Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-021-03621-y

ORIGINAL RESEARCH

A novel control factor and Brownian motion-based improved Harris Hawks Optimization for feature selection

K. Balakrishnan[1] · **R. Dhanalakshmi**[1] · **Utkarsh Mahadeo Khaire**[2]

Received: 8 April 2021 / Accepted: 25 November 2021

Abstract
The massive growth in data size has prompted proliferation in need for Feature Selection (FS). Hence, FS has become an imperative method for dealing with high-dimensional data. This research critique proposes an enhanced feature selection of Harris Hawks Optimization (HHO) based on the novel control factor and Brownian motion. The Brownian motion augments the exploitation of foragers. It also replicates the deceptive movement of prey, allowing predators to correct their location and direction according to the prey's position. At the same time, the novel control factor imitates the exact behavior of the prey's escaping energy. The comparative analysis with the existing technique using six real high-dimensional microarray datasets highlights the impact of the proposed Improved Harris Hawks Optimization (iHHO). The experimental results of FS and classification accuracy vividly depict how the proposed model outperforms the existing techniques.

Keywords Feature selection · Harris Hawks Optimization · Meta-heuristic optimization · Microarray dataset

1 Introduction

Feature selection (FS) is an established preprocessing stage used in machine learning tasks to deal with high-dimensionality issues. The primary aim of FS is to classify main features and eliminate extraneous and insignificant features from the raw dataset for knowledge discovery. As illustrated in Fig. 1, FS is the process of selecting a smaller subset of features without radically decreasing classification accuracy. The first step to generate a subset from the input dataset is a valid search procedure. The second step compares the evaluated optimal subset and the antecedent subset. The newly updated subset is intensely recommended, assuming it will reinstate the older one compared to the existing subset. The loop will continue until the criteria of termination are satisfied. FS has assisted in the development of a wide range of algorithms in various fields of Engineering like Sentiment analysis (Madasu and Elango 2020), image processing applications (Bolón-Canedo and Remeseiro 2020), medical applications (Tuba et al. 2019), power systems (Abedinpourshotorban et al. 2016), text classification (Kou et al. 2020), Pattern recognition (Gunal and Edizkan 2008), drug design (Houssein et al. 2020a), wireless sensor networks (Houssein et al. 2020b), information retrieval (Lew 2001), job scheduling (Gao et al. 2020a) and many other real-world applications.

Myriads of scholars have defined Feature selection as the ability of feature subsets to identify targets, maximize prediction accuracy, or alter the composition of the original data group. The three types of function subset search strategies are global optimal, sequence, and random. Discovering the current optimal subset of the initial feature sets is the aim of global optimal search. The three types of sequence search algorithms are forward search, backward search, and bidirectional search (Marcano-Cedeño et al. 2010). Forward search refers to the greedy method of applying the element with the best score to the Selected Feature Subset (SFS). Backward search denotes eliminating a component from the selected subset of the feature at a time using Sequence Backward Search (SBS). Bidirectional search is a forward and backward search technique that paves the way for adding and removing features (Gu et al. 2015). The random search strategy's feature selection is chaotic, with uncertainty,

✉ Utkarsh Mahadeo Khaire
utkarshkhaire@gmail.com

[1] Department of Computer Science and Engineering, Indian Institute of Information Technology, Tiruchirappalli, Tamil Nadu 620012, India

[2] Department of Data Science and Intelligent Systems, Indian Institute of Information Technology, Dharwad, Karnataka 580009, India

Published online: 11 January 2022

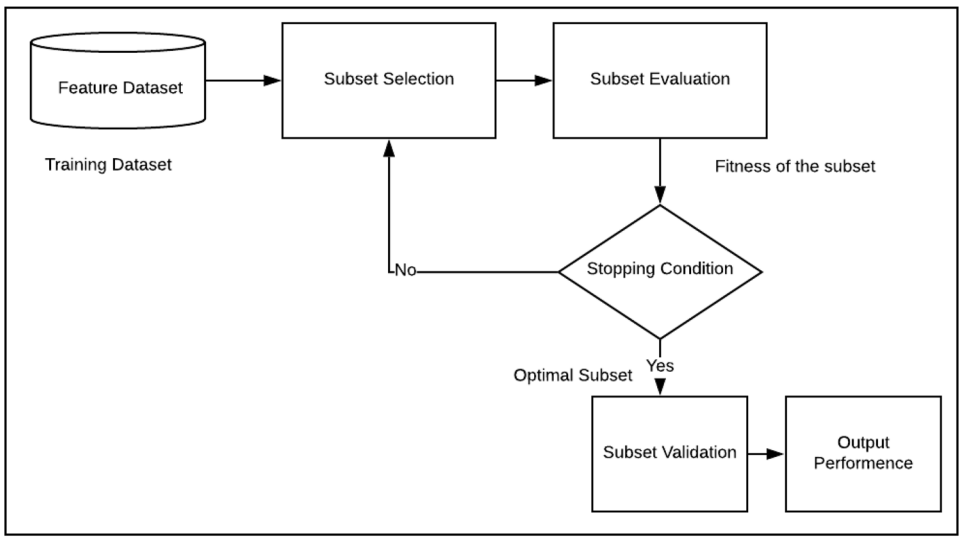

Fig. 1 General feature selection procedure

forcing the algorithm to escape the local optimum, allowing the algorithm to find the estimated optimal solution (Dash and Liu 2003). As a result, random search strategies such as Simulated Annealing (SA) (Liu et al. 2018) and Genetic Algorithm (GA) (Dong et al. 2018) selection of the feature subset outperforms sequence search in most cases.

This research pericope focuses on the FS using Meta-Heuristic (MH) optimization techniques. Traditional optimization techniques have several limitations when it comes to FS problems. Population-based methods are of five types (Heidari et al. 2019): Evolutionary Algorithm (EA), Physics-based Algorithm (PBA), Swarm-based Algorithm (SBA), Human-based Algorithm (HBA), and Mathematical-based Algorithm (MBA). Figure 2 depicts numerous well-known population-based algorithms. In population-based algorithms, two major characteristics are typically found: intensification and diversification.

Harris Hawks Optimization (HHO) is a nature-inspired MH algorithm developed by Heidari et al. (2019). HHO imitates the attacking strategy such as besiege, perching, and surprise pounce strategies of Harris hawks to find an optimal solution. HHO is split into two exploration stages and four exploitation stages. In many real-world applications, HHO outperforms traditional meta-heuristic algorithms such as Whale Optimization Algorithm (WOA) and Genetic Algorithm (GA). However, through this research, two drawbacks associated with traditional HHO are identified. First, the representation of prey escape energy is inefficient in conventional HHO. Second, Levy flight represents the imitation of leapfrog movements of prey while fleeing from a predator. The initial random solution's more petite and extended step sizes hurdle the algorithm emerging from local optima. As a result, a novel control factor that accurately imitates prey's escape energy over time is developed through this research

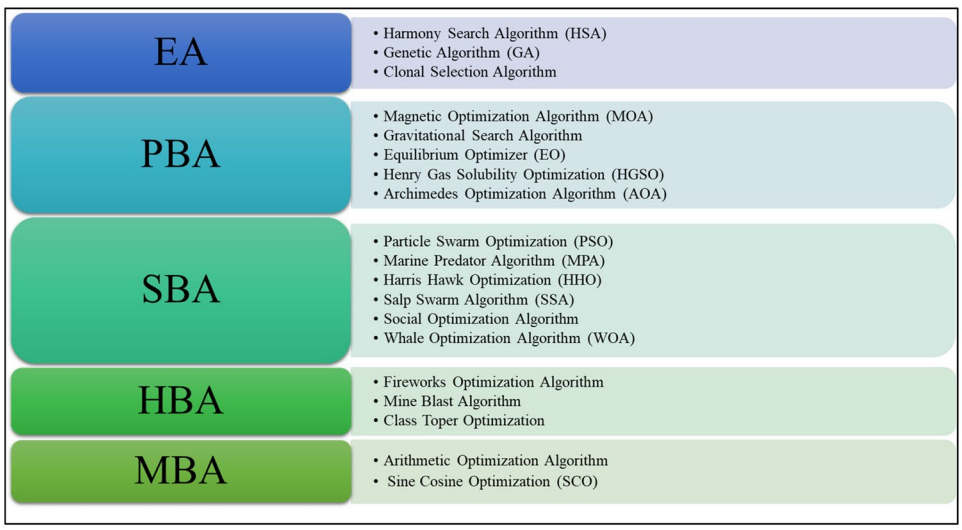

Fig. 2 Population-based MH optimization algorithms

critique. This research connoisseur has employed Brownian motion rather than Levy flight to attain the optimum solution in the stipulated period. In brief, the following are the significant contributions of this research:

1. This research pericope has proposed a robust iHHO for selecting significant features from the high-dimensional microarray datasets.
2. A novel control factor is used to express the prey's escape energy accurately.
3. Considering randomization is the key to any MH algorithm, zigzag movement of escaping prey and surprise attacks of predators are represented using Brownian motion instead of Levy flight.
4. The proposed iHHO's performance is assessed using six real high-dimensional microarray datasets and several unimodal and multimodal functions.
5. The outcomes of the proposed iHHO are compared with six well-established optimization techniques.

Section 2 of this research article covers the literature survey. Section 3 stresses the motivation of the proposed iHHO. The background and mathematical approach of iHHO are discussed in Sect. 4. Section 5 deals with the implementation of the proposed method. Section 6 focuses on the simulation findings, and discussions are elaborated in Sect. 7. Section 8 emphasizes the conclusion and the future scope.

2 Related work

The reliability factor has resulted in the stochastic optimizers' interest in solving problems in various fields such as the manufacturing industry, environmental quality, solar systems, power systems, and other engineering areas (Alabool et al. 2021). Compared to traditional methods, nature-inspired MH algorithms have a promising result for complex problems. Feature Selection (FS) is one of the most popular applications of stochastic optimization algorithms. Owing to the slow convergence of MH algorithms, researchers use different techniques to improve the MH algorithm by adding one or more techniques to the existing algorithms. The well-known techniques are levy flight, Brownian motion, binary variant, and opposition-based learning. Sihwail et al. proposed an improved version of HHO based on Elite Opposition Based Learning (EOBL) and Three Search Strategies (TSS) to boost the global and local searches of HHO (Sihwail et al. 2020). EOBL has increased the diversity of the HHO population, and TSS search strategies assisted the algorithm in its search for global optima by avoiding traps in local optima. From the results of the experiments, it could be understood that the proposed approach outperforms the other algorithms in all the metrics.

Elminaam et al. proposed a novel technique to the FS problem based on the Marine Predator Algorithm (MPA) (Elminaam et al. 2021). This study has encompassed the hybridization of MPA with KNN. The proposed MPA-KNN adapts the basic exploratory and exploitative procedures to choose the best relevant features for the most accurate identification. Five different evaluation criteria are carried out on 18 UCI datasets to explore the suggested approach's performance. The obtained results enunciate that the proposed model provides superior performance than the conventional MH algorithms in selecting the optimal features. Elgamal et al. proposed an improved variant of HHO based on Simulated Annealing (SA) for FS in the medical field (Elgamal et al. 2020). The suggested model handles concerns such as population diversity and local optima of conventional HHO. In this proposed method, SA is used to improve the exploitation capability of HHO. The SA algorithm is also used by Abdel et al. in their study and has proposed a variant of HHO termed the Chaotic HHO (CHHO) algorithm. CHHO is developed using chaotic maps and incorporates the SA algorithm to enhance the population diversity and converging ability to avoid the local optima (Abdel-Basset et al. 2021).

Neggaz et al. proposed an improved variant of the Salp Swarm Algorithm (SSA) for FS problems using Sine–Cosine Algorithm (SCA) and disrupted operator (Neggaz et al. 2020a). The SCA facilitates exploration and prevents stagnation. In addition to that, disrupt operator is used to enhancing the population diversity. Neggaz et al. also suggest a novel approach based on Henry Gas Solubility Optimization (HGSO) to select the optimal features and improve classification accuracy (Neggaz et al. 2020b). Ahmed et al., in their research, has used SSA with four different chaotic maps to balance exploration and exploitation (Ahmed et al. 2018). Twelve real-world datasets are used to test the proposed techniques. The findings indicate that chaotic maps improve the proposed model's performance considerably as compared to conventional methods. In a study, Zhang et al. proposed the binary variant of HHO for global optimization and FS problems (Zhang et al. 2020). The SSA mechanism is merged into traditional HHO to enhance the exploitation and exploration behavior of HHO. The suggested HHO is tested on 23 classical functions using statistical metrics and convergence rate.

Houssein et al. balanced exploration and exploitation of HHO using genetic operators and two methods, such as Opposition Based Learning (OBL) and Random Opposition Based Learning (ROBL) (Houssein et al. 2021). Monoamine Oxidase and QSAR Biodegradation datasets are used to assess the efficacy of the proposed system. The findings indicate that the three variants of the proposed model outperform the conventional techniques in determining the best subset of chemical descriptors. Houssein et al. also

suggested a hybrid version of HHO based on Cuckoo Search Optimization (CSO) and chaotic maps to boost the efficiency of the original HHO (Houssein et al. 2020a). Furthermore, the proposed model was paired with the Support Vector Machine (SVM) as a machine learning classifier for performing chemical descriptor collection and chemical compound operations.

Hussein et al. proposed an improved variant of HHO with three strategies OBL, Chaotic Local Search, and self-adaptive technique (Hussien and Amin 2021). The OBL strategy is incorporated in the initialization phase, and the other two strategies are embedded with the update phase to enhance the converging ability of the proposed model. Hussain et al. suggested a unique hybrid approach for numerical methods and FS issues by combining two algorithms: SCA and HHO (Hussain et al. 2021). The primary goal of this research is to balance exploration and exploitation capabilities. Ismael et al. suggested a novel hybrid model that optimizes the hyper parameters of v-SVR when simultaneously embedding feature selection using OBL (Ismael et al. 2020).

Gao et al. discovered a tent map that could be used to improve the converging capabilities of HHO (Gao et al. 2019). The tent maps can initialize the population distribution and escape the equal distribution by generating the chaos using non-period, non-converged, and bounded random numbers. The chaos is also applied to the algorithm to substitute the random numbers. The proposed algorithm is evaluated using 18 benchmarks of unimodal and multimodal functions. Gao et al. also proposed a binary variant of Evolutionary Optimization (EO) for FS problems (Gao et al. 2020b). The model employs the Sigmoid transfer function (S-Shaped) and V-shaped transfer functions to transform the binary version of EO and change the particle's current position vector. The suggested method is tested using nineteen UCI benchmark datasets. Based on the experimental findings, the binary EO technique performs well compared to other approaches for addressing FS issues. Zhang et al. introduced a return cost-based Firefly Algorithm (FFA), which uses the binary movement operator to change firefly positions (Zhang et al. 2017). Zhang et al. also presented an improved HHO based on SSA, assuming that SSA's powerful explorative capacity would facilitate exploring the original HHO (Zhang et al. 2020). Initialize and update are the two stages of the proposed method. The proposed approach clearly outperforms other approaches in terms of average feature length and error rate measures according to the experimental findings.

To solve the issue of feature selection in classification tasks, Too et al. proposed two improved variants of HHO, namely Binary HHO (BHHO) and Quadratic Binary HHO (QBHHO) (Too et al. 2019). The BHHO has a built-in S-shaped or V-shaped conversion mechanism that converts continuous search agents to binary. The quadratic transfer function is used in QBHHO to rejuvenate BHHO's feature selection efficiency. The best fitness value, mean fitness value, standard deviation of fitness value, classification accuracy, and feature size are used to assess the results of the proposed algorithms. The findings of this analysis indicate that the quadratic transfer function algorithm provides the best results. Too and Mirjalili have also introduced a general learning method to aid search agents in avoiding local optima and improving their capacity to find a promising region (Too and Mirjalili 2021). Sixteen biological datasets were used to test the proposed General Learning Equilibrium Optimizer (GLEO).

Mafarja and Mirjalili developed a new hybrid metaheuristic technique combining SA and WOA. The SA scheme improves the optimum solution found by WOA (Mafarja and Mirjalili 2017). The key objective of this hybridization is to improve SA's exploitation ability, represented by WOA. The performance assessment demonstrates an improvement in classification accuracy and produces better results than wrapper-based techniques. 18 standard datasets were taken from the UCI repository to compare the proposed method. Hussein et al. suggested a novel binary variant of WOA (BWOA) to choose the best function subset for the FS problem (Hussien et al. 2019). An S-shaped transfer function is used to control the novel strategy. Over eleven different datasets, a series of parameters is used to evaluate and equate the proposed model with the existing one. Emary et al. has used a threshold value to solve function selection problems, encompassing the first binary variant of the Firefly Algorithm (FFA) (Emary et al. 2015). The suggested algorithm had a high level of investigation and was able to find a simple solution to the problem. Kanimozhi et al. proposed an image retrieval strategy based on SVM classifiers and FFA (Kanimozhi and Latha 2015). The fundamental goal was to improve the algorithm's accuracy by using optimal functionality, and the algorithm was put to the test on a variety of image datasets.

3 Motivation

The existing Harris Hawks Optimization presents the escaping energy of the prey using Eq. (1).

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

In Eq. (1), E denotes the prey's escaping energy, T is the maximum number of iterations, E_0 is the energy's initial state, which varies randomly across the interval $(-1, 1)$. The graphical representation of the behavior of E is shown in Fig. 3. The figure imitates the deceptive zigzag movement of the prey. Figure 3 represents an unexpected hike in the

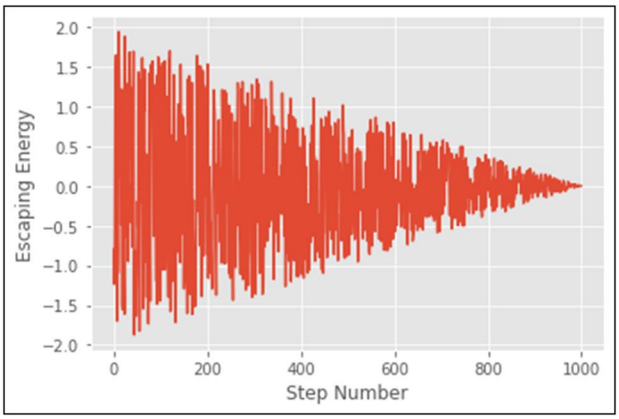

Fig. 3 Behavior of escaping energy (E) in traditional HHO during 1000 iterations

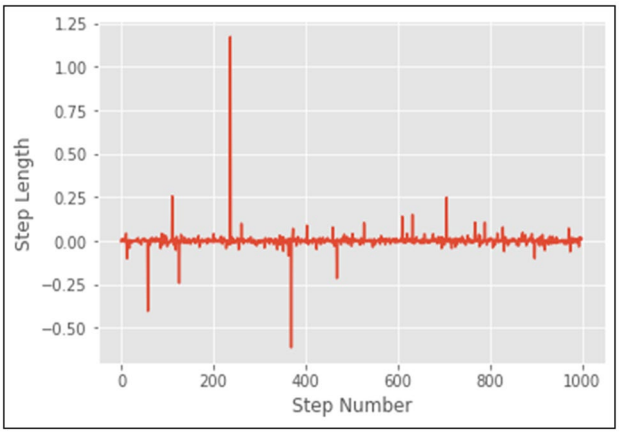

Fig. 5 Movement of Harris hawks using Levy flight

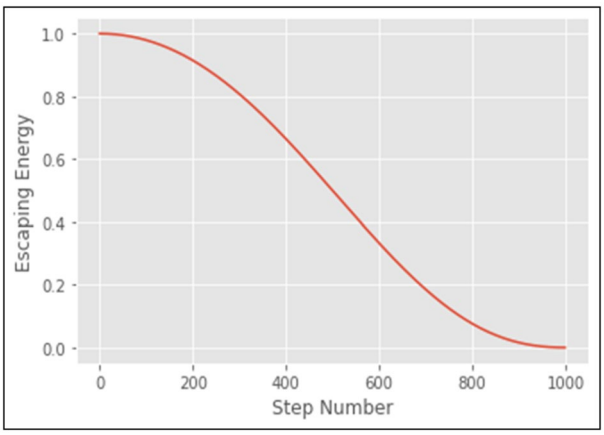

Fig. 4 The behaviour of novel control factor-based escaping energy (Ee) in improved HHO during 1000 iterations

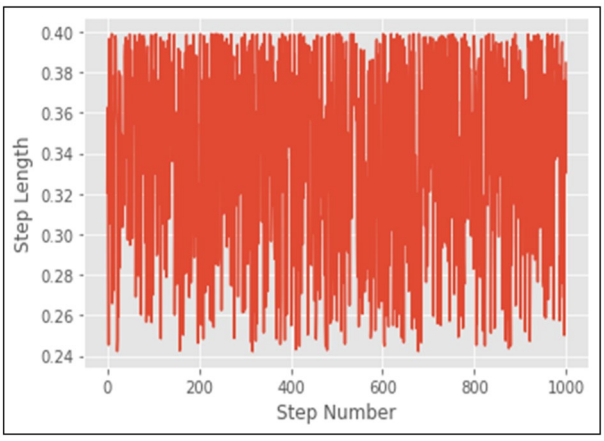

Fig. 6 Movement of search agent using Brownian motion in the improved Harris Hawks

prey's energy as it gradually loses its energy while running away from its predator.

The accurate behaviour of the escaping energy of the prey is depicted in Eq. (2), which is the proposed novel Control Factor (CF) that accurately imitates the behaviour of the prey's escaping energy.

$$CF = \left(1 - \frac{t}{T}\right)^{\left(2\times\frac{t}{T}\right)} \tag{2}$$

The graphical representation of the CF in Fig. 4 shows the smooth deprivation of escaping energy of the prey as the model progresses. The CF gradually decreases from 1 to 0, the exact imitation of prey losing its energy while escaping from the predator.

The conventional Harris Hawks use Levy flight with progressive dives into the exploitation phase in the hard besiege stage. Figure 5 shows the nature of movements of Harris Hawks using the Levy flight. Due to the smaller and occasional larger step size, the algorithm sometimes fails to recover from local optima.

The aforementioned drawback of existing Harris Hawks Optimization is overcome by the proposed movement of Harris hawks using Brownian motion, as stated in Eq. (3). The graphical representation of Brownian motion is shown in Fig. 6.

$$\text{Brownian Motion} = f_B(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) \tag{3}$$

In Eq. (3), x denotes Harris Hawk's current location in the iteration. The Brownian motion allows hawks to take longer step size, which eventually precludes the algorithm from stagnation at local minima.

4 Improved Harris Hawks Optimization (iHHO)

The modified HHO is divided into three phases: initialization, update, and classification. The following subsection elucidates each component.

4.1 Initialization phase

The initialization of iHHO reflects other meta-heuristic techniques where it initializes the random search agents and identifies the current best solution using a defined objective function.

4.2 Update phase

In this phase, to achieve the optimum solution, the proposed algorithm performs both intensification and diversification. Figure 7 represents the specific conditions that contribute to the diversification and intensification of the algorithm. This procedure is categorized into four stages: soft besiege, soft besiege with progressive rapid dives, hard besiege, and hard besiege with progressive rapid dives.

4.2.1 Exploration

In this phase, the predator (Harris Hawks) monitors the grander search space to discover the location of prey (Rabbit). Let q denote the probability of success for the perching strategy. For the condition $q < 0.5$, the predator perches based on the position of other family members and the rabbit. Otherwise, a random search agent changes its position if $q \geq 0.5$. The mathematical representation of the perching strategy of Harris hawks is modeled in Eq. (4).

$$Y(t+1) = \begin{cases} Y_{rand}(t) - r_1 \left| Y_{rand}(t) - 2r_2 Y(t) \right| & q \geq 0.5 \\ \left(Y_{rabbit}(t) - Y_m(t) \right) - r_3 \left(lb + r_4 (ub - lb) \right) & q \leq 0.5 \end{cases} \tag{4}$$

where $Y(t)$ and $Y(t+1)$ is a position of the search agent in the iteration t and $t+1$, respectively. r_1, r_2, r_3, and r_4 are the arbitrary value in the range of [0, 1]. ub and lb denotes the upper and lower bound value of every individual search agent. $Y_{rabbit}(t)$ and $Y_{rand}(t)$ are the position of the prey and the current random position, respectively. $Y_m(t)$ is the average location of the new search agents obtained using Eq. (5).

$$Y_m(t) = \frac{1}{N} \sum_{i=1}^{N} Y_i(t) \tag{5}$$

where N is the number of search agents, $Y_i(t)$ epitomizes the position of the search agent in iteration t.

4.2.2 Transition from exploration to exploitation

Generally, the performance of MH algorithms is determined by the factor that balances exploration and exploitation. The evolution of iHHO from exploration to exploitation is focused on the prey's escape energy using the following equation:

$$Escaping\_Energy\,(Ee) = 2 \times CF \times E_0 \tag{6}$$

In Eq. (6), E_0 is a random initial state, and its values range from -1 to 1. $|Ee| \geq 1$ indicates the exploration process, and $|Ee| < 1$ represents the exploitation process, and CF denotes control factor.

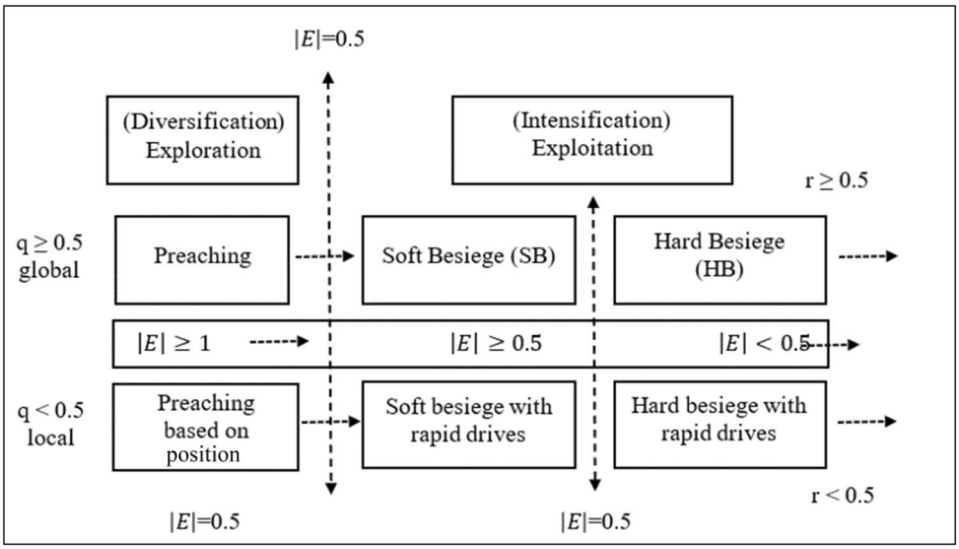

Fig. 7 Phases of improved HHO

4.2.3 Exploitation phase

In this phase, iHHO uses four strategies to attack the prey.

4.2.3.1 Soft besiege During this stage, the prey has enough energy to flee from the predator ($r \geq 0.5$). However, the predator surrounds the prey tactfully and attacks unexpectedly. The behavior of unpredictable attack is modeled as follows:

$$Y(t+1) = \Delta Y(t) - (Ee)|J \times Y_{rabbit}(t) - Y(t)| \tag{7}$$

$$\Delta Y(t) = Y_{rabbit}(t) - Y(t) \tag{8}$$

In Eq. (7), $J = 2*(1-r_5)$ represents the random jump strength of the rabbit throughout the escaping procedure where r_5 is an arbitrary number in the range of (0, 1). The J value changes randomly in each iteration to simulate the nature of rabbit motions.

4.2.3.2 Hard besiege The escaping energy of the prey is low, when $r \geq 0.5$ and $|Ee| < 0.5$, the predator encircles the victim fiercely and strikes randomly as formulated in Eq. (9).

$$Y(t+1) = Y_{rabbit}(t) - (Ee)|\Delta Y(t)| \tag{9}$$

4.2.3.3 Soft besiege with progressive rapid dives The prey has sufficient energy to flee from the predator when $|Ee| \geq 0.5$ and $r < 0.5$. The prey performs leapfrog movement to escape from the predator. Compared to other previous techniques, performing soft besiege before a sudden attack is a better approach. The respective zigzag movements of prey and attacking strategy of predator are modeled using the succeeding equations:

$$X = Y_{rabbit}(t) - (Ee)|J \times Y_{rabbit}(t) - Y(t)| \tag{10}$$

According to the movement of prey, predators make more abrupt dives that suit the prey's deceptive movements. This research critique has espoused the Brownian motion to imitate the predator's most irregular and rapid plunge. Equation (11) envisages the procedure of a predator changing its position based on the Brownian motion.

$$\begin{aligned} stepsize &= f_B \otimes \left(Y_{rabbit} - f_B \otimes Y(t)\right) \\ Z &= X + P \times R \otimes stepsize \end{aligned} \tag{11}$$

In Eq. (11), $P = 0.5$ is a constant number. The final procedure for updating the position of the search agent can be performed by Eq. (12), in which F is the fitness function.

$$Y(t+1) = \begin{cases} X \text{ if } F(X) < F(Y(t)) \\ \text{Z if } F(Z) < F(Y(t)) \end{cases} \tag{12}$$

4.2.3.4 Hard besiege with progressive rapid dives When $|Ee| < 0.5$ and $r < 0.5$ denotes that the prey cannot flee the hawks owing to its insufficient energy. The hard besiege of hawks with progressive dive is formulated using the following equations:

$$X = Y_{rabbit}(t) - (Ee)|J \times Y_{rabbit}(t) - Y_m(t)| \tag{13}$$

$$\begin{aligned} stepsize &= f_B \otimes \left(Y_{rabbit} - f_B \otimes Y(t)\right) \\ Z &= X + P \times R \otimes stepsize \end{aligned} \tag{14}$$

Final position updating of Harris hawks is done by using Eq. (15)

$$Y(t+1) = \begin{cases} X \text{ if } F(X) < F(Y(t)) \\ \text{Z if } F(Z) < F(Y(t)) \end{cases} \tag{15}$$

4.3 Classification phase

The standard sigmoid function is used to classify the data sample to verify whether the algorithm progresses correctly. The sigmoid function is a customary binary classifier that assures the data sample category using a predefined threshold value. As shown in Eq. (16), the binary cross-entropy is used to calculate the error rate. Where y_i is the actual value, and $p(y_i)$ is the predicted value of the data sample.

$$1 = -\frac{1}{N}\sum_{i=1}^{N} y_i \times \log(p(y_i)) + (1-y_i) \times \log(1-p(y_i)) \tag{16}$$

The improved Harris Hawks Optimization (iHHO) algorithm's pseudo-code is as follows:

Algorithm 1: Pseudo-code of iHHO

Initial random solutions Y_i (I = 1, 2, 3, 4…., n)
Compute log loss for each initial random solution using Eqn. (16) to identify Y_{rabbit} (best location)
While (current iteration t < T)
  for each search agent (Yi)
    Update the escaping energy-based Eq. (6)
    if ($|Ee| \geq 1$)
      Update the position vector based on Eq. (4)
    if ($|Ee| < 1$)
      if ($|Ee| \geq 0.5$ and $r \geq 0.5$)
        Update position vector based on Eq. (7)
      else if ($|Ee| < 0.5$ and $r \geq 0.5$)
        Update position vector based on Eq. (9)
      else if ($|Ee| \geq 0.5$ and $r < 0.5$)
        Update position vector based on Eq. (12)
      else if ($|Ee| < 0.5$ and $r < 0.5$)
        Update position vector based on Eq. (15)
Return Y_{rabbit}

5 Implementation

This section exemplifies the experimental setup, process flow, and overview of the high-dimensional datasets used in the present research.

5.1 Dataset

The proposed model is evaluated using six publicly available high-dimensional microarray datasets. Table 1 provides an overview of the datasets used in the study. The microarray dataset contains the gene expression level in a continuous data format, which helps analyze a multitude of genes in a limited period. Analyzing microarray data helps in the diagnosis and prognosis of life-threatening diseases. Microarray data is the combination of significant and noisy features. Irrelevant and insignificant features in microarray data encumber the diagnosis of diseases. All extraneous columns of the input datasets (such as Patient id, Name, Address, etc.) are removed before starting the experiment. All the datasets used in the process have two classes viz., normal and malignant, in the target variable.

Table 1 Overview of datasets

Dataset name	Number of features	Number of samples
Breast cancer	24,481	97
Central nervous system (CNS)	7129	60
Colon Cancer	2000	60
Leukemia	7129	72
OSCC	41,003	50
Ovarian cancer	15,154	253

5.2 Process flow

Figure 8 depicts the implementation of the proposed technique. In the initial phase, the data balancing step uses the SMOTE-tomek algorithm to manage the dis-proportion of data samples in input data. The SMOTE-tomek algorithm is a combination of over-sampling and under-sampling techniques that aid in the elimination of data imbalances. Due to the addition of new data samples, data balancing typically increases the size of the original input data. In the initial stage, the input data has been normalized using Min–Max scaling. Along with the data preprocessing, the maximum number of iterations has been set to 100. The position matrix containing 100 random search agents is also generated. The fitness value of the logistics function is determined using an unbiased cross-entropy function for each row of the Position matrix, which aids in choosing the best location of prey (*Yrabbit*).

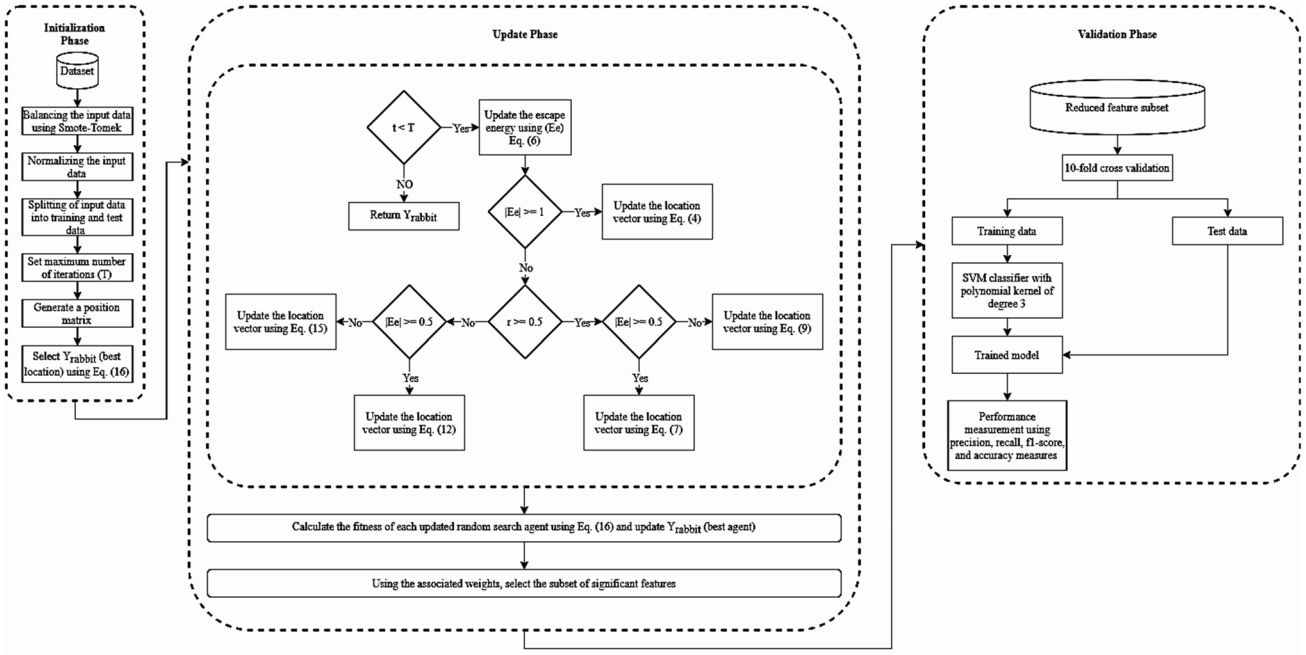

Fig. 8 The proposed approach's process flow

During the updating phase, the iHHO updates the *Yrabbit* and evaluates the fitness value associated with it. The proposed model returns the best search agent with the lowest objective function value at the end of each iteration. The weight of each feature of input data is indicated by the unsurpassed search agent, which aids in determining the significance of the corresponding feature. The Support Vector Machine (SVM) classifier with the polynomial kernel function of degree three verifies the validity of the selected feature subset during the classification phase. Training and testing are the two divisions of the reduced feature set used in the tenfold cross-validation method (CV). Using a tenfold CV minimizes the probability of overfitting the predictive model. For each epoch, nine out of ten input data blocks are used as a training group, and one set is used as a testing set. All results are recorded and compared based on the average performance of optimizers over 30 independent runs in which each run includes 100 iterations. The performance of the proposed approach is measured using precision, recall, f1-score, and accuracy measures, in addition to the ROC-AUC curve.

6 Experimental results and discussion

From Sect. 6.1, the experimental results and an intricate discussion highlight the proposed technique's effectiveness.

6.1 Performance on high dimensional microarray datasets

This phase focuses on how the suggested iHHO is applied to six real high-dimensional microarray datasets. The outcomes of iHHO are paralleled to numerous existing optimization techniques as proposed in earlier researches. Particulars of several microarray datasets used in this research critique are presented in Table 1.

6.1.1 Comparison based on converging ability

The unique potentialities and efficacies of the suggested model are assessed employing the cross-entropy objective function, which calculates the error rate in each iteration. Every outcome of this research experiment is documented and paralleled, considering the midline performance of optimizers over 30 independent runs in which each run includes 100 iterations. The ability of the proposed model to converge to global minima is demonstrated by the decrease in error rate with each iteration. Figure 9 compares the converging ability of iHHO to that of conventional HHO for three microarray datasets. The results and performance of the proposed iHHO is compared with other well-established optimization techniques such as the MFO (Mirjalili 2015), MPA (Faramarzi et al. 2020), SCA (Mirjalili 2016), SSA (Mirjalili et al. 2017), WOA (Mirjalili and Lewis 2016), and HHO algorithms based on six different microarray datasets. The proposed iHHO's optimum convergence rate against

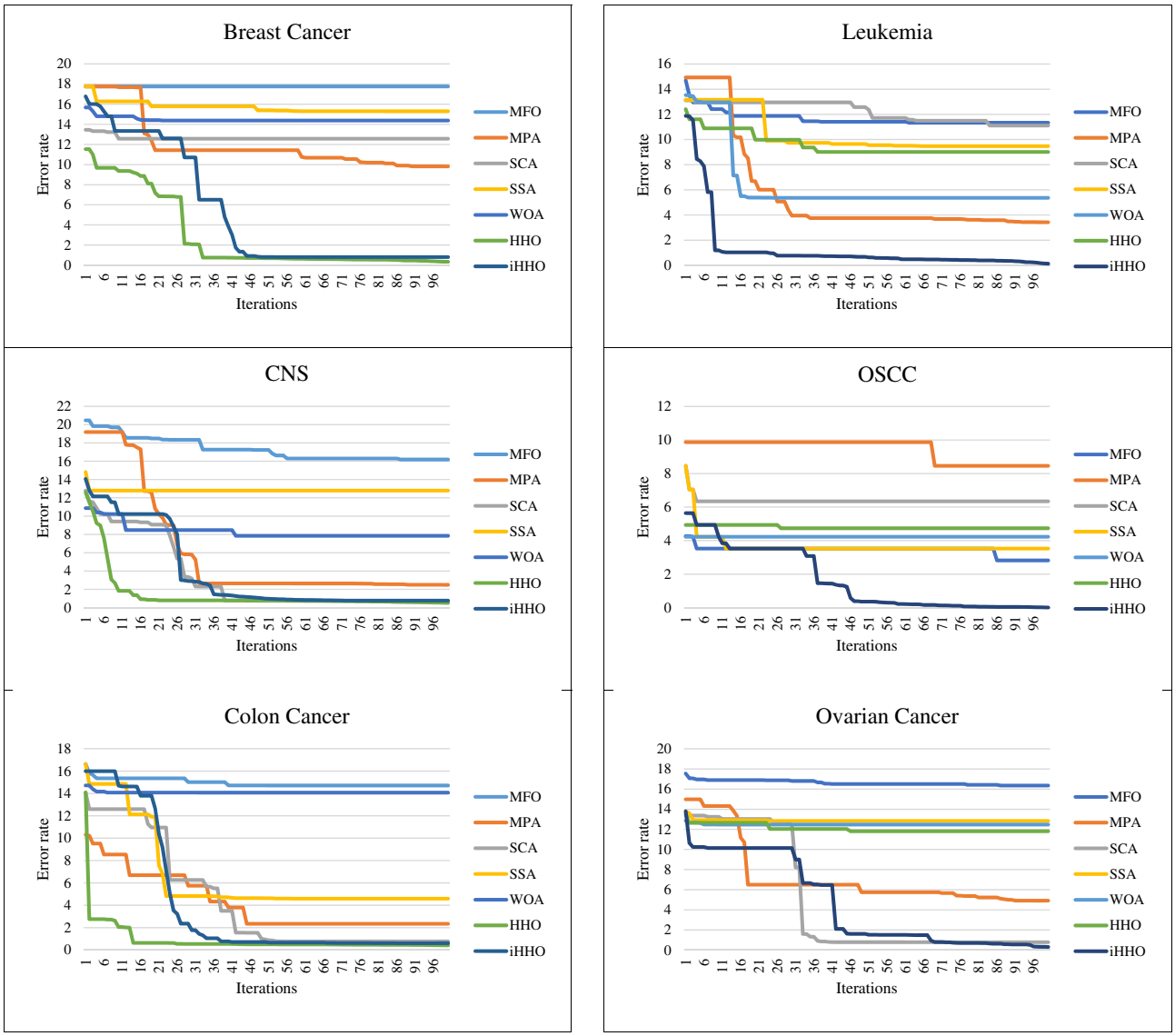

Fig. 9 Convergence curve

the global minimum demonstrates its efficacy in discovering and locating a better solution. Following the convergence curves in Fig. 9, it could be analyzed that the proposed iHHO reaches the global minima. In contrast, the other conventional techniques failed in providing an optimum solution even at the 100th iteration. The projected movement that the iHHO transits from exploration to exploitation can also be detected. Also, it is noted that the iHHO can expose an augmented convergence trend. It could be analyzed that the proposed iHHO is free from premature convergence, which is an ailment of MH optimization which avoid algorithm from achieving the optimum output. The proposed iHHO requires more than 50 epochs to reach the global minima, whereas traditional HHO requires only 20 iterations in the majority of the datasets. This requirement of epochs indicates that HHO is experiencing premature convergence, whereas iHHO is devoid of the same.

6.1.2 Comparison based on training accuracy

The comparison based on training accuracy between the proposed iHHO and other conventional optimizers is illustrated in Fig. 10. In contrast to the different approaches, the proposed model has a higher potential for increased accuracy in the early stages. From Fig. 10, it can be understood that compared to the other algorithms, there is a gradual increase and noteworthy development

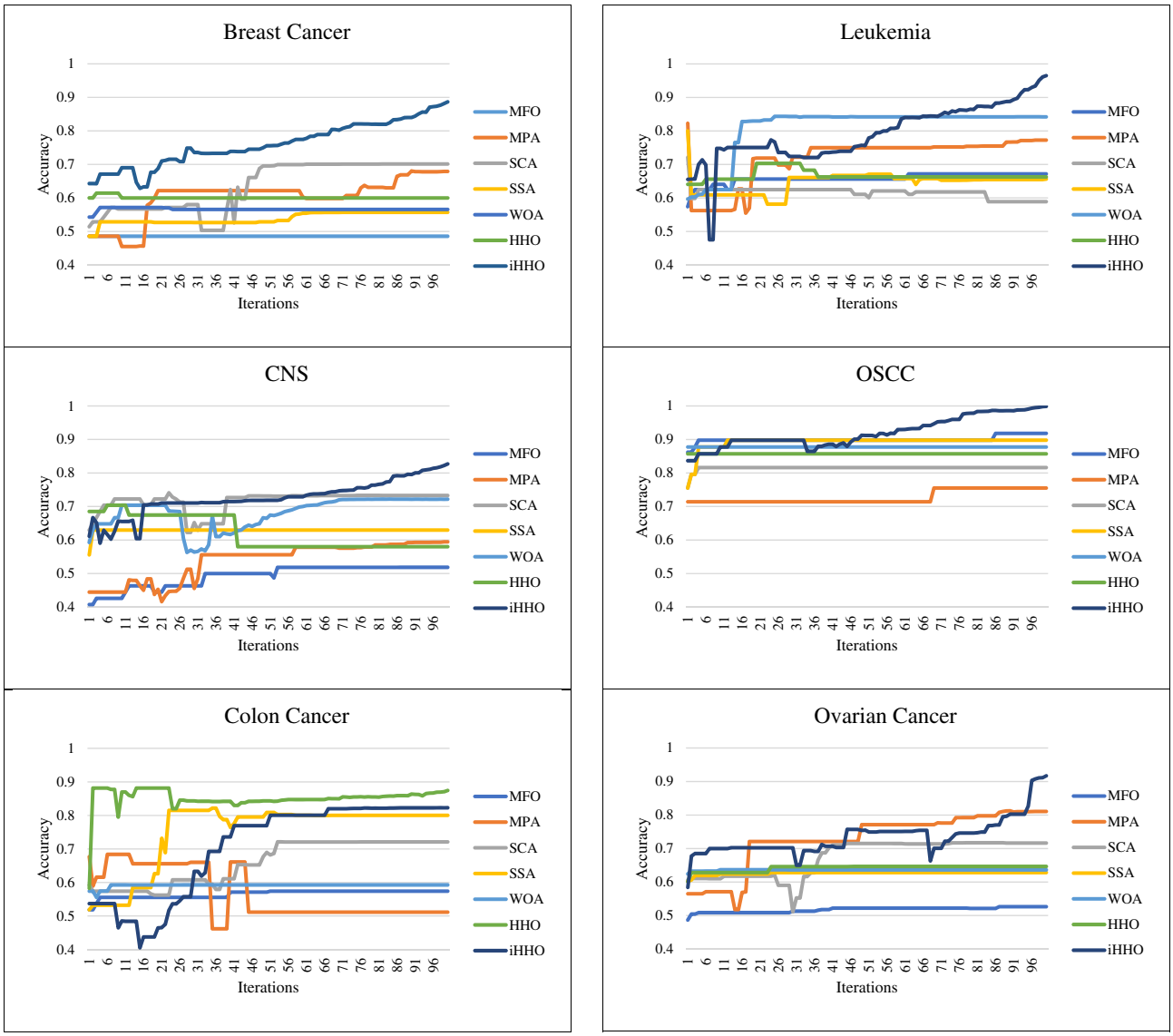

Fig. 10 Training accuracy throughout 100 epochs

in the accurateness of the proposed iHHO. The precision value of the suggested model ranges between (90–100%), whereas values of other techniques range around (50–70%). In the majority of the dataset, the proposed iHHO shows significant improvement. However, in the case of colon cancer, though the traditional HHO demonstrates higher accuracy over the proposed approach, it is to be noted that there is no increase in the accuracy value after the initial iterations.

6.1.3 Comparison based on unseen test data

In this phase, an assessment of the unseen data is utilized to investigate the credibility of the iHHO in handling the unknown data. The enactment of the proposed iHHO is equated with other accustomed optimization techniques based on the box plots, violin plots, heat maps, Standard Deviation (STD), and an average of the test results (AVG). The non-parametric Wilcoxon sum rank test, commonly known as the Mann Whitney U statistical test with a 5% degree of significance, is carried out alongside the experimental evaluations to trace the substantial variances among the obtained outcomes of varied nuances. The research hypothesis of the Wilcoxon sum rank test states that there is a noteworthy difference amongst the two groups.

The performance of the proposed iHHO for several microarray test datasets is demonstrated in Figs. 11, 12, and 13. A box plot in Fig. 11 utilizes boxes and lines to illustrate the distributions of one or more groups of numeric data. Box limits point to the central 50% of the data range, with a

Table 2 Comparison of average accuracy results over 30 runs

Dataset	Metric	MFO	MPA	SCA	SSA	WOA	HHO	iHHO
Breast Cancer	AVG	0.4929	0.6123	0.6367	0.4487	0.5313	0.5640	**0.7236**
	STD	0.0000	0.0534	0.0755	0.0236	0.0047	0.0121	0.0531
CNS	AVG	0.4614	0.5229	**0.6535**	0.5853	0.5862	0.6964	0.6352
	STD	0.0197	0.0535	0.0566	0.0037	0.0188	0.0404	0.0532
Colon Cancer	AVG	0.5745	0.5607	0.5970	**0.7504**	0.5234	0.8117	0.6646
	STD	0.0057	0.0282	0.1013	0.1052	0.0105	0.0333	0.1034
Leukemia	AVG	0.5674	0.6569	0.6112	0.5392	0.7500	0.6651	**0.7717**
	STD	0.0171	0.0485	0.0175	0.0331	0.0683	0.0573	0.0965
OSCC	AVG	0.8553	0.6894	0.8116	0.8476	0.7959	0.8281	**0.9521**
	STD	0.0039	0.0016	0.0083	0.0304	0.0000	0.0315	0.0208
Ovarian Cancer	AVG	0.5408	0.7141	0.6765	0.6346	0.6193	0.6091	**0.7566**
	STD	0.0100	0.1006	0.0515	0.0025	0.0024	0.0099	0.0479

Bold values represent the significant difference between the proposed iHHO and other conventional techniques

Table 3 p values of the Mann Whitney U test with 5% significance

Datasets	Metric	MFO	MPA	SCA	SSA	WOA	HHO
Breast Cancer	U val	0	639	1799	0	0	0
	p val	5.60E−39	1.02 E−26	5.02 E−15	2.28 E−34	2.01 E−34	1.03 E−37
	RBC	1	0.8722	0.6402	1	1	1
CNS	U val	0	457	6001.5	2037.5	2131	8892
	p val	1.77 E−34	1.06 E−28	0.014378	1.01 E−14	1.13 E−12	1.90 E−21
	RBC	1	0.9086	− 0.2003	0.5925	0.5738	− 0.7784
Colon Cancer	U val	100	86	63	2559	0	115
	p val	1.43 E−33	6.14 E−34	1.26 E−33	2.43 E−09	2.22 E−34	7.31 E−33
	RBC	0.98	0.9828	0.9874	0.4882	1	0.977
Leukemia	U val	450	839	541	453	4326	739
	p val	7.75 E−29	2.11 E−24	6.99 E−28	9.94 E−29	0.099742	3.33 E−26
	RBC	0.91	0.8322	0.8918	0.9094	0.1348	0.8522
OSCC	U val	6260	300	300	300	2822	10,000
	p val	4.94 E−07	7.98 E−38	2.52 E−38	6.94 E−35	9.44 E−09	1.62 E−37
	RBC	− 0.252	0.94	0.94	0.94	0.4356	− 1
Ovarian Cancer	U val	9981	1698	1158	0	77	0
	p val	5.92 E−35	2.74 E−16	2.01 E−21	3.46 E−37	1.94 E−34	4.11 E−35
	RBC	− 0.9962	0.6604	0.7684	1	0.9846	1

central line denoting the median value. Lines outspread from each box to seize the range of the balanced data, with dots placed past the line edges to specify outliers. Box plots are the best source for a comparative analysis between the two groups, and they are precise in data summarization. Compared to other optimizers, the accuracy value produced by the proposed iHHO in 100 iterations over 30 independent runs has symmetric distribution in the majority of datasets. Except for CNS and colon cancer, the proposed optimizer evenly distributes outliers on either side of the box. The average and standard deviation values in Table 2 show that iHHO performs significantly superior to other algorithms in handling with 66.66% of six datasets, demonstrating the optimizer's superior performance. Concerning *p* values in Table 3, it is identified that the elucidations of iHHO are suggestively better than those analyzed by further procedures in almost all cases.

Off to the side of the box median rather than in the middle, as well as an inequity in whisker lengths, where one cross is short with no outliers, and the other has a long tail with myriads of outliers of other optimizers, this indicates that the distribution is skewed. Box plots provide only a high-level data summary and cannot present the particulars of a data distribution's shape. A box plot's effortlessness also contributes to the limits on the data density that it can demonstrate. It blindfolds the capability to detect the meticulous

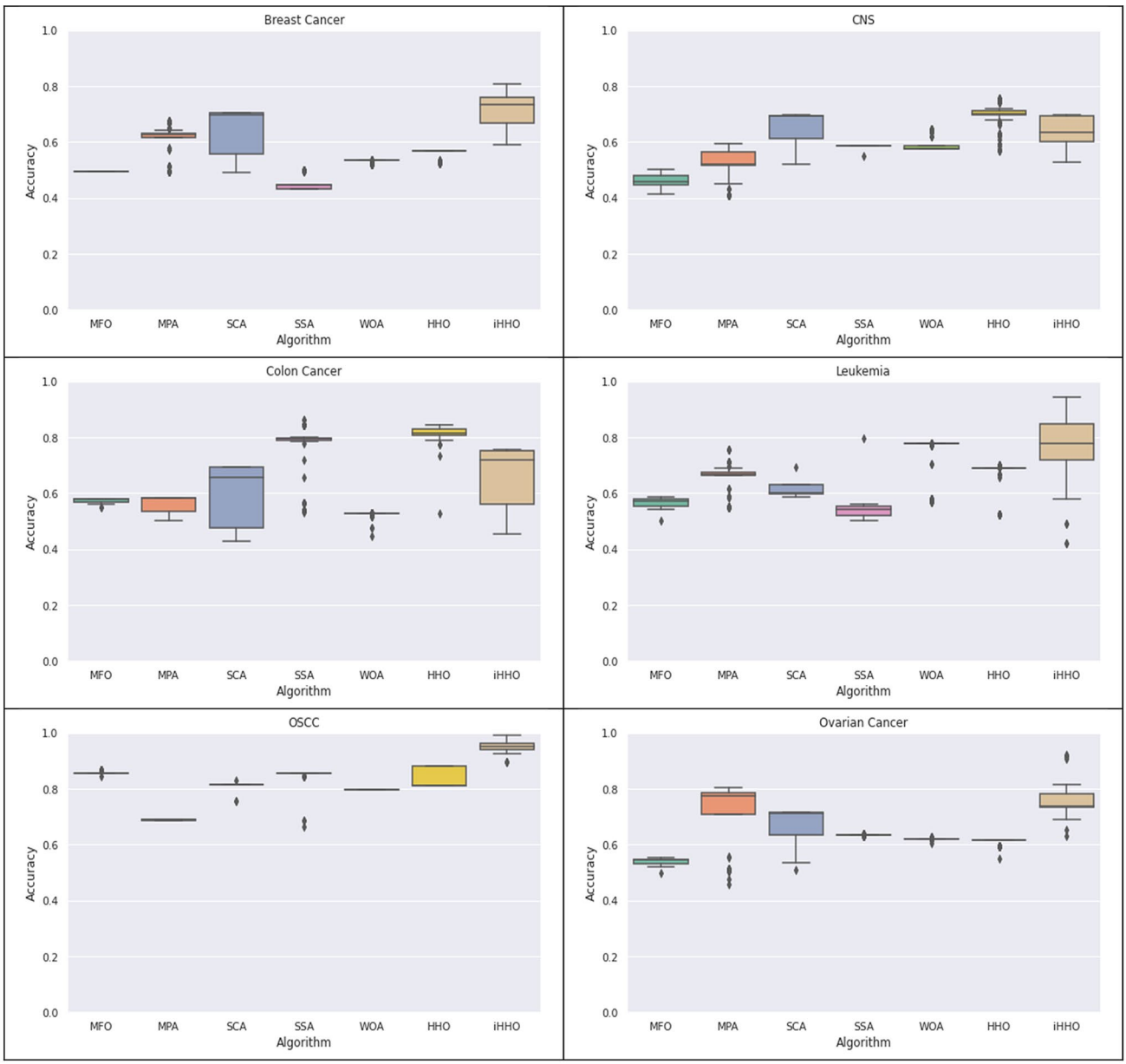

Fig. 11 Test accuracy assessment using box plots

shape of the dissemination, like, if there are oddities in a distribution's modality (number of 'humps' or peaks) and skew.

The violin plots are used to assess the proposed iHHO in Fig. 12 to overcome the drawback of box plots. A violin plot illustrates distributions of numeric data for one or more groups using density curves. The width of each curve corresponds with the approximate frequency of data points in each region. Densities are frequently accompanied by an overlaid chart type, such as a box plot, to provide additional information. In the middle of each density curve is a small box plot, with the rectangle showing the ends of the first and third quartiles and the central dot the median. The plots in Fig. 12 show that as more data points are added to a region, the height of the density curve in that area of proposed iHHO increases in the majority of datasets. The proposed optimizer has a much more curtailed distribution compared to the other conventional techniques.

The heatmaps in Fig. 13 indicate the variations in the accuracy value throughout the 100 iterations. A sequential color ramp between value and color shows that the lighter colors correspond to larger values and darker colors to smaller values. In support of the evidence provided by Tables 2 and 3, the range of accuracy values produced by the proposed iHHO in Breast cancer, Leukemia, OSCC, and

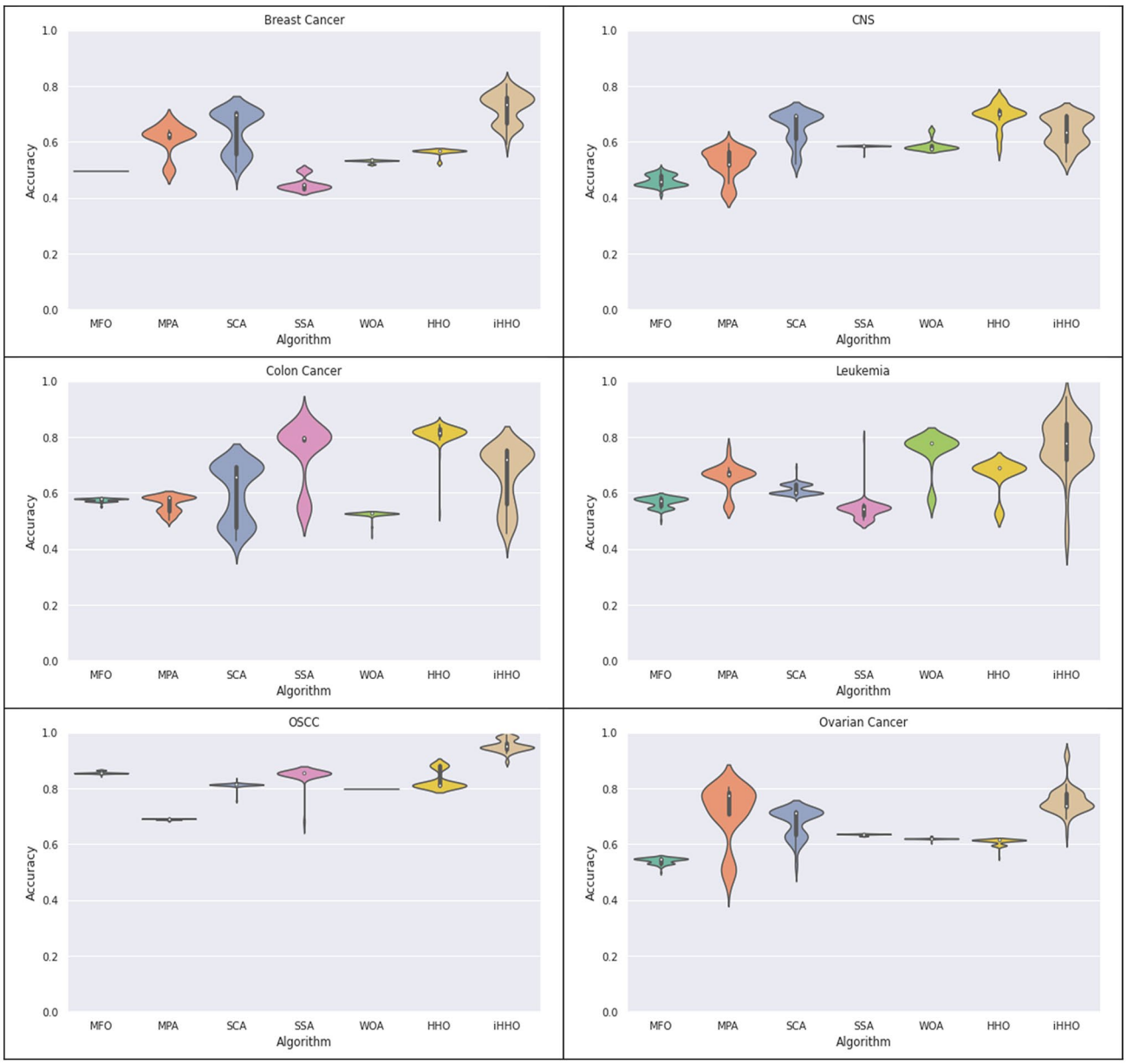

Fig. 12 Test accuracy assessment using violin plots

Ovarian cancer are much higher than other well-established optimizers.

6.1.4 Performance analysis of the selected features subset

The Receiver Operating Characteristic (ROC) curve scrutinizes the performance of a classification model at various threshold settings to determine the effectiveness of a predictive model in evaluating the data sample class. Figure 14 compares the performance of the features selected by the proposed iHHO with other optimizers for different microarray datasets. The proposed optimizer shows significant improvement over other methods in CNS, Colon Cancer, Leukemia, and OSCC. In addition to the ROC-AUC score, the precision, recall, f1-score, and accuracy metrics are used to validate the performance of the selected feature, as shown in Table 4. In the case of Breast and Ovarian cancer, the proposed iHHO yields 50% and 60% accuracy values, respectively. Even though the outcome in Breast and Ovarian cancer is not significant, it is the highest accuracy value shown in that category. However, the proposed iHHO produces precision scores ranging from 81 to 96% in most datasets. A Higher ROC-AUC score of iHHO indicates greater confidence in the predictive potential of selected

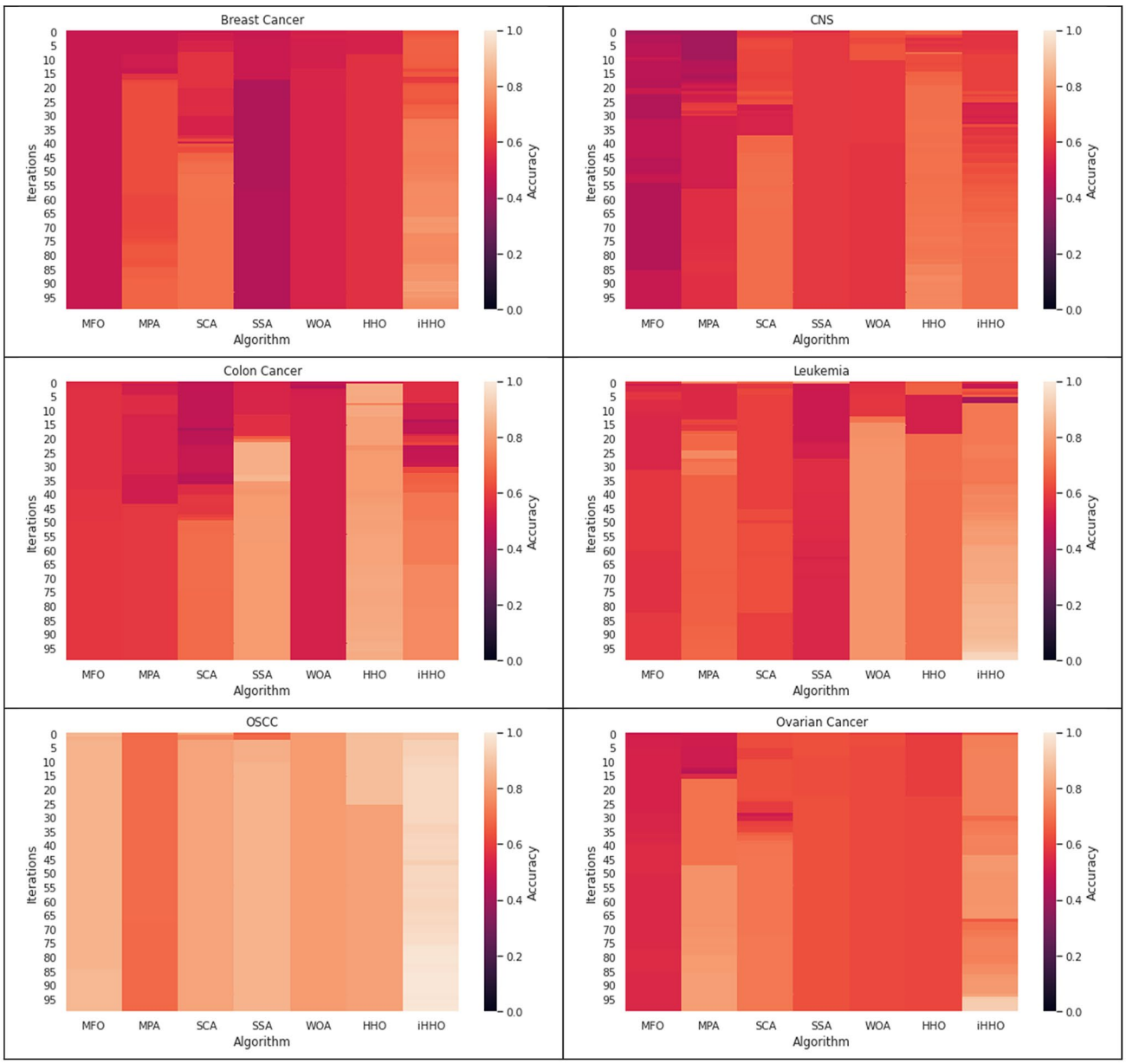

Fig. 13 Test accuracy assessment using heatmaps

features. The higher the AUC value, the less likely the proposed model will reverse the effects and have a high level of separability. The features chosen by iHHO provide greater confidence in disease prediction and higher accuracy across all six datasets.

6.2 Performance-based on unimodal and multimodal functions

The effectiveness of the suggested iHHO optimizer is thoroughly analyzed using a set of varied benchmark functions. The recommended benchmark functions encompass a set of Unimodal (UM), and Multimodal (MM) functions. The UM functions (F1–F4) can disclose the intensification abilities of diverse optimizers. The MM functions (F5–F8) can divulge diversification. Tables 5 and 6 represent the mathematical formulation of UM and MM problems. The outcomes of the suggested iHHO are compared with the existing optimization techniques such as the MFO, MPA, SCA, SSA, WOA, and HHO algorithms.

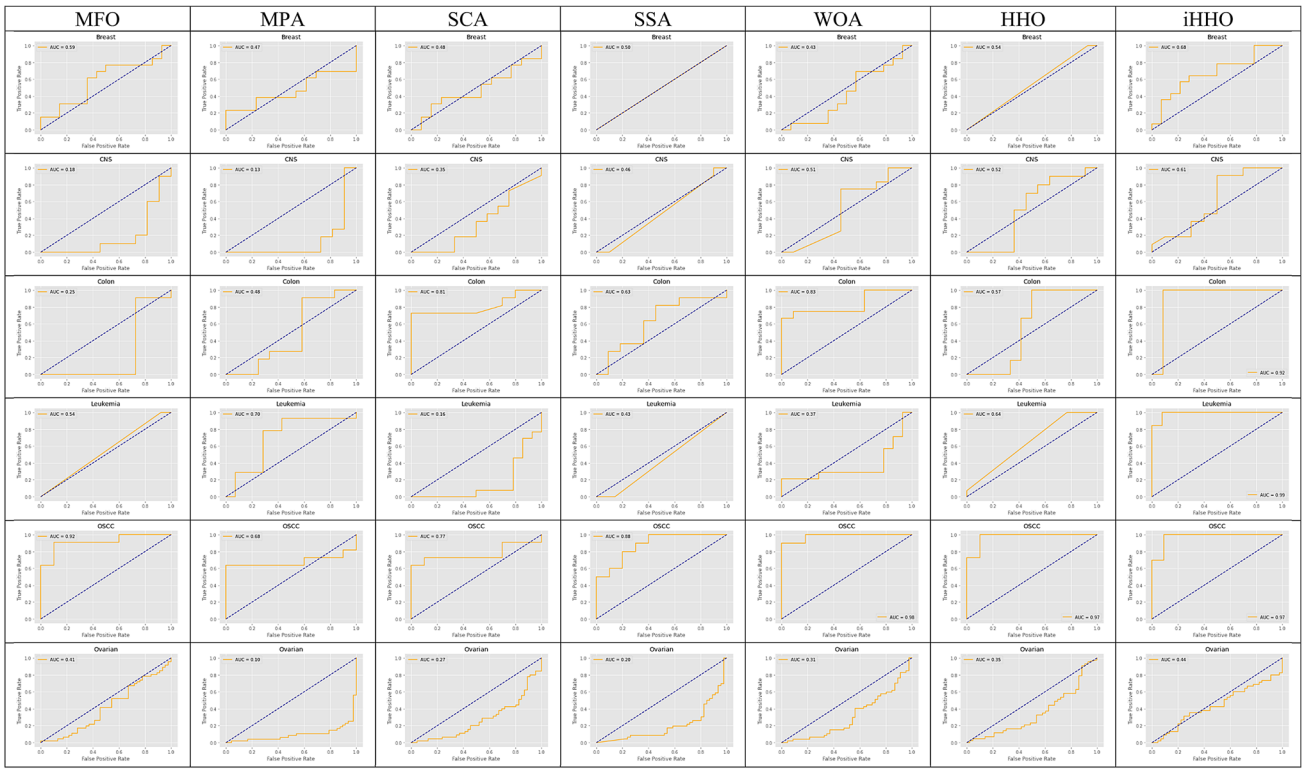

Fig. 14 ROC-AUC Curve of Selected Features

6.2.1 Scalability analysis

Scalability evaluation is employed to reconnoiter the effect of dimension on the outcome of iHHO. This test can reveal the impact of dimensions on the quality of solutions for the iHHO optimizer to identify its effectiveness for problems with lower and higher dimension tasks. This research employs the iHHO to handle the scalable UM and MM F1–F8 test cases with 30, 100, 500, and 1000 dimensions. Table 7 represents the analysis outcomes of the iHHO in handling F1-F8 problems with different dimensions. Table 7 also depicts that the iHHO can expose exceptional results, which remain consistent in all dimensions.

6.2.2 Comparative analysis

A comparative analysis for diverse dimensions of F1–F8 benchmark functions is done between the proposed iHHO and the conventional optimizers to highlight the significant performance of the iHHO, which is demonstrated in Tables 8, 9, 10, and 11. According to Table 8, the iHHO achieves the best results in dealing with 87.5% of 30-dimensional functions (F1–F3 and F5–F8), demonstrating the superior performance of the iHHO compared to existing optimizers. According to Table 9, the iHHO can substantially outperform other techniques and attain the best results for 87.5% (F1–F6 and F8) of 100-dimensional search space problems. Table 10 demonstrates that the iHHO achieves the paramount results in AVG and STD in dealing with all 8 test cases with 500 dimensions. Table 11 presents that iHHO performs significantly superior when dealing with F1–F8 test functions with 1000 dimensions. It could be observed from the results of the formulated tables that the performance of the conventional methods diminishes with the increase of dimensions which reveals the potentiality of the iHHO in consistently balancing investigative and intensification tendencies.

6.3 Comparative analysis of HHO variants

The comparison of accuracy between traditional HHO, HHO with Brownian motion, HHO with proposed control factor, and HHO with Brownian and proposed control factor is illustrated in Fig. 15. Compared to other techniques, the suggested model has a greater chance of improving accuracy in the initial stages. As depicted in Fig. 15, the accuracy of the suggested iHHO algorithm has been steadily increasing and developing in comparison to the other methods. From Fig. 15, it is vividly understood that the proposed iHHO outperforms other variants of HHO in terms of classification accuracy by producing accuracy in the range of [0.6 to 0.9]. The combination of HHO with Brownian motion and HHO

Table 4 Validation of Selected Features

Dataset	Breast Cancer				CNS			Colon Cancer			Leukemia			OSCC			Ovarian Cancer			
Algorithm	Class	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
MFO	0	0.00	0.00	0.00	0.4805	0.75	0.30	0.43	0.6258	1.00	0.55	0.71	0.7726	1.00	0.62	0.76	0.8109	0.89	0.80	0.84	0.8669	0.50	1.00	0.67	0.5051
	1	0.48	1.00	0.65		0.59	0.91	0.71		0.69	1.00	0.81		0.74	1.00	0.85		0.83	0.91	0.87		0.00	0.00	0.00	
MPA	0	0.00	0.00	0.00	**0.5092**	0.75	0.82	0.78	0.7709	0.56	0.42	0.48	0.5295	0.73	0.57	0.64	0.6807	0.71	1.00	0.83	0.8154	0.80	0.17	0.28	0.5695
	1	0.50	1.00	0.67		0.80	0.73	0.76		0.50	0.64	0.56		0.65	0.79	0.71		1.00	0.64	0.78		0.53	0.96	0.69	
SCA	0	0.00	0.00	0.00	**0.5092**	0.25	0.08	0.13	0.3925	0.77	1.0	0.87	0.8620	0.91	0.71	0.80	0.8198	0.71	1.00	0.83	0.8154	0.00	0.00	0.00	0.4965
	1	0.50	1.00	0.67		0.42	0.73	0.53		1.00	0.73	0.84		0.75	0.92	0.83		1.00	0.64	0.78		0.49	1.00	0.66	
SSA	0	0.48	1.00	0.65	0.4809	1.00	0.40	0.57	0.7139	0.71	0.45	0.56	0.6429	0.92	0.79	0.85	0.8678	0.67	1.00	0.80	0.7555	0.50	1.00	0.67	0.5094
	1	0.00	0.00	0.00		0.65	1.00	0.79		0.60	0.82	0.69		0.81	0.93	0.87		1.00	0.50	0.67		0.00	0.00	0.00	
WOA	0	0.00	0.00	0.00	0.4839	0.63	0.45	0.53	0.6118	0.77	0.91	0.83		0.69	0.79	0.73	0.7120	0.92	1.00	0.96	**0.9509**	0.49	1.00	0.66	0.4937
	1	0.48	1.00	0.65		0.60	0.75	0.67		0.90	0.75	0.82	0.8398	0.75	0.64	0.69		1.00	0.90	0.95		0.00	0.00	0.00	
HHO	0	0.00	0.00	0.00	0.4572	0.60	0.55	0.57	0.5857	1.00	0.50	0.67	0.7507	0.80	0.62	0.70	0.7404	0.90	0.90	0.90	0.9095	0.51	1.00	0.67	0.5158
	1	0.50	1.00	0.67		0.55	0.60	0.57		0.67	1.00	0.80		0.71	0.86	0.77		0.91	0.91	0.91		1.00	0.02	0.05	
iHHO	0	0.46	0.93	0.62	**0.5092**	0.89	0.73	0.80	**0.8189**	1.00	0.92	0.96	**0.9609**	1.00	0.92	0.96	**0.9636**	1.00	0.91	0.95	**0.9509**	0.76	0.28	0.41	**0.6084**
	1	0.50	1.00	0.67		0.75	0.90	0.82		0.92	1.00	0.96		0.93	1.00	0.96		0.91	1.00	0.95		0.56	0.91	0.69	

Table 5 Description of unimodal benchmark functions

Functions	Dimensions	Range	f_{min}
$f_1(x) = \sum_{i=1}^{n} x_i^2$	30, 100, 500, 1000	[100, 100]	0
$f_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$	30, 100, 500, 1000	[10, 10]	0
$f_3(x) = \sum_{i=1}^{n} \left(\sum_{j-1}^{i} x_j \right)^2$	30, 100, 500, 1000	[100, 100]	0
$f_4(x) = \max_i \left\{ \lvert x_i, 1 \le i \le n \rvert \right\}$	30, 100, 500, 1000	[100, 100]	0

with novel control factor performs almost symmetrical but shows meagre accuracy compared to iHHO. The traditional HHO performs measly as compared to other approaches in all datasets.

6.4 Complexity analysis

The time complexity of the suggested method is analysed in this section. O(n*d) time–space is required to construct the criteria and initialize the random population. The population size is *n*, and the number of dimensions based on the dataset is *d*. The update phase requires O(n*d). Moreover, the outer loop has an O(t) time complexity, where *t* is the maximum number of iterations. Furthermore, the proposed approach complexity is O(t*n*d).

7 Results discussion

From the results, it can be inferred that the results of iHHO are significantly superior for multi-dimensional F1–F8 problems. Six real high-dimensional microarray datasets are juxtaposed with existing and renowned optimizers such as MFO, MPA, SCA, SSA, WOA, and HHO methods. The efficiency of other methods significantly degrades as the algorithm progresses the converging ability and accuracy. Figures 9, 10 represents how iHHO can retain a balanced equilibrium amidst the exploratory and exploitative propensities related to topographies with a plethora of variables. As recorded in F1–F8 in Tables 8, 9, 10, 11, a vast significant fissure can be observed in the outcome of varied nuances such as the MFO, MPA, SCA, SSA, WOA, and HHO, with high-quality solutions based on iHHO. This analysis endorses the progressive exploitative merits of the suggested iHHO. From the solution noted for unseen data in Figs. 11–13 and Tables 2, 3, it can be traced that iHHO discovers superior and competitive results grounded on an equivocal balance amidst exploration and exploitation predispositions and an unwavering shift linking the searching modes. The outcomes of the experiment exemplify that the suggested iHHO has a multitude of exploratory and exploitative strategies. In

Table 6 Description of multimodal benchmark functions

Functions	Dimensions	Range	f_{min}
$f_5(x) = \sum_{i=1}^{n} -x_i \sin\left(\sqrt{\lvert x_i \rvert}\right)$	30, 100, 500, 1000	[500, 500]	$-418.98 \times n$
$f_6(x) = \sum_{i=1}^{n} \left[x_i^2 - 10\cos(2\Pi x_i) + 10\right]$	30, 100, 500, 1000	[5.12, 5.12]	0
$f_7(x) = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\Pi x_i)\right) + 20 + e$	30, 100, 500, 1000	[32, 32]	0
$f_8(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30, 100, 500, 1000	[600, 600]	0

Table 7 Results of iHHO for different dimensions of F1–F8 functions

Problem/dimension		Metric	30	100	500	1000
Unimodal	F1	Avg	3.56E−85	2.93 E−79	1.92 E−62	1.49 E−44
		Std	1.02 E−84	5.53 E−79	6.10 E−62	3.26 E−44
	F2	Avg	7.19 E−92	6.85 E−87	4.66 E−72	3.32 E−56
		Std	9.35 E−91	1.43 E−86	1.06 E−71	2.91 E−55
	F3	Avg	3.92 E−89	5.90 E−92	4.53 E−88	2.35 E−63
		Std	8.37 E−89	6.52 E−92	3.62 E−87	1.09 E−62
	F4	Avg	7.32 E−51	5.40 E−69	4.95 E−62	3.69 E−58
		Std	1.09 E−50	2.77 E−69	1.92 E−62	1.62 E−57
Multimodal	F5	Avg	− 2.05E + 16	− 7.45E + 16	− 1.52E + 20	− 1.06E + 17
		Std	6.68E + 15	1.35E + 15	2.31E + 20	6.38E + 16
	F6	Avg	4.68 E−38	6.28 E−16	5.89 E−22	4.61 E−13
		Std	7.61 E−38	1.64 E−16	2.67 E−21	6.62 E−12
	F7	Avg	5.21 E−09	3.69 E−05	2.34 E−24	1.62 E−17
		Std	4.02 E−08	9.26 E−05	4.72 E−24	3.92 E−17
	F8	Avg	1.28 E−65	6.89 E−21	3.62 E−18	1.62 E−08
		Std	7.09 E−64	4.02 E−21	1.94 E−18	6.69 E−08

Table 8 Results and comparisons of unimodal and multimodal benchmark functions with 30 dimensions

Problem/dimension		Metric	MFO	MPA	SCA	SSA	WOA	HHO	iHHO
Unimodal	F1	Avg	2.05 E−67	9.98E + 01	1.73E + 04	1.46E + 02	2.74E + 01	9.12 E−05	**3.56** E−**85**
		Std	1.68 E−66	5.67E + 01	6.89E + 03	4.90E + 01	3.67E + 01	3.07 E−04	1.02 E−84
	F2	Avg	3.95 E−23	8.83E + 01	8.30E + 00	4.82E + 01	7.15 E−01	1.56 E−51	**7.19** E−**92**
		Std	6.98 E−22	1.07E + 01	2.89E + 03	1.90E + 01	6.28 E−02	8.55 E−50	9.35 E−91
	F3	Avg	2.39 E−04	6.94 E−02	1.25E + 02	5.83 E−71	3.97E + 05	5.96 E−88	**3.92** E−**89**
		Std	1.08 E−03	2.95 E−02	7.89E + 01	9.31 E−71	2.12E + 05	2.36 E−87	8.37 E−89
	F4	Avg	5.01 E−02	**2.98** E−**53**	7.98 E−23	1.36E + 01	6.06 E−27	2.91 E−49	7.32 E−51
		Std	1.53 E−02	7.62 E−53	2.67 E−22	1.10E + 01	1.92 E−26	5.62 E−48	1.09 E−50
Multimodal	F5	Avg	− 1.07E + 02	− 1.09E + 05	− 5.98 E−04	− 7.29E + 02	− 2.32E + 04	− 1.96E + 10	− **2.05E + 16**
		Std	3.05E + 01	8.38E + 01	7.09 E−04	3.95E + 01	1.96E + 03	4.52E + 11	6.68E + 15
	F6	Avg	1.42 E−09	3.65 E−02	1.96E + 04	6.62 E−09	2.35 E−29	3.36 E−04	**4.68** E−**38**
		Std	6.52 E−08	9.64 E−02	8.72E + 03	1.35 E−08	7.03 E−28	8.96 E−03	7.61 E−38
	F7	Avg	5.69 E−04	8.45E + 01	9.01 E−07	6.75 E−06	2.98E + 04	1.32E + 02	**5.21** E−**09**
		Std	3.81 E−04	1.92E + 01	6.34 E−06	1.68 E−03	3.39E + 04	5.21E + 02	4.02 E−08
	F8	Avg	3.97E + 06	2.39E + 04	2.95 E−09	9.45 E−24	1.94 E−05	5.60 E−08	**1.28** E−**65**
		Std	1.92E + 06	2.09E + 04	1.56 E−08	6.65 E−23	5.50 E−05	4.94 E−07	7.09 E−64

Bold values represent the significant difference between the proposed iHHO and other conventional techniques

Table 9 Results and comparisons of unimodal and multimodal benchmark functions with 100 dimensions

Problem/dimension		Metric	MFO	MPA	SCA	SSA	WOA	HHO	iHHO
Unimodal	F1	Avg	7.86 E−39	3.65E + 03	1.73E + 04	7.91 E−09	4.14 E−03	1.72 E−07	**2.93 E−79**
		Std	3.68 E−38	1.29E + 03	5.36E + 03	2.39 E−08	8.63 E−03	6.91 E−06	5.53 E−79
	F2	Avg	8.65 E−04	2.35E + 01	7.38 E−07	9.92E + 02	3.67E + 01	5.02 E−61	**6.85 E−87**
		Std	9.61 E−04	5.91E + 01	5.60 E−06	4.27E + 04	7.39E + 02	6.94 E−61	1.43 E−86
	F3	Avg	7.35 E−06	2.98E + 02	6.09 E−32	3.36 E−05	1.25 E−30	5.69 E−06	**5.90 E−92**
		Std	3.65 E−07	5.66E + 05	4.95 E−31	9.42 E−04	5.68 E−29	8.32 E−06	6.52 E−92
	F4	Avg	2.37E + 01	6.62 E−04	3.09 E−02	7.50 E−28	7.56 E−59	6.96 E−64	**5.40 E−69**
		Std	6.87E + 01	1.25 E−03	8.08 E−01	3.45 E−27	3.54 E−62	2.35 E−63	2.77 E−69
Multimodal	F5	Avg	− 1.24E + 05	− 7.56E + 15	-6.68E + 10	− − 5.23E + 04	− 1.29E + 07	1.18E + 05	**− 7.45E + 16**
		Std	6.32E + 04	5.94E + 11	8.97E + 09	3.62E + 03	4.26E + 06	1.28E + 04	1.35E + 15
	F6	Avg	9.65E + 01	2.91 E−09	4.29 E−04	2.22E + 02	2.06 E−01	5.19 E−10	**6.28 E−16**
		Std	5.94E + 01	7.91 E−09	1.93 E−03	3.62E + 02	8.39 E−03	7.95 E−08	1.64 E−16
	F7	Avg	2.95 E−01	8.62E + 02	1.62 E−03	6.56E + 01	**4.50 E−09**	6.72E + 01	3.69 E−05
		Std	6.54 E−01	9.65E + 02	7.32 E−02	5.06E + 01	2.62 E−08	2.55E + 01	9.26 E−05
	F8	Avg	2.84 E−01	5.02 E−14	5.24 E−18	4.02E + 02	6.30 E−03	4.92 E−06	**6.89 E−21**
		Std	8.92 E−01	4.93 E−14	2.93 E−18	6.83E + 02	2.96 E−03	1.97 E−05	4.02 E−21

Bold values represent the significant difference between the proposed iHHO and other conventional techniques

Table 10 Results and comparisons of unimodal and multimodal benchmark functions with 500 dimensions

Problem/dimension		Metric	MFO	MPA	SCA	SSA	WOA	HHO	iHHO
Unimodal	F1	Avg	5.92E + 01	3.51 E−04	8.23E + 02	2.38 E−23	6.24 E−29	7.55E + 01	**1.92 E−62**
		Std	1.25E + 01	6.56 E−03	7.19E + 02	5.09 E−22	4.95 E−28	3.66E + 02	6.10 E−62
	F2	Avg	3.51 E−04	6.16 E−02	5.57 E−32	8.56 E−15	2.69E + 00	6.98E + 02	**4.66 E−72**
		Std	2.60 E−04	3.58 E−02	4.92 E−31	7.95 E−14	4.62E + 00	2.69E + 01	1.06 E−71
	F3	Avg	1.09 E−19	5.91 E−02	6.82E + 01	8.11 E−02	7.62 E−04	4.01 E−27	**4.53 E−88**
		Std	5.25 E−18	3.71 E−01	4.92E + 01	6.62 E−02	4.69 E−04	3.62 E−26	3.62 E−87
	F4	Avg	7.62 E−02	5.36E + 01	3.93 E−29	4.87 E−24	6.02E + 02	5.68 E−13	**4.95 E−62**
		Std	5.62 E−02	2.26E + 01	1.78 E−29	8.66 E−24	5.91E + 01	3.69 E−12	1.92 E−62
Multimodal	F5	Avg	− 1.24 E−14	− 1.65 E−24	− 2.26E + 01	− 4.92E + 12	− 1.35 E−15	− 1.24 E−18	**− 1.52E + 20**
		Std	4.25 E−14	3.26 E−24	7.55E + 01	2.15E + 12	6.56 E−15	5.62 E−18	2.31E + 20
	F6	Avg	8.26E + 01	7.52 E−05	6.90E + 02	1.25 E−16	7.62 E−14	8.10E + 02	**5.89 E−22**
		Std	5.56E + 02	2.98 E−04	9.58E + 02	4.56 E−15	3.25 E−13	2.25E + 02	2.67 E−21
	F7	Avg	2.32E + 01	7.93 E−02	1.46 E−19	3.74 E−01	7.24E + 04	3.56 E−02	**2.34 E−24**
		Std	5.92E + 01	3.48 E−01	4.74 E−18	8.27 E−01	3.89E + 04	6.40 E−01	4.72 E−24
	F8	Avg	6.98 E−04	5.38E + 02	8.34 E−06	9.65 E−04	6.62E + 01	8.26 E−02	**3.62 E−18**
		Std	2.49 E−04	4.28E + 02	3.62 E−06	5.62 E−04	2.95E + 01	5.26 E−02	1.94 E−18

Bold values represent the significant difference between the proposed iHHO and other conventional techniques

addition, handling diverse classes of problems has tactfully eluded LO and immature convergence. The recommended iHHO has demonstrated a better ability in leaping out of local optimum solutions in any Local Optima stagnation. Features stated below would assist in comprehending various theoretical reasons that substantiate the constructive nature of the recommended iHHO in exploring or exploiting the search space of a given optimization problem:

- Modified escaping energy (*Ee*) using novel control factor (*CF*) parameter represents smooth deprivation of escaping energy of the prey as it tries to escape from the predator. It not only requires an even shift between exploration and exploitation patterns of iHHO but also enhances it.
- When piloting a local search, diversification mechanisms such as Brownian motion (f_B) eventually preclude the algo-

Table 11 Results and comparisons of unimodal and multimodal benchmark functions with 1000 dimensions

Problem/dimension		Metric	MFO	MPA	SCA	SSA	WOA	HHO	iHHO
Unimodal	F1	Avg	6.89 E−04	3.81 E−02	7.82E+01	5.62 E−35	4.98E+09	1.09E+02	**1.49 E−44**
		Std	7.26 E−03	5.26 E−02	2.54E+01	6.29 E−34	2.95E+07	6.25E+01	3.26 E−44
	F2	Avg	2.61E+02	6.23E+04	3.64 E−02	9.56E+01	5.62 E−13	7.29E+10	**3.32 E−56**
		Std	5.65E+01	4.92E+04	1.95 E−02	5.62E+01	3.62 E−12	2.32E+10	2.91 E−55
	F3	Avg	8.92 E−18	3.26 E−04	6.26E+02	4.69 E−10	3.52E+10	9.55E+01	**2.35 E−63**
		Std	2.55 E−18	1.98 E−03	4.55E+02	6.28 E−09	8.14E+10	8.77E+01	1.09 E−62
	F4	Avg	3.62E+10	2.92 E−01	9.56 E−15	2.62E+10	6.29 E−42	7.92 E−56	**3.69 E−58**
		Std	2.60E+10	1.59 E−01	4.29 E−14	7.69E+10	3.27 E−42	5.62 E−55	1.62 E−57
Multimodal	F5	Avg	− 5.34E+15	− 3.68 E−09	− 1.35 E−17	− 7.69 E−04	− 6.25 E−24	− 7.75E+01	**− 1.06E+17**
		Std	9.62E+13	1.64 E−08	4.69 E−16	2.78 E−03	5.62 E−24	3.62E+01	6.38E+16
	F6	Avg	5.16E+01	8.13E+03	7.23 E−09	2.30 E−03	6.29E+10	8.26 E−05	**4.61 E−13**
		Std	2.30E+01	7.33E+03	6.62 E−08	1.18 E−02	5.69E+10	5.95 E−04	6.62 E−12
	F7	Avg	4.26 E−13	6.78E+03	1.16E+12	7.95 E−15	8.69 E−04	5.22E+01	**1.62 E−17**
		Std	8.96 E−11	9.21E+02	2.36E+11	5.26 E−14	1.96 E−03	3.26E+01	3.92 E−17
	F8	Avg	3.25 E−01	8.65E+04	6.62E+10	2.62 E−01	9.28E+10	5.91 E−03	**1.62 E−08**
		Std	6.23 E−01	5.00E+00	1.95E+09	7.28 E−01	4.25E+10	1.56 E−02	6.69 E−08

Bold values represent the significant difference between the proposed iHHO and other conventional techniques

rithm from stagnation at local minima and improves the exploitative nature of iHHO.

- There is a constructive impact on the exploitation ability of iHHO as it employs a sequence of searching stratagems centered on *Ee* and *r* parameters and then selects the best movement step.
- The randomized jump (*J*) potentiality can support optimum solutions in maintaining the tendencies of intensification and diversification.

8 Conclusion

Harris Hawks Optimization is a population-based optimizer enthused by the cooperative nature and hurtling skills of predatory birds, Harris hawks, in nature. This research has stressed the two setbacks associated with traditional HHO. Firstly, the conventional HHO representation of prey escape energy is inefficient. Secondly, Levy flight's smaller and occasionally more extended step sizes in the initial random solution prevented the algorithm from emerging from local optima.

Randomization is a core component of swarm intelligence-based optimization techniques. In this research critique, the randomization is instilled in traditional HHO using the Brownian motion rather than levy flight. Using Brownian motion to update the best location of the initial random solution, the proposed model avoids stagnation in local minima and successfully converges to global optima. In addition, the novel control factor to efficiently imitate the escape energy of the prey is also incorporated. The performance of iHHO is assessed using six publically accessible real high-dimensional microarray datasets. The performance measures such as precision, recall, f1-score, and classification accuracy are used to evaluate the confidence of selected features. From the outcome of this research evaluation presented by different performance measures, it is understood that the feature chosen by the proposed iHHO provides profound insight in detecting life-threatening diseases.

Various unimodal and multimodal problems were employed to scrutinize the evasion of exploitative, exploratory, and local optima by the proposed iHHO. The outcomes of iHHO illustrate that iHHO is proficient in discovering exceptional results paralleled to archetypal optimizers. Furthermore, the effects of six high-dimensional microarray datasets exposed that the iHHO can demonstrate superior results to other optimizers. The proposed randomization using Brownian motion improves the convergence ability and helps to avoid premature convergence. However, like other optimization methods, iHHO has a few shortcomings. The proposed model experiments exclusively for FS problems with a high-dimensional dataset. It can be applied for constrained engineering design tasks for various engineering applications. To conclude, the performance of the Brownian motion has demonstrated a novel way to improve the outcomes of other optimization approaches in the future.

Future works can utilize other evolutionary schemes such as mutation and crossover schemes, evolutionary updating structures, and chaos-based phases to develop binary and multi-objective versions of iHHO. In addition, it can be employed to tackle the problems in image segmentation,

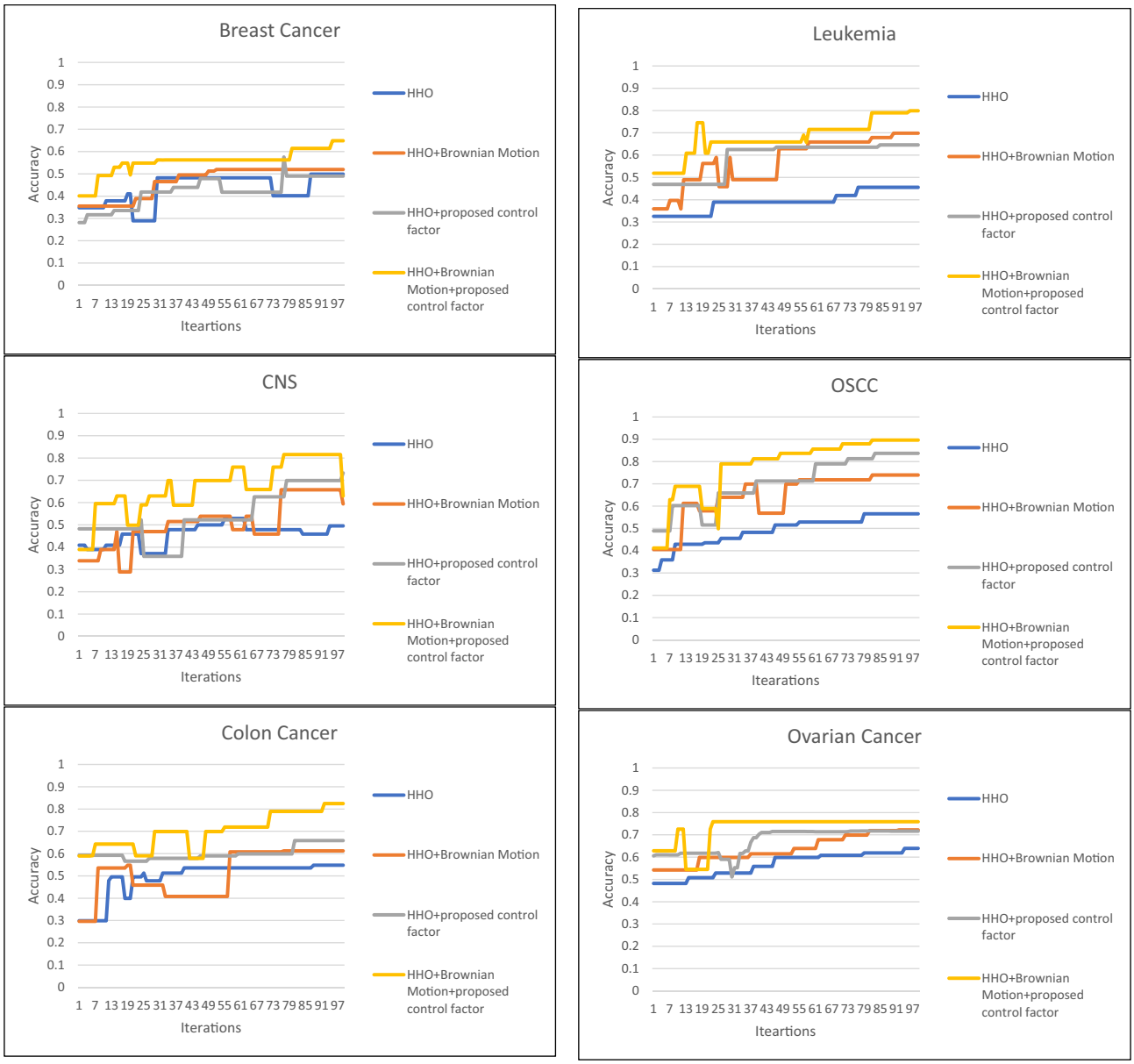

Fig. 15 Comparative analysis of the variants of HHO

sentiment analysis, signal processing, fuzzy system, and different engineering applications.

Acknowledgements The authors sincerely thank the Department of Science and Technology (DST), Government of India, for funding this research project work under the Interdisciplinary Cyber-Physical Systems (ICPS) scheme (Grant No. T-54).

References

Abdel-Basset M, Ding W, El-Shahat D (2021) A hybrid Harris Hawks Optimization algorithm with simulated annealing for feature selection. Artif Intell Rev 54:593–637. https://doi.org/10.1007/s10462-020-09860-3

Abedinpourshotorban H, Mariyam Shamsuddin S, Beheshti Z, Jawawi D (2016) Electromagnetic field optimization: a physics-inspired metaheuristic optimization algorithm. Swarm Evol Comput 26:8–22. https://doi.org/10.1016/j.swevo.2015.07.002

Ahmed S, Mafarja M, Faris H, Aljarah I (2018) Feature selection using salp swarm algorithm with chaos. In: ACM International Conference Proceeding Series, pp 65–69

Alabool HM, Alarabiat D, Abualigah L, Heidari AA (2021) Harris Hawks Optimization: a comprehensive review of recent variants and applications. Neural Comput Appl 33:8939–8980. https://doi.org/10.1007/s00521-021-05720-5

Bolón-Canedo V, Remeseiro B (2020) Feature selection in image analysis: a survey. Artif Intell Rev 53:2905–2931. https://doi.org/10.1007/s10462-019-09750-3

Dash M, Liu H (2003) Consistency-based search in feature selection. Artif Intell 151:155–176. https://doi.org/10.1016/S0004-3702(03)00079-1

Dong H, Li T, Ding R, Sun J (2018) A novel hybrid genetic algorithm with granular information for feature selection and optimization. Appl Soft Comput J 65:33–46. https://doi.org/10.1016/j.asoc.2017.12.048

Elgamal ZM, Yasin NBM, Tubishat M et al (2020) An improved Harris Hawks Optimization algorithm with simulated annealing for feature selection in the medical field. IEEE Access 8:186638–186652. https://doi.org/10.1109/ACCESS.2020.3029728

Elminaam DSA, Nabil A, Ibraheem SA, Houssein EH (2021) An efficient marine predators algorithm for feature selection. IEEE Access 9:60136–60153. https://doi.org/10.1109/ACCESS.2021.3073261

Emary E, Zawbaa HM, Ghany KKA, et al (2015) Firefly optimization algorithm for feature selection. In: ACM International Conference Proceeding Series

Faramarzi A, Heidarinejad M, Mirjalili S, Gandomi AH (2020) Marine predators algorithm: a nature-inspired metaheuristic. Expert Syst Appl 152:113377. https://doi.org/10.1016/j.eswa.2020.113377

Gao D, Wang GG, Pedrycz W (2020a) Solving fuzzy job-shop scheduling problem using de algorithm improved by a selection mechanism. IEEE Trans Fuzzy Syst 28:3265–3275. https://doi.org/10.1109/TFUZZ.2020.3003506

Gao Y, Zhou Y, Luo Q (2020b) An efficient binary equilibrium optimizer algorithm for feature selection. IEEE Access 8:140936–140963. https://doi.org/10.1109/ACCESS.2020.3013617

Gao ZM, Zhao J, Hu YR, Chen HF (2019) The improved harris hawk optimization algorithm with the tent map. In: 2019 IEEE 3rd International Conference on Electronic Information Technology and Computer Engineering, EITCE 2019

Gu N, Fan M, Du L, Ren D (2015) Efficient sequential feature selection based on adaptive eigenspace model. Neurocomputing 161:199–209. https://doi.org/10.1016/j.neucom.2015.02.043

Gunal S, Edizkan R (2008) Subspace based feature selection for pattern recognition. Inf Sci (NY) 178:3716–3726. https://doi.org/10.1016/j.ins.2008.06.001

Heidari AA, Mirjalili S, Faris H et al (2019) Harris Hawks Optimization: algorithm and applications. Fut Gen Comput Syst 97:849–872. https://doi.org/10.1016/j.future.2019.02.028

Hussien AG, Amin M (2021) A self-adaptive Harris Hawks Optimization algorithm with opposition-based learning and chaotic local search strategy for global optimization and feature selection. Int J Mach Learn Cybern. https://doi.org/10.1007/s13042-021-01326-4

Houssein EH, Hosney ME, Elhoseny M et al (2020a) Hybrid Harris Hawks Optimization with cuckoo search for drug design and discovery in chemoinformatics. Sci Rep 10:1–22. https://doi.org/10.1038/s41598-020-71502-z

Houssein EH, Saad MR, Hussain K et al (2020b) Optimal sink node placement in large scale wireless sensor networks based on harris' hawk optimization algorithm. IEEE Access 8:19381–19397. https://doi.org/10.1109/ACCESS.2020.2968981

Houssein EH, Neggaz N, Hosney ME et al (2021) Enhanced Harris Hawks Optimization with genetic operators for selection chemical descriptors and compounds activities. Neural Comput Appl. https://doi.org/10.1007/s00521-021-05991-y

Hussain K, Neggaz N, Zhu W, Houssein EH (2021) An efficient hybrid sine-cosine Harris Hawks Optimization for low and high-dimensional feature selection. Expert Syst Appl 176:114778. https://doi.org/10.1016/J.ESWA.2021.114778

Hussien AG, Hassanien AE, Houssein EH, et al (2019) S-shaped binary whale optimization algorithm for feature selection. In: Recent trends in signal and image processing. Springer Singapore, pp 79–87

Ismael OM, Qasim OS, Algamal ZY (2020) Improving Harris Hawks Optimization algorithm for hyperparameters estimation and feature selection in v-support vector regression based on opposition-based learning. J Chemom. https://doi.org/10.1002/cem.3311

Kanimozhi T, Latha K (2015) An integrated approach to region based image retrieval using firefly algorithm and support vector machine. Neurocomputing 151:1099–1111. https://doi.org/10.1016/j.neucom.2014.07.078

Kou G, Yang P, Peng Y et al (2020) Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods. Appl Soft Comput J. https://doi.org/10.1016/j.asoc.2019.105836

Lew MS (2001) Principles of visual information retrieval

Liu C, Wu J, Mirador L, et al (2018) Classifying DNA methylation imbalance data in cancer risk prediction using SMOTE and tomek link methods. In: International Conference of Pioneering Computer Scientists, Engineers and Educators. Springer, Singapore, pp 1–9

Madasu A, Elango S (2020) Efficient feature selection techniques for sentiment analysis. Multimed Tools Appl 79:6313–6335. https://doi.org/10.1007/s11042-019-08409-z

Mafarja MM, Mirjalili S (2017) Hybrid whale optimization algorithm with simulated annealing for feature selection. Neurocomputing 260:302–312. https://doi.org/10.1016/j.neucom.2017.04.053

Marcano-Cedeño A, Quintanilla-Domínguez J, Cortina-Januchs MG, Andina D (2010) Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network. In: IECON Proceedings (Industrial Electronics Conference). pp 2845–2850

Mirjalili S (2015) Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowl Based Syst 89:228–249. https://doi.org/10.1016/j.knosys.2015.07.006

Mirjalili S (2016) SCA: a sine cosine algorithm for solving optimization problems. Knowl Based Syst 96:120–133. https://doi.org/10.1016/j.knosys.2015.12.022

Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

Mirjalili S, Gandomi AH, Mirjalili SZ et al (2017) Salp swarm algorithm: a bio-inspired optimizer for engineering design problems. Adv Eng Softw 114:163–191. https://doi.org/10.1016/j.advengsoft.2017.07.002

Neggaz N, Ewees AA, Elaziz MA, Mafarja M (2020a) Boosting salp swarm algorithm by sine cosine algorithm and disrupt operator for feature selection. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2019.113103

Neggaz N, Houssein EH, Hussain K (2020b) An efficient henry gas solubility optimization for feature selection. Expert Syst Appl 152:113364. https://doi.org/10.1016/j.eswa.2020.113364

Sihwail R, Omar K, Ariffin KAZ, Tubishat M (2020) Improved Harris Hawks Optimization using elite opposition-based learning and novel search mechanism for feature selection. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3006473

Too J, Mirjalili S (2021) General learning equilibrium optimizer: a new feature selection method for biological data classification. Appl Artif Intell 35:247–263. https://doi.org/10.1080/08839514.2020.1861407

Too J, Abdullah AR, Saad NM (2019) A new quadratic binary harris hawk optimization for feature selection. Electron. https://doi.org/10.3390/electronics8101130

Tuba E, Strumberger I, Bezdan T et al (2019) Classification and feature selection method for medical datasets by brain storm optimization algorithm and support vector machine. Procedia Comput Sci 162:307–315

Zhang Y, Song XF, Gong DW (2017) A return-cost-based binary firefly algorithm for feature selection. Inf Sci (NY) 418–419:561–574. https://doi.org/10.1016/j.ins.2017.08.047

Zhang Y, Liu R, Wang X et al (2020) Boosted binary Harris hawks optimizer and feature selection. Eng Comput. https://doi.org/10.1007/s00366-020-01028-5