Computers, Materials & Continua
DOI:10.32604/cmc.2022.022177
Article

Tech Science Press

Plant Identification Using Fitness-Based Position Update in Whale Optimization Algorithm

Ayman Altameem[1], Sandeep Kumar[2], Ramesh Chandra Poonia[3] and Abdul Khader Jilani Saudagar[4,*]

[1]Department of Computer Science, College of Applied Studies, King Saud University, Riyadh, 11495, Saudi Arabia
[2]Department of Computer Science and Engineering, CHRIST (Deemed to be University), Bangalore, 560074, India
[3]Department of Computer Science, CHRIST (Deemed to be University), Bangalore, 560029, India
[4]Information Systems Department, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, 11432, Saudi Arabia
*Corresponding Author: Abdul Khader Jilani Saudagar. Email: aksaudagar@imamu.edu.sa

Received: 30 July 2021; Accepted: 22 September 2021

Abstract: Since the beginning of time, humans have relied on plants for food, energy, and medicine. Plants are recognized by leaf, flower, or fruit and linked to their suitable cluster. Classification methods are used to extract and select traits that are helpful in identifying a plant. In plant leaf image categorization, each plant is assigned a label according to its classification. The purpose of classifying plant leaf images is to enable farmers to recognize plants, leading to the management of plants in several aspects. This study aims to present a modified whale optimization algorithm and categorizes plant leaf images into classes. This modified algorithm works on different sets of plant leaves. The proposed algorithm examines several benchmark functions with adequate performance. On ten plant leaf images, this classification method was validated. The proposed model calculates precision, recall, F-measurement, and accuracy for ten different plant leaf image datasets and compares these parameters with other existing algorithms. Based on experimental data, it is observed that the accuracy of the proposed method outperforms the accuracy of different algorithms under consideration and improves accuracy by 5%.

Keywords: Bag-of-features; feature optimization; plant leaf classification; swarm intelligence; nature-inspired algorithm

1 Introduction

In the agricultural sector, plants play a vital role for all living things. Plant identification with greater precision is a complex issue that requires interdisciplinary research. The multidisciplinary approach combines computer vision and plant/botanical taxonomy. This approach helps in the automated classification and identification of plants. These techniques use leaf images as an input and perform classification/identification using machine learning and deep learning approaches. The main obstacle to automated identification is the lack of a suitable dataset. With the advancement in computer vision, many sophisticated models have been proposed to automate agricultural activities, including plant identification, soil classification, disease identification, weeds plant identification, and crop row detection. The timely and efficient identification of plants helps in management of crop related activities, including timely irrigation, supply of fertilizer, and removal of weed plants.

The process of plant identification includes feature extraction, optimization, and classification. Generally, feature extraction is performed by speed-up robust features (SURF) [1], scale-invariant feature transform (SIFT) [2] etc. The feature optimization phase is an important step where an individual can deploy an algorithm to select the best feasible set of features. Recent research used spider monkey optimization (SMO) [3], whale optimization algorithm (WOA) [4], bat algorithm [5], differential evolution (DE) [6], salp swarm algorithm (SSA) [7], sine cosine algorithm (SCA) [8], and many other swarms and evolutionary algorithms for solving this combinatorial optimization problem. Most of the recent research in the domain of plant identification used artificial neural network (ANN) and support vector machine (SVM) [9], deep residual network [10], convolutional neural network (CNN) [11], deep learning models (AlexNet, VGGNet, and GoogLeNet) [12] CNN models (Xception, ResNet50, InceptionV3, and MobileNet) and recurrent neural networks (RNN) including simple RNN, gated recurrent unit (GRU), and long short term memory (LSTM) models [13], and convolutional siamese network (CSN) [14] for efficient classification.

The automated identification of plants is based on the analysis of leaf images. Therefore, leaves are critical sources of information about the plant. However, this task is challenged by many hurdles like similarity in plant leaves, background variation, and the colour of leaves. Moreover, natural images require an efficient segmentation approach for further processing. Thus, developing machine learning and deep learning approaches to identify plants with higher accuracy is highly desirable. This study presents a new method for plant identification with improved WOA-based feature selection, leading to efficient classification. The significant research contribution of this paper is as follows:

1. A fitness-based WOA (FWOA) was proposed, and its performance was evaluated over a set of benchmark functions.
2. Feature extraction was performed using SIFT and SVM classifier used for classification.

Following is a breakdown of the rest of the paper. Section 2 discusses some recent developments in plant identification and WOA. Feature extraction, feature selection, and classification are discussed in Section 3. Section 4 discusses the experimental results of FWOA and its application to plant leaf identification. Section 5 concludes the paper.

2 Preliminaries

Computational intelligence-based techniques can solve a complex optimization problem with fewer resources. These techniques are classified into different classes based on their source of inspiration, like swarm-based, evolutionary, and bio-inspired. These algorithms start with a set of randomly generated populations. Subsequently, each update their position shares details with other individuals and selects the best one for the next iteration. This section discusses recent development in plant identification and the basics of WOA [15].

2.1 Plant Identification

Machine learning and deep learning are becoming more popular nowadays for the identification of plants. Some of the recent research contributions for the identification of plants using these techniques are discussed here. Pankaja and Suma deployed WOA to reduce dimensions and classified using Random Forest (RF) [15]. The author extracted texture, shape, and color features from the leaf image dataset. The WOA-based approach selects a set of optimal features. They used Flavia and

Swedish leaf datasets for this experiment. Results reported that the WOA-based strategy outperformed other considered algorithms for feature extraction, feature selection, and plant identification. Sun et al. deployed a deep residual network with 26 layers on the BJFU100 dataset collected from their university campus [16]. The new approach was first validated on the Flavia leaf dataset with a 99.65% recognition rate. The main feature of this work is that Sun et al. acquired this data set on a mobile device [16]. Ghazi et al. employed transfer learning with a deep neural network [12]. Here, the author performed fine-tuning of the pre-trained model and used AlexNet, VGGNet, and GoogLeNet. The new model gives significantly improved results. Zhu et al. deployed a deep CNN with a set of five max-pooling layers, five soft-max layers, three fully connected layers, and sixteen convolutional layers [11]. This study concluded that the use of ReLUs along with these layers improved overall performance. Finally, Rzanny et al. studied various image acquisition and preprocessing techniques to identify plants with varying backgrounds [17]. Kho et al. focused on intact leaves [9] and used ANN and SVM to identify Ficus species plants. Original images were preprocessed by detecting edges and segmentation. While extracting shape and texture features, they reported 83% accuracy. Bodhwani et al. deployed a 50-layer deep residual network for plant identification and achieved 93% accuracy [10]. Liu et al. combined pre-trained CNN (Exception, ResNet50, InceptionV3, and Mobile Net) and RNN (simple RNN, GRU, and LSTM) models [13]. Some of the combinations archived very high accuracy. A summary of some of the recent development in plant identification is illustrated in Tab. 1.

Table 1: Recent development in plant identification

Year	Author(s) [ref.]	Approach used	Remark
2017	Sun et al. [16]	Deep learning	Deep residual network employed with 26 layers
2017	Ghazi et al. [12]	Deep learning	AlexNet, VGGNet and GoogLeNet used with transfer learning
2017	Rzanny et al. [17]	ResNet-50 CNN	Considered images with natural, plain and back-light backgrounds.
2017	Kho et al. [9]	ANN and SVM	Ficus species plant dataset used for the identification where ANN outperformed SVM
2018	Liu et al. [13]	CNN and RNN	RNN with CNN-based backbone implemented.
2018	Zhu et al. [11]	Very deep CNN	Use of ReLUs improved overall performance for hand-crafted features
2019	Bodhwani et al. [10]	Deep residual network	A 50-layer deep residual network employed for plant identification

(Continued)

Table 1: Continued

Year	Author(s) [ref.]	Approach used	Remark
2020	Pankaja and Suma [15]	Machine learning, WOA and RF	Dimensionality reduction performed using WOA
2020	Chen et al. [18]	Deep transfer learning	Performed disease identification in rice and maze leaves using pre-trained models with higher accuracy
2020	Figueroa and Montero [14]	Convolutional Siamese Network	The proposed approach suitable for a small dataset
2021	Reddy et al. [19]	Convolutional Neural Network	CNN-based approach analyzed five plant leaf datasets

2.2 Whale Optimization Algorithm

The WOA is a new nature-inspired algorithm developed by Mirjalili and Lewis in 2016 [4]. This nature-inspired optimization algorithm is used to solve many complex real-world optimization problems. WOA is inspired by the bubble-net hunting approach used by humpback whales during foraging. This method mimics the hunting style by using the fittest search agent to hunt the prey, and the spiral method is used to model the bubble-net attacking mechanism. The hunting method is an exciting mechanism for humpback whales. This approach of hunting is recognized as the bubble-net feeding strategy [20]. The mathematical model of this optimization algorithm majorly consists of three steps. The first step is encircling prey, the second step is a bubble net attacking method (exploitation phase), and the last step is the search for prey (exploration phase) [4]. Each phase is illustrated in subsequent sections.

2.2.1 Encircling Prey

Humpback whales locate the target and encircle it. Initially, the optimal design is unknown; hence, the WOA method assumes the target prey as the present ideal candidate solution, or it can be close to the optimum. Once the optimal search agent is well-defined, some other agent will update the location of the existing best search agent.

$$\vec{D} = |\vec{C} \cdot \vec{X}^{*}(t) - \vec{X}(t)| \tag{1}$$

$$\vec{X}(t+1) = \vec{X}^{*}(t) - \ \vec{A} \cdot \vec{D} \tag{2}$$

where $\vec{A}$ and $\vec{C}$ both are coefficient vectors, $\vec{X}$ and $\vec{X}^{*}$ denotes position vector and the best solution respectively, t denotes the iteration counter and updated in every iteration if found the improved solution. The vectors $\vec{A}$ and $\vec{C}$ are calculated as given below:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{3}$$

$$\vec{C} = 2 \cdot \vec{r} \tag{4}$$

where $\vec{a}$ is linearly decreased from 2 to 0 throughout iterations and $\vec{r}$ is a random vector whose values lie between 0 and 1.

2.2.2 Bubble-Net Attacking Method

The exploitation phase in WOA is simulated by the bubble-net behaviour of humpback whales with two steps.

1. **Shrinking encircling mechanism:** The behaviour of the humpback whale is accomplished by decreasing the value in Eq. (3) from 2 to 0. As the value of $\vec{a}$ decreases, the variation range of $\vec{A}$ decreases. The new location of the individual is defined between the current best and the original location by setting a random variable for $\vec{A}$ in $[-1, 1]$.
2. **Spiral updating position:** In this step, compute the distance between the whale situated at ($\vec{X}$, $\vec{Y}$) and prey found at ($\vec{X}^*$, $\vec{Y}^*$). A spiral equation is formulated for the position of whale and prey to impersonate. The humpback whale's spiral-shaped movement is shown in Eq. (5)

$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \tag{5}$$

where d is the distance of prey from i^{th} whale to take the best solution obtained so far, l is a random number, b is a constant, which defines the shape of the spiral. Thus, humpback whales continuously swim in a spiral-shaped path within a decreasing circle around the prey. To model this synchronized behaviour, assume a 50% probability of selecting either a shrinking circle or a spiral model to update the whale's location. The calculated model is shown in Eq. (6).

$$\vec{X}(t+1) = \begin{cases} \vec{X}(t) - \vec{A}\vec{D} & if\ p < 0.5 \\ \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) & if\ p \geq 0.5 \end{cases} \tag{6}$$

where p is a random number within the range [0, 1].

2.2.3 Search for Prey

In the bubble net technique, humpback whales hunt prey randomly according to each other's location. Vector A is used to search for prey in the exploration phase, calculated in the first phase. In this step, update the location of the search agent by using the randomly selected search agent. This method sheds light on exploration and allows this algorithm to perform global searches. The calculated model is shown in Eqs. (7) and (8).

$$\vec{D} = |\vec{C} \cdot \overrightarrow{X_{rand}} - \vec{X}| \tag{7}$$

$$\vec{X}(t+1) = \overrightarrow{X_{rand}} - \vec{A} \cdot \vec{D} \tag{8}$$

where X_{rand} is an|| arbitrarily chosen whale from the current population. The detailed pseudo code for WOA is given in Algorithm 1.

Algorithm 1: Whale Optimization Algorithm

```
    Initialize population of $N$ whales
    Evaluate the fitness of all individuals and choose the best search agent
    while Termination criteria is not meet do
        for every individual do
        Update value of A, $C, a, 1, p$
            if $p < 0.5$ then
                if $|A| < 1$ then
                     Update position using Eq. (2)
                else
                     Update position using Eq. (8)
                end if
            else
                 Update position using Eq. (5)
            end if
        end for
        Evaluate the fitness of new solution and update the best search agent
  End while
```

Initially, the WOA starts with some arbitrary solutions. Then, individuals update the positions using the best answer ever found on each iteration or an arbitrarily picked individual. Using vector, $\vec{A}$ update the location of an individual with the condition if $|\overrightarrow{A}|>1$ selects a random search agent, and if $|\overrightarrow{A}|<1$ selects the best solution. WOA includes the exploration and exploitation phase. Hence it is considered a global optimizer. Moreover, the proposed method describes a search space in the locality. WOA mainly includes two vector parameters, namely $|\overrightarrow{A}|$ and $|\overrightarrow{C}|$. However, modification and additional evolutionary procedures are included in WOA formulation to mimic the behaviour of humpback whales.

Algorithm 2: Fitness-based solution update strategy

```
    Input: Current solution, $p_{r}. = 0, A, C, a, 1, p$
    Compute $p_r$ using $p_r = 0.9$ x(Fitness$_i$/Fitness$_{max}$)+0.1
    if $pr < 0.2$ then
        if $|A| < 1$ then
            Update position using Eq. (2)
        else
            Update position using Eq. (8)
        end if
    else
        Update position using Eq. (5)
    end if
```

3 Fitness-Based Whale Optimization Algorithm

Exploitation and exploration are two significant phases in all the meta-heuristic algorithms for accomplishing the precise solution. The performance of an algorithm is strongly reliant on balancing these two opposing processes. In the WOA algorithm, bubble-net attacking is responsible for exploitation and search for prey phase perform exploration. They are essential phases in the WOA algorithm and affect the convergence behavior of WOA. The exploration phase searches whales' property for renovating position; this selection uses the random function for updating to recognize the best whale [4]. To improve the performance of WOA, a new version of WOA is proposed here and named fitness-based status update WOA (FWOA). The new variant update uses highly fitted solutions and explores the search space for a solution with low fitness. The introduced concept works on the principle that solutions in the proximity of higher fitness solutions are also highly fitted and try to exploit the best solution. In the case of low fitness, it updates its position according to the search for prey phase. Detailed pseudo-code for the new strategy is given in Algorithm 2.

Additionally, a fitness-based method is used to compute the value of $\overrightarrow{f_1}$ and $\overrightarrow{f_2}$ instead of a random function, which improves the performance of the current method. In addition, the accuracy of the proposed model is increased by using the fitness function. Calculate the values of $\overrightarrow{f_1}$ and $\overrightarrow{f_2}$ to surround the hunting stage according to Eqs. (9) and (10).

$$\overrightarrow{f_1} = Fitness_{Max} + Fitness_{Min}/2 \tag{9}$$

$$\overrightarrow{f_2} = Fitness_{Max} - Fitness_{Min}/2 \tag{10}$$

The vectors $\vec{A}$ and $\vec{C}$ are considered new variables $\overrightarrow{f_1}$ and $\overrightarrow{f_2}$ in random variables r_1 and r_2. Hence vector $\vec{A}$ and $\vec{C}$ is calculated as follows:

$$\vec{A} = 2\vec{a} \cdot \overrightarrow{f_1} - \vec{a} \tag{11}$$

$$\vec{C} = 2 \cdot \overrightarrow{f_2} \tag{12}$$

$\vec{A}$ is decreased from 2 to 0 throughout iterations; $\overrightarrow{f_1}$ and $\overrightarrow{f_2}$ are calculated using fitness-based position update.

The new approach takes advantage of a highly fitted solution. It assumes that the proximity of highly-fitted solutions may be a feasible solution for the considered problem. As a result, the swarm always moves in the direction of the solution with good fitness with self-organizing characteristics, and it improves the convergence speed and avoids skipping real solutions.

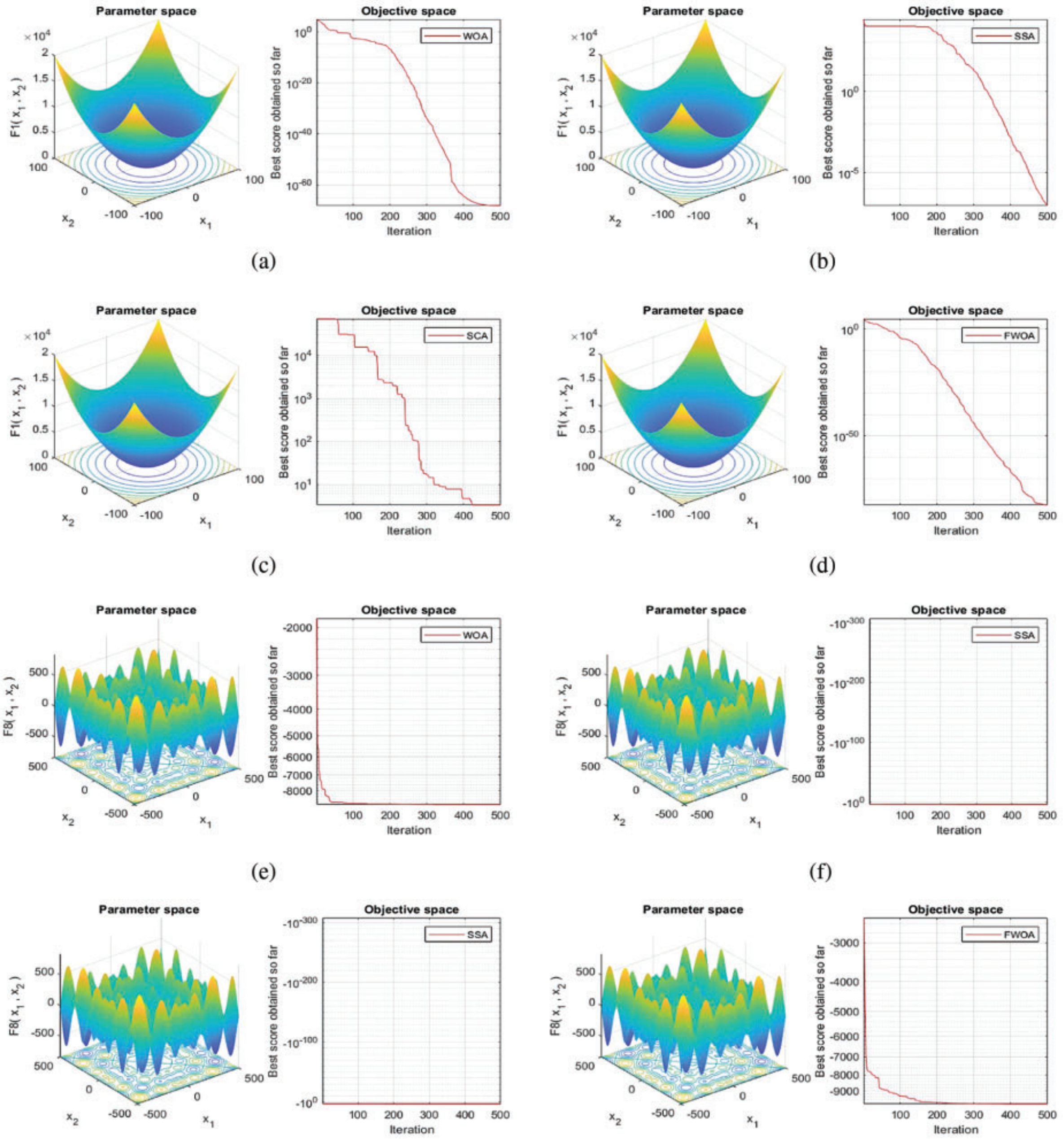
The performance of the newly proposed FWOA is evaluated over a set of thirteen benchmark problems [4]. The selected problems are uni-modal and multi-modal optimization problems with known solutions and search. Performance of FWOA and other competitive algorithms compared in terms of the average function value (Avg), standard deviation (SD), and optimal function value. All the algorithms are implemented in MATLAB R2020b on an Intel Core i7 machine with 16 GB RAM and 8 GB GeForce GTX1650Ti Graphics processor to measure these parameters. Tabs. 2 and 3 illustrate results for FWOA, WOA [4], SSA [16], SCA [8]. Tab. 2 illustrates the efficiency and robustness of FWOA in comparison to other algorithms. Graphical representations of results for functions F1, F8, and F12 are depicted in Fig. 1. These results proved that FWOA outperformed considered algorithms in terms of best function value, as shown in Tab. 3.

Table 2: Results for benchmark problems

Function	FWOA		WOA		SSA		SCA	
	Avg	SD	Avg	SD	Avg	SD	Avg	SD
F_{1}	4.22E+02	3.79E+03	7.66E+02	5.66E+03	2.90E+03	3.48E+03	1.04E+04	1.52E+04
F_{2}	3.40E+09	7.35E+10	3.79E+08	8.04E+09	2.35E+01	1.12E+02	5.37E+05	1.16E+07
F_{3}	1.01E+05	5.45E+04	1.01E+05	4.87E+04	1.42E+04	1.40E+04	4.24E+04	3.27E+04
F_{4}	4.72E+01	4.29E+01	1.86E+01	2.42E+01	2.65E+01	7.85E+00	7.04E+01	2.34E+01
F_{5}	1.88E+06	2.38E+07	1.40E+06	1.26E+07	7.07E+05	1.37E+06	7.03E+07	1.18E+08
F_{6}	4.06E+02	3.56E+03	7.69E+02	5.78E+03	3.03E+03	3.55E+03	1.04E+04	1.36E+04
F_{7}	1.07E+00	9.05E+00	1.15E+00	9.94E+00	1.04E+00	1.16E+00	1.56E+01	2.28E+01
F_{8}	−8.84E+03	5.43E+02	−8.94E+03	5.12E+02	−5.67E+03	1.79E+03	−3.66E+03	3.58E+02
F_{9}	2.98E+01	7.97E+01	1.24E+01	5.55E+01	1.08E+02	8.95E+01	1.91E+02	1.31E+02
F_{10}	6.37E−01	2.67E+00	7.14E−01	2.67E+00	8.22E+00	5.79E+00	2.02E+01	9.04E−01
F_{11}	4.54E+00	4.38E+01	4.70E+00	3.91E+01	2.54E+01	2.98E+01	1.04E+02	1.86E+02
F_{12}	5.02E+06	5.71E+07	4.14E+06	4.68E+07	6.27E+04	7.28E+05	2.03E+08	2.97E+08
F_{13}	7.92E+06	7.79E+07	6.65E+06	5.91E+07	3.16E+06	6.34E+06	3.97E+08	4.64E+08

Table 3: Comparison of optimal value for benchmark problems

Function	FWOA	WOA	SSA	SCA
F_{1}	4.12E−84	7.91E−67	1.40E−07	4.82E−01
F_{2}	1.03E−57	2.89E−40	3.78E−01	1.85E−02
F_{3}	2.38E+04	5.46E+04	1.81E+03	1.06E+04
F_{4}	4.55E−01	5.35E+00	1.82E+01	3.86E+01
F_{5}	2.81E+01	2.72E+01	1.01E+03	1.29E+03
F_{6}	1.85E−01	3.47E−01	4.76E−08	5.37E+01
F_{7}	5.22E−05	4.89E−03	3.38E−01	6.04E−02
F_{8}	−9.01E+03	−8.99E+03	−7.44E+03	−3.93E+03
F_{9}	0.00E+00	0.00E+00	1.39E+01	5.29E+01
F_{10}	8.88E−16	4.44E−15	2.41E+00	2.02E+01
F_{11}	0.00E+00	0.00E+00	2.07E−02	9.19E−01
F_{12}	1.03E−02	2.61E−02	1.01E+01	2.10E+00
F_{13}	1.15E−01	4.70E−01	1.25E+00	5.30E+00


Parameter space
Objective space
F1(x1 , x2)
×10^4
x1
x2
Best score obtained so far
Iteration
WOA
SSA
SCA
FWOA
F8(x1 , x2)
10^0
10^-20
10^-40
10^-60
10^-5
10^4
10^3
10^2
10^1
10^-50
-2000
-3000
-4000
-5000
-6000
-7000
-8000
-9000
-10^-300
-10^-200
-10^-100
-10^0
100
200
300
400
500

(a)

(b)

(c)

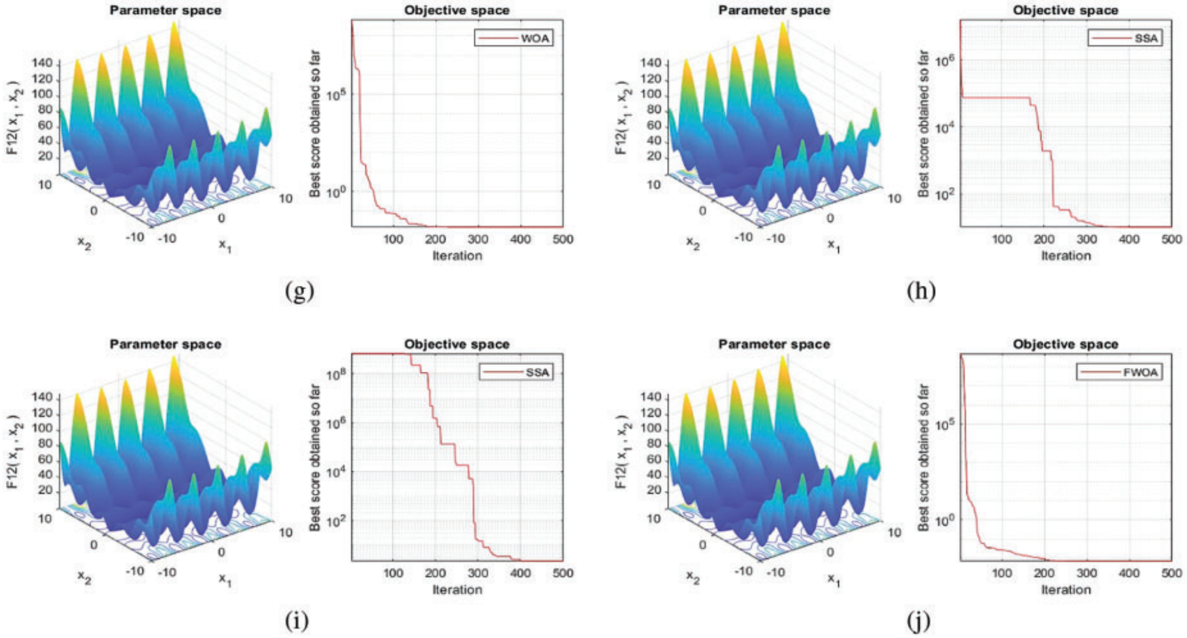
(d)

(e)

(f)

(g)

(h)



(g)

(h)

(i)

(j)

Figure 1: Performance assessment for selected functions. (a) F_1 WOA (b) F_1 SSA (c) F_1 SCA (d) F_1 FWOA (e) F_8 WOA (f) F_8 SSA (g) F_8 SCA (h) F_8 FWOA (i) F_{12} WOA (j) F_{12} SSA (k) F_{12} SCA (l) F_{12} FWOA

4 FWOA-Based Plant Identification System

The proposed model introduced a fitness-based WOA to classify plants based on the leaf image dataset. The suggested model has three significant steps: feature extraction using the SIFT algorithm, feature selection, histogram generation using the modified WOA, and classification of plants based on their leaf image using the SVM classifier, as shown in Fig. 2. A detailed description is given in later sections. The detailed process of the plant leaf classification model is shown in Fig. 3. To validate the proposed model, images of apple, banana, borages, maize, grapes, mint, orange, pepper, potato, and tomato leaf were used in this research. This dataset is used as a sample dataset for validating this model. Some sample images from each category are depicted in Fig. 4.

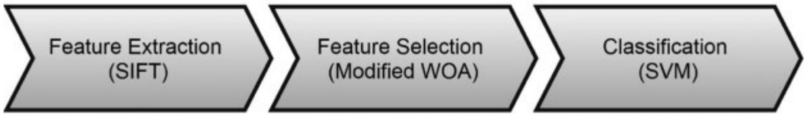


Figure 2: Process of the proposed model

4.1 Feature Extraction

In image processing and computer vision, a feature is an information in a picture [21]. Objects, edges, and points, for example, have extraordinary quality and distinct structure. Feature extraction is a process of classifying essential features of an image, classifying common themes from a broad collection of images, and pattern recognition [22]. The proposed model's first step is to extract all image features and group them into corresponding groups. This extraction is one of the leading steps for image analysis relating to their features. Similar and different image features have to be extracted and stored in respective clusters for practical analysis. The SIFT algorithm is used to extract the features in the proposed method. SIFT is a feature detection method that detects and defines local features of plant leaf images. These local features are essential points in the image that aid in identifying the object of the image [23]. This method can rotate and select an image of a different scale and handle the noise points. Therefore, it is a practical algorithm for feature extraction.

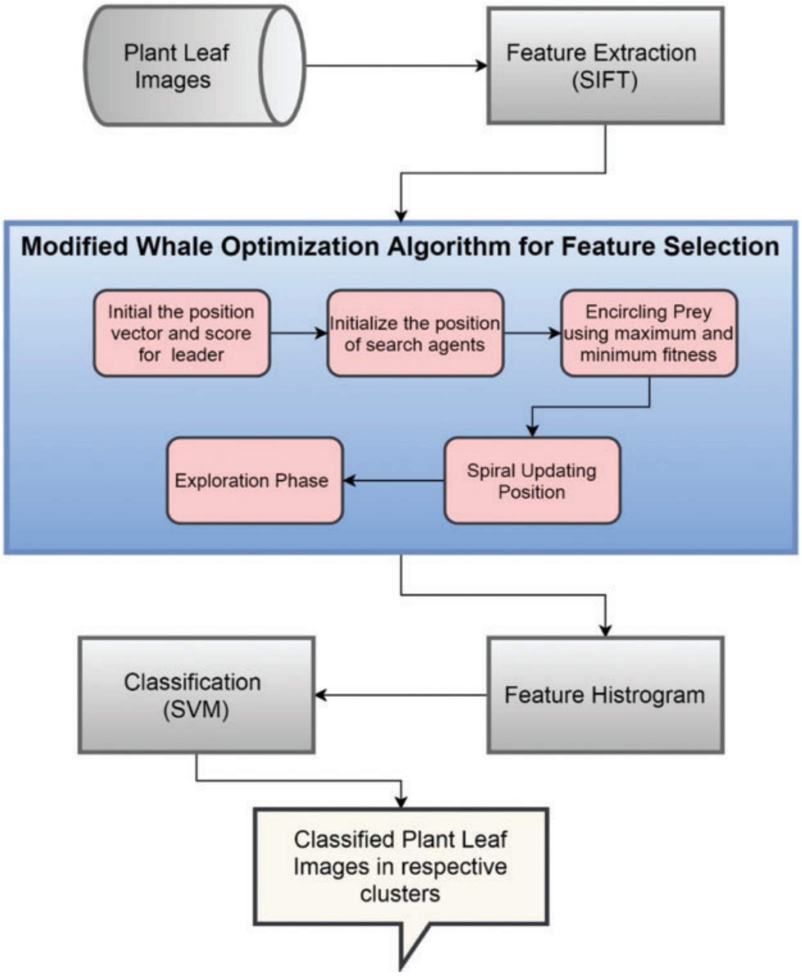

Plant Leaf Images
Feature Extraction (SIFT)
Modified Whale Optimization Algorithm for Feature Selection
Initial the position vector and score for leader
Initialize the position of search agents
Encircling Prey using maximum and minimum fitness
Spiral Updating Position
Exploration Phase
Feature Histrogram
Classification (SVM)
Classified Plant Leaf Images in respective clusters

Figure 3: Proposed plant leaf classification model

4.2 *Feature Selection*

Feature selection is a technique that significantly affects the performance of the proposed classification model. A combinatorial optimization problem is selecting an optimal collection of features from a vast set of extracted features. Thus, it is highly desired to solve this problem with a non-conventional optimization algorithm. This step chooses the most relevant elements that will aid in estimating the class of each leaf image. Next, extracted features from the previous step are used to select the optimal features and create clusters using selected features, increasing accuracy and decreasing overfitting. This paper used the modified WOA for clustering to select optimal features.

Finally, a histogram is plotted using selected features by the proposed model. This histogram shows the fundamental frequency distribution of the selected features. In addition, the histogram allows the review of the selected features in terms of outliers and skewness. The graphical representation is depicted in Fig. 5.

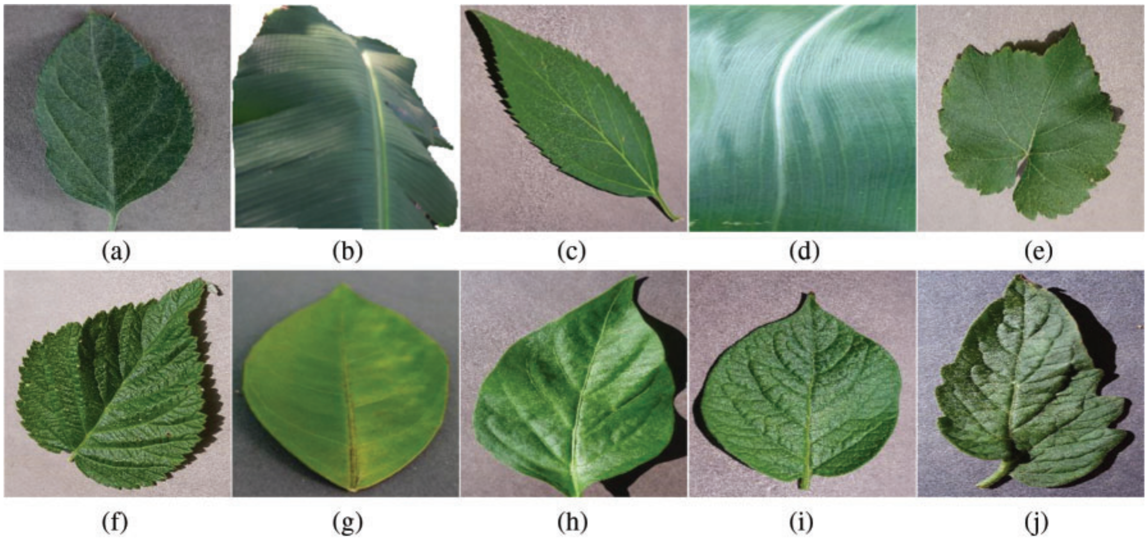


Figure 4: Sample plant leaf image set for considered plants. (a) Apple (b) Banana (c) Borages (d) Corn (e) Grapes (f) Mint (g) Orange (h) Pepper (i) Potato (j) Tomato

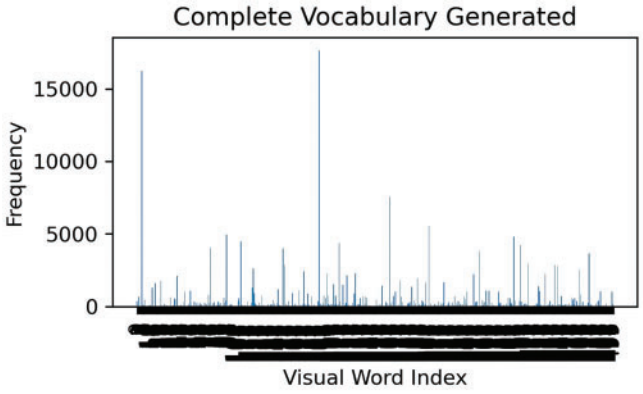


Figure 5: Histogram of the proposed model

4.3 *Plant Leaf Classification*

Classification of plants using a leaf image dataset is the final step of this proposed method. In the previous step, the histogram is generated based on the selected features and passed to SVM classifier along with their labels [24]. SVM is a high-performance binary classifier, which creates a hyperplane in ample feature space for separate leaf images into their respective classes [25]. In this step, the SVM classifier predicts the class labels of each plant leaf image based on training. Hence labelling, training and testing plant leaf image dataset confirm the accuracy of this model. Experimental results are discussed in the next section.

5 Experimental Results for Plant Identification

Three steps are used to analyze the proposed plant classification using a leaf image dataset based on FWOA. The first step represents plant leaf dataset description, the second step shows the performance of benchmark functions, and the third step analyses the result of FWOA based plant leaf classification.

5.1 Dataset Description

This dataset consists of more than 10000 images; 200 images from each category are used for training and testing this model. This dataset is categorized into ten different classes named apple, banana, borages, corn, grapes, mint, orange, pepper, potato, and tomato. This dataset is taken from Plant Village [26] and Kaggle [27]. This dataset is used to measure the performance of the proposed method in terms of the accuracy of classification of each class using a leaf image dataset. These images are divided into a 70%–30% train cross-test split for each class.

5.2 Experimental Results for Plant Leaf Image Classification

The proposed model has been predicted outputs using Python programming. In this section, the proposed approach is described using experimental results based on the input image dataset. Tab. 4 shows some of the parameters and best fitness values. The value of these parameters is decided with exhaustive experiments. The proposed modified WOA has been compared with SCA, BAT, SSA, DE, and WOA. An equal number of image sets have been from each class for these algorithms. Create a confusion matrix concerning each class for performance analysis. The confusion matrix for each class is depicted in Fig. 6. These matrixes show the comparison of actual data and predicted data. The performance of all the considered algorithms for classification is illustrated in Fig. 6. It is important as the considered data set has ten classes. In the case of three or more categories, it is better to visualize results with confusion matric as accuracy can be misleading. The results are measured by calculating the F1 score, precision, recall, and accuracy.

$$Sensitivity\ or\ Recall = Sum\ of\ all\ TP\ for\ each\ class/(Sum\ of\ all\ TP + Sum\ of\ all\ FN) \tag{13}$$

$$Precision = Sum\ of\ all\ TP\ for\ each\ class/(Sum\ of\ all\ TP + Sum\ of\ all\ FP) \tag{14}$$

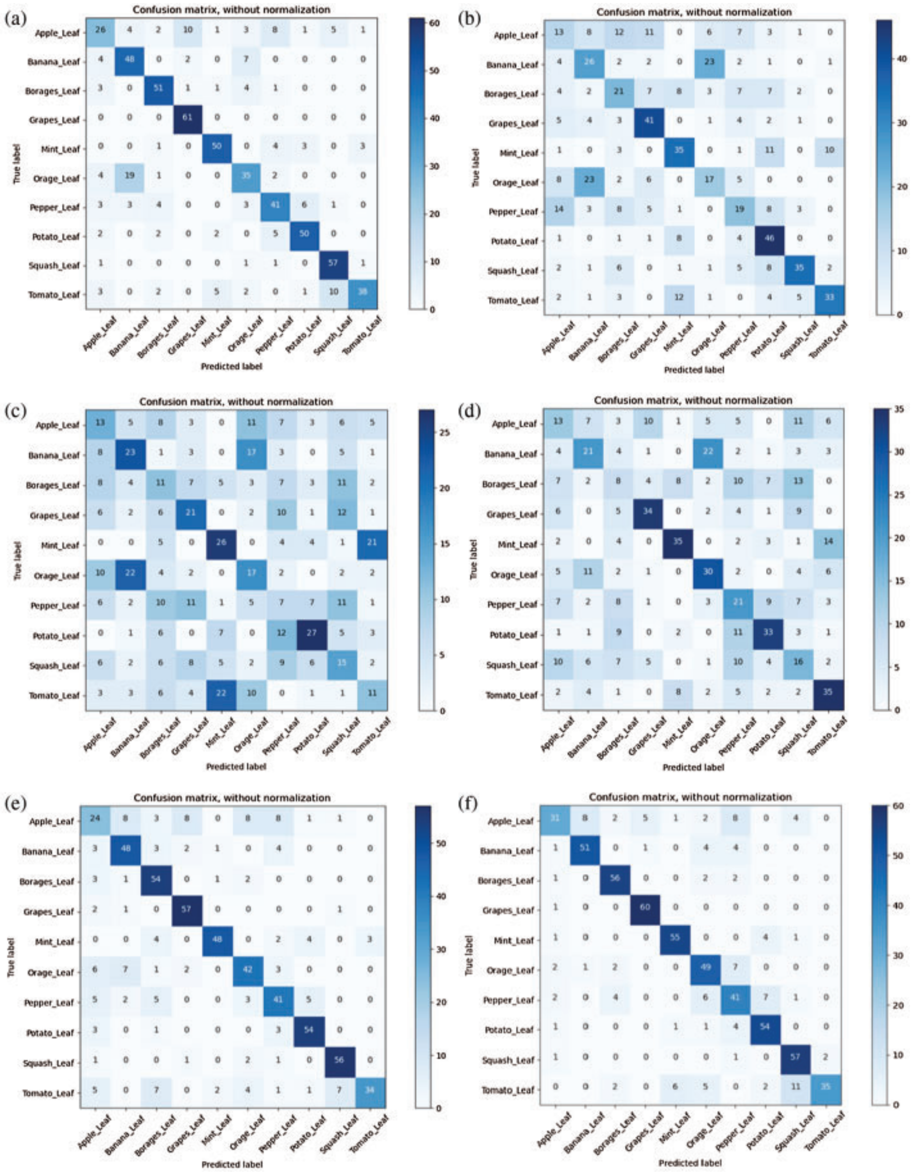
$$F_1\ Score = 2 \times (Precision \times Recall)/(Precision + Recall) \tag{15}$$

$$Accuracy = \frac{Sum\ of\ all\ TP\ for\ each\ class + Sum\ of\ all\ TN\ for\ each\ class}{Sum\ of\ all\ TP + Sum\ of\ all\ TN + Sum\ of\ all\ FP + Sum\ of\ all\ FN} \tag{16}$$

where TP stands for true positive, TN for true negative, FP for false positive, and FN for false negative. The measured performance and comparisons are shown in Tab. 5. This table summarizes accuracy and other matrices for considered algorithms. The proposed method shows better performance when compared with another existing algorithm. For example, compare the accuracy of the modified WOA algorithm with some other algorithms, which is depicted in Fig. 7. Hence, it can be stated that the proposed modified WOA classification method is better than the existing algorithms.

Table 4: Parameter setting for WOA

Name of parameter	Values
Number of iterations	100
Number of run	30
Population size	100



	Apple_Leaf	Banana_Leaf	Borages_Leaf	Grapes_Leaf	Mint_Leaf	Orage_Leaf	Pepper_Leaf	Potato_Leaf	Squash_Leaf	Tomato_Leaf
Apple_Leaf	26	4	2	10	1	3	8	1	5	1
Banana_Leaf	4	48	0	2	0	7	0	0	0	0
Borages_Leaf	3	0	51	1	1	4	1	0	0	0
Grapes_Leaf	0	0	0	61	0	0	0	0	0	0
Mint_Leaf	0	0	1	0	50	0	4	3	0	3
Orage_Leaf	4	19	1	0	0	35	2	0	0	0
Pepper_Leaf	3	3	4	0	0	3	41	6	1	0
Potato_Leaf	2	0	2	0	2	0	5	50	0	0
Squash_Leaf	1	0	0	0	0	1	1	0	57	1
Tomato_Leaf	3	0	2	0	5	2	0	1	10	38

	Apple_Leaf	Banana_Leaf	Borages_Leaf	Grapes_Leaf	Mint_Leaf	Orage_Leaf	Pepper_Leaf	Potato_Leaf	Squash_Leaf	Tomato_Leaf
Apple_Leaf	13	8	12	11	0	6	7	3	1	0
Banana_Leaf	4	26	2	2	0	23	2	1	0	1
Borages_Leaf	4	2	21	7	8	3	7	7	2	0
Grapes_Leaf	5	4	3	41	0	1	4	2	1	0
Mint_Leaf	1	0	3	0	35	0	1	11	0	10
Orage_Leaf	8	23	2	6	0	17	5	0	0	0
Pepper_Leaf	14	3	8	5	1	0	19	8	3	0
Potato_Leaf	1	0	1	1	8	0	4	46	0	0
Squash_Leaf	2	1	6	0	1	1	5	8	35	2
Tomato_Leaf	2	1	3	0	12	1	0	4	5	33

	Apple_Leaf	Banana_Leaf	Borages_Leaf	Grapes_Leaf	Mint_Leaf	Orage_Leaf	Pepper_Leaf	Potato_Leaf	Squash_Leaf	Tomato_Leaf
Apple_Leaf	13	5	8	3	0	11	7	3	6	5
Banana_Leaf	8	23	1	3	0	17	3	0	5	1
Borages_Leaf	8	4	11	7	5	3	7	3	11	2
Grapes_Leaf	6	2	6	21	0	2	10	1	12	1
Mint_Leaf	0	0	5	0	26	0	4	4	1	21
Orage_Leaf	10	22	4	2	0	17	2	0	2	2
Pepper_Leaf	6	2	10	11	1	5	7	7	11	1
Potato_Leaf	0	1	6	0	7	0	12	27	5	3
Squash_Leaf	6	2	6	8	5	2	9	6	15	2
Tomato_Leaf	3	3	6	4	22	10	0	1	1	11

	Apple_Leaf	Banana_Leaf	Borages_Leaf	Grapes_Leaf	Mint_Leaf	Orage_Leaf	Pepper_Leaf	Potato_Leaf	Squash_Leaf	Tomato_Leaf
Apple_Leaf	13	7	3	10	1	5	5	0	11	6
Banana_Leaf	4	21	4	1	0	22	2	1	3	3
Borages_Leaf	7	2	8	4	8	2	10	7	13	0
Grapes_Leaf	6	0	5	34	0	2	4	1	9	0
Mint_Leaf	2	0	4	0	35	0	2	3	1	14
Orage_Leaf	5	11	2	1	0	30	2	0	4	6
Pepper_Leaf	7	2	8	1	0	3	21	9	7	3
Potato_Leaf	1	1	9	0	2	0	11	33	3	1
Squash_Leaf	10	6	7	5	0	1	10	4	16	2
Tomato_Leaf	2	4	1	0	8	2	5	2	2	35

	Apple_Leaf	Banana_Leaf	Borages_Leaf	Grapes_Leaf	Mint_Leaf	Orage_Leaf	Pepper_Leaf	Potato_Leaf	Squash_Leaf	Tomato_Leaf
Apple_Leaf	24	8	3	8	0	8	8	1	1	0
Banana_Leaf	3	48	3	2	1	0	4	0	0	0
Borages_Leaf	3	1	54	0	1	2	0	0	0	0
Grapes_Leaf	2	1	0	57	0	0	0	0	1	0
Mint_Leaf	0	0	4	0	48	0	2	4	0	3
Orage_Leaf	6	7	1	2	0	42	3	0	0	0
Pepper_Leaf	5	2	5	0	0	3	41	5	0	0
Potato_Leaf	3	0	1	0	0	0	3	54	0	0
Squash_Leaf	1	0	0	1	0	2	1	0	56	0
Tomato_Leaf	5	0	7	0	2	4	1	1	7	34

	Apple_Leaf	Banana_Leaf	Borages_Leaf	Grapes_Leaf	Mint_Leaf	Orage_Leaf	Pepper_Leaf	Potato_Leaf	Squash_Leaf	Tomato_Leaf
Apple_Leaf	31	8	2	5	1	2	8	0	4	0
Banana_Leaf	1	51	0	1	0	4	4	0	0	0
Borages_Leaf	1	0	56	0	0	2	2	0	0	0
Grapes_Leaf	1	0	0	60	0	0	0	0	0	0
Mint_Leaf	1	0	0	0	55	0	0	4	1	0
Orage_Leaf	2	1	2	0	0	49	7	0	0	0
Pepper_Leaf	2	0	4	0	0	6	41	7	1	0
Potato_Leaf	1	0	0	0	1	1	4	54	0	0
Squash_Leaf	1	0	0	0	0	0	1	0	57	2
Tomato_Leaf	0	0	2	0	6	5	0	2	11	35

Figure 6: Confusion matrix for considered algorithms. (a) SCA (b) BAT (c) SSA (d) DE (e) WOA (f) MWOA

Table 5: Performance comparison of modified WOA with BAT, SSA, SCA, DE and WOA

Algorithm	Accuracy	Plant Leaf Class	n Truth	Precision	Recall	F – measure
BAT	46.89%	Apple	54	0.21	0.24	0.23
		Banana	68	0.43	0.38	0.4
		Borages	61	0.34	0.34	0.34
		Corn	73	0.67	0.56	0.61
		Grapes	65	0.57	0.54	0.56
		Mint	52	0.28	0.33	0.3
		Orange	54	0.31	0.35	0.33
		Pepper	90	0.75	0.51	0.61
		Potato	47	0.57	0.74	0.65
		Tomato	46	0.54	0.72	0.62
DE	40.33%	Apple	57	0.21	0.23	0.22
		Banana	54	0.34	0.39	0.37
		Borages	51	0.13	0.16	0.14
		Corn	56	0.56	0.61	0.58
		Grapes	54	0.57	0.65	0.61
		Mint	67	0.49	0.45	0.47
		Orange	72	0.34	0.29	0.32
		Pepper	60	0.54	0.55	0.55
		Potato	69	0.26	0.23	0.25
		Tomato	70	0.57	0.5	0.53
SCA	74.92%	Apple	46	0.43	0.57	0.49
		Banana	74	0.79	0.65	0.71
		Borages	63	0.84	0.81	0.82
		Corn	74	1	0.82	0.9
		Grapes	59	0.82	0.85	0.83
		Mint	55	0.57	0.64	0.6
		Orange	62	0.67	0.66	0.67
		Pepper	61	0.82	0.82	0.82
		Potato	73	0.93	0.78	0.85
		Tomato	43	0.62	0.88	0.73
SSA	28.03%	Apple	60	0.21	0.22	0.21
		Banana	64	0.38	0.36	0.37
		Borages	63	0.18	0.17	0.18
		Corn	59	0.34	0.36	0.35
		Grapes	66	0.43	0.39	0.41
		Mint	67	0.28	0.25	0.27
		Orange	61	0.11	0.11	0.11
		Pepper	52	0.44	0.52	0.48
		Potato	69	0.25	0.22	0.23

(Continued)

Table 5: Continued

Algorithm	Accuracy	Plant Leaf Class	n Truth	Precision	Recall	F – measure
		Tomato	49	0.18	0.22	0.2
WOA	75.08%	Apple	52	0.39	0.46	0.42
		Banana	67	0.79	0.72	0.75
		Borages	78	0.89	0.69	0.78
		Corn	70	0.93	0.81	0.87
		Grapes	52	0.79	0.92	0.85
		Mint	61	0.69	0.69	0.69
		Orange	63	0.67	0.65	0.66
		Pepper	65	0.89	0.83	0.86
		Potato	65	0.92	0.86	0.89
		Tomato	37	0.56	0.92	0.69
MWOA	80.16%	Apple	41	0.51	0.76	0.61
		Banana	60	0.84	0.85	0.84
		Borages	66	0.92	0.85	0.88
		Corn	66	0.98	0.91	0.94
		Grapes	63	0.9	0.87	0.89
		Mint	69	0.8	0.71	0.75
		Orange	67	0.67	0.61	0.64
		Pepper	67	0.89	0.81	0.84
		Potato	74	0.93	0.77	0.84
		Tomato	37	0.57	0.95	0.71

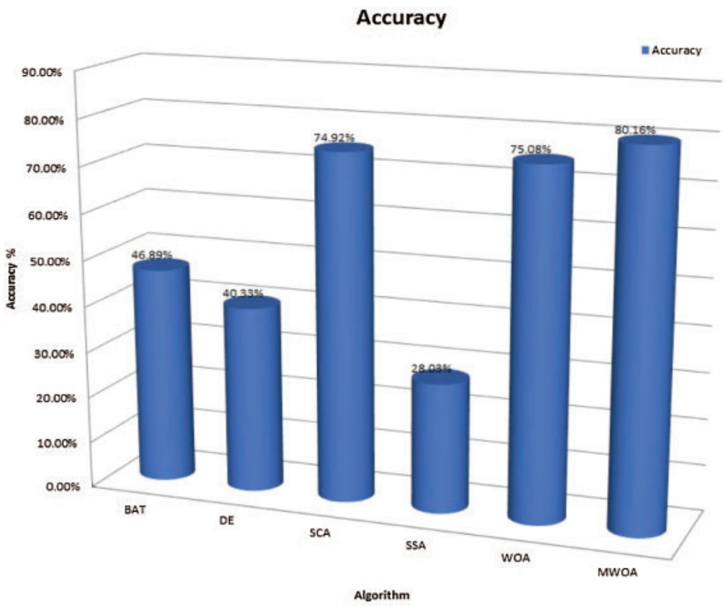


Figure 7: Comparison of overall accuracy

6 Conclusion

Using a plant leaf image dataset, this study presents a new plant classification method. The new version of WOA uses a fitness-based status update method instead of random numbers. This method shows the effectiveness of the results by estimating the maximum accuracy value. In this study, we primarily used three steps: feature extraction using the SIFT method, feature selection using the modified WOA method, and classification using the SVM classifier. The proposed method achieves maximum recall, precision, F1 scores, and accuracy with 80.16%. We analyze the experimental results, and it was found that the WOA with the fitness function increased the efficiency of the proposed algorithm. WOA is employed to handle the problem of feature selection and clustering in this study. The proposed algorithm results are compared with well-known stochastic algorithms such as BAT, DE, WOA, SCA, and SSA.

Furthermore, when compared to other algorithms, the proposed method's results were effective, practical, and simple to implement. In the future, the proposed method can be applied to various plant classifications utilizing different plant leaf image datasets. Besides this, the WOA can be combined with another clustering approach to improve performance.

Acknowledgement: The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this research.

Funding Statement: This work was supported by the Deanship of Scientific Research, King Saud University, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] H. Bay, A. Ess, T. Tuytelaars and L. V. Gool, "Speeded-up robust features (SURF)," *Computer Image and Vision Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[2] H. R. Kher and V. K. Thakar, "Scale invariant feature transform-based image matching and registration," in *Proc. 2014 Fifth Int. Conf. Signal Image Processing*, Bangalore, India, pp. 50–55, 2014.

[3] J. C. Bansal, H. Sharma, S. S. Jadon and M. Clerc, "Spider monkey optimization algorithm for numerical optimization," *Memetic Computing*, vol. 6, no. 1, pp. 31–47, 2014.

[4] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.

[5] X. S. Yang and A. H. Gandomi, "Bat algorithm: A novel approach for global engineering optimization," *Engineering Computations*, vol. 29, no. 5, pp. 464–483, 2012.

[6] R. Storn and K. Price, "Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, pp. 341–359, 1997.

[7] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris *et al.,* "Salp swarm algorithm: A bio-inspired optimizer for engineering design problems," *Advanced Engineering Software*, vol. 114, pp. 163–191, 2017.

[8] S. Mirjalili, "SCA: A sine cosine algorithm for solving optimization problems," *Knowledge-Based Systems*, vol. 96, pp. 120–133, 2016.

[9] S. J. Kho, S. Manickam, S. Malek, M. Mosleh and S. K. Dhillon, "Automated plant identification using artificial neural network and support vector machine," *Frontiers in Life Sciences*, vol. 10, pp. 98–107, 2017.

[10] V. Bodhwani, D. P. Acharjya and U. Bodhwani, "Deep residual networks for plant identification," *Procedia Computer Science*, vol. 152, pp. 186–194, 2019.

[11] H. Zhu, Q. Liu, Y. Qi, X. Huang, F. Jiang *et al.,* "Plant identification based on very deep convolutional neural networks," *Multimedia Tools and Applications*, vol. 77, pp. 29779–29797, 2018.

[12] M. M. Ghazi, B. Yanikoglu and E. Aptoula, "Plant identification using deep neural networks via optimization of transfer learning parameters," *Neurocomputing*, vol. 235, pp. 228–235, 2017.

[13] X. Liu, F. Xu, Y. Sun, H. Zhang and Z. Chen, "Convolutional recurrent neural networks for observation-centered plant identification," *Journal of Electrical and Computer Engineering*, vol. 2018, pp. 1–7, 2018.

[14] G. Figueroa-Mata and E. Mata-Montero, "Using a convolutional siamese network for image-based plant species identification with small datasets," *Biomimetics*, vol. 5, no. 1, pp. 1–17, 2020.

[15] K. Pankaja and V. Suma, "Plant leaf recognition and classification based on the whale optimization algorithm (WOA) and random forest (RF)," *Journal of the Institution of Engineers (India): Series B*, vol. 101, pp. 597–607, 2020.

[16] Y. Sun, Y. Liu, G. Wang and H. Zhang, "Deep learning for plant identification in natural environment," *Computer Intelligence and Neuroscience*, vol. 2017, pp. 1–6, 2017.

[17] M. Rzanny, M. Seeland, J. Wäldchen and P. Mäder, "Acquiring and preprocessing leaf images for automated plant identification: Understanding the tradeoff between effort and information gain," *Plant Methods*, vol. 13, pp. 1–11, 2017.

[18] J. Chen, J. Chen, D. Zhang, Y. Sun and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," *Computers and Electronics in Agriculture*, vol. 173, pp. 1–11, 2020.

[19] S. R. G. Reddy, G. P. S. Varma and R. L. Davuluri, "Optimized convolutional neural network model for plant species identification from leaf images using computer vision," *International Journal of Speech Technology*, pp. 1–28, 2020. https://doi.org/10.1007/S10772-021-09843-X.

[20] W. A. Watkins and W. E. Schevill, "Aerial observation of feeding behavior in four baleen whales: Eubalaena glacialis, balaenoptera Borealis, megaptera novaeangliae, and balaenoptera physalus," *Journal of Mammalia*, vol. 60, no. 1, pp. 155–163, 1979.

[21] J. Sklansky, "Image segmentation and feature extraction," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, no. 4, pp. 237–247, 1978.

[22] G. Kumar and P. K. Bhatia, "A detailed review of feature extraction in image processing systems," in *Proc. 2014 Fourth Int. Conf. on Advanced Computing & Communication Technologies*, Rohtak, India, pp. 5–12, 2014. https://doi.org/10.1109/ACCT.2014.74.

[23] Z. N. Khudhair, F. Mohamed and K. A. Kadhim, "A review on copy-move image forgery detection techniques," *Journal of Physics: Conference Series*, vol. 1892, pp. 1–11, 2021.

[24] M. A. Chandra and S. S. Bedi, "Survey on SVM and their application in image classification," *International Journal of Information Technology*, 2018. https://doi.org/10.1007/s41870-017-0080-1.

[25] X. Li, L. Wang and E. Sang, "Multi-label SVM active learning for image classification," in *Proc. 2004 Int. Conf. on Image Processing*, vol. 4, pp. 2207–2210, 2004.

[26] Plant village dataset, https://mavyn.in/webapp/pdf/googledrive/okjocta/3e3833-plant-village-dataset (accessed on July 26, 2021), 2021.

[27] S. P. Mohanty, D. P. Hughes and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, pp. 1–10, 2016.