IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 29, 2019, accepted January 21, 2020, date of publication January 30, 2020, date of current version February 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2970438

Handwritten Arabic Optical Character Recognition Approach Based on Hybrid Whale Optimization Algorithm With Neighborhood Rough Set

AHMED TALAT SAHLOL[1], MOHAMED ABD ELAZIZ[2], MOHAMMED A. A. AL-QANESS[3], AND SUNGHWAN KIM[4]

[1]Computer Teacher Preparation Department, Faculty of Specific Education, Damietta University, Damietta 34511, Egypt
[2]Department of Mathematics, Faculty of Science, Zagazig University, Zagazig 44519, Egypt
[3]School of Computer Science, Wuhan University, Wuhan 430072, China
[4]School of Electrical Engineering, University of Ulsan, Ulsan 44610, South Korea

Corresponding author: Sunghwan Kim (sungkim@ulsan.ac.kr)

This work was supported by the Research Program through the National Research Foundation of Korea under Grant NRF-2019R1A2C1005920.

ABSTRACT Accomplishing high recognition performance is considered one of the most important tasks for handwritten Arabic character recognition systems. In general, Optical Character Recognition (OCR) systems are constructed from four phases: pre-processing, feature extraction, feature selection, and classification. Recent literature focused on the selection of appropriate features as a key point towards building a successful and sufficient character recognition system. In this paper, we propose a hybrid machine learning approach that utilizes neighborhood rough sets with a binary whale optimization algorithm to select the most appropriate features for the recognition of handwritten Arabic characters. To validate the proposed approach, we used the CENPARMI dataset, which is a well-known dataset for machine learning experiments involving handwritten Arabic characters. The results show clear advantages of the proposed approach in terms of recognition accuracy, memory footprint, and processor time than those without the features of the proposed method. When comparing the results of the proposed method with other recent state-of-the-art optimization algorithms, the proposed approach outperformed all others in all experiments. Moreover, the proposed approach shows the highest recognition rate with the smallest consumption time compared to deep neural networks such as VGGnet, Resnet, Nasnet, Mobilenet, Inception, and Xception. The proposed approach was also compared with recently published works using the same dataset, which further confirmed the outstanding classification accuracy and time consumption of this approach. The misclassified failure cases were studied and analyzed, which showed that they would likely be confusing for even Arabic natives because the correct interpretation of the characters required the context of their appearance.

INDEX TERMS Machine learning approach, feature selection, optimization, Arabic handwritten character recognition, whale optimization, neighborhood rough set, optical character recognition (OCR).

I. INTRODUCTION

In character recognition systems, many solutions have been constructed for different languages, such as English, Japanese, and Chinese; however, relatively little progress has been made for the Arabic language. As such, the recognition of handwritten Arabic characters is still a current and relatively unaddressed research problem. The digitization of Arabic documents can open windows for the processing (indexing, searching,...) of historic and Islamic documents [1]. Earlier efforts to digitize Arabic languages have encountered several issues. First, the alphabet system consists of 28 characters, with several types and numbers of dots; one, two, or three dots. There are several writing styles for each

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

TABLE 1. Characters variations in Arabic handwritten characters.

Row	challenges	Description			
1	Different shapes	ع ـعـ ـعـ ـع		هـ ـهـ ـه ه	ك ک
2	Secondaries exists or not	ظ ط	غ ع	ق ف	
		ش س	ض ص	ز ر	ذ د
3	Number and position of secondaries	خ ج ح		يـ تـ نـ بـ	
4	Secondary types	شـ نـ يـ	يٍ ي ت ٮ	ـش ش ش	أ ك

character based on its location in a word (beginning, middle, end). Therefore, each character has about 80 written shapes or styles. Table 1 shows different shapes for the same Arabic character, as shown in the first row. The second row presents each pair of characters with different styles of dots. The third row of the table presents characters with different positions of dots. The last row shows combinations of the same character with different shapes based on its dot style.

Table 2 presents some recent machine learning studies on Arabic handwritten character recognition.

Most of the proposed methods that deal with the handwritten Arabic characters utilize the dimension reduction of extracted features. In those methods (such as PCA), the original features must be transformed to another domain [2]. Although those methods are not reliable for narrowing down the most efficient features, feature selection methods are used to avoid this limitation. With regard to feature selection methods, several swarm techniques have been introduced to improve the process of determining the most informative features [6], [9]. However, those methods do not take into consideration the impact of the dependency between features which may lead to the selection of irrelevant features. Therefore, to avoid this limitation, a rough set can be used to improve the performance of feature selection by finding relationships between features depending on the degree of dependency [11]. Most recent trends and application of swarm intelligence, such as Bat Algorithm, Grey Wolf Optimization, Whale optimization Algorithm, Particle Swarm Optimization and Genetic Algorithm as a feature selector for Arabic recognition systems was presented in [7].

The Rough Sets (RS) method is mainly used to reduce the dimension of the data by selecting the most relevant features [11]. This approach only deals with the data itself and needs no other external information. However, the current techniques for feature selection based on RS are not sufficient due to the high computational time requirements. Therefore, to tackle this problem, optimization algorithms are combined with RS [15], [16], such as genetic algorithms (GA), PSO [13], [14], the cuckoo search algorithm [55], the social-spider algorithm [54] and other optimization algorithms [56].

In the same context, the Whale Optimization Algorithm (WOA) was proposed in [17] as a new swarm-based

TABLE 2. Previous works on Arabic handwritten character recognition.

Study name	Methodology	Results
Abandah et. al. [2]	After the feature extraction phase from Arabic letters, PCA was applied to choose the most informative features.	The best-performing classifier was LDA, which outperformed all other classifiers, correctly classifying 87% of letters
Sahlol et. al. [3]	A feature selection method based on a swarm intelligence algorithm called the Bat algorithm was applied to select the best features from the entirety of the extracted features. Classification algorithms were then used like RF, naive Bayes, and KNN.	About a 50% feature reduction was successfully performed based on the proposed approach, with an overall classification performance of 84%.
Sahlol et. al. [4]	A feature selection approach based on PSO to reduce the whole feature set by choosing the best of the produced features	The proposed feature selection approach was applied to several classifiers; RF outperformed all the other classifiers, with a performance reaching 91.66%.
Sahlol et. al. [5]	An Arabic character recognition approach using the Moth-Flame algorithm was applied to select a feature set with a high performance. Several classification algorithms (KNN, RF, LDA) were tested to validate the approach.	This approach narrowed down the feature set size by half, while at the same time preserving the performance.
El-Sawy et. al. [6]	A Convolutional Neural Network (CNN) was optimized to improve the recognition accuracy of a database that contained 16,800 handwritten Arabic characters. It was used for both feature extraction and as a classifier trainer.	This approach achieved a misclassification error rate as low as 5.1% on the testing data.
Elleuch et. al. [8]	A CNN was applied for the reduction of the dimension for textual images.	The approach showed promising results compared to other relevant Arabic OCR works.
Ben Ahmed et. al. [9]	A CNN was used for Arabic scene text recognition using a variant filter size, with stride values of 1 and 2.	This approach achieved encouraging results for recognizing Arabic characters when applied to segmented scene images.

optimization technique that emulates the behavior of the humpback whales for hunting preys. This behavior includes the process of searching for prey through different strategies, such as random movement in a spiral that emulates the bubble-net attraction process. This behavior is converted into an algorithm to find the optimal solution for global optimization problems [17]. Moreover, Yan et al. [52] proposed a Multi-Objective whale for the allocation of water resources. Additionally, Elaziz et al. proposed another multi-objective version of the WOA and used it to determine an optimal threshold level for image segmentation [60]. For feature selection purposes, Majdi et al. improved the performance of the WOA by using Simulated Annealing (SA) as a local search method and applied it to different UCI datasets [58]. In [53], they presented a binary version of the Whale Optimization Algorithm (BWOA) using an S-shaped family for different feature selection applications. Majdi and Mirjalili in [57] proposed a modified version of the WOA using two stages: 1) Tournament Selection and 2) Roulette Wheel Selection. The main difference between the traditional WOA and its binary version (BWOA) is as follows: First, the WOA deals with continuous problems, whereas the BWOA is applied to discrete problems, such as feature selection. Second, the BWOA depends on a function that converts a real solution to a binary solution, but the WOA does not require these functions.

Moreover, those modified versions of the WOA have yet to be evaluated on Arabic handwritten character recognition systems. Also, they have not been combined with RS or its extensions. This motivates us to propose an alternative method for Arabic handwritten character recognition systems by combining the Binary Whale Optimization Algorithm with the Neighborhood Rough Set (BWOA-NRS). In this approach, the NRS can be considered an extended version of the RS that has the ability to deal with heterogeneous datasets [12].

The proposed approach consists of four stages. The first is preprocessing of the dataset, which aims to remove noise and clean the data. The second is feature extraction, which aims to extract features from the data, such as gradient features, vertical and horizontal projection features, vertical/horizontal/diagonal projection features, and other features. The important third stage is feature selection, which is considered the main contribution of this paper. In this stage, the feature selection approach starts by generating a random population that represents a set of solutions. Then, each solution is converted into a binary version, where the features that correspond to 1's are considered relevant features, while the other features are ignored. Thereafter, the quality of the selected feature (based on the current solution) is evaluated through computing the objective function that consists of two parts: 1) the degree of dependency, based on NRS. 2) the ratio of selected features. The next step is to determine the best solution and update the other solutions based on the operators of the traditional WOA and the best solution. In this stage, the previous steps are repeated until the stop conditions are met. The best solution is the output of the feature selection stage and it passed to the next stage, which is the classification stage. In the classification stage, each feature vector was exposed to some popular classifiers. Five-fold cross-validation was applied, in which the classifier learned based on the training set for each fold. After that, the testing set is applied to the learned classifier to evaluate the performance based on the average of the five split results.

Therefore, the main contributions of this paper can be summarized as follows:

1) Proposed a new approach for Arabic handwritten character recognition.
2) The proposed model combines the Whale optimization algorithm with NRS for performing the feature selection work.
3) Evaluated the performance of the proposed model using the CENPARMI dataset.

The organization of this paper is as follows: Section II presents the preliminaries of the rough set, neighborhood rough set, and the whale optimization algorithm. The proposed approach is given in Section III. The experimental results are given in Section IV. Finally, the conclusions and research directions for future work are presented in Section V.

II. PRELIMINARIES

A. ROUGH SETS

This section introduces the mathematical formulation of Rough sets (RS). First, assume the system $S =< U, A, V, f >$ where, $U = \{u_1, \ldots, u_n\}$ and A represent the instances and features, respectively [22]. Further, $V = \bigcup_{a\in A} V_a$, and V_a is a vector comprised of the respective values of the a-th feature. In addition, $f : U \times A \rightarrow V$ is the information function used to find the relationship between U and A.

The main step in using the RS is to find the indiscernibility relation, with a mathematical definition given by the following equation:

$$IN(B) = \{(q, z) \in U \times U : \forall b \in B, f(b, q) = f(b, z)\} \tag{1}$$

where $B \in A$ is a subset of the features. If two instances belong to $IN(B)$, then these instances are indiscernible with respect to B. The groups of $IN(B)$ can be formulated as:

$$\begin{aligned} U/IN(B) &= \bigotimes \{b \in B : U/IN(\{b\})\} \\ A \bigotimes B &= \{Q \bigcap Z : \forall Q \in A, \forall Z \in B, \\ &\quad Q \bigcap Z \neq \emptyset\} \end{aligned} \tag{2}$$

If $(w, z) \in U/IN(B)$, then w and z are indiscernible by the features of B. Assuming that $X \subseteq U$, the upper and lower approximations of X are defined as:

$$\overline{B}(X) = \{[w]_B | [w]_B \bigcap X \neq \emptyset\} \tag{3}$$

$$\underline{B}(X) = \{[w]_B | [w]_B \subseteq X\} \tag{4}$$

where $[\cdot]_B$ represent the equivalence classes of the B-indiscernibility relation. Then, the positive region POS can be formulated as:

$$POS_C(d) = \bigcup \underline{B}(X), x \in U/d \tag{5}$$

where C represents the condition features and d is the decision feature [23]. The degree of dependency $\gamma_C(d)$ is applied to determine the reduced set and is computed using the following equation [11]:

$$\gamma_C(d) = |POS_C(d)| / |U| \tag{6}$$

In this study, $\gamma_C(d)$ is used to represent the fitness function to assess the solutions.

B. NEIGHBORHOOD ROUGH SETS

The mathematical definition of Neighborhood Rough Sets (NRS) is illustrated in this section. In general, the neighborhood $\Phi_P(z_i)$ of z_i is based on the threshold ϵ, which is defined as [11]:

$$\Phi_P(z_i) = \left\{ z_j \,\middle|\, z_j \in U,\, \Delta^P\left(z_i, z_j\right) \leq \epsilon,\ z_i \in U \right\} \tag{7}$$

In Equation (7), $P \subseteq C$ and Δ represents the distance. Considering the metric space $<U, \Phi_P(z_i)>$, then the set of neighborhood granules $\{\Phi_P(z_i) \,|z_i \in U\}$ covers the universal space, rather than partitioning it in the manner used in RS.

Considering that $< U, N >$ is the neighborhood approximation space and its lower and upper approximations of $Z \subset U$, are formulated as:

$$\underline{N}Z = \{z_i | \Phi_P(z_i) \subseteq Z, z_i \in U\} \tag{8}$$

$$\overline{N}Z = \{z_i | \Phi_P(z_i) \cap Z \neq \emptyset, z_i \in U\} \tag{9}$$

C. WHALE OPTIMIZATION ALGORITHM

Swarm intelligence has been applied to several machine learning problems in medical fields [18], [19], biology [20], and drug design [21]. In [17], the authors presented the basic steps of the whale optimization algorithm, which emulates the social behavior of whales when hunting prey. This behavior can be performed either through encircling or the bubble-net method. Whales update their locations when encircling, according to the best position W^* as in [17]:

$$Dis = \left|A_W \odot W^*(t) - W(t)\right| \tag{10}$$

$$W(t+1) = \left|W^*(t) - S_W \odot Dis\right| \tag{11}$$

where $W(t)$ represents the location of a whale, Dis represents the distance between W and W^*, and A_W and S_W are coefficient vectors, which are determined as follows:

$$S_W = 2s \odot r - s \tag{12}$$

$$A_W = 2r \tag{13}$$

where r and s are random values (note that the value of s decreases with each iteration from 2 to 0).

The behavior of a bubble-net is emulated through two methods: 1) Shrinking encircling. 2) A Spiral. In the shrinking encircling process, the parameter s decreases from 2 to 0. Simultaneously, S_W is decreased. Meanwhile, in the spiral method, the location of the whale is updated using its helix-shaped movement around W^* as:

$$W(t+1) = Dis' \odot e^{bl} \odot cos(2\pi l) + W^*(t), \tag{14}$$

$$Dis' = |W^*(t) - W(t)| \tag{15}$$

where b represents a constant that defines the logarithmic spiral shape and $l \in [-1, 1]$ represents a random number. $\odot$ represents element-wise multiplication.

These whales have the ability simultaneously swim around the prey in a long a spiral-shaped path, approximated by a shrinking circle.

$$W(t+1) = \begin{cases} W^*(t) - S_W \odot Dis & \text{if } pr \geq 0.5 \\ Dis' \odot ECS + W^*(t) & \text{otherwise} \end{cases} \tag{16}$$

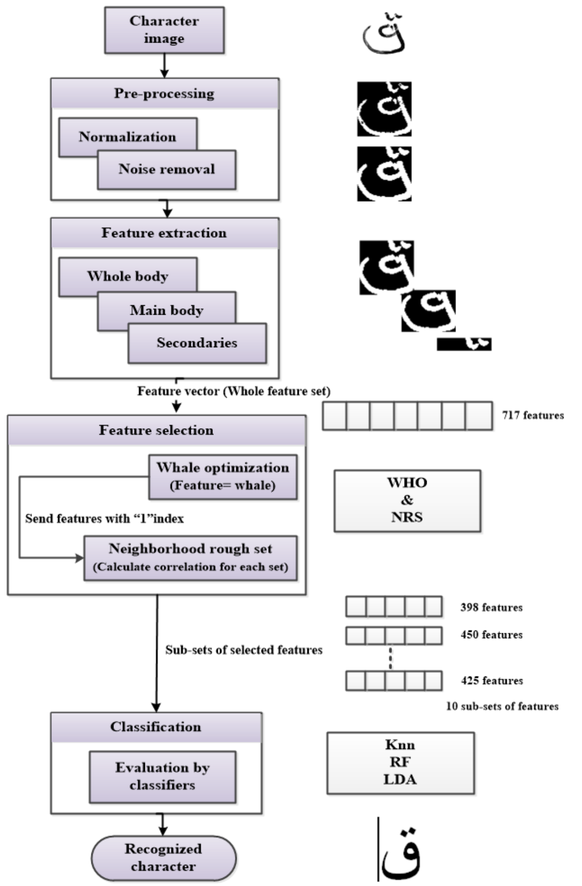


FIGURE 1. Proposed Arabic character recognition approach.

where $ECS = e^{bl} \odot cos(2\pi l)$ and the random number $pr \in [0, 1]$ is the selection probability (which chooses either the spiral model or the shrinking encircling mechanism).

In addition, the whales can randomly search for W^*, therefore, their positions are updated by selecting a random whale W_r:

$$Dis = |A_W \odot W_r(t) - W(t)| \tag{17}$$

$$W(t+1) = |W_r(t) - S_W \odot Dis| \tag{18}$$

III. THE PROPOSED APPROACH

The framework of the proposed approach is given in Fig. 1, which consists of four stages, preprocessing, feature extraction, feature selection, and classification. The details of these four stages are given in the following subsections.

A. PREPROCESSING STAGE

In this stage, the binarization method is employed, which was proposed by Otsu [25] and used in [24]. Further, some noise removal techniques from our previous work [26] were applied in this work, which include dilation and median filtering. The main purpose of this stage is to enhance the character images, which implies the removal of non-informative pixels. These tasks can be vital for better classification, as demonstrated in our previous works [4], [26].

B. FEATURE EXTRACTION STAGE

Several types of features were extracted; each of them represents aspects of Arabic characters. This is vital also to solve some vexing problems, like those involving identical or nearly-identical shapes of different characters. Further, this approach addresses differences in the shape of the same character. Thus, we adopted some types of features that were used before in our previous work [26], which consist of main body features and secondary features. The extracted features include:

1) Gradient features. Gradient features are direction-based features, which imposes the decomposition into two components by determining the number of directions. Each pixel measures the change in intensity of the same point in the original image in a given direction. Gradient images can be created by convolution with a filter, such as the Sobel filter [27] or Prewitt operator [28].

$$g_x = I(x+1, y+1) - I(x, y) \tag{19}$$
$$g_y = I(x+1, y) - I(x, y+1) \tag{20}$$

For each pixel I(x,y), strength and direction of the gradient can be calculated as follows:
Strength:

$$\theta = I(X) \tag{21}$$

Direction:

$$I\,(X) = \tan^{-1}\left(g_y \big/ g_x\right) \tag{22}$$

where θ is 0 for a vertical edge, which is lighter on the right-hand side. The Sobel operator can be used for computing a gradient, so there are two masks for computing the horizontal (g_x) and vertical gradient (g_y) components, which can be calculated by:

$$\begin{aligned} g_x &= I(u-1, v+1) - 2I(u, v+1) \\ &\quad + I(u+1, v+1) - I(u-1, v-1) - 2I(u, v-1) \\ &\quad - I(u+1, v-1) \end{aligned} \tag{23}$$

$$\begin{aligned} g_y &= I(u-1, v-1) - 2I(u-1, v) \\ &\quad + I(u-1, v+1) - I(u+1, v-1) - 2I(u+1, v) \\ &\quad - I(u+1, v+1) \end{aligned} \tag{24}$$

2) Vertical horizontal projection features.
A projection is the transformation of a shape's points in one plane onto another plane; this can be performed by connecting the corresponding points on the two planes. The projection of a vector a onto a vector b is given by:

$$|proj_u a| = \frac{a.u}{|u|^2} u \tag{25}$$

where a.u is the dot product, and the length of this projection is given by:

$$|proj_u a| = \frac{a.u}{|u|} \tag{26}$$

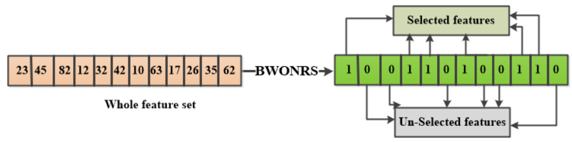


FIGURE 2. WOA working mechanism.

3) Vertical and horizontal diagonal projection features. To obtain a vector d with components from the diagonal of A, the following approach is used:

$$d = \sum_{i=1}^{n} (e_i^T A e_i) e_i \tag{27}$$

where e_i is the i-th unit vector of the standard basis.

4) Other types of features. Based on the recommendations of the literature [26], we decided to adopt more convenient features for OCR systems. These are the height to width ratio of a character, the number of holes of a character, and the number/position of secondaries.

Finally, we perform feature vector normalization, where the "min-max normalization" method is used [33].

C. FEATURE SELECTION STAGE

The main purpose of this stage is to select the most appropriate features. This must be performed by excluding inappropriate features. In this stage, a modified version of the WOA is presented to choose the relevant features from the entire feature set. The search domain is represented as an n-dimensional Boolean space and the NRS is used as a fitness function. Therefore, this approach is called the Binary Whale Optimization Algorithm- Neighborhood Rough Set (BWOA-NRS). The output of the BWOA-NRS method is a binary solution, where the features, corresponding to values of 1, are considered the relevant features, while other features are irrelevant features, so they are kept at the initial zero values and ignored.

The BWOA-NRS approach generates a random location for N whales ($W_i, i = 1, 2, \ldots, N$), which represents a solution for the given problem. Then each solution is converted into binary solutions based on a random value (ε), as in the following equation:

$$W_i = \begin{cases} 1 & if\ W_i \geq \varepsilon \\ 0 & if\ W_i < \varepsilon \end{cases} \tag{28}$$

Fig. 2 represents the mechanism of the feature selection stage in BWOA-NRS.

Thereafter, the current whale location (solution) is evaluated through computing the fitness function, which is defined as:

$$F\,(W_i) = \alpha \gamma_{W_i}\,(d) + (1-\alpha)\left(1 - \frac{|W_i|}{N_F}\right), \tag{29}$$

where $\gamma_{W_i}\,(d)$ is defined in (6), where the lower approximation of the NRS is used instead the RS and $|W_i|$ is the number

TABLE 3. Extracted features number.

Features	Whole feature set	BWOA-NRS
Num. of features	717	261: 573
Percentage of features%	100	35%: 80%

of selected features from the total number of features N_F. α represents a random variable that strikes a balance between the degree of dependency and $|W_i|$. For each value of decreasing a from 2 to 0, the solution is updated based on the value of the probability p.

Algorithm 1 BWO-NRS Algorithm

```
Input: Dim, the dimension of each whale, the feature vector,
Determine the number of whales N, and the maximum number iterations t_max
t = 1 the current iteration
Compute the fitness function F(W_i), ∀i.
Determine the global best fitness function value
Determine the best position
while t < t_max do
    for each integer decrement of a from 2 to 0 do
        for each W_i do
            Update A, S and the selection probability pr.
            if pr < 0.5 then
                Update the current W_i, as in Equation (14)
            else
                if |A| < 0.5 then
                    Update the current W_i, as in Equations (10)-(11)
                else
                    Update the current W_i, as in Equations Eq. (17)-(18)
    t = t + 1
Output: The best position
```

If the value of p is less than 0.5, then equation (14) is used, otherwise, based on the value of $|A|$, equations (17)-(18) are used if $|A| \geq 0.5$ or equations (10)-(11) are used if $|A| < 0.5$. The global best fitness and the best solution are updated at each iteration. This process is performed until the stopping criterion is reached. We present the steps of the BWOA-NRS algorithm in Algorithm 1. Table 3 shows the number and percentage of features before and after applying the proposed feature selection approach.

D. CLASSIFICATION STAGE

In this stage, the dataset is divided internally into training and testing sets, where the training set is used to learn the classifier, which updates its weights and parameters. While the testing set is used to evaluate the performance of the learned classifier. The testing mechanism was measured by performing the classification job on the produced feature sets from the BWOA-NRS algorithm (25 rounds represents 25 feature sets). Several experiments were performed after applying the BWOA-NRS algorithm. Classification work was performed using well-known classifiers, such as K-nearest neighborhood (KNN) [61], multilayer perceptron (MLP) [40], support vector machine (SVM) [41] and linear discriminant analysis (LDA) [42]. These classifiers showed advantages compared to others in relevant image classification works [3], [5], [29], [30] and in other machine learning works [4], [31].

The steps of the proposed model are given in Algorithm 2.

TABLE 4. Experimental parameters SSA, GWO, ABC, SCA.

Algorithm	Parameter	Value
WOA	s	2
	b	1
	l	0.5
ABC	Acceleration Coefficient	1
SCA	a	2
GWO	a	2
SSA	l	1

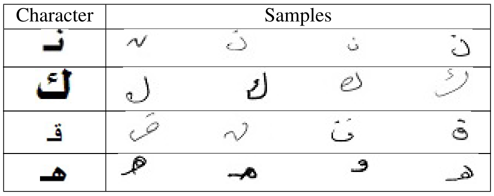
Algorithm 2 Proposed Approach

```
Input character image.
Put the initial value for the set of all extracted features (EX_F) from the images to Ø.
for each image do
    Convert the current image to binary form using the preprocessing stage.
    Extract the Feature EF_I from the binary image.
    EX_F = [EX_F; EF_I].
select the relevant features from EX_F using BWO-NRS method.
Classification with KNN/RF/LDA
Output: Selected features, model performance
```

IV. EXPERIMENTAL RESULTS

The proposed approach was implemented in MATLAB with preprocessing, feature extraction, and feature selection by the BWOA-NRS algorithm, while classification was performed using Python. The proposed algorithm is compared with other algorithms where the parameter value of each algorithm is given in Table 4. In addition, there are common parameters between the algorithms, such as the total number of iterations and population size. To determine the suitable value for these parameters, several experiments have been executed. It has been found that the performance improves with an increasing the number of solutions. However, finding more solutions requires more computational time and we observed that when the number of solutions set to 5, the proposed model achieves good results in a short time. Moreover, the same observation is noticed for simulations when the total number of iterations is set to 100.

TABLE 5. Samples from the CENPARMI dataset.



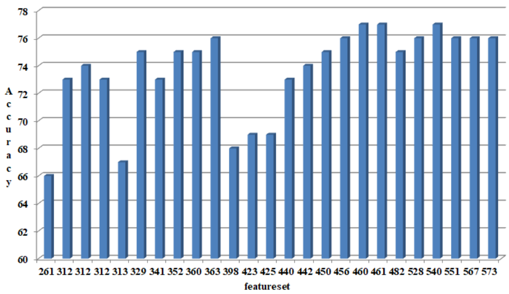


FIGURE 3. Recognition accuracy for each produced feature set by KNN.

A. DATASET

To validate the proposed BWOA-NRS approach, we apply it on the CENPARMI dataset [34], which is published by Concordia University in Canada. This dataset contains the handwritten Arabic alphabet, composed by hundreds of writers with a very wide variety of shapes and styles to generate many variations of Arabic characters. In this work, we select from the dataset the 28 basic characters and ignored the 4 Farsi characters, which are rarely used in modern life. Some characters from the dataset are shown in Table 5. As seen in Table 5, the dataset is very challenging; each character's appearance is totally changed based on dot style and also changes depending on whether or not it is connected to the main shape. The shape of a character also changes based on personal writing style. All of these factors make classification very challenging.

B. PERFORMANCE MEASURE

The performance of the proposed approach was validated using accuracy,recall, precision and F1 as defined in Equations (30:33):

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{30}$$

$$Recall(sensitivity) = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{31}$$

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{32}$$

$$F_1 = 2 \times \frac{\mathrm{Recall} \times \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}} \tag{33}$$

where *TP* represents the true positive samples, *TN* represents the true negative samples, *FP* represents the false positive samples, and *FN* represents the false-negative samples.

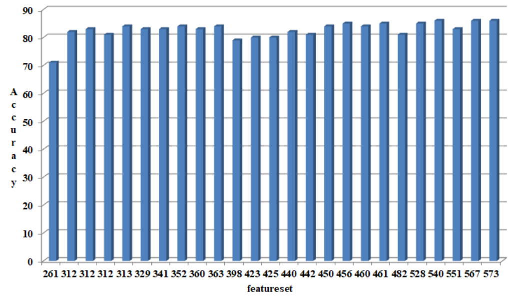


FIGURE 4. Recognition accuracy for each produced feature set by MLP.

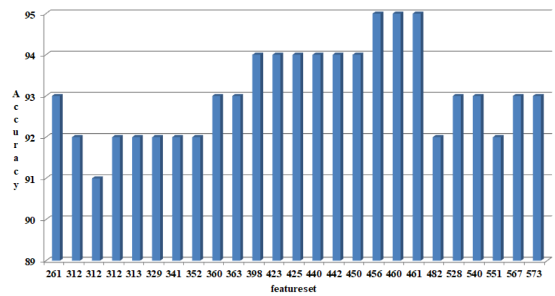


FIGURE 5. Recognition accuracy for each produced feature set by SVM.

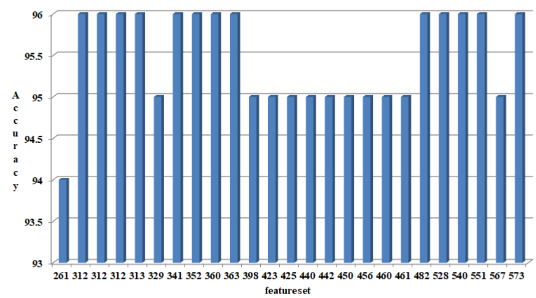


FIGURE 6. Recognition accuracy for each produced feature set by LDA.

TABLE 6. Performance of BWOA-NRS.

Classifier	Accuracy (%)	Recall (%)	precision (%)	F1(%)
LDA	**96**	86	**97**	90
KNN	73	78	96	86
MLP	86	80	93	76
SVM	94	**93**	96	**94**

C. RESULTS AND DISCUSSION

The extracted feature set from the feature extraction stage is a relatively large feature matrix, as it contains 717 features, so we apply our hybrid approach (BWOA-NRS) to select the most relevant features. BWOA-NRS works randomly, so it produces varied feature sets. The size of the produced feature sets is varied from 261 to 573, which means a minimizing capacity around of 2 times.

Figures 3-6 represent the performance of the proposed BWOA-NRS approach for 25 runs. Each produced a set of feature sets (25 feature sets) that was tested by using KNN [61], MLP [40], SVM [41], and LDA [42].

TABLE 7. Comparison with recent feature selection algorithms.

Classifier	ABC		SCA		GWO		ALO		SSA		BWOA-NRS	
	Accuracy	Time	Accuracy	Time	Accuracy	Time	Accuracy	Time	Accuracy	Time	Accuracy	Time
KNN	76	55	76	57	76	59	75	35	74	35	77	42
MLP	86	48	83	39	86	49	86	51	82	33	86	44
LDA	96	4.14	96	4.29	96	4.58	96	2.31	96	2.06	**96**	**1.91**
SVM	93	55	92	31	93	57	93	55	92	33	95	38
Average	87.75	40.53	86.75	32.8	87.75	42.39	87.5	35.82	86	**25.76**	**88.5**	31.47

For all classifiers, the same parameters were used in all runs, and a five-fold cross-validation approach was applied during the classification phase. It was observed that the performance of the proposed approach is acceptable. The highest recognition accuracy was achieved by LDA, which means that it is able to correctly classify 96% of the testing samples. LDA and SVM have an advantage over KNN and MLP. Even the lowest classification accuracy by LDA was (91%), which is considered quite high for handwritten Arabic characters. This means that the proposed BWOA-NRS approach is highly capable of extracting the most efficient features that can lead to better classification with the smallest number of features. Additionally, in most cases, it can be observed that there is a positive relationship between the number of features and the classification accuracy, so the more features extracted, the better the performance of the proposed approach. However, in some cases, some selected feature sets achieve a lower accuracy, in particular with a feature vector with a size of 567 for both SVM and LDA. Full performance measure details including all used classifiers and classification metrics are shown in Table 6.

1) COMPARISON WITH OTHER FEATURE SELECTION ALGORITHMS

The proposed BWOA-NRS approach is also compared with other well-known and very recent feature selection algorithms, such as the Artificial Bee Colony (ABC) [35], Sine Cosine Algorithm (SCA) [36], Grey Wolf Optimization (GWO) [37], Ant Lion Optimization (ALO) [38], and the Salp Swarm Algorithm (SSA) [39].

This comparison is shown in Table 7, where all optimization algorithms depend on NRS as the objective function. Each of the optimization algorithms also produced several feature sets, and only the best performance for each is presented in this table.

Table 7 shows that the BWOA-NRS algorithm achieves equal or better classification accuracy compared to the other feature selection methods. Moreover, the BWOA-NRS approach achieves the classification task in the smallest amount of time. Except SSA and ALO with some classifiers, but it outcomes also ALO in the average consumption time. LDA has remarkable success relative to the other classifiers in the smallest amount of time, for all experiments. For this task, KNN obtained the lowest performance for all optimizers.

TABLE 8. Comparison with deep neural networks.

Model	Number of features	Accuracy (%)	Recall	Precision	Time
VGGnet	25 K	91	**100**	91	40:35
Resnet	100 K	53	74	85	180:7
Inception	51 K	86	95	96	49
Xception	100 K	93	96	87	60:40
Mobilenet	50 K	93	95	**100**	48:33
BWOA-NRS	261	**96**	86	97	**0:91**

It is obvious that the proposed approach is more robust than the other feature selection algorithms (such as ABC, GWO, SSA, SCA, and ALO). Generally, the Whale Optimization Algorithm (WOA) performs successfully, with a good balance between both exploitation and exploration, while other feature selection algorithms still suffer from premature convergence when they become stuck at a local optimal value. This research also confirms that the Neighborhood Rough Set itself is a powerful feature selection algorithm.

2) COMPARISON WITH DEEP NEURAL NETWORKS (DNN)

In this subsection, we compare the efficiency of the proposed BWOA-NRS approach to deep neural networks like VGGnet [43], Resnet [44], Nasnet [45], Mobilenet [46], Inception [48], and Xception [47]. Since DNNs have demonstrated advantages in machine learning, especially in image classification tasks, DNNs have been employed in many recent studies for domain-specific tasks. A comparison between the BWOA-NRS algorithm and a DNN is shown in Table 8.

From Table 8 it can be observed that the DNN achieves high recognition accuracy in most cases, but still underperforms the proposed approach (with accurate classifications of 91% and 93% compared to 96%). Additionally, DNNs produce huge feature matrices (some of them are 3000 times larger than ours), which can be inefficient in terms of time and memory resources consumption. For example, some of the DNNs require about 800 times more memory than our proposed approach, wasting time, training effort, and memory resources. Table 8 also emphasizes that classic machine learning approaches, which imply handcrafted feature extraction

TABLE 9. Comparison with related works.

Previous work	Classifier	Feature selection algorithm	Accuracy (%)	Time (secs)
Khedher, et.al [49]	—	—	73.4	—
Aburas et.al [50]	Haar Wavelet	—	70	—
Al-Taani et.al [51]	Tree	—	75.3	—
Sahlol et.al [3]	Knn	BA	80.05	3.40
Sahlol et.al [4]	RF	PSO	91.66	58.83
El-Sawy et. al [6]	Convolutional neural networks	—	94.9	—
BWOA-NRS	LDA	WOA-NRS	**96**	**1.91**

and selection methods, are still able to provide a more broadly robust and efficient solution for current machine learning tasks than deep neural networks.

3) COMPARISON WITH RELEVANT WORKS

This section provides a comparison with relevant works that have tackled the notably difficult task of classifying handwritten Arabic characters.

Table 9 shows the most recent works on the CENPARMI dataset and other manually-created Arabic handwritten datasets.

From Table 9, the proposed BWOA-NRS approach outperforms all other works in both performance and time consumption. The performance of our approach is better than other algorithms by about 5%. Our approach also overcomes other works in terms of the time consumption, which means reduced memory and processor burdens. Other successful works, such as [4] and [3], also applied other feature selection algorithms such as Particle Swarm Optimization (PSO), and the Bat Algorithm (BA) (with the same classifier) and the Random Forest (RF), which indicates that using an appropriate optimization algorithm can lead to a successful OCR model. When building a machine learning approach, it is necessary to narrow down the features because this both improves the processing time and the overall high performance.

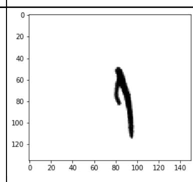
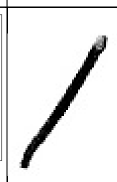
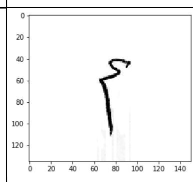
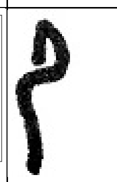
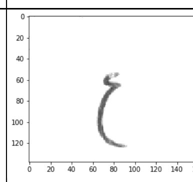
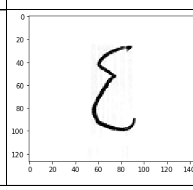
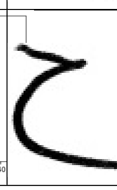
It should be emphasized that building a character recognition model is a challenging task, especially for Arabic characters. Arabic characters have varieties of drawing styles for a particular character, which, in turn may have a shape that is similar to that of other characters, as mentioned earlier in Table 1. Finally, building a model with 28 classes (the number of basic Arabic alphabetic characters) is a difficult task for any machine learning model.

4) FAILURE CASES ANALYSIS

To investigate our model's behavior, we tested the misclassified characters. With the proposed approach, there are some misclassified characters (failure cases), as can be seen in Table 10.

In Table 10, the first column shows the character's name, the second column presents the model's input character image, the third column shows the predicted character class, and the last column presents the classification confidence of the model. In the first failure case, the input character

TABLE 10. Model's failure cases.

Character	Model's input	Classified as	confidence (%)
Alif			89
Alif			99
Ayn			80
Ayn			62

is "Alif", which is classified as the "Raa" character with a high confidence level of 89%. Actually, in some writing styles, both look identical, even for Arabic natives, as shown in the second and the third columns. In the second case, although the "Alif" character is written properly, the secondary component, which is called "Hamza," is linked to the main character's shape, which changes the character's general appearance. Accordingly, this makes it looks like the character "Meem" in the third column. In the third case, the "Ayn" character was misclassified as "Meem" because they have the same shape, especially in the lower part. The third column shows a sample image of a "Meem" character taken from the CENPARMI dataset, which we used to perform our experiment. In the fourth case, the "Ayn" character does not seem to be written professionally (or whiten by a non-native), so it looks different than the "Ayn" character in real-life writing, thus it looks more similar to the "Meem" character.

The only current advantage of humans over a computer-based OCR system is that those confusing characters can be recognized by understanding the meaning of their context.

However, the proposed method still has some limitations, particularly that it works only with the basic Arabic Alphabetic characters, ignoring the 4 characters which are rarely used. Also, the current applications requiring the recognition of handwritten Arabic characters are very limited and not widely used because the Arabic language is not supported in many systems and needs further support from governments and national institutions.

V. CONCLUSION AND FUTURE WORK

In this paper, a hybrid approach based on the Binary Whale Optimization Algorithm with the Neighborhood Rough Set (BWOA-NRS) was proposed for handwritten Arabic character recognition. The main purpose of this paper is to build a hybrid approach that can select sufficient features that improve the performance of handwritten Arabic characters in the smallest amount of time with a low memory footprint. The results show that the BWOA-NRS approach can select the most appropriate features, which demonstrably improves the classification performance. The results were compared to the most recent feature selection algorithms such as ABC, SCA, GWA, ALO, and SSA, it can be observed that the BWOA-NRS algorithm outperforms other approaches based on swarm techniques. Further, by comparing the proposed approach with deep neural networks (i.e., VGGnet, Resnet, Nasnet, Mobilenet, Inception, and Xception) the proposed model demonstrated the highest recognition rate while requiring very little time or memory compared to the DNNs. Thereafter, studying some failure cases, it was ascertained that some of these cases are confusing for Arabic natives, and can only be recognized correctly by understanding the context. Encouraged by the achieved results, this work can be extended in the near future to automating the mailing systems by extracting and digitizing addresses from envelopes or houses.

REFERENCES

[1] A. Husni Al-Muhtaseb, S. Mahmoud, and S. Rami Qahwaji, "Statistical analysis for the support of Arabic text recognition," in *Proc. Int. Symp. Comput. Arabic Lang.*, Riyadh, Saudi Arabia, Nov. 2007, pp. 1–16.

[2] G. A. Abandah and T. M. Malas, "Feature selection for recognizing handwritten arabic letters," *Dirasat, Eng. Sci. J.*, vol. 37, no. 2, pp. 242–256, 2010.

[3] A. T. Sahlol, C. Y. Suen, H. M. Zawbaa, A. E. Hassanien, and M. A. Elfattah, "Bio-inspired BAT optimization algorithm for handwritten Arabic characters recognition," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Vancouver, BC, Canada, Jul. 2016, pp. 1749–1756.

[4] T. Ahmed Sahlol, M. A. Elfattah, Y. Ching Suen, and A. E. Hassanien, "Particle swarm optimization with random forests for handwritten arabic recognition system," in *Proc. Int. Conf. Adv. Intell. Syst. Inform. (AISI)*, Nov. 2016, pp. 437–446.

[5] A. A. Ewees, A. T. Sahlol, and M. A. Amasha, "A bio-inspired moth-flame optimization algorithm for arabic handwritten letter recognition," in *Proc. Int. Conf. Control, Artif. Intell., Robot. Optim. (ICCAIRO)*, Prague, Czech Republic, May 2017, pp. 154–159.

[6] A. El-Sawy and H. EL-Bakry, "Arabic handwritten characters recognition using convolutional neural network," *WSEAS Trans. Comput. Res.*, vol. 5, pp. 11–19, Jan. 2017.

[7] A. T. Sahlol and A. Hasanain, "Bio-inspired optimization algorithms for arabic handwritten characters," in *Handbook of Research on Machine Learning Innovations and Trends*. Hershey, PA, USA: IGI Global, 2017, pp. 897–914.

[8] M. Elleuch, N. Tagougui, and M. Kherallah, "Arabic handwritten characters recognition using Deep Belief Neural Networks," in *Proc. IEEE 12th Int. Multi-Conf. Syst., Signals Devices (SSD)*, Mahdia, Tunisia, Mar. 2015, pp. 1–5.

[9] S. B. Ahmed, S. Naz, M. I. Razzak, and R. Yousaf, "Deep learning based isolated Arabic scene character recognition," in *Proc. 1st Int. Workshop Arabic Script Anal. Recognit. (ASAR)*, Apr. 2017, pp. 46–51.

[10] M. T. Parvez and S. A. Mahmoud, "Offline arabic handwritten text recognition: A Survey," *J. Assoc. Comput. Machinery*, vol. 45, no. 2, pp. 1–35, Feb. 2013.

[11] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning About Data*. Dordrecht, The Netherlands: Kluwer, 1991.

[12] Q. Hu, D. Yu, J. Liu, and C. Wu, "Neighborhood rough set based heterogeneous feature subset selection," *Inf. Sci.*, vol. 178, no. 18, pp. 3577–3594, Sep. 2008.

[13] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Nov. 1995, pp. 1942–1948.

[14] A. E. Hassanien and E. Emary, *Swarm Intelligence: Principles, Advances, and Applications*. New York, NY, USA: Taylor & Francis, 2015.

[15] S.-Y. Jing, "A hybrid genetic algorithm for feature subset selection in rough set theory," *Soft Comput.*, vol. 18, no. 7, pp. 1373–1382, Jul. 2014.

[16] Y. Liu, G. Wang, H. Chen, H. Dong, X. Zhu, and S. Wang, "An improved particle swarm optimization for feature selection," *J. Bionic Eng.*, vol. 8, no. 2, pp. 191–200, Jun. 2011.

[17] S. Mirjalili and A. Lewisa, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016.

[18] A. T. Sahlol, M. A. Abdeldaim, and A. E. Hassanien, "Automatic acute lymphoblastic leukemia classification model using social spider optimization algorithm," *Soft Comput.*, vol. 23, no. 15, pp. 6345–6360, 2019.

[19] A. T. Sahlol, F. H. Ismail, A. Abdeldaim, and A. E. Hassanien, "Elephant herd optimization with neural networks: A case study on acute lymphoblastic leukemia diagnosis," in *Proc. 12th Int. Conf. Comput. Eng. Syst. (ICCES)*, Dec. 2017, pp. 657–662.

[20] A. T. Sahlol, A. A. Ewees, A. M. Hemdan, and A. E. Hassanien, "Training feedforward neural networks using Sine-Cosine algorithm to improve the prediction of liver enzymes on fish farmed on nano-selenite," in *Proc. 12th Int. Comput. Eng. Conf. (ICENCO)*, Dec. 2016, pp. 35–40.

[21] A. T. Sahlol, Y. S. Moemen, A. A. Ewees, and A. E. Hassanien, "Evaluation of cisplatin efficiency as a chemotherapeutic drug based on neural networks optimized by genetic algorithm," in *Proc. 12th Int. Conf. Comput. Eng. Syst. (ICCES)*, Dec. 2017, pp. 682–685.

[22] O. Adjei, L. Chen, H.-D. Cheng, D. H. Cooley, R. J. Cheng, and X. Twombly, "A fuzzy search method for rough sets in data mining," in *Proc. IFSA World Congr. 20th NAFIPS Int. Conf.*, vol. 2, Jul. 2001, pp. 980–985.

[23] U. Maulik and D. Chakraborty, "Fuzzy preference based feature selection and semisupervised SVM for cancer classification," *IEEE Trans. Nanobiosci.*, vol. 13, no. 2, pp. 152–160, Jun. 2014.

[24] A. T. Sahlol, C. Y. Suen, M. R. Elbasyoni, and A. A. Sallam, "A proposed OCR algorithm for cursive handwritten arabic character recognition," *J. Pattern Recognit. Intell. Syst.*, vol. 2, no. 1, pp. 90–104, 2014.

[25] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[26] A. T. Sahlol, C. Y. Suen, M. R. Elbasyoni, and A. A. Sallam, "Investigating of preprocessing techniques and novel features in recognition of handwritten arabic characters," in *Artificial Neural Networks in Pattern Recognition*. Cham, Switzerland: Springer, 2014, pp. 264–276.

[27] I. Sobel, "An isotropic gradient operator," in *Machine Vision for Three-Dimensional Scenes*, H. Freeman, Ed. New York, NY, USA: Academic, 1990, pp. 376–379.

[28] J. M. S. Prewitt, "Object enhancement and extraction," in *Picture Processing and Psychopictorics*. New York, NY, USA: Academic, 1970.

[29] A. T. Sahlol, M. Elhoseny, E. Elhariri, and A. E. Hassanien, "Arabic handwritten characters recognition system, towards improving its accuracy," in *Proc. IEEE Int. Conf. Intell. Techn. Control, Optim. Signal Process. (INCOS)*. Srivilliputhur, India: Kalasalingam Univ., Mar. 2017, pp. 1–7.

[30] A. T. Sahlol and A. E. Hassanien, "Bio-inspired optimization algorithms for Arabic handwritten characters," in *Handbook of Research on Machine Learning Innovations and Trends*. Hershey, PA, USA: IGI Global, 2017.

[31] A. A. E. Ewees and A. Talaat, ''Bio-inspired optimization algorithms for improving artificial neural networks: A case study on handwritten letter recognition,'' *Comput. Linguistics, Speech Image Process. Arabic Lang.*, vol. 4, pp. 249–266, Nov. 2018.

[32] M. W. Sagheer, ''Novel word recognition and word spotting systems for off-line urdu handwriting,'' M.S. thesis, Dept. Comput. Sci. Softw. Eng., Concordia Univ., Montreal, QC, Canada, 2010.

[33] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, 3rd ed. San Mateo, CA, USA: Morgan Kaufmann, 2011.

[34] H. Alamri, J. Sadri, C. Y. Suen, and N. Nobile, ''A novel comprehensive database for arabic off line handwriting recognition,'' in *Proc. 11th Int. Conf. Frontiers Handwriting Recognit.*, 2008, pp. 664–669.

[35] D. Karaboga and B. Basturk, ''A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm,'' *J. Global Optim.*, vol. 39, no. 3, pp. 459–471, Oct. 2007.

[36] S. Mirjalili, ''SCA: A Sine Cosine algorithm for solving optimization problems,'' *Knowl.-Based Syst.*, vol. 96, pp. 120–133, Mar. 2016.

[37] S. Mirjalili, S. M. Mirjalili, and A. Lewis, ''Grey wolf optimizer,'' *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.

[38] S. Mirjalili, ''The Ant Lion Optimizer,'' *Adv. Eng. Softw.*, vol. 83, pp. 80–98, May 2015.

[39] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, ''Salp swarm algorithm: A bio-inspired optimizer for engineering design problems,'' *Adv. Eng. Softw.*, vol. 114, pp. 163–191, Dec. 2017.

[40] S. Haykin, *Neural Networks: A Comprehensive Foundation*. New York, NY, USA: Macmillan, 1994.

[41] M. A. Hearst, ''Support vector machines,'' *IEEE Intell. Syst.*, vol. 13, no. 4, pp. 18–28, July/Aug. 1998.

[42] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. Mullers, ''Fisher discriminant analysis with kernels,'' in *Proc. Neural Netw. Signal Process. IX, IEEE Signal Process. Soc. Workshop*, Jan. 2003, pp. 41–48.

[43] K. Simonyan and A. Zisserman, ''Very deep convolutional networks for large-scale image recognition,'' Sep. 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[44] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[45] Blog, GR, Google Research. (2017). *AutoML for Large Scale Image Classification and Object Detection*. [Online]. Available: https://research.googleblog.com/2017/11/automl-for-large-scaleimage.html

[46] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, ''MobileNets: Efficient convolutional neural networks for mobile vision applications,'' Apr. 2017, *arXiv:1704.04861*. [Online]. Available: https://arxiv.org/abs/1704.04861

[47] F. Chollet, ''Xception: Deep learning with depthwise separable convolutions,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258.

[48] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, ''Going deeper with convolutions,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[49] M. Z. Khedher, G. A. Abandah, and A. M. Al-Khawaldeh, ''Optimizing feature selection for recognizing handwritten arabic characters,'' *World Acad. Sci., Eng. Technol.*, vol. 1, no. 4, pp. 1023–1026, 2005.

[50] A. Abdurazzag, S. M. A. Aburas, and S. M. Rehiel, ''Off-line Omni-style handwriting Arabic character recognition system based on wavelet compression,'' *Arab Res. Inst. Sci. Eng.* vol. 3, pp. 123–135, Jan. 2007.

[51] A. T. Al-Taani and S. Al-Haj, ''Recognition of on-line arabic handwritten characters using structural features,'' *J. Pattern Recognit. Res.*, vol. 5, no. 1, pp. 23–37, 2010.

[52] Z. Yan, J. Sha, B. Liu, W. Tian, and J. Lu, ''An ameliorative whale optimization algorithm for multi-objective optimal allocation of water resources in Handan, China,'' *Water*, vol. 10, no. 1, p. 87, 2018.

[53] A. G. Hussien, A. E. Hassanien, E. H. Houssein, S. Bhattacharyya, and M. Amin, ''S-shaped binary whale optimization algorithm for feature selection,'' in *Recent Trends in Signal and Image Processing* (Advances in Intelligent Systems and Computing), vol. 727, S. Bhattacharyya, A. Mukherjee, H. Bhaumik, S. Das, and K. Yoshida, Eds. Singapore: Springer, 2019.

[54] M. A. E. Aziz and A. E. Hassanien, ''An improved social spider optimization algorithm based on rough sets for solving minimum number attribute reduction problem,'' *Neural Comput. Appl.*, vol. 30, no. 8, pp. 2441–2452, 2017.

[55] M. A. E. Aziz and A. E. Hassanien, ''Modified cuckoo search algorithm with rough sets for feature selection,'' *Neural Comput. Appl.*, vol. 29, pp. 925–934, Feb. 2018.

[56] J. R. Anaraki and M. Eftekhari, ''Rough set based feature selection: A review,'' in *Proc. 5th Conf. Inf. Knowl. Technol.*, Shiraz, Iran, May 2013, pp. 301–306.

[57] M. Mafarja and S. Mirjalili, ''Whale optimization approaches for wrapper feature selection,'' *Appl. Soft Comput.*, vol. 62, pp. 441–453, Jan. 2018.

[58] M. M. Mafarja and S. Mirjalili, ''Hybrid whale optimization algorithm with simulated annealing for feature selection,'' *Neurocomputing*, vol. 260, pp. 302–312, Oct. 2017.

[59] M. A. E. Aziz, A. A. Ewees, and A. E. Hassanien, ''Whale optimization algorithm and moth-flame optimization for multilevel thresholding image segmentation,'' *Expert Syst. Appl.*, vol. 83, pp. 242–256, Oct. 2017.

[60] M. A. E. Aziz, A. A. Ewees, A. E. Hassanien, M. Mudhsh, and S. Xiong, ''Multi-objective whale optimization algorithm for multilevel thresholding segmentation,'' in *Advances in Soft Computing and Machine Learning in Image Processing*. Cham, Switzerland: Springer, 2018, pp. 23–39.

[61] K. Thearling, ''An introduction to data mining,'' in *Direct Marketing Magazine*. Feb. 1999, pp. 28–31.

AHMED TALAT SAHLOL was born in Damietta, Egypt, in 1983. He received the B.S. and M.S. degrees in computer teacher preparation from the Mansoura University, in 2004 and 2010, respectively, and the Ph.D. degree as a channel system betaeem from Damietta University, Egypt, and Concordia University, Canada, in 2015. Since 2015, he has been a Lecturer with the Computer Teacher Preparation Department, Damietta University. In March 2019, he has awarded a Postdoctoral Fellowship in Wurzburg, Germany, funded by DAAD. He has authored more than 20 articles. His research interests include deep learning, machine learning, and image processing.

MOHAMED ABD ELAZIZ was born in Zagazig, Egypt, in 1987. He received the B.S. and M.S. degrees in computer science from Zagazig University, in 2008 and 2011, respectively, and the Ph.D. degree in mathematics and computer science from Zagazig University, in 2014. From 2008 to 2011, he was an Assistant lecturer with the Department of Computer Science. Since 2014, he has been a Lecturer with the Mathematical Department, Zagazig University. He has authored more than 30 articles. His research interests include machine learning, signal processing, and image processing.

MOHAMMED A. A. AL-QANESS received the B.S., M.S., and Ph.D. degrees from Wuhan University of Technology, in 2010, 2014, and 2017, respectively, all in information and communication engineering. He is currently an Assistant Professor with the School of Computer Science, Wuhan University, Wuhan, China. His current research interests include wireless sensing, mobile computing, natural language processing (NLP), machine learning, and signal & image processing.

SUNGHWAN KIM received the B.S., M.S., and Ph.D. degrees from Seoul National University, South Korea, in 1999, 2001, and 2005, respectively. He was a Postdoctoral Visitor with the Georgia Institute of Technology (GeorgiaTech), from 2005 to 2007 and a Senior Engineer with Samsung Electronics, from 2007 to 2011. He is currently a Professor with the School of Electrical Engineering, University of Ulsan, South Korea. His main research interests are channel coding, modulation, massive MIMO, visible light communication, and quantum information.