.2	1232.0	1153.5	1103.8	1258.4	1161.9
	平均值	1551.2	1618.5	1630.2	1646.4	1661.5	1671.0	1675.9	1536.6	1561.6	1577.8
	标准差	160.3	287.6	188.8	229.3	231.1	221.6	230.6	259.1	187.4	220.4
CEC3	最优值	723.2	724.7	718.3	718.8	719.6	731.9	720.0	725.3	720.6	724.8
	平均值	745.4	743.9	743.2	743.5	739.3	746.4	746.1	747.0	744.3	746.0
	标准差	9.6	12.2	11.0	12.5	12.6	9.0	13.4	14.2	12.5	12.9
CEC4	最优值	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0
	平均值	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0	1900.0
	标准差	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CEC5	最优值	2028.5	2119.2	1809.8	1869.9	2435.3	2123.3	2048.8	1913.4	2092.7	1970.9
	平均值	5330.1	7386.2	5655.3	6092.9	6393.3	5495.3	6926.5	5867.0	7002.3	5760.0
	标准差	3231.3	2787.7	2716.4	3202.9	2865.2	2931.2	3421.7	3303.6	3816.8	2720.7
CEC6	最优值	1601.3	1601.0	1601.4	1601.5	1600.6	1601.5	1601.2	1601.0	1601.3	1601.2
	平均值	1693.3	1668.3	1690.0	1687.7	1700.1	1694.9	1676.8	1676.2	1698.4	1695.3
	标准差	83.6	70.6	73.7	75.7	87.7	79.7	64.7	80.1	65.9	63.9
CEC7	最优值	2279.7	2120.7	2202.0	2139.1	2234.9	2222.1	2146.2	2158.8	2281.3	2118.3
	平均值	3860.1	3902.3	3284.8	3576.4	3600.5	3266.2	3562.7	3306.9	3167.7	3430.9
	标准差	1637.4	1759.9	1297.6	1359.5	1767.2	1115.0	1691.5	1247.3	1003.6	1138.7
CEC8	最优值	2238.3	2230.6	2217.9	2228.3	2301.0	2245.5	2226.2	2222.5	2211.4	2234.4
	平均值	2301.5	2302.3	2295.6	2299.3	2305.2	2302.7	2303.3	2298.8	2299.7	2303.3
	标准差	17.2	18.0	23.7	23.2	4.0	11.2	14.9	23.5	21.8	13.3
CEC9	最优值	2500.0	2500.0	2500.0	2500.0	2500.0	2500.0	2500.0	2500.0	2500.0	2500.0
	平均值	2749.1	2748.2	2724.3	2722.6	2749.8	2728.1	2742.2	2739.7	2740.7	2739.7
	标准差	48.0	47.7	89.9	89.2	48.0	78.1	66.5	65.8	66.3	65.9
CEC10	最优值	2897.7	2897.9	2897.9	2898.1	2600.1	2600.2	2897.9	2897.9	2898.7	2897.9
	平均值	2933.1	2934.5	2940.6	2934.4	2913.0	2901.9	2928.8	2935.0	2929.6	2929.7
	标准差	25.0	23.9	25.5	23.8	64.7	85.5	25.5	22.2	23.3	23.4