IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 30, 2021, accepted August 14, 2021, date of publication August 16, 2021, date of current version August 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3105485

An Efficient Slime Mould Algorithm Combined With K-Nearest Neighbor for Medical Classification Tasks

YASER M. WAZERY[1], EMAN SABER[1], ESSAM H. HOUSSEIN[1], ABDELMGEID A. ALI[1], AND ESLAM AMER[2]

[1]Faculty of Computers and Information, Minia University, Minia 61519, Egypt
[2]Faculty of Computer Science, Misr International University, Cairo 11111, Egypt

Corresponding author: Essam H. Houssein (essam.halim@mu.edu.eg)

ABSTRACT Growing science and medical technologies have produced a massive amount of knowledge on different scales of biological systems. By processing various amounts of medical data, these technologies will increase the quality of disease detection and enhance the usability of health information systems. The integration of machine learning in computer-based diagnostic systems facilitates the early detection of diseases, enabling more productive treatments and prolonged survival rates. The slime mould algorithm (SMA) may have drawbacks, such as being trapped in minimal local regions and having an unbalanced exploitation and exploration phase. To overcome these limitations, this paper proposes ISMA, an improved version of the slime mould algorithm (SMA) hybridized with the opposition-based learning (OBL) strategy based on the k-nearest neighbor (kNN) classifier for the classification approach. Opposition-based learning improves global exploratory ability while avoiding premature convergence. The experimental results revealed the superiority of the proposed ISMA–kNN in various classification evaluation metrics, including accuracy, sensitivity, specificity, precision, F-score, G-mean, computational time, and feature selection (FS) size compared with the tunicate swarm algorithm (TSA), the marine predators algorithm (MPA), the chimp optimization algorithm (ChOA), the moth–flame optimization (MFO) algorithm, the whale optimization algorithm (WOA), the sine cosine algorithm (SCA), and the original SMA algorithm. Performance tests were run on the same maximum number of function evaluations (FEs) on nine UCI benchmark disease data sets with different feature sizes.

INDEX TERMS Medical classification, feature selection (FS), machine learning (ML), slime mould algorithm (SMA), opposition-based learning (OBL).

I. INTRODUCTION

Artificial intelligence techniques have enhanced the outcomes of medical diagnoses, decreasing the risk of accidental errors by inexperienced physicians. Internet-based remote techniques have also reduced the costs of patient monitoring. Providing accurate, dependable diagnoses at the early stage of disease will positively affect patients' lives. Knowledge discovery techniques based on data mining, search for hidden patterns among various amounts of data, thus, extracting useful information [1].

Analyzing the data and extracting its useful information is difficult because data are increasingly being created, exchanged, and shared. Before learning, the dimensionality of the data must be reduced by feature selection (FS) or dimensionality reduction. FS plays an important role in machine learning techniques because high-dimensional data sets contain redundant, noisy, and irrelevant data to improve the precision of classification methods [2], [3].

Although FS reduces the amount of data, it is a sophisticated and computationally demanding process, especially for high-dimensional data sets [4]–[6]. The primary objective of FS is extracting a small subset of features from a specific problem domain to improve the classification performance. The purposes of FS are summarized as; 1) Simplify data for users, especially for researchers, 2) Reduce the computational time by selecting the necessary features only, 3) Avoid the

The associate editor coordinating the review of this manuscript and approving it for publication was Mouloud Denai.

dimensionality curse, and 4) Enhance generalization by eliminating over-fitting.

Machine learning (ML) [7] is often effective in solving these problems, improving system performance, and machine design. ML algorithms use the same collection of features to represent an instance in any data set. A crucial step is deciding which particular learning algorithm to use. The major problem is the increase in exponentially the search space with features numbers in the data set [7]. k-nearest neighbor (kNN) [8] is among the most widely used methods in classification systems [9]–[13]. Swarm intelligence (SI) algorithms can successfully solve several problems. In difficult optimization problems, the dynamic searching behaviors of SI algorithms are mandatory. Many SI algorithms have recently resolved the FS problem, either alone or combined into hybrid approaches. Swarm intelligence algorithms have been used to predict the results of major diseases such as cancer, heart disease, and cardiology. Swarm Intelligence methods have been used in the diagnosis and treatment of diseases. Various SI algorithms have been used for the diagnosis of various diseases. The collected datasets in many medical domain problems, such as medical imaging, biomedical signal processing, and DNA microarray data, typically have very large feature dimensions. To deal with this high dimensionality problem, related literature has shown that considering feature selection on various medical domain datasets has a positive impact [14]–[17].

Opposition-Based Learning (OBL) [18] technique has been developed to improve the efficiency of metaheuristic optimization algorithms by overcoming premature convergence and slow movement. The candidate solutions obtained using a stochastic iteration strategy, as well as their opposite solutions found in opposite parts of the search space, are significant to OBL because they are closer to the global optimum than a random solution. The OBL strategy has been used with several bio-inspired optimization approaches to produce expected distances to the global optimum that are shorter than those obtained by randomly sampling solution pairings [19].

Although many existing techniques provide accurate and reliable diagnostics, our proposed technique is innovative in the integration of OBL with SMA in the exploration phase, as well as improving the SMA exploitation strategy with the assistance of extra dynamic solutions to avoid stagnancy issues, using the hybridization technique ISMA. The experimental analysis and comparative study performed in the next sections, using numerical optimization and feature selection issues, indicate that the proposed methodology has been proven to be effective. ISMA approach based on kNN classifier improves the local search-ability. In this study, The ISMA approach is innovative because ISMA-kNN improves various classification evaluation metrics of the original kNN, including accuracy, sensitivity, specificity, precision, F-score, G-mean, computational time, and feature selection (FS) size compared with the tunicate swarm algorithm (TSA) [20], marine predators algorithm (MPA) [21], chimp optimization algorithm (ChOA) [22], moth–flame optimization (MFO) [23], whale optimization algorithm (WOA) [24], sine cosine algorithm (SCA) [25], and the original slime mould algorithm (SMA) [26] based on the kNN classifier in terms of performance evaluation metrics for classification on nine UCI benchmark disease data sets with different feature sizes.

Our present work was driven by two principal reasons: 1) The No Free Lunch (NFL) principle states that no optimization technique can solve every optimization problem. An optimizer delivering superior performance on a specific set of problems may not provide the same performance on another set of problems. We propose ISMA-kNN for solving the FS problem. And 2) To the authors' knowledge, the FS problem has never been solved by a combined ISMA and kNN classifier. The main objectives and contributions of the current study are summarized as:

- A modified SMA was proposed as an alternate feature selection approach.
- An efficient classification approach called ISMA-kNN to improve SMA convergence, OBL is being used as a local search approach.
- We implemented comprehensive experiments with nine UCI disease data sets with medium and high dimensionality features.
- The proposed ISMA achieved superior results compared to seven existing state-of-the-art constrained optimization algorithms: TSA, MPA, ChOA, MFO, WOA, SCA, and the original SMA.

The paper is structured as follows; Section II presents an overview of the related work. Section III includes a brief description of the principles and mathematical foundations of the used techniques. The proposed ISMA-kNN approach is explained in Section IV, and then the experimental evaluation introduced in Section V. Section V-E includes discussions on the limitations and motivations of the proposed method. The Conclusion and future work are demonstrated in Section VI.

II. RELATED WORKS

Abdel-Basset *et al.* [27] presented four binary versions of the slime mould algorithm (SMA) for feature selection: A binary version (BSMA), BSMA integrates with two-phase mutation (TM) (TMBSMA), BSMA combines with a novel attacking-feeding strategy (AF) (AFBSMA), TM and AF are integrated with BSMA (FMBSMA). Furthermore, the FMBSMA version is shown to be the best as compared to the other three versions and six state-of-the-art feature selection algorithms after comparing the four versions. Ewees *et al.* [28] proposed the Slime Mold Algorithm (SMA), which is based on the Firefly Algorithm (FA) (SMAFA). FA is used to improve SMA exploration since it has a high ability to find feasible regions with optimal solutions. This will improve convergence by improving the quality of the final result. We used twenty UCI datasets and compared them to different MH algorithms to see how well the SMAFA performed. The SMAFA surpasses all compared

algorithms in most experiments in terms of fitness function, classification accuracy, and standard deviation, according to the results of the evaluation. Wang and Chen [29] proposed an improved whale optimization algorithm (CMWOA), which combines chaotic and multi-swarm strategies to perform parameter optimization and feature selection simultaneously for support vector machine (SVM). The proposed SVM model, called CMWOAFS-SVM, was compared to several competitor's SVM models based on other optimization algorithms like the initial algorithm, particle swarm optimization, bacterial foraging optimization, and genetic algorithms using several well-known medicals diagnose problems such as breast cancer, and diabetes.

Diego and Elaziz [30] proposed a combination of chaotic maps, opposition-based learning, and disruption operator to boost brainstorm optimization exploration abilities by expanding the diversity of the population. A set of benchmark functions was used to evaluate the suggested method, and it was also utilized for feature selection in data mining. The results demonstrate the suggested method's great efficacy in determining the optimal solutions to the evaluated functions.

Choubey *et al.* [31] proposed an indigenous diagnostic instrument for diabetes detection. The proposed methodology has two phases. Phase-I acquires the Lung Cancer data set and interprets the data by two separate methods. Phase-II performs the FS uses principal component analysis and particle swarm optimization (PSO). Classification is then performed by C4.5 DT, naive Bayes, ID3 DT, logistic regression, and k-Nearest Neighbor. The proposed method consumes less computational time and delivers higher accuracy than conventional classifications. Moreover, it is potentially extendible to the early detection of medical disorders other than diabetes.

Chowdhury *et al.* [32] developed potential candidates for Alzheimer's disease (AD) biomarkers in blood cells. To examine how the differentially expressed genes correlate in these cells, they studied two microarray gene expression data sets from the peripheral blood and brain cells of AD patients. Their study was the first attempt to classify candidate AD factors from blood cells by this technique. Although the usefulness of the candidate factors must be confirmed in clinical investigations, the technique opens a new window for early detection of AD progression. The biomarkers and their related molecular pathways might provide new insights into AD progression. As these biomarkers exist in tissues outside the central nervous system, they may be targeted for therapeutic development after assessing their functional utility.

Tubishat *et al.* [19] proposed using Opposition Based Learning (OBL) at the initialization phase of the Salp Swarm Algorithm (SSA) termed (ISSA) to improve population diversity in the search space. In the majority of the 18 datasets from the UCI repository, the ISSA outperforms all baseline algorithms in terms of fitness values, accuracy, convergence curves, and feature reduction.

Kewat *et al.* [33] proposed particle swarm optimization (PSO), a genetic search algorithm, and a greedy searching technique for dimensional reduction, and combined naive Bayes, k-Nearest Neighbor, and C4.5 as the subset evaluator for medical data sets. They reported that selecting the wrapper-based features increased the classification accuracy of their chosen data sets.

More related works are summarized in Table.1.

III. PRELIMINARIES

This section briefly overviews the principles and mathematical foundations of the used techniques.

A. K-NEAREST NEIGHBOR

The kNN is a popular technique for machine learning and is mostly applied to classification on benchmark data sets. This simple and easy-to-use algorithm delivers excellent results in many domains; even in comparisons with the most advanced machine-learning approaches [11], [51], [52]. Recent interest in kNN has been sparked by the increasing availability of data presented in new forms, such as free text, images, audio, and video. However, the performance of kNN is sensitive to several factors, primarily the distance metric selection and choice of the k parameter. The most popular distance measure is the Euclidean distance, determined as the square root of the summed squared differences between a new point (y) and an existing point (y_i) overall attributes of input j. The calculation is given by Eq. (1).

$$d(y, yi) = \sqrt{\sum_{j=1}^{n} (y_j - y_{ij})^2} \tag{1}$$

Algorithm 1 outlines the k-nearest neighbor (kNN) algorithm for classifying sample s.

Algorithm 1 Pseudocode of the kNN Algorithm

Inputs: Load the training and test data.
Outputs: Assign a class to the test point.
For each point in the test data, choose the value of k.
while the stopping state is not reached **do**
    Find the Euclidean distance as shown in Eq. (1).
    Store the Euclidean distances in a list and sort it.
    Choose the first k points.
end while
Return *Accuracy*.

B. SLIME MOULD ALGORITHM

Slime mould algorithm (SMA) [26] is based on the oscillation mode of slime mould in nature. Governed by a novel mathematical model, the proposed SMA has several new features, including adaptive weights for simulating the generation of positive and negative propagation waves that shape the optimal path of the slime mould.

The SMA was influenced by slime mould behavior and morphological changes. Individuals in the swarms will be

TABLE 1. Summary of the papers for MH algorithms for feature selections.

Ref.	Data sets used	Algorithms used	Performance evaluation
[34]	Lung cancer, brain image, and Alzheimer's disease data sets.	Proposed an Opposition-based Crow Search (OCS) algorithm to enhance the performance of the Optimal Deep Learning (DL) classifier.	Accuracy = 95.22%, Sensitivity = 86.45%, and Specificity=100%.
[35]	10 diverse data sets.	Proposed CBCSEM, a feature selection method based on the cuckoo search algorithm with Elitist Preservation and Uniform Mutation.	CBCSEM eliminated traditional approaches' early convergence and the tendency to fall into local optima.
[36]	45 texture features from Region Of Interests (ROI) were extracted from 322 images from mini-mias datasets.	Proposed an Opposition based enhanced version of Grey Wolf Optimization (OEGWO) Algorithm on thirteen different benchmark functions.	OEGWO outperformed the other algorithms for solving the feature selection problem in breast density classification.
[37]	COVID-19 data sets.	COVID-19 Patients Detection Strategy (CPDS) based on hybrid feature selection and enhanced KNN classifier to detect the disease at an early stage.	Precision = 0.72, Recall = 0.69, Accuracy = 0.93, Error = 0.07.
[38]	10 UCI data sets.	Proposed opposition-based moth-flame optimization improved by differential evolution noted as OMFODE.	OMFODE was superior to the state-of-the-art metaheuristic algorithms in terms of performance measures.
[39]	23 UCI benchmark data sets.	An improved binary version of SSA for FS tasks in the wrapper approach.	Accuracy = 0.84, Execution time = 82.78, Feature reduction ratio = 19.4, Best fitness = 0.0743.
[40]	Large patient no-show data sets.	Proposed Opposition-based Self-Adaptive Cohort Intelligence (OSACI).	OSACI outperformed the other algorithms in terms of dimensionality reduction and convergence speed, as well as AUC, sensitivity, and specificity.
[41]	21 diverse data set.	Proposed Binary Multi-Verse Optimization Algorithm (BMVO).	BMVO outperformed the other algorithms in terms of the number of features selected and the accuracy of the classification process.
[42]	20 high-dimensional data sets.	Proposed chaotic cuckoo optimization algorithm with levy flight, disruption operator, and opposition-based learning (CCOALFDO).	CCOALFDO outperformed the other algorithms in terms of classification accuracy rate and the number of features selected for classification tasks.
[43]	23 benchmarks mathematical functions.	Proposed an improved version of the Woodpecker Mating Algorithm (WMA), which is based on opposition-based learning (OWMA).	OWMA outperformed the other algorithms and promising performance in solving extremely difficult optimization problems.
[44]	5 benchmark data sets.	A novel chaotic multi-verse optimizer algorithm (CMVO) to avoid MVO drawbacks.	The logistic chaotic map was the best chaotic map for improving MVO performance.
[45]	18 disease data sets.	Proposed a new FS approach based on Henry Gas Solubility Optimization (HGSO).	HGSO with SVM accuracy = 88.88%.
[46]	Heart failure patients data sets.	Predicted the survival of patients with heart failure from serum creatinine and ejection fraction alone.	Matthews correlation coefficient (MCC) = +0.384, Accuracy = 0.740, Area Under the Curve (AUC) = 0.800.
[47]	24 benchmark high-dimensional data sets.	Proposed an improved version of the SCA (ISCA) for solving high-dimensional problems.	ISCA outperformed both classic SCA and other population-based algorithms in terms of avoiding local optima and achieving faster convergence.
[31]	Pima Indian Diabetes data sets.	Proposed an important indigenous diagnostic approach for diabetes detection using Principle component analysis (PCA) and particle swarm optimization (PSO).	C4.5 DT accuracy = 95.58%, PCA_C4.5 DT accuracy =95.58%, PSO_C4.5 DT accuracy =94.01%.
[48]	Soybean data sets.	Improved the performance of the kNN classifier, called Particle Optimized Scored K-Nearest Neighbor (POSkNN).	POSkNN Accuracy = 94.21%.
[49]	19 mathematical benchmark test functions of varying complexities.	Proposed a novel quasi-oppositional chaotic antlion optimizer (ALO) (QOCALO) for solving global optimization problems.	Ensured better convergence speed of the proposed algorithm and provided a better exploration of the search space.
[50]	23 classical benchmark functions, 10 modern CEC2019 benchmark test functions and two engineering optimization problems.	Proposed an improved elephant herding optimization (EHOI) with the help of the position updating mechanism of sine–cosine algorithm (SCA) and opposition-based learning (OBL).	The proposed EHOI outperformed most of the selected MAs in terms of solution quality.

divided into three groups, some of which would be picked from the beginning to be resurrected with a proportional number z as shown in Eq. (8) carry on exploration. Based on their present positions, some of them would continue their exploration, and the rest of them would be directed towards the best candidate. The detailed mathematical expressions of the SMA approach are as follows:

1) Initialization process

The following rule is suggested to model the SMA approach mathematically, as shown in Eq. (2):

$$\overrightarrow{S(it+1)} = \begin{cases} \overrightarrow{S_b(it)} + \overrightarrow{vb}.(\overrightarrow{W}.\overrightarrow{S_A(it)} - \overrightarrow{S_B(it)}), & r < p \\ \overrightarrow{vc}.\overrightarrow{S(it)}, & r \geq p \end{cases} \tag{2}$$

where $\overrightarrow{vb}$ as shown in Eq. (4), $\overrightarrow{vc}$ decreases linearly from one to zero. $\overrightarrow{S_b}$ stands for the positions with the greatest accuracy, it stands for the current iteration, $\overrightarrow{S}$ represents the location of slime mould, $\overrightarrow{S_A}$, and $\overrightarrow{S_B}$ represent two randomly selected individuals from the swarm, $\overrightarrow{W}$ represents the weight of slime mould, and p as shown in Eq. (3):

$$p = tanh|F(j) - bF| \tag{3}$$

where $F(j)$ represents the fitness of $\overrightarrow{S}$ where $j \in 1, 2, \ldots, n$ and bF reflects the best fitness in all iterations. The $\overrightarrow{vb}$ is identified as follows:

$$\overrightarrow{vb} = [a, -a] \tag{4}$$

where a calculated as shown in Eq. (5)

$$a = arctanh(-(\frac{it}{max_{it}}) + 1) \tag{5}$$

where it is the current iteration and max_{it} is the max-iterations.

2) Fitness Evaluation
The $\overrightarrow{W}$ definition is identified as follows:

$$\overrightarrow{W(Fsort(j))} = \begin{cases} 1 + r.log\left(\frac{bF - F(j)}{bF - wF} + 1\right), & cond. \\ 1 - r.log\left(\frac{bF - F(j)}{bF - wF} + 1\right), & others \end{cases} \quad (6)$$

where wF stands for the worst fitness, bF stands for the best fitness, *cond.* indicates that $F(j)$ ranks the first half of the population, $Fsort$ stands for the sorted fitness values, $Fsort$ stands for sorted fitness and calculated as shown in Eq. (7), and r stands for the random value in the [0,1] interval.

$$Fsort = sort(F) \quad (7)$$

This section mathematically simulates the contraction mode of the slime mould's venous tissue structure while looking. The uncertainty of the mode of venous contraction in Eq. (6) is simulated by r. The log is used to decrease the numerical value change rate so that the contraction frequency value does not change too much.

3) Location Update
Slime moulds are programmed to change their search patterns according to food quality. The greater the weight near the region is when the food concentration is content; when the food concentration is poor, the region's weight will be decreased, thereby turning to other regions to explore. The mathematical formula for slime mould position updating is 8:

$$\overrightarrow{S^*} = \begin{cases} rand.(ub - lb) + lb, & rand < z \\ \overrightarrow{S_b(it)} + \overrightarrow{vb}.\left(W.\overrightarrow{S_A(it)} - \overrightarrow{S_B(it)}\right), & r < p \\ \overrightarrow{vc}.\overrightarrow{S(it)}, & r \geq p \end{cases} \quad (8)$$

where lb and ub denote the lower and upper boundaries of the search respectively, it stands for iteration, and r and $rand$ stand for the random value of [0,1].

C. OPPOSITION-BASED LEARNING

Opposition-Based Learning (OBL) [18] is used to define the opposite solution to the present solution, and it then compares the value of the fitness function (f_{obj}) to the present solution to see if the opposite is better. The OBL assuming the opposite value $\overline{x_i}$ for the true value $x \in$ [u, l] in N–dimensions, which may be derived from Eq. (9). So OBL finds a solution and its corresponding opposite solution simultaneously to approximate the global optima. In the ISMA, if the solution is x_i, its corresponding opposite solution can be defined as $\overline{x_i}$

$$\overline{x_i} = u_i + l_i - x_i, \quad i = 1, 2, \ldots, N \quad (9)$$

where u_i and l_i are the minimum and the maximum values in the i dimension of the search space of the current population, respectively. Furthermore, the two solutions (x and $\overline{x}$) are compared during the optimization process, with the better of these solutions being saved and the other being deleted by comparing the fitness function. If $f(x) \leq f(\overline{x})$ (for minimization), for example, x is saved; Otherwise, $\overline{x}$ is stored.

IV. THE PROPOSED ISMA-kNN APPROACH

In general, the SMA and metaheuristic optimization algorithms shortcomings are dependent on the issue to be solved which include slow convergence, being trapped in sub-optimal regions, and an improper balance between the exploration and exploitation phases, particularly in high-dimension problems. To prevent the algorithm from getting stuck into local minimal regions, the SMA is hybridized with OBL based on the kNN classifier to optimize the classification and feature selection. In ISMA, the parameter S_{best} defines the best location that optimizes the parameters in the selected feature set for all cross-validation folds. Fig. 1 is a flowchart of the proposed ISMA–kNN approach. The three phases of the approach shown as: (1) Preprocessing, (2) Feature selection and optimization, and (3) Classification. Algorithm 2 presents the pseudocode of the ISMA–kNN approach.

A. FITNESS FUNCTION

The fitness function (f_{obj}) determines the FS subset and assesses the quality of the obtained features. The target value of the ISMA–kNN classification approach is computed as shown in Eq. (10).

$$Fit_i = \alpha * Err_i + \beta * d_i/D \quad (10)$$

where $\alpha = 0.9$ and $\beta = 1 - \alpha$. The factor α balances the classification error rate Err_i and the number d_i of selected features. In Eq. (10), D is the attribute size of the used data set.

Before the fitness assessment process that selects a feature subset, we require an intermediate phase called a binary conversion. The FS process employs the kNN classifier as an expert system. The error rate of the test set computed by kNN is Err_i (see Algorithm 1).

B. FEATURE SELECTION

We now present the execution of ISMA in FS. The three phases of the proposed FS solution are outlined below:

- *Initialization Phase*
  The SMA generates an initial population of N candidate solutions, where each entity covers a range of features to be chosen for assessment. This step critically affects the convergence and quality of the optimal solution. The population S_0 is randomly generated by Eq.2, and the fitness function (f_{obj}) is computed by Eq. 10.
- *Update Phase*
  Based on a fitness function, the quality of the solution is evaluated for each new position. The current position is then updated if the new position's solution quality is better than the current position's solution quality. Eqs. (6) and (8) are used to update the position of each search

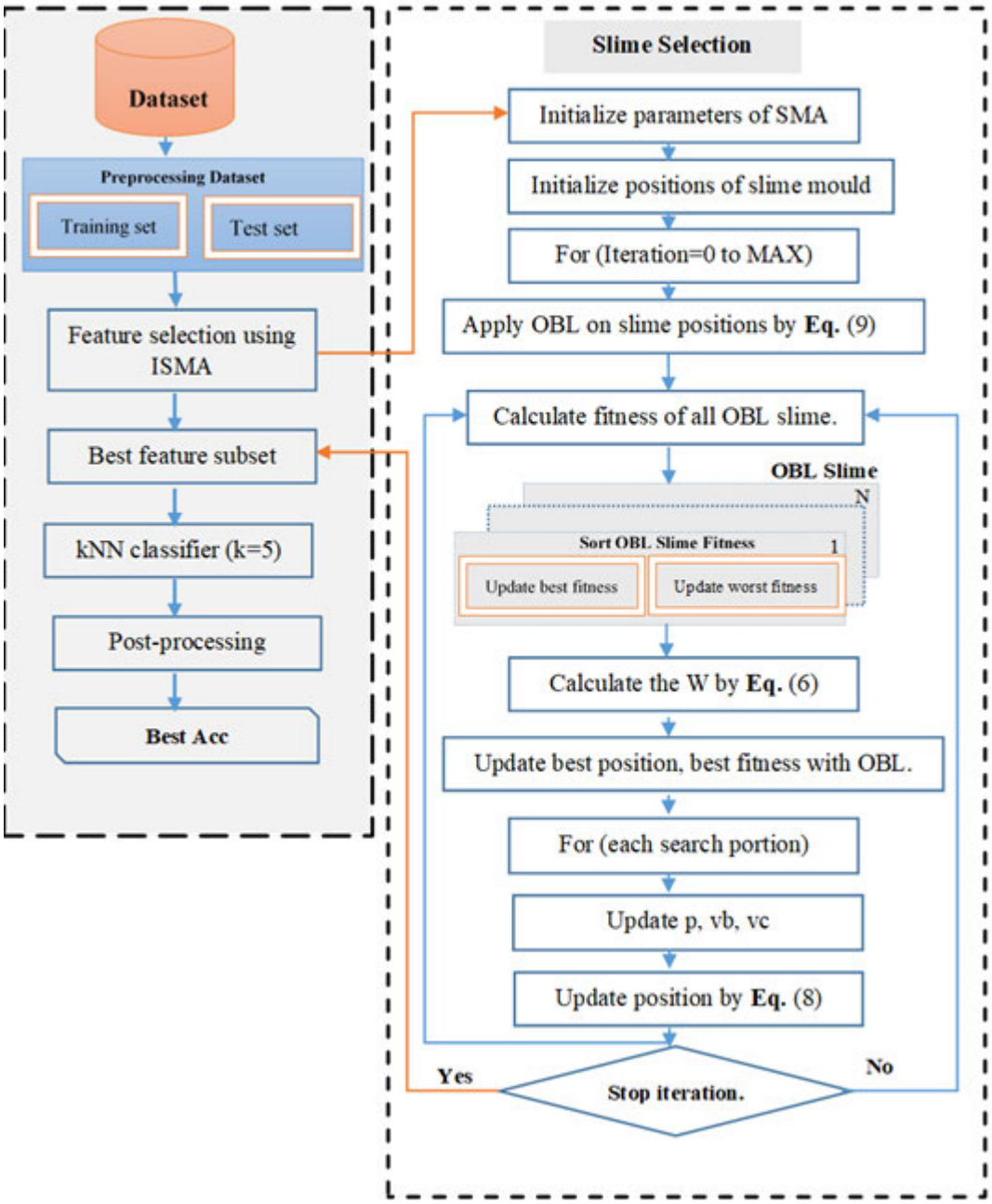


FIGURE 1. Flowchart of the proposed ISMA-kNN approach.

agent with opposite-based learning as shown in Eq. (9) and adjust the updated search agent which goes beyond the boundary in a given search space in Algorithm 2.

To improve the search process by exploring new regions in quest of the optimal solution, increasing algorithm diversity, avoiding local optima, and confirming whether the new solution is better than the old one, the basic principle of OBL is to consider a solution and its matching opposite solution simultaneously. If a solution is x_i in the ISMA, the equivalent opposite solution is $\overline{x_i}$ as shown in Eq. (9). The steps are outlined in Fig. 1 and Algorithm 2, which includes pseudocode.

The best solution is then determined by calculating the fitness values of the new population. This process is repeated until the maximum number of function evaluations (i.e., the termination condition) is reached. In this process, we can determine the consistency of the ISMA approach.

- *Classification phase*

The ISMA process returns the best solution obtained in the previous step. Only the features valued as equal to one is in S_{best} are retained from the original data. We use a classification holdout strategy, in which the data set is randomly divided into two sets, one for training

Algorithm 2 Pseudocode of the Proposed ISMA–kNN Approach

Inputs: Initialize random popsize using Eq. (2), max_{it}, ub, lb, f_{obj}, dim.
Outputs: *BestAccuracy*.
Initialize slime positions S_j (j=1,2,...,n).
while $it \leq max_{it}$ **do**
  Apply OBL on slime positions by Eq. (9).
  Calculate OBL slime mould fitness.
  Sort OBL slime mould fitness.
  Update bF, S_b with OBL.
  Select the best feature selection.
  Call kNN classifier.
  Calculate W using Eq. (6).
  for each search portion **do**
    Update p, vb, vc.
    Update positions by Eq. (8).
    Select the best feature selection.
    Call kNN classifier.
  end for
  it=it+1.
end while
Return *BestAccuracy*.

and the other for testing. After reviewing the literature for a reasonable analogy, the number of classifiers ($k = 5$) was selected [53]–[56]. As a result, high classification accuracy equates to a high fitness value, and fitness has been used to represent classification accuracy. To achieve meaningful results, all experiments were run 20 times with 2,000 FEs in each run.

- *Termination phase*
  The maximum number of function evaluations of the proposed algorithm are performed until the halting criteria are satisfied. The best viable option is then discovered. Algorithm 2 explains each step in detail.

C. COMPUTATIONAL COMPLEXITY

For initialization apply OBL on slime positions, the computation complexity is O(I), sorting is O(I + I log I), the location update is O(I x dim), and weight update is O(I x dim). Consequently, the complete complexity of ISMA is O(I * (1 + *FEs* * I * (2 + log I + 2 * dim))) where I stands for the instance number, *dim* stands for function dimension, and *FEs* is the maximum number of function evaluations [26].

V. EXPERIMENTAL EVALUATION AND DISCUSSION

The performance of the proposed ISMA–kNN classification approach was experimentally evaluated on nine data sets extracted from University of California, Irvine (UCI) [57] machine learning repository. All state-of-the-art constrained optimization algorithms were executed in the same programming language (for machine specification, as in Table 2). Note that the device specifications influence only the speed of the calculations and not the accuracy of the method. Hardware specifics will clarify the computational times of different methods if the elapsed time is important.

TABLE 2. Details of the hardware system employed in the comparison experiments.

Name	Settings
Operating system	Windows 10(64 bits)
Programming language	MATLAB R2014a
RAM	8GB
HDD	500GB
CPU	Intel Core (TM) i5-6400 processor

A. ALGORITHM CONFIGURATIONS AND DATASETS

Table 5 defines nine benchmark disease data sets extracted from the UCI repository [57] employed in the present experiments. Listed are the total numbers of features and patients, data categories, and feature types. The feasibility of the proposed ISMA–kNN approach can be extensively validated in these problems, as they cover a wide range of feature and instance numbers.

Table 3 lists the parameter sets employed in using state-of-the-art constrained optimization algorithms. The selected parameters have been widely used by various researchers [58], [59]. As mentioned in [60], the default parameter values are a fair parametrization. Moreover, employing default values, reduce comparison bias risks as no algorithm could be advantaged with a better parametrization. Our experiments are described in Table 4. Note that the ISMA, TSA, MPA, ChOA, MFO, WOA, SCA, and the original SMA algorithm must be replicated sufficiently to find all potential high-quality solutions. For this purpose, we performed the 20 runs, setting the maximum number of function evaluations (FEs) to 2,000 in each run to ensure a fair benchmarking comparison. Qualitative and quantitative metrics measure algorithm performance are introduced in Subsection V-B.

B. PERFORMANCE MEASURES OF THE ISMA-kNN APPROACH

To determine the best SI algorithm, we executed each algorithm 20 times as mentioned above (i.e., $M = 20$) under the same conditions. In the performance evaluation, the terms ''Patient'' and ''Healthy'' represented disease-positive and disease-negative, false-positive (*FP*) and true-positive (*TP*) denoted the numbers of cases incorrectly and correctly recognized as patients, respectively, and true-negative (*TN*) and false-negative (*FN*) denoted the numbers of cases correctly and incorrectly recognized as healthy, respectively. The assessment metrics of the ISMA–kNN approach were computed as follows:

- *Mean accuracy* (μ_{ACC})
  This metric assesses the model's ability to distinguish patients from healthy cases. To estimate the accuracy, we must determine the proportions of *TP* and *TN* cases

among all measured cases as shown in Eq. (11)

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{11}$$

The accuracy metric reflects the rate of proper data classification. Therefore, μ_{Acc} is calculated as shown in Eq. (12):

$$\mu_{ACC} = \frac{1}{M}\sum_{j=1}^{M} ACC_*^j \tag{12}$$

where M is the number of runs and ACC_*^j is the accuracy in the j^{th} run.

- *Mean best fitness ($\mu_{Fitness}$)*
The fitness metric, which assesses the efficiency of the algorithm, relates the minimization of the classification error rate to a reduction of the FS ratio (Eq. (10)). The lower value represents the best one as shown in Eq. (13):

$$\mu_{Fitness} = \frac{1}{M}\sum_{j=1}^{M} Fitness_*^j \tag{13}$$

where M is the number of runs and $Fitness_*^j$ is the best fitness value in the j^{th} run.

- *Mean Feature Selection ratio (μ_{FS})*
This metric defines the average size of FS and is expressed as shown in Eq. (14):

$$\mu_{FS} = \frac{1}{M}\sum_{j=1}^{M} f_*^j \tag{14}$$

To determine the overall feature selection ratio, we computed the ratio of the feature selection size f_* to the total size of the features F in the original data set as shown in Eq. (15):

$$Overall_{FS} = \frac{1}{M}\sum_{j=1}^{M} \frac{f_*^j}{F} \tag{15}$$

where M is the number of runs, f_*^j is the feature selection size in the j^{th} run.

- *Mean Sensitivity (μ_{SE})*
The sensitivity (or recall) is a statistical measure of binary classification performance. It defines the percentage of identifying positive cases within the positive population of disease diagnose as shown in Eq. (16):

$$Sensitivity = \frac{TP}{(TP + FN)} \tag{16}$$

The sensitivity reflects the true positive rate. The μ_{SE} metric is computed as shown in Eq. (17):

$$\mu_{SE} = \frac{1}{M}\sum_{j=1}^{M} SE_*^j \tag{17}$$

where M is the number of runs and SE_*^j is the sensitivity value in the j^{th} run.

- *Mean Specificity (μ_{SP})*
The specificity measures the percentage of identifying negatives among the negative population of medical diagnoses as shown in Eq. (18):

$$Specificity = \frac{TN}{(TN + FP)} \tag{18}$$

The specificity metric reflects the true negative rate. μ_{SP} is then computed as shown in Eq. (19):

$$\mu_{SP} = \frac{1}{M}\sum_{j=1}^{M} SP_*^j \tag{19}$$

where M is the number of runs and SP_*^j is the specificity value in the j^{th} run.

- *Mean Precision (μ_{PPV})*
The precision or positive predictive value (PPV) defines the proportion of true positives among all individuals that are expected to test positive in the model's medical diagnoses. It is computed as shown in Eq. (20):

$$Precision = \frac{TP}{(TP + FP)} \tag{20}$$

Precision represents the accuracy of the predicted positive outcome. Therefore, the μ_{PPV} metric is calculated as shown in Eq. (21):

$$\mu_{PPV} = \frac{1}{M}\sum_{j=1}^{M} PPV_*^j \tag{21}$$

where M is the number of runs and PPV_*^j is the precision value in the j^{th} run.

- *Mean F-score (μ_{F1})*
F-score or F-measure (F1) is the harmonic mean of precision Eq. (20) and sensitivity Eq. (16) as shown in Eq. (22):

$$F\text{-}score = 2 * \left(\frac{Precision * Sensitivity}{Precision + Sensitivity}\right) \tag{22}$$

F-score reflects the accuracy of the model. The μ_{F1} metric is calculated as shown in Eq. (23):

$$\mu_{F1} = \frac{1}{M}\sum_{j=1}^{M} F1_*^j \tag{23}$$

where M is the number of runs and $F1_*^j$ is the F-score value in the j^{th} run.

- *Mean G-mean (μ_G)*
The G-mean calculates the balance between the results of the majority and minority groups in the classification. A lower G-mean indicates the poor performance of the positive-case classification, even when the negative cases are correctly categorized, as shown in Eq. (24):

$$G\text{-}Mean = \sqrt{Sensitivity * Precision} \tag{24}$$

TABLE 3. Parameter settings of the algorithms.

Algorithm	Parameter	Value
Slime Mould Algorithm (SMA) [26]	$\overrightarrow{vc}$	Decreased linearly from one to zero.
	$\overrightarrow{vb}$	With a range of [-a, a] as shown in Eq. (5).
	α	$\alpha = 0.9$ as shown in Eq. (10)
	β	$\beta = 1 - \alpha$ as shown in Eq. (10)
Tunicate Swarm Algorithm (TSA) [20].	P_{min}	1.
	P_{max}	4.
	r, c_1, c_2, c_3	rand() function.
Marine Predators Algorithm (MPA) [21]	P	0.5.
	$FADs$	0.2.
	β	1.5.
Chimp Optimization Algorithm (ChOA) [61]	$\overrightarrow{a}$	Decreased linearly from 2 to 0.
Moth-flame Optimization (MFO) [23]	P	Decreased linearly from -1 to -2.
Whale Optimization Algorithm (WOA) [24]	$\overrightarrow{a_1}$	Decreased linearly from 2 to 0.
	$\overrightarrow{a_2}$	Linearly increased from -1 to 1
Sine Cosine Algorithm (SCA) [62]	$\overrightarrow{b}$	Decreased linearly from 2 to 0.

TABLE 4. Details of the experimental runs.

Item	Setting
Splitting criteria	Dividerand () function.
dim	Total number of features.
Lower bound (lb)	0*ones (1, dim).
Upper bound (ub)	1*ones (1, dim).
Distance metric	Euclidean distance as shown in Eq. (1).
max_{it}	Number of iterations (100).
N	Number of search agents (20).
M	Total number of runs (20).
FEs	Max function evaluations (2,000).
k	Number of nearest neighbors(5)

A good G-mean measure avoids overfitting of the negative class and under-fitting of the positive class. The μ_G metric is calculated as shown in Eq. (25):

$$\mu_G = \frac{1}{M}\sum_{j=1}^{M} G_*^j \tag{25}$$

where M is the number of runs and G_*^j is the G-Mean value in the j^{th} run.

- *Standard deviation (σ)*
  The Standard deviation (STD) determines the variations among the outputs of all used algorithms over the various executions. It is calculated as shown in Eq. (26). Note that σ_x was computed for all measures: Accuracy, best fitness, time-consuming, sensitivity, specificity, precision, F-score, G-mean, and feature selection ratio.

$$\sigma_x = \sqrt{\frac{1}{M}\sum_{j=1}^{M}(S_*^j - \mu_x)^2} \tag{26}$$

- *Mean Time-consumption (μ_{Time})*
  The average consumption time (in seconds) of each SI algorithm was estimated as shown in Eq. (27):

$$\mu_{Time} = \frac{1}{M}\sum_{j=1}^{M} Time_*^j \tag{27}$$

where M is the number of runs and $Time_*^j$ is the time-consuming value in the j^{th} run.

- *Function Evaluations (FEs)*
  Function Evaluations (FEs) are the most relevant criteria for comparison. To ensure that all algorithms sample the search space an equal number of times, the maximum number of FEs is considered a fair measure to compare algorithms. Once the FEs has been executed, the evaluation number should increase to make sure that all algorithms sample the search space an equal number of times. FEs calculated as shown in Eq. (28).

$$FEs = Max_{it} * N \tag{28}$$

where Max_{it} is the maximum iteration number and N is the number of search agents.

C. COMPARATIVE ANALYSIS OF ISMA-kNN AND OTHER SI ALGORITHMS

The subsection performs a comparative analysis of ISMA–kNN ($k = 5$) and the other SI algorithms (TSA, MPA, ChOA, MFO, WOA, SCA, and the original SMA; see Table 3) based on the kNN classifier on the nine disease data sets shown in Table 5. The population size and the maximum number of FEs for all algorithms evaluated are 20 and 2,000, respectively. Furthermore, each algorithm executes 20 runs for each function, with the results based on the average of these runs. The comparison assessment was based on the metrics described in the previous subsection V-B.

- *Best-fitness evaluation*
  Table 6 compares the best fitness results of the competing SI algorithms and their STDs. In all nine databases, the ISMA–kNN approach achieved lower fitness values than the other algorithms with competitive STDs. The second-best optimizer (MPA) achieved lower fitness values with competitive means. The ISMA–kNN approach converged most rapidly on all data sets (see Fig.2).
- *Accuracy (ACC) evaluation*
  Table 7 compares the accuracy performances of ISMA–kNN and the other SI algorithms based on

TABLE 5. Descriptions of the disease data sets used in classification.

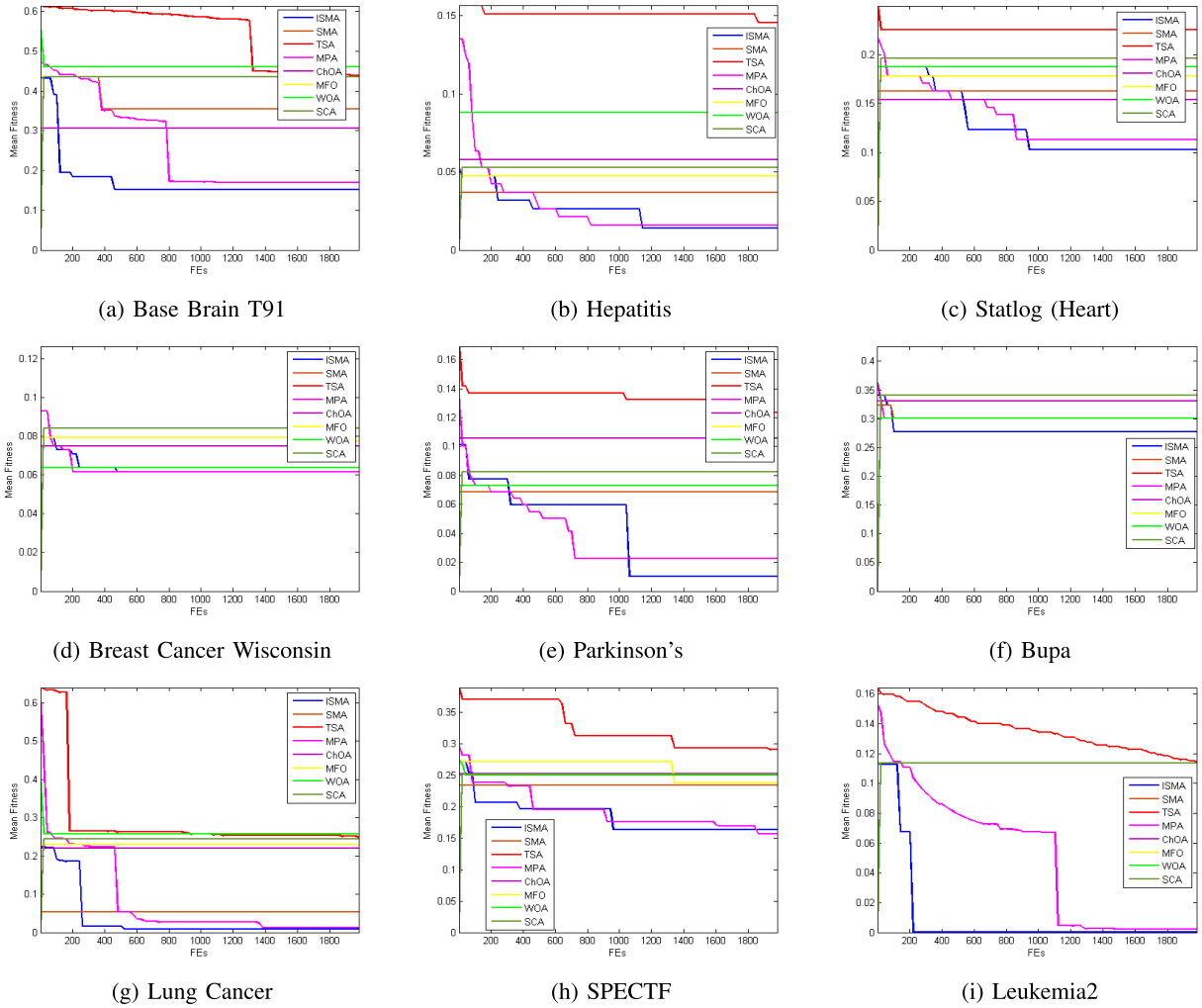
Data set	Total features	Total patients	Feature types
Base Brain T91	5727	60	Categorical, Integer, Real
Hepatitis	20	155	Categorical, Integer, Real
Statlog (Heart)	13	270	Categorical, Real
Breast Cancer Wisconsin (Diagnostic)	9	699	Real
Parkinson's	23	197	Real
Liver Disorders (Bupa)	7	345	Categorical, Integer, Real
Lung Cancer	32	56	Integer
SPECTF Heart (SPECTF)	44	267	Integer
Leukemia2	11226	72	Integer, Real

the kNN classifier under the same conditions. The ISMA–kNN approach outperforms the other algorithms on seven out of nine data sets. The MPA achieved the best results only on the Breast Cancer Wisconsin data set and SCA on the SPECTF data set. The ISMA–kNN approach achieved excellent classification accuracy 99.5% on Hepatitis, accuracies ranging from 97.1% to 97.2% on the Parkinson's and Breast Cancer Wisconsin data sets, and accuracies ranging from 90.4% to 90.6% on Statlog (Heart) and Leukemia2 data sets. The ISMA–kNN approach achieved the highest overall accuracy (89.33%), followed by MPA with 87.90%. From Fig.3 we observe that the ISMA–kNN approach yielded the highest boxplots for all data sets except Breast Cancer Wisconsin and SPECTF data sets.

- *Feature selection evaluation*
  Table 8 compares the average number of feature selections in the ISMA–kNN approach and the other SI algorithms based on the kNN classifier on the same UCI disease data sets. In terms of the average number of selected features, the proposed approach outperformed the other optimizers on four out of the nine data sets. Meanwhile, the TSA achieved the best results on the Statlog (Heart), Breast Cancer Wisconsin, and the Parkinson's data sets, and MPA on Hepatitis and Lung Cancer data sets. Meanwhile, the highest overall FS ratio was 46.90% in the ISMA–kNN approach, versus 47.14% in the second-best optimizer (MPA).
- *Sensitivity (SE) evaluation*
  Table 9 compares the sensitivity (SE) results of the ISMA–kNN approach and the other algorithms under the same conditions. ISMA–kNN achieved the best results in eight out of nine data sets; whereas MPA obtained the best results only on the Lung Cancer data set. Note that ISMA–kNN, MPA, and MFO achieved the same results on the Hepatitis data set with a classification rate was 100%. Nevertheless, the ISMA–kNN approach achieved the most superior output, as its classification rate was 100% on the Hepatitis and Breast Cancer Wisconsin data sets (Table 9). ISMA–kNN also achieves the highest overall sensitivity (93.25%) versus 92.47% in the second-best optimizer (SCA).
- *Specificity (SP) evaluation*
  Table 10 compares the specificity (SP) results of ISMA–kNN and the other algorithms under the same conditions. The ISMA–kNN approach outperformed the other algorithms on all data sets. Note that MPA, and MFO provided the same results as ISMA–kNN on the Hepatitis data set. The ISMA–kNN achieved the highest overall specificity (86.26%; see Table 10), whereas the second-best optimizer (MPA) achieved only 85.31% overall specificity. It should be remembered that Table 10 and the subsequent tables were acquired after 2,000 maximum number of function evaluations (FEs) in each of the 20 runs of each algorithm.
- *Precision (PPV) evaluation*
  Table 11 compares the precision results of ISMA–kNN and the other algorithms executed under the same conditions. The ISMA–kNN achieved the top results in eight out of the nine data sets; the ChOA obtained superior results on the Parkinson's data set. Note that ISMA–kNN provides the same results as the MPA and MFO on the Hepatitis data set. As evidenced in Table 11, the overall precision was highest in ISMA-kNN (87.83%), followed by the MPA optimizer (86.78%).
- *F-score (F1) evaluation*
  Table 12 compares the F-scores (F-measures) of the ISMA–kNN approach and the other algorithms evaluated under the same conditions. ISMA–kNN achieved the best results in eight out of the nine data sets but was outperformed by ChOA on the Lung Cancer data set. The ISMA–kNN approach provides the same results as MPA and MFO on the Hepatitis data set. ISMA–kNN obtained the highest overall F-score (79.12%; see Table 12), versus 78.01% in the second-best optimizer (ChOA).
- *G-Mean evaluation*
  Table 13 compares the G-mean results of the ISMA–kNN approach and the other algorithms

TABLE 6. Best-fitness values of improved slime mould algorithm (ISMA) and other algorithms based on k-nearest neighbor (kNN) classifier.

data sets	ISMA		SMA		TSA		MPA		ChOA		MFO		WOA		SCA	
	$\mu Fitness$	$\sigma Fitness$	$\mu Fitness$	$\sigma Fitness$	$\mu Fitness$	$\sigma Fitness$	$\mu Fitness$	$\sigma Fitness$	$\mu Fitness$	$\sigma Fitness$	$\mu Fitness$	$\sigma Fitness$	$\mu Fitness$	$\sigma Fitness$	$\mu Fitness$	$\sigma Fitness$
Base Brain T91	**0.5002**	**0.0000**	0.7242	**0.0000**	0.7492	0.0134	0.5227	0.0915	0.6497	**0.0000**	0.7242	**0.0000**	0.7247	**0.0000**	0.7176	0.0725
Hepatitis	**0.0274**	**0.0000**	0.0474	**0.0000**	0.1370	0.0199	0.0450	0.0184	0.1089	**0.0000**	0.1141	**0.0000**	0.0579	**0.0000**	0.0573	0.0058
Statlog (Heart)	**0.2062**	**0.0006**	0.2400	0.0069	0.2832	0.0077	0.2450	0.0079	0.2295	0.0009	0.2385	**0.0006**	0.2295	0.0009	0.2742	0.0277
Breast Cancer Wisconsin	**0.0698**	0.0032	0.0727	**0.0000**	0.1372	0.0070	0.0751	0.0066	0.0859	**0.0000**	0.0904	0.0002	0.0793	**0.0000**	0.0895	0.0090
Parkinson's	**0.0273**	**0.0000**	0.1027	0.0057	0.1149	0.0307	0.0695	0.0109	0.0962	0.0020	0.0853	0.0024	0.1098	0.0005	0.0997	0.0101
Bupa	**0.3036**	**0.0000**	0.3353	0.0086	0.3939	0.0027	0.3104	0.0150	0.3275	**0.0000**	0.3275	0.0030	0.3276	0.0009	0.3243	0.0328
Lung Cancer	**0.0064**	0.0091	0.0357	0.0030	0.2463	0.0452	0.0094	0.0168	0.0375	0.0004	0.0482	0.0020	0.0482	**0.0000**	0.0442	0.0045
SPECTF	**0.1414**	0.0228	0.1872	0.0013	0.2896	0.0125	0.1559	0.0190	0.2051	0.0030	0.1813	**0.0000**	0.2084	0.0010	0.2131	0.0215
Leukemia2	**0.0273**	**0.0000**	0.1137	**0.0000**	0.1391	0.0134	0.0393	0.0433	0.1133	0.0050	0.1133	0.0020	0.1135	0.0010	0.1127	0.0114
Overall best finess	**0.1455**	0.0039	0.2065	0.0028	0.2767	0.0169	0.1635	0.0254	0.2059	0.0012	0.2136	0.0011	0.2109	**0.0004**	0.2147	0.0217



(a) Base Brain T91 (b) Hepatitis (c) Statlog (Heart)

(d) Breast Cancer Wisconsin (e) Parkinson's (f) Bupa

(g) Lung Cancer (h) SPECTF (i) Leukemia2

FIGURE 2. Convergence curves of improved slime mould algorithm (ISMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

evaluated under the same conditions. The ISMA-kNN achieved the best results in eight of the nine data sets; MPA outperformed the other algorithms on the Bupa data set. The ISMA–kNN provides the same results as MPA and MFO on the Hepatitis data set and the same results as MPA on the Base Brain T91 data set. Meanwhile, the overall G-mean was highest in ISMA-kNN (80.67%; see Table 13), followed by ChOA (80.05%).

- *Time-consumption evaluation*

Table 14 compares the mean computational times, results of the ISMA–kNN approach and the other SI algorithms based on the kNN classifier. The proposed ISMA–kNN consumed the lowest computational time on three out of nine data sets; whereas TSA achieved superior results on the Hepatitis, Statlog (Heart), Breast Cancer Wisconsin, and Parkinson's data sets, and ChOA obtained superior results only on the Lung Cancer data set. The MPA achieved the best results only on the Leukemia2 data set. It is obvious that ISMA faster than SMA because of integrating OBL in the SMA exploration phase, as well as improving SMA exploitation strategy with the assistance of extra dynamic solutions

TABLE 7. Accuracy comparisons of improved slime mould algorithm (ISMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

data sets	ISMA		SMA		TSA		MPA		ChOA		MFO		WOA		SCA	
	μ_{ACC}	σ_{ACC}	μ_{ACC}	σ_{ACC}	μ_{ACC}	σ_{ACC}	μ_{ACC}	σ_{ACC}	μ_{ACC}	σ_{ACC}	μ_{ACC}	σ_{ACC}	μ_{ACC}	σ_{ACC}	μ_{ACC}	σ_{ACC}
Base Brain T91	**0.5730**	**0.0073**	0.3497	0.0334	0.3041	0.0398	0.3399	0.0228	0.3458	0.0399	0.3227	0.0395	0.3291	0.0321	0.4085	0.0861
Hepatitis	**0.9942**	**0.0085**	0.9741	0.0366	0.9265	0.0621	0.9988	0.0093	0.9779	0.0359	0.9635	0.0711	0.9806	0.0292	0.9888	0.0998
Statlog (Heart)	**0.9041**	**0.0014**	0.8684	0.0165	0.8219	0.0255	0.9045	0.0202	0.8703	0.0203	0.8755	0.0439	0.8720	0.0131	0.9034	0.0923
Breast Cancer Wisconsin	0.9726	0.0044	0.9667	0.0042	0.9640	**0.0018**	**0.9745**	0.0040	0.9668	0.0037	0.9683	0.0048	0.9689	0.0032	0.9617	0.0967
Parkinson's	**0.9719**	**0.0088**	0.9457	0.0183	0.9076	0.0126	0.9688	0.0152	0.9397	0.0203	0.9380	0.0323	0.9448	0.0168	0.9691	0.0989
Bupa	**0.7073**	**0.0120**	0.6904	0.0205	0.6519	0.0296	0.7064	0.0124	0.6717	0.0206	0.6878	0.0392	0.6876	0.0237	0.6966	0.0716
Lung Cancer	**0.8226**	**0.0043**	0.7800	0.0600	0.6198	0.0598	0.7954	0.0300	0.7200	0.0980	0.6172	0.0561	0.7400	0.0917	0.7874	0.0847
SPECTF	0.8680	**0.0139**	0.8461	0.0198	0.7892	0.0396	0.8740	0.0274	0.8339	0.0165	0.8411	0.0788	0.8292	0.0164	**0.8922**	0.0948
Leukemia2	**0.9059**	0.0059	0.8000	0.0286	0.7857	**0.0000**	0.8097	0.0337	0.8036	0.0309	0.7857	**0.0000**	0.8036	0.0309	0.7980	0.0864
Overall Accuracy	**0.8933**	**0.0074**	0.8589	0.025563	0.8083	0.0288	0.8790	0.0190	0.8479	0.0307	0.8346	0.0407	0.8533	0.0281	0.8746	0.0906

TABLE 8. Number of feature selections (FSs) in improved slime mould algorithm (ISMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

data sets	ISMA		SMA		TSA		MPA		ChOA		MFO		WOA		SCA	
	μ_{FS}	σ_{FS}	μ_{FS}	σ_{FS}	μ_{FS}	σ_{FS}	μ_{FS}	σ_{FS}	μ_{FS}	σ_{FS}	μ_{FS}	σ_{FS}	μ_{FS}	σ_{FS}		
Base Brain T91	**0.4323**	**0.4078**	0.4875	0.4892	0.4529	0.4978	0.4516	0.4977	0.4611	0.4985	0.4586	0.4983	0.4581	0.4983	0.4402	0.4897
Hepatitis	0.5023	0.3218	0.5148	0.3298	0.5529	0.3658	**0.4921**	**0.3001**	0.5411	0.4215	0.5486	0.4083	0.4581	0.4700	0.5412	0.4148
Statlog (Heart)	0.4938	0.3201	0.5538	0.4981	**0.4844**	**0.2746**	0.5077	0.5009	0.4992	0.4244	0.5500	0.4985	0.4954	0.4998	0.5346	0.4998
Breast Cancer Wisconsin	0.3856	**0.2900**	0.3802	0.4701	**0.3326**	0.3137	0.3667	0.4832	0.3611	0.4817	0.3578	0.4707	0.3778	0.4862	0.3500	0.4783
Parkinson's	0.4523	0.3983	0.4702	0.473	**0.4148**	**0.3238**	0.4500	0.4981	0.4368	0.4918	0.4864	0.5004	0.4318	0.4959	0.4750	0.4999
Bupa	**0.5417**	**0.2815**	0.5952	0.2345	0.6089	0.3916	0.5417	0.5004	0.5583	0.4987	0.6833	0.4671	0.6333	0.4839	0.6250	0.4862
Lung Cancer	0.4415	0.4001	0.4688	0.5006	0.4750	0.5009	**0.4242**	**0.2798**	0.5062	0.5015	0.4938	0.5015	0.4562	0.4996	0.4938	0.5015
SPECTF	**0.4307**	**0.4493**	0.4901	0.5021	0.4719	0.3205	0.4648	0.4990	0.5159	0.5000	0.5034	0.5003	0.5375	0.4989	0.5091	0.5002
Leukemia2	**0.4819**	**0.32010**	0.5148	0.4401	0.5230	0.4957	0.49253	0.4978	0.4858	0.4981	0.5160	0.5129	0.4972	0.5216	0.5232	0.5019
Overall FS ratio	**0.4624**	**0.3543**	0.4972	0.4375	0.4796	0.3871	0.4657	0.4507	0.4850	0.4795	0.5108	0.4842	0.4828	0.4949	0.4991	0.4858

TABLE 9. Sensitivity comparisons of improved slime mould algorithm (ISMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

data sets	ISMA		SMA		TSA		MPA		ChOA		MFO		WOA		SCA	
	μ_{SE}	σ_{SE}	μ_{SE}	σ_{SE}	μ_{SE}	σ_{SE}	μ_{SE}	σ_{SE}	μ_{SE}	σ_{SE}	μ_{SE}	σ_{SE}	μ_{SE}	σ_{SE}	μ_{SE}	σ_{SE}
Base Brain T91	**0.9459**	**0.0002**	0.9300	0.0780	0.9449	0.0009	0.9449	0.0012	0.9355	0.0940	0.9321	0.0016	0.9449	0.0006	0.9394	0.0403
Hepatitis	**1.0000**	**0.0000**	0.9900	0.0995	0.9903	**0.0000**	**1.0000**	**0.0000**	0.9900	0.0995	**1.0000**	**0.0000**	0.9930	0.0995	0.9900	0.0995
Statlog (Heart)	**0.9292**	**0.0003**	0.9004	0.0256	0.9278	0.0012	0.9279	0.0007	0.9186	0.0924	0.9279	0.0008	0.9045	0.0032	0.9234	0.0344
Breast Cancer Wisconsin	**1.0000**	**0.0003**	0.9986	0.0107	0.9999	0.0006	0.9999	**0.0003**	0.9900	0.0995	0.9999	0.0008	0.9989	0.0004	0.9986	0.0107
Parkinson's	**0.9969**	**0.0238**	0.9900	0.0414	0.9960	0.0236	0.9964	0.0225	0.9874	0.0325	0.9894	0.0999	0.9922	0.0324	0.9847	0.0604
Bupa	**0.7954**	**0.0302**	0.7645	0.0510	0.7939	0.0348	0.7919	0.0390	0.7856	0.0865	0.7524	0.0501	0.7821	0.0610	0.7530	0.1010
Lung Cancer	**0.9271**	**0.0023**	0.9015	0.0230	0.9214	0.0200	0.9220	0.0141	0.9264	0.0119	0.9056	0.0925	0.9198	0.0164	0.9189	0.0252
SPECTF	0.7995	0.0621	0.7496	0.0862	0.7608	0.0828	0.6808	0.0968	0.6854	**0.0351**	0.7456	0.1102	0.6955	0.1048	**0.8257**	0.1744
Leukemia2	**0.9989**	**0.0004**	0.9704	0.0305	0.9814	0.0305	0.9849	0.0214	0.9898	0.0061	0.9823	0.0741	0.9899	0.0140	0.9891	0.0325
Overall Senstivity	**0.9325**	**0.0132**	0.9105	0.0495	0.9240	0.0216	0.9144	0.0217	0.9120	0.0619	0.9139	0.0478	0.9134	0.0369	0.9247	0.0642

TABLE 10. Specificity comparisons of improved slime mould algorithm (ISMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

data sets	ISMA		SMA		TSA		MPA		ChOA		MFO		WOA		SCA	
	μ_{SP}	σ_{SP}	μ_{SP}	σ_{SP}	μ_{SP}	σ_{SP}	μ_{SP}	σ_{SP}	μ_{SP}	σ_{SP}	μ_{SP}	σ_{SP}	μ_{SP}	σ_{SP}	μ_{SP}	σ_{SP}
Base Brain T91	**0.6342**	**0.0007**	0.6095	0.0077	0.6055	0.0940	0.6095	0.0077	0.6095	0.0077	0.6037	0.0607	0.6096	0.0055	0.5972	0.0587
Hepatitis	**1.0000**	**0.0000**	0.9898	0.0110	0.9909	0.0025	**1.0000**	**0.0000**	0.9900	0.0995	**1.0000**	**0.0000**	0.9908	0.0210	0.9890	0.0995
Statlog (Heart)	**0.5421**	0.0039	0.5127	0.0027	0.5012	0.0132	0.5128	**0.0011**	0.5128	0.0016	0.5077	0.0510	0.5127	0.0028	0.5081	0.0292
Breast Cancer Wisconsin	**0.9815**	**0.0025**	0.9801	0.0041	0.9805	0.0050	0.9814	0.0028	0.9718	0.0977	0.9687	0.0035	0.9799	0.0057	0.9780	0.0203
Parkinson's	**0.9983**	**0.0042**	0.9883	0.0084	0.9970	0.0079	0.9966	0.0109	0.9989	0.0067	0.9863	0.0997	0.9980	0.0089	0.9925	0.0257
Bupa	**0.9112**	**0.0056**	0.8902	0.0071	0.9002	0.0065	0.9102	0.0081	0.9108	0.0073	0.9023	0.0908	0.9088	0.0107	0.9004	0.0428
Lung Cancer	**0.6982**	**0.0016**	0.6774	0.0126	0.6761	0.0226	0.6775	0.0132	0.6774	0.0120	0.6688	0.0683	0.6684	0.0146	0.6670	0.0485
SPECTF	**0.9994**	**0.0047**	0.9801	0.0087	0.9990	0.0054	0.9993	0.0049	0.9992	0.0055	0.9884	0.0996	0.9976	0.0080	0.9922	0.0221
Leukemia2	**0.9989**	**0.0329**	0.9812	0.0425	0.9814	0.0454	0.9914	0.0074	0.9620	0.1013	0.9834	0.1027	0.9754	0.0874	0.8795	0.2143
Overall Specificity	**0.8626**	**0.0062**	0.8454	0.0116	0.8479	0.0225	0.8531	0.0062	0.8480	0.0377	0.8454	0.0640	0.8490	0.0182	0.8337	0.0623

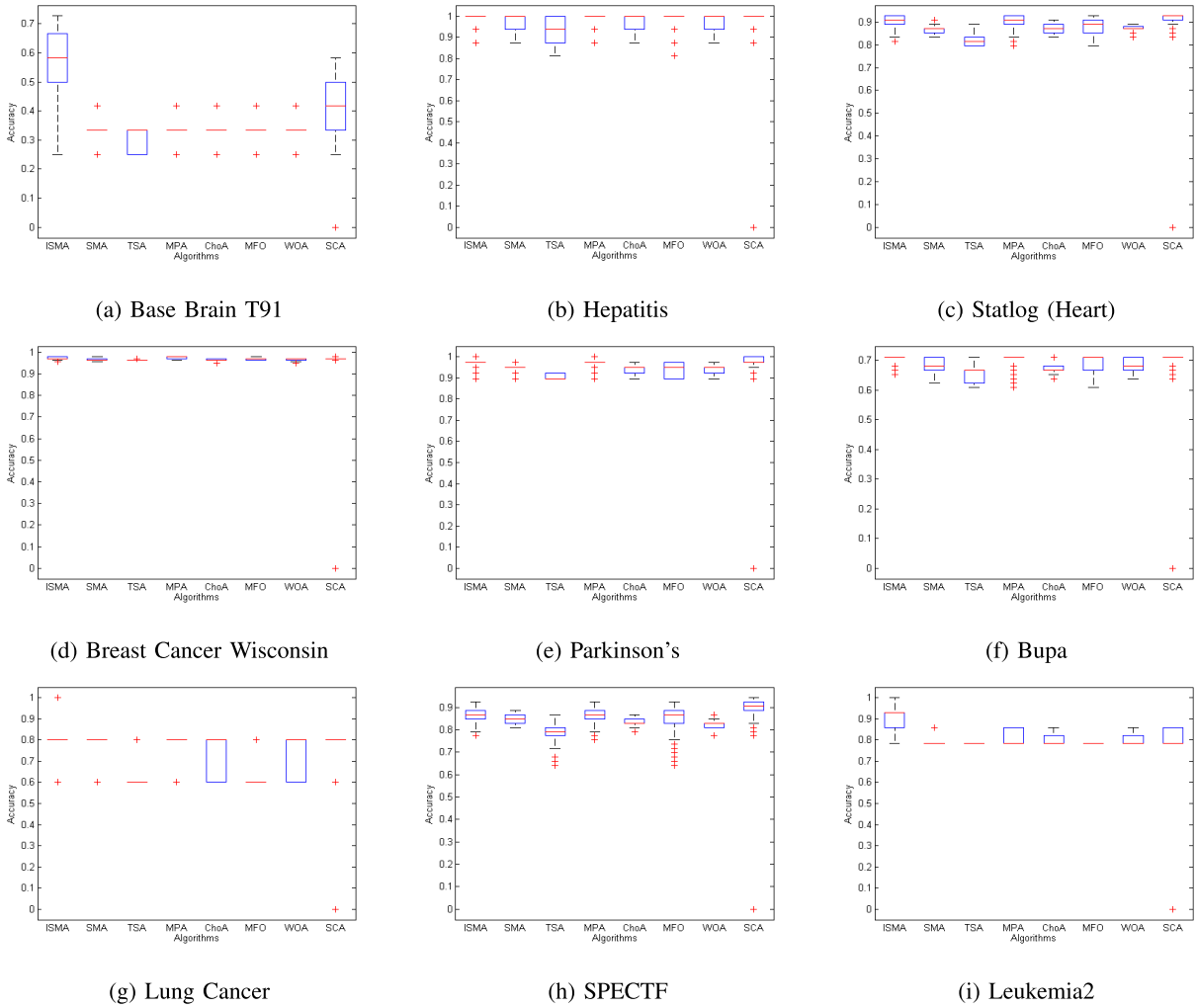
TABLE 11. Precision comparisons of improved slime mould algorithm (ISMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

data sets	ISMA		SMA		TSA		MPA		ChOA		MFO		WOA		SCA	
	μ_{PPV}	σ_{PPV}	μ_{PPV}	σ_{PPV}	μ_{PPV}	σ_{PPV}	μ_{PPV}	σ_{PPV}	μ_{PPV}	σ_{PPV}	μ_{PPV}	σ_{PPV}	μ_{PPV}	σ_{PPV}	μ_{PPV}	σ_{PPV}
Base Brain T91	**0.7321**	**0.0024**	0.7045	0.0090	0.7231	0.0049	**0.7141**	0.0054	0.7068	0.0715	0.7140	0.0075	0.7139	0.0032	0.7058	0.0504
Hepatitis	**1.0000**	**0.0000**	0.9930	0.0991	0.9890	0.0745	**1.0000**	**0.0000**	0.9900	0.0995	**1.0000**	**0.0000**	0.9990	0.0100	0.9990	0.0100
Statlog (Heart)	**0.6108**	**0.0002**	0.5908	0.0008	0.5808	0.0124	0.5909	0.0006	0.5908	0.0013	0.5848	0.0588	0.5907	0.0011	0.5860	0.0333
Breast Cancer Wisconsin	**0.9915**	**0.0025**	0.9603	0.0063	0.9795	0.0033	0.9805	0.0050	0.9814	0.0028	0.9718	0.0977	0.9799	0.0057	0.9780	0.0203
Parkinson's	0.9990	0.0084	0.9893	0.0046	0.9992	0.0054	0.9982	0.0074	**0.9997**	**0.0034**	0.9895	0.0996	0.9987	0.0063	0.9956	0.0152
Bupa	**0.8825**	**0.0061**	0.8616	0.0100	0.8542	0.0125	0.8616	0.0111	0.8625	0.0073	0.8502	0.0865	0.8607	0.0092	0.8502	0.0511
Lung Cancer	**0.7939**	**0.0183**	0.7829	0.0209	0.7800	0.0256	0.7834	0.0207	0.7849	0.0193	0.7629	0.0793	0.7761	0.0235	0.7683	0.0457
SPECTF	**0.9000**	**0.0236**	0.8955	0.0296	0.8905	0.0265	0.8923	0.0286	0.8917	0.0291	**0.9000**	0.0936	0.8967	0.0229	0.9247	0.0476
Leukemia2	**0.9940**	**0.0133**	0.9899	0.0310	0.9721	0.0325	0.9898	0.0193	0.9809	0.0321	0.9801	0.1003	0.9742	0.0399	0.9595	0.0421
Overall Precision	**0.8783**	**0.0083**	0.8631	0.0234	0.8631	0.0219	0.8678	0.0109	0.8654	0.0295	0.8615	0.0692	0.8655	0.0135	0.8630	0.0350

to avoid stagnancy issues, using the hybridization technique ISMA. The TSA was the fastest optimizer, with an overall consumption time of 14.33 seconds, followed by ISMA (14.50 seconds).

D. CONVERGENCE ANALYSIS

This paper evaluated different disease data sets by the proposed ISMA–kNN approach. The best fitness results of the optimization algorithms as well as their means are compared in Table 6. Through these analyses, we can obtain the convergence curves of ISMA-kNN and estimate the algorithm's convergence potential based on the kNN classifier. Fig. 2 compares the convergence curves of the ISMA, TSA, MPA, ChOA, MFO, WOA, SCA, and the original SMA based on the kNN classifier under the same conditions (population size, maximum number



(a) Base Brain T91 (b) Hepatitis (c) Statlog (Heart)

(d) Breast Cancer Wisconsin (e) Parkinson's (f) Bupa

(g) Lung Cancer (h) SPECTF (i) Leukemia2

FIGURE 3. Boxplots of slime mould algorithm (SMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

TABLE 12. F-score comparisons of improved slime mould algorithm (ISMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

data sets	ISMA		SMA		TSA		MPA		ChOA		MFO		WOA		SCA	
	μ_{F1}	σ_{F1}	μ_{F1}	σ_{F1}	μ_{F1}	σ_{F1}	μ_{F1}	σ_{F1}	μ_{F1}	σ_{F1}	μ_{F1}	σ_{F1}	μ_{F1}	σ_{F1}	μ_{F1}	σ_{F1}
Base Brain T91	**0.5128**	**0.0010**	0.4838	0.0015	0.4355	0.0320	0.4838	**0.0010**	0.4838	0.0016	0.4789	0.0482	0.4837	0.0032	0.4800	0.0243
Hepatitis	**1.0000**	**0.0000**	0.9989	0.0045	0.9998	0.0075	**1.0000**	**0.0000**	0.9979	0.0065	**1.0000**	**0.0000**	0.9976	0.0047	0.9978	0.0065
Statlog (Heart)	**0.5144**	0.0023	**0.4994**	0.0075	0.5022	0.0035	0.5143	**0.0016**	0.5142	0.0023	0.5091	0.0512	0.5142	0.0018	0.5098	0.0236
Breast Cancer Wisconsin	**0.9615**	**0.0025**	0.9401	0.0075	0.9515	0.0045	0.9605	0.0050	0.9614	0.0028	0.9518	0.0977	0.9599	0.0057	0.9580	0.0203
Parkinson's	**0.9933**	**0.0080**	0.9833	0.0093	0.9763	0.0107	0.9793	0.0097	0.9852	0.0096	0.9795	0.0991	0.9789	0.0120	0.9845	0.0198
Bupa	**0.6868**	**0.0057**	0.6721	0.0123	0.6842	0.0065	0.6861	0.0074	0.6866	0.0066	0.6794	0.0684	0.6838	0.0109	0.6728	0.0425
Lung Cancer	0.7100	0.0160	0.7107	0.0159	0.7025	0.0156	0.7096	0.0179	**0.7137**	**0.0136**	0.6970	0.0719	0.6984	0.0201	0.6922	0.0458
SPECTF	**0.9389**	**0.0087**	0.9144	0.0102	0.9123	0.0245	0.9156	0.0145	0.9159	0.0122	0.9096	0.0922	0.9139	0.0127	0.9279	0.0287
Leukemia2	**0.7999**	0.0625	0.7701	0.0705	0.7425	0.0421	0.7504	0.0321	0.7622	**0.0202**	0.7513	0.0789	0.7519	0.0373	0.7562	0.0301
Overall F-score	**0.7912**	0.0115	0.7747	0.0154	0.7674	0.0163	0.7777	0.0099	0.7801	**0.0083**	0.7729	0.0675	0.7758	0.0120	0.7754	0.0268

of function evaluations (FEs), and runs). As shown in the figure, the ISMA–kNN approach rapidly converged on all data sets. Observing the convergence behavior of ISMA–kNN, we find that the optimal fitness values exactly corresponded to the optimal accuracies. Fig. 2 presents the convergence curves of the proposed ISMA approach based on the kNN classifier and for the nine disease data sets with different feature sizes. The proposed algorithm reached a stable point for most of the data sets, which implies that the algorithm converged. In comparison with other algorithms, the proposed algorithm reached the lowest average of the best solutions thus far, higher for all disease data sets than the original SMA. On the disease data sets, the ISMA-kNN approach converged the fastest (see Fig. 2). Rapid convergence to the (near-) optimal solution was observed, which makes the proposed ISMA-kNN approach a promising optimization algorithm to solve problems that require rapid computation.

E. DISCUSSION

The purpose of this research is to provide an efficient search technique for the feature selection problem in low and high-dimensional data sets. The study advocated integrating

TABLE 13. G-mean comparisons of improved slime mould algorithm (ISMA) and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

data sets	ISMA		SMA		TSA		MPA		ChOA		MFO		WOA		SCA	
	μ_G	σ_G	μ_G	σ_G	μ_G	σ_G	μ_G	σ_G	μ_G	σ_G	μ_G	σ_G	μ_G	σ_G	μ_G	σ_G
Base Brain T91	**0.5880**	**0.0006**	0.5879	0.0009	0.5866	0.0010	**0.5880**	**0.0006**	0.5879	0.0008	0.5821	0.0585	0.5879	0.0012	0.5856	0.0174
Hepatitis	**1.0000**	**0.0000**	0.9980	0.0975	0.9890	0.0841	**1.0000**	**0.0000**	0.9900	0.0995	**1.0000**	**0.0000**	0.9904	0.0011	0.9898	0.0095
Statlog (Heart)	**0.6383**	**0.0020**	0.6382	0.0029	0.6374	0.0032	0.6300	0.0039	0.6383	0.0022	0.6318	0.0636	0.6382	0.0030	0.6344	0.0199
Breast Cancer Wisconsin	**0.9705**	**0.0025**	0.9615	0.0039	0.9532	0.0042	0.9605	0.0050	0.9614	0.0028	0.9518	0.0977	0.9599	0.0057	0.9580	0.0203
Parkinson's	**0.9804**	**0.0076**	0.9716	0.0176	0.9532	0.0167	0.9752	0.0177	0.9727	0.0157	0.9762	0.0992	0.9674	0.0206	0.9733	0.0443
Bupa	0.7080	0.0046	0.7090	0.0056	0.6999	0.0066	**0.7092**	**0.0045**	0.6987	0.0055	0.7015	0.0707	0.7066	0.0088	0.7010	0.0356
Lung Cancer	**0.7725**	**0.0100**	0.7658	0.0110	0.7428	0.0130	0.7644	0.0117	0.7675	**0.0100**	0.7550	0.0766	0.7605	0.0143	0.7535	0.0358
SPECTF	**0.8116**	**0.0037**	0.7716	0.0467	0.7652	0.0525	0.7913	0.0468	0.8054	0.0451	0.8088	0.0903	0.7961	0.0534	0.8044	0.1098
Leukemia2	**0.7897**	**0.0025**	0.7888	0.0321	0.7688	0.0321	0.7897	**0.0165**	0.7831	0.0201	0.7809	0.0647	0.7800	0.0200	0.7645	0.0624
Overall G-mean	**0.8067**	**0.0037**	0.7992	0.0242	0.7884	0.0237	0.8009	0.0118	0.8005	0.0224	0.7987	0.0690	0.7985	0.0142	0.7961	0.0394

TABLE 14. Time-consuming of ISMA and other SI algorithms based on the k-nearest neighbor (kNN) classifier.

data sets	ISMA	SMA	TSA	MPA	ChOA	MFO	WOA	SCA
	μ_{Time}	μ_{Time}	μ_{Time}	μ_{Time}	μ_{Time}	μ_{Time}	μ_{Time}	μ_{Time}
Base Brain T91	**19.11**	23.88	22.39	21.08	19.11	23.82	19.98	21.29
Hepatitis	18.94	20.02	**16.39**	17.1743	16.43	17.39	19.03	16.92
Statlog (Heart)	17.65	23.2	**15.59**	16.19	16.00	20.43	15.60	17.90
Breast Cancer Wisconsin	9.11	9.83	**8.99**	19.64	9.78	9.73	9.93	9.37
Parkinson's	7.48	8.33	**7.39**	15.83	8.08	7.75	7.83	8.04
Bupa	**8.22**	8.67	8.43	17.44	9.65	8.53	8.64	8.47
Lung Cancer	15.34	17.01	14.99	14.77	**14.35**	15.74	14.56	14.58
SPECTF	**19.14**	21.20	18.96	20.17	21.22	18.96	15.18	18.46
Leukemia2	15.52	18.24	15.81	**14.88**	17.85	21.07	21.76	16.70
Overall elapsed time	14.50	16.71	**14.33**	17.46	14.72	15.93	14.72	14.63

OBL in the SMA exploration phase, as well as improving the SMA exploitation strategy with the assistance of extra dynamic solutions to avoid stagnancy issues, using the hybridization technique ISMA. The experimental analysis and comparative study performed in the previous section, using numerical optimization and feature selection issues, indicate that the proposed methodology has been proven to be effective.

The proposed ISMA algorithm, being an optimization method, presents certain advantages:

- ISMA can perform an efficient search of optimization landscapes with varying levels of difficulty and complexity. Table 6 shows how ISMA delivers optimization solutions with better fitness values than original and other population-based optimization approaches. The proposed hybridization also proves to enhance the convergence ability of the algorithm, see Fig. 2. However, as shown in Table 6, the proposed approach could be enhanced further to outperform state-of-the-art constrained optimization algorithms.
- ISMA has lower computational complexity than the original SMA and can yield more efficient solutions than SMA and other related optimization methods, as shown in Table 14.
- The data sets used for this study provide enough testing environment for an optimization technique, with feature sizes ranging from 7 to more than 11226 features. Table 8 shows that ISMA reduced feature size by up to 46.24% on average across all data sets, which is better than the original and other approaches utilized in this study.
- In terms of accuracy, ISMA-kNN approach achieve an average accuracy up to 89.33% across all classification data sets, as shown in Table 7 and Fig. 3.
- The performance evaluation metrics generated by ISMA are statistically significantly different from those generated by the original and other state-of-the-art constrained optimization algorithms (see Table 9 for sensitivity evaluation metrics, Table 10 for specificity evaluation metrics, Table 11 for precision evaluation metrics, Table 12 for F-score evaluation metrics, Table 13 for G-mean evaluation metrics, and Table 14 for time-consuming evaluation).
- OBL is utilized to avoid premature convergence as well as to improve global exploratory abilities.
- ISMA's design is straightforward, therefore any potential to improve the method is simple to implement.

In addition to its advantages, the proposed ISMA has some limitations, which are detailed below:

- ISMA's selected features may change each time it is performed because it is a randomization-based optimization technique. As a result, there is no guarantee that the features subset selected in one run will be found in another, potentially leading to user confusion.
- Because of its simplicity and low computing cost, this work used kNN as a learning algorithm in a wrapper-based feature selection strategy. However, kNN has certain limitations, such as being a slow learner and being vulnerable to noisy data.

VI. CONCLUSION AND FUTURE WORK

This paper proposed an efficient classification approach that combined the Opposition-Based learning (OBL) and the slime mould algorithm (SMA) based on k-nearest neighbor (kNN) called ISMA–kNN for reducing the feature selection (FS) and classification purpose. The proposed algorithm ISMA aims to overcome the drawbacks of the original SMA, which are getting trapped into local minimum regions and

suffering from an inadequate balance between exploitation and exploration, particularly when solving high-dimension problems. The performances of ISMA–kNN and various existing swarm intelligence (SI) algorithms based on the kNN classifier were quantified by various assessment metrics on nine disease data sets with medium and high dimensionality features. The proposed ISMA–kNN approach generally outperformed the well-known state-of-the-art constrained optimization algorithms such as TSA, MPA, ChOA, MFO, WOA, SCA, and the original SMA algorithm based on the kNN classifier. On most of the data sets, the ISMA–kNN classification approach has been achieved the lowest number of feature selection with the highest classification accuracy within a reasonable period.

In future work, the ISMA algorithm can be proposed as a filter FS method with other classifiers such as support vector machines (SVM), neural networks, and logistic regression. These studies would assess the generality of the selected features and the classification accuracy in different scenarios. In addition, promisingly, the proposed ISMA–kNN approach can be viewed as an efficient and effective strategy for more complex optimization scenarios and the intelligent optimization field's theoretical work as well.

DECLARATION OF COMPETING INTEREST

The authors declare that there is no conflict of interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

COMPLIANCE WITH ETHICAL STANDARDS

This article does not contain any studies with human participants or animals performed by any of the authors.

CREDIT AUTHOR STATEMENT

All authors contributed equally to this paper, where; Yaser M. Wazery: Conceptualization, Formal analysis, Writing - review & editing. Eman Saber: Software, Resources, Data curation, Writing - original draft. Essam H. Houssein: Supervision, Methodology, Formal analysis, Software, Writing - review & editing. Abdelmgeid A. Ali: Supervision, Methodology, Formal analysis. Eslam Amer: Conceptualization, Formal analysis, Writing - review & editing. All authors read and approved the final paper.

REFERENCES

[1] K. J. Cios, W. Pedrycz, and R. W. Swiniarski, *Data Mining Methods for Knowledge Discovery*, vol. 458. Springer, 2012.
[2] K. Hussain, N. Neggaz, W. Zhu, and E. H. Houssein, "An efficient hybrid sine-cosine Harris hawks optimization for low and high-dimensional feature selection," *Expert Syst. Appl.*, vol. 176, Aug. 2021, Art. no. 114778.
[3] E. H. Houssein, I. E. Ibrahim, N. Neggaz, M. Hassaballah, and Y. M. Wazery, "An efficient ECG arrhythmia classification method based on manta ray foraging optimization," *Expert Syst. Appl.*, vol. 181, Nov. 2021, Art. no. 115131.
[4] Q. Al-Tashi, H. M. Rais, S. J. Abdulkadir, S. Mirjalili, and A. Alhussian, "A review of grey wolf optimizer-based feature selection methods for classification," in *Evolutionary Machine Learning Techniques*. Springer, 2020, pp. 273–286.
[5] D. S. A. Elminaam, A. Nabil, S. A. Ibraheem, and E. H. Houssein, "An efficient Marine Predators Algorithm for feature selection," *IEEE Access*, vol. 9, pp. 60136–60153, 2021.
[6] E. H. Houssein, M. Kilany, and A. E. Hassanien, "ECG signals classification: A review," *Int. J. Intell. Eng. Inform.*, vol. 5, no. 4, pp. 376–396, 2017.
[7] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerg. Artif. Intell. Appl. Comput. Eng.*, vol. 160, no. 1, pp. 3–24, 2007.
[8] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.
[9] S. Jadhav, H. He, and K. Jenkins, "Information gain directed genetic algorithm wrapper feature selection for credit rating," *Appl. Soft Comput.*, vol. 69, pp. 541–553, Aug. 2018.
[10] A. Wang, N. An, G. Chen, L. Li, and G. Alterovitz, "Accelerating wrapper-based feature selection with K-nearest-neighbor," *Knowl.-Based Syst.*, vol. 83, pp. 81–91, Jul. 2015.
[11] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, "Efficient kNN classification with different numbers of nearest neighbors," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1774–1785, May 2018.
[12] E. H. Houssein, M. M. Emam, and A. A. Ali, "An efficient multilevel thresholding segmentation method for thermography breast cancer imaging based on improved chimp optimization algorithm," *Expert Syst. Appl.*, vol. 185, Dec. 2021, Art. no. 115651.
[13] E. H. Houssein, K. Hussain, L. Abualigah, M. A. Elaziz, W. Alomoush, G. Dhiman, Y. Djenouri, and E. Cuevas, "An improved opposition-based Marine Predators Algorithm for global optimization and multilevel thresholding image segmentation," *Knowl.-Based Syst.*, vol. 229, Oct. 2021, Art. no. 107348.
[14] C. Huang, J. Du, B. Nie, R. Yu, W. Xiong, and Q. Zeng, "Feature selection method based on partial least squares and analysis of traditional Chinese medicine data," *Comput. Math. Methods Med.*, vol. 2019, pp. 1–11, Jul. 2019.
[15] B. Remeseiro and V. Bolon-Canedo, "A review of feature selection methods in medical applications," *Comput. Biol. Med.*, vol. 112, Sep. 2019, Art. no. 103375.
[16] M. M. Ahmed, E. H. Houssein, A. E. Hassanien, A. Taha, and E. Hassanien, "Maximizing lifetime of large-scale wireless sensor networks using multi-objective whale optimization algorithm," *Telecommun. Syst.*, vol. 72, no. 2, pp. 243–259, Oct. 2019.
[17] E. H. Houssein, A. A. Ewees, and M. A. Elaziz, "Improving twin support vector machine based on hybrid swarm optimizer for heartbeat classification," *Pattern Recognit. Image Anal.*, vol. 28, no. 2, pp. 243–253, Apr. 2018.
[18] H. R. Tizhoosh, "Opposition-based learning: A new scheme for machine intelligence," in *Proc. Int. Conf. Comput. Intell. Modelling, Control Automat.*, vol. 1, Nov. 2005, pp. 695–701.
[19] M. Tubishat, N. Idris, L. Shuib, M. A. M. Abushariah, and S. Mirjalili, "Improved Salp Swarm Algorithm based on opposition based learning and novel local search algorithm for feature selection," *Expert Syst. Appl.*, vol. 145, May 2020, Art. no. 113122.
[20] S. Kaur, L. K. Awasthi, A. L. Sangal, and G. Dhiman, "Tunicate Swarm Algorithm: A new bio-inspired based metaheuristic paradigm for global optimization," *Eng. Appl. Artif. Intell.*, vol. 90, Apr. 2020, Art. no. 103541.
[21] A. Faramarzi, M. Heidarinejad, S. Mirjalili, and A. H. Gandomi, "Marine Predators Algorithm: A nature-inspired metaheuristic," *Expert Syst. Appl.*, vol. 152, Aug. 2020, Art. no. 113377.
[22] M. Khishe and M. R. Mosavi, "Classification of underwater acoustical dataset using neural network trained by chimp optimization algorithm," *Appl. Acoust.*, vol. 157, Jan. 2020, Art. no. 107005.
[23] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm," *Knowl.-Based Syst.*, vol. 89, pp. 228–249, Nov. 2015.
[24] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016.
[25] S. M. Mirjalili, S. Z. Mirjalili, S. Saremi, and A. Mirjalili, "Sine cosine algorithm: Theory, literature review, and application in designing bend photonic crystal waveguides," in *Nature-Inspired Optimizers*. Springer, 2020, pp. 201–217.
[26] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: A new method for stochastic optimization," *Future Gener. Comput. Syst.*, vol. 111, pp. 300–323, Oct. 2020.

[27] M. Abdel-Basset, R. Mohamed, R. K. Chakrabortty, M. J. Ryan, and S. Mirjalili, "An efficient binary slime mould algorithm integrated with a novel attacking-feeding strategy for feature selection," *Comput. Ind. Eng.*, vol. 153, Mar. 2021, Art. no. 107078.

[28] A. A. Ewees, L. Abualigah, D. Yousri, Z. Y. Algamal, M. A. A. Al-qaness, R. A. Ibrahim, and M. A. Elaziz, "Improved slime mould algorithm based on firefly algorithm for feature selection: A case study on QSAR model," *Eng. with Comput.*, vol. 1, pp. 1–15, Mar. 2021.

[29] M. Wang and H. Chen, "Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis," *Appl. Soft Comput.*, vol. 88, Mar. 2020, Art. no. 105946.

[30] D. Oliva and M. A. Elaziz, "An improved brainstorm optimization using chaotic opposite-based learning with disruption operator for global optimization and feature selection," *Soft Comput.*, vol. 24, no. 18, pp. 14051–14072, 2020.

[31] D. K. Choubey, P. Kumar, S. Tripathi, and S. Kumar, "Performance evaluation of classification methods with PCA and PSO for diabetes," *Netw. Model. Anal. Health Informat. Bioinf.*, vol. 9, no. 1, p. 5, Dec. 2020.

[32] U. N. Chowdhury, M. B. Islam, S. Ahmad, and M. A. Moni, "Systems biology and bioinformatics approach to identify gene signatures, pathways and therapeutic targets of Alzheimer's disease," *Informat. Med. Unlocked*, vol. 21, 2020, Art. no. 100439.

[33] A. Kewat, P. Srivastava, and D. Kumhar, "Performance evaluation of wrapper-based feature selection techniques for medical datasets," in *Advances in Computing and Intelligent Systems*. Singapore: Springer, 2020, pp. 619–633.

[34] R. J. S. Raj, S. J. Shobana, I. V. Pustokhina, D. A. Pustokhin, D. Gupta, and K. Shankar, "Optimal feature selection-based medical image classification using deep learning model in Internet of Medical Things," *IEEE Access*, vol. 8, pp. 58006–58017, 2020.

[35] L. Wang, Y. Gao, J. Li, and X. Wang, "A feature selection method by using chaotic cuckoo search optimization algorithm with elitist preservation and uniform mutation for data classification," *Discrete Dyn. Nature Soc.*, vol. 2021, pp. 1–19, Jun. 2021.

[36] R. Hans and H. Kaur, "Opposition-based enhanced grey wolf optimization algorithm for feature selection in breast density classification," *Int. J. Mach. Learn. Comput.*, vol. 10, no. 3, pp. 458–464, May 2020.

[37] W. M. Shaban, A. H. Rabie, A. I. Saleh, and M. A. Abo-Elsoud, "A new COVID-19 patients detection strategy (CPDS) based on hybrid feature selection and enhanced KNN classifier," *Knowl.-Based Syst.*, vol. 205, Oct. 2020, Art. no. 106270.

[38] M. A. Elaziz, A. A. Ewees, R. A. Ibrahim, and S. Lu, "Opposition-based moth-flame optimization improved by differential evolution for feature selection," *Math. Comput. Simul.*, vol. 168, pp. 48–75, Feb. 2020.

[39] A. E. Hegazy, M. Makhlouf, and G. S. El-Tawel, "Improved Salp Swarm Algorithm for feature selection," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 32, no. 2, pp. 335–344, 2018.

[40] M. Aladeemy, L. Adwan, A. Booth, M. T. Khasawneh, and S. Poranki, "New feature selection methods based on opposition-based learning and self-adaptive cohort intelligence for predicting patient no-shows," *Appl. Soft Comput.*, vol. 86, Jan. 2020, Art. no. 105866.

[41] R. Hans and H. Kaur, "Binary multi-verse optimization (BMVO) approaches for feature selection," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 6, no. 1, p. 91, 2020.

[42] M. Kelidari and J. Hamidzadeh, "Feature selection by using chaotic cuckoo optimization algorithm with levy flight, opposition-based learning and disruption operator," *Soft Comput.*, vol. 25, no. 4, pp. 2911–2933, Feb. 2021.

[43] M. K. Parizi, F. Keynia, and A. K. Bardsiri, "OWMA: An improved self-regulatory woodpecker mating algorithm using opposition-based learning and allocation of local memory for solving optimization problems," *J. Intell. Fuzzy Syst.*, vol. 40, no. 1, pp. 919–946, Jan. 2021.

[44] A. A. Ewees, M. A. El Aziz, and A. E. Hassanien, "Chaotic multi-verse optimizer-based feature selection," *Neural Comput. Appl.*, vol. 31, no. 4, pp. 991–1006, Apr. 2019.

[45] N. Neggaz, E. H. Houssein, and K. Hussain, "An efficient Henry gas solubility optimization for feature selection," *Expert Syst. Appl.*, vol. 152, Aug. 2020, Art. no. 113364.

[46] D. Chicco and G. Jurman, "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone," *BMC Med. Informat. Decis. Making*, vol. 20, no. 1, p. 16, Dec. 2020.

[47] W. Long, T. Wu, X. Liang, and S. Xu, "Solving high-dimensional global optimization problems using an improved sine cosine algorithm," *Expert Syst. Appl.*, vol. 123, pp. 108–126, Jun. 2019.

[48] R. Kadry and O. Ismael, "A new hybrid KNN classification approach based on particle swarm optimization," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 11, pp. 1–6, 2020.

[49] S. Saha and V. Mukherjee, "A novel quasi-oppositional chaotic antlion optimizer for global optimization," *Int. J. Speech Technol.*, vol. 48, no. 9, pp. 2628–2660, Sep. 2018.

[50] H. Muthusamy, S. Ravindran, S. Yaacob, and K. Polat, "An improved elephant herding optimization using sine–cosine mechanism and opposition based learning for global optimization problems," *Expert Syst. Appl.*, vol. 172, Jun. 2021, Art. no. 114607.

[51] A. Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, "A hybrid eeg signals classification approach based on grey wolf optimizer enhanced svms for epileptic detection," in *Proc. Int. Conf. Adv. Intell. Syst. Inform.* Cham, Switzerland: Springer, 2017, pp. 108–117.

[52] E. H. Houssein, M. E. Hosney, D. Oliva, W. M. Mohamed, and M. Hassaballah, "A novel hybrid Harris hawks optimization and support vector machines for drug design and discovery," *Comput. Chem. Eng.*, vol. 133, Feb. 2020, Art. no. 106656.

[53] M. Mafarja, I. Aljarah, A. A. Heidari, H. Faris, P. Fournier-Viger, X. Li, and S. Mirjalili, "Binary dragonfly optimization for feature selection using time-varying transfer functions," *Knowl.-Based Syst.*, vol. 161, pp. 185–204, Dec. 2018.

[54] A. Darwish, G. Sayed, and A. Hassanien, "Meta-heuristic optimization algorithms based feature selection for clinical breast cancer diagnosis," *J. Egyptian Math. Soc.*, vol. 26, no. 2, pp. 321–336, Apr. 2018.

[55] A. Subasi, "Use of artificial intelligence in alzheimer's disease detection," in *Artificial Intelligence in Precision Health*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 257–278.

[56] B. Boukenze, H. Mousannif, and A. Haqiq, "Predictive analytics in healthcare system using data mining techniques," *Comput Sci Inf Technol*, vol. 1, pp. 1–9, 2016.

[57] Frank. (2010). *Machine Learning Repository*. [Online]. Available: https://archive.ics.uci.edu/ml/index.php

[58] F. Yan, X. Xu, and J. Xu, "Grey wolf optimizer with a novel weighted distance for global optimization," *IEEE Access*, vol. 8, pp. 120173–120197, 2020.

[59] C. M. Rahman and T. A. Rashid, "Dragonfly algorithm and its applications in applied science survey," *Comput. Intell. Neurosci.*, vol. 2019, pp. 1–21, Dec. 2019.

[60] A. Arcuri and G. Fraser, "Parameter tuning or default values? An empirical investigation in search-based software engineering," *Empirical Softw. Eng.*, vol. 18, pp. 594–623, Jun. 2013.

[61] M. Khishe and M. R. Mosavi, "Chimp optimization algorithm," *Expert Syst. Appl.*, vol. 149, Jul. 2020, Art. no. 113338.

[62] S. Mirjalili, "SCA: A sine cosine algorithm for solving optimization problems," *Knowl.-Based Syst.*, vol. 96, pp. 120–133, Mar. 2016.

YASER M. WAZERY received the Ph.D. degree in information technology. He is currently working as an Associate Professor with the Information Technology Department, Faculty of Computers and Information, Minia University, Egypt. His research interests include network security, cryptography, biological inspired clustering, multimedia communications and processing, big data, cloud computing, and data mining.

EMAN SABER is currently pursuing the M.Sc. degree in computer science with the Faculty of Computers and Information, Minia University, Egypt. She is also working as a System Engineer with the Information Technology Center, Minia University. Her research interests include machine learning, data mining, and optimization.

ESSAM H. HOUSSEIN received the Ph.D. degree in computer science, in 2012. He currently works as an Associate Professor with the Faculty of Computers and Information, Minia University, Egypt. He is the Founder and the Chair of the Computing and Artificial Intelligence Research Group (CAIRG), Egypt. He has more than 100 scientific research articles published in prestigious international journals in the topics for instance meta-heuristics optimization, artificial intelligence, image processing, the IoT, and its applications. His research interests include WSNs, the IoT, AI, bioinformatics and biomedical, image processing, data mining, and meta-heuristics optimization techniques. He serves as a reviewer for more than 40 journals (Elsevier, Springer, and IEEE).

ABDELMGEID A. ALI is currently a Professor with the Computer Science Department, Minia University, El Minia, Egypt. He has published over 80 research articles in prestigious international journals, and conference proceedings. He has supervised over 60 Ph.D. and M.Sc. students. His research interests include information retrieval, software engineering, image processing, data security, metaheuristics, the IoT, digital image steganography, and data warehousing. He is a member of the *International Journal of Information Theories and Applications* (ITA).

ESLAM AMER received the Ph.D. degree in computer science from Helwan University, in 2012. He held two postdoctoral positions at top universities in Europe. He currently works as an Associate Professor of computer science with Misr International University. He is also working on proposing new paradigms for malware analysis using NLP and deep learning. His research interests include NLP and machine learning, along with their applications.