Date of publication JUN-30, 2022, date of current version JUN-07, 2022.
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.21.2.2584

A Performant Clustering Approach Based on An Improved Sine Cosine Algorithm

LAHBIB KHRISSI[1], NABIL EL AKKAD[2], HASSAN SATORI[1], KHALID SATORI[1]
[1]LISAC, Department of Computer Science, Faculty of Science, Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco
[2]Laboratory of Engineering, Systems and Applications, ENSA of Fez, Sidi Mohamed Ben Abdellah University, Fez, Morocco

Corresponding author: Lahbib Khrissi (email: lahbib.khrissi@usmba.ac.ma).

ABSTRACT Image segmentation is a fundamental and important step in many computer vision applications. One of the most widely used image segmentation techniques is clustering. It is a process of segmenting the intensities of a non-homogeneous image into homogeneous regions based on their similarity property. However, clustering methods require a prior initialization of random clustering centers and often converge to the local optimum, thanks to the choices of the initial centers, which is a major drawback. Therefore, to overcome this problem, we used the improved version of the sine-cosine algorithm to optimize the traditional clustering techniques to improve the image segmentation results. The proposed method provides better exploration of the search space compared to the original SCA algorithm which only focuses on the best solution to generate a new solution. The proposed ISCA algorithm is able to speed up the convergence and avoid falling into local optima by introducing two mechanisms that take into account the first is the given random position of the search space and the second is the position of the best solution found so far to balance the exploration and exploitation. The performance of the proposed approach was evaluated by comparing several clustering algorithms based on metaheuristics such as the original SCA, genetic algorithms (GA) and particle swarm optimization (PSO). Our evaluation results were analyzed based on the best fitness values of several metrics used in this paper, which demonstrates the high performance of the proposed approach that gives satisfactory results compared to other comparison methods.

KEYWORDS Clustering; Image Segmentation; Improved Sine Cosine Algorithm (ISCA); Classification; Optimization.

I. INTRODUCTION

THE purpose of image analysis is to extract the characteristic information contained in an image. The result of such an analysis is often called the structural description. This can take the form of an image or any data structure allowing a description of the entities contained in the image. In contrast to the interpretation phase, the analysis tries, as much as possible, to ignore the context (i.e., the application). Essentially, image analysis uses segmentation where we try to associate a label to each pixel of the image based on the information carried (grayscale or color), its spatial distribution on the image medium, simple models.

The objective of segmentation is to simplify the image into segments that have a strong correlation with real-world objects in order to facilitate its analysis. In other words, segmenting an image means changing it into a machine-readable representation. Thus, image segmentation has been of great interest to imaging researchers for the last three decades. However, until today, no universal method for image segmentation exists. Each proposed method is often effective for a given type of image, and for a given application. Researchers are constantly opening new horizons and exploring different techniques in order to create more and more efficient approaches. Segmentation is mainly used as a pre-processing step to annotate, enhance, analyze, classify, categorize, and/or summarize the image information set. Generally, image segmentation techniques are based either on the search for local discontinuities (edge detection) or on the detection of image areas (region extraction).

Many image processing applications cannot do without the segmentation step. This is the case in the various applications of [1] remote surveillance for the detection of mobile objects, in the medical field [2], [3] and in the field of security and cryptography [4].

A multitude of segmentation methods are proposed in the literature, but it is not easy to choose a best technique for image segmentation. As for many segmentation tasks in image processing, the variability from one image to another represents a limit to the performance of segmentation subtlety. Several methods exist, including: split/merge [5], region growing [6], [7], active contouring [8], [9], clustering [9], [10], graphical clipping [11], genetic algorithms [12], etc. Among the many existing methods, image segmentation by clustering appears to be one of the most used segmentation techniques because of its efficiency and speed. This method can be used directly or easily adapted for dimensional data, its use for images was practically an intuitive choice. Thus, clustering is defined as the process of grouping similar objects into a single class (cluster), and dissimilar objects into different classes. This according to a given similarity criterion. The clustering techniques proposed in the literature for image segmentation can be classified into two categories. Partitioning methods (non-hierarchical) [13]; or hierarchical methods [14], which are often based on similarity or distance criteria between pixels. Among the most used non-hierarchical clustering methods is the k-means.

The evaluation of image segmentation poses a major problem in distinguishing between the different segmentation methods found in the literature, which involves a fundamental conflict between objectivity and generality [15].

Segmentation methods are often based on the minimization of a criterion or the estimation of parameters (models). When confronted with very difficult cases, these methods use either simplifying assumptions on the models or segmentation criteria and attributes of low complexity. Contrary to these so-called classical methods, metaheuristics have the advantage of treating the image segmentation problem as a whole, thanks to their ability to find sub-optimal solutions close to the global optima in reasonable computation times. Thus, the segmentation problem can be treated under a different angle, which does not require the use of a priori simplifications

The strength of metaheuristics lies in their stochastic character allowing them to exploit very large search spaces and their acceptance of degraded solutions allowing them to get rid of the blocking problem in local minima. These methods are, for the most part, inspired by from physics (simulated annealing), biology (evolutionary algorithms) or ethology (particle swarms, ant colonies, etc.).

Optimization metaheuristics have been widely exploited by several authors to improve the quality of image segmentation according to several well-recognized segmentation evaluation criteria in the literature. Simulated annealing has been used in works [16]. Ant colony optimization has been used by Khorram et al [17], and particle swarm optimization by Tan et al [18]. Genetic algorithms have been exploited by Khrissi et al [19]. Moreover, other metaheuristics and their hybridization have been proposed by Yue et al [20], and Karthikeyan et al [21]

SCA is among the recent promising population-based metaheuristic algorithms described by Mirjalili in 2016. This algorithm is based on the properties of sine and cosine trigonometric functions to perform mining and exploration of the search space. SCA has been tested on many benchmark functions, and has shown its performance and high efficiency compared to several well recognized metaheuristics existing in the literature. It is used to solve various optimization problems such as image processing, robot trajectory planning, feature selection, economic dispatching, radial distribution networks, etc. [22].

The original SCA suffers from some limitations, such as locking in local optima, slow convergence, unbalanced operation, and it is time consuming. In this paper, we present an improved version of SCA that improves the performance of the traditional SCA to handle the clustering problem.

However, like other population-based algorithms, the original SCA suffers from some limitations, such as blocking in local optima, slow convergence and unbalanced operation. In this paper, we present an improved version of SCA, called ISCA (improved sine cosine algorithm). The use of ISCA improves the performance of the traditional SCA to deal with the clustering problem whose objective is to increase the quality of segmented images. The improvement of the SCA is done at the level of the update of the modified position by integrating inertia weights to accelerate the convergence and of course to avoid falling in local optima. The performance of the proposed technique has been evaluated on different types of reference images and compared by several metaheuristics such as the original SCA and others. ISCA shows an efficiency of clustering to obtain the best image segmentation result.

The rest of this paper is organized as follows: in the second part, we describe the previous work. The third part will be devoted to the theoretical framework. The clustering method proposed by ISCA is treated in the fourth part. The fifth part will be devoted to the discussion of the obtained results, and the conclusion will be addressed in the last part.

II. RELATED WORKS

Image segmentation by clustering has been the subject of various studies that have given rise to a multitude of approaches. These approaches can be differentiated according to the similarity measure used during the classification of the pixels. Any function that evaluates the degree of similarity between pixels can be used for this purpose. The most efficient data clustering algorithms are the K-means algorithm and the fuzzy C-means algorithm (FCM), but the solution of these methods depends on the initial random state and always converges to the local optimum [23]. This motivates authors to integrate metaheuristics to optimize traditional algorithms. Nowadays, several metaheuristics have been used to solve the problem of image segmentation. Among these works we can cite:

Singh [24] describes a method of segmentation of sunflower leaf images performed by particle swarm optimization algorithm. this technique has an important aspect for disease classification, the results obtained are very satisfactory which have given by the experiments performed on leaf images. The average classification accuracy of the proposed technique is up to 98,0 %, but the most recent methods do not reach this accuracy.

Tongbram et al [25] proposed a new image segmentation approach based on Whale Optimization Algorithm (WOA) and Fuzzy c-means (FCM) with the use of noise detection and reduction mechanism. The number of iterations is almost equal between the exploration and exploitation phases because they are performed separately. The WOA algorithm showed higher convergence speed and better avoidance of local optima simultaneously. The effectiveness of the technique proposed by the authors was tested using synthetic images and medical resonance imaging (MRI) images by taking different types of noise. The obtained results are validated by comparison with other existing segmentation methods and evaluated by different evaluation indices well known in the literature.

Bosco [27] described a new image segmentation algorithm based on genetic algorithms that allow us to consider the segmentation problem as a global optimization problem (GOP). The fitness function is based on the similarity between images. The results obtained from real images, show a good performance of the proposed method.

Jia et al [26] presented a new efficient image segmentation approach based on the hybridization between the Grasshopper

Optimization Algorithm (GOA) and Minimum Cross Entropy (MCE). A series of experiments is performed on various satellite images to evaluate the efficiency of the proposed algorithm and a comparison has been made to standard methods, as well as to advanced satellite image thresholding techniques based on different criteria. The superiority of the proposed algorithm is presented in different aspects and metrics well recognized in the literature, such as average fitness function value, peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), standard deviation (STD), and others.

Farmer et al [27] described a segmentation method that uses the classification subsystem as an integral part of the segmentation, which provides contextual information about the objects to be segmented. The performance of the segmentation wrapper technique is shown by the genetic algorithm (GA) on real complex images of occupants of motor vehicles. the results obtained are very satisfactory either at the segmentation level are extremely accurate, or at the classification level which arrives an accuracy of 88%.

Recently, Mirjalili et al [21] proposed a new population-based meta-heuristic optimization algorithm for solving optimization problems, called Sine Cosine Algorithm (SCA). The SCA creates several random initial candidate solutions and requires them to scale outward or to the best solution using a mathematical model based on sine and cosine functions. The sine-cosine algorithm has the advantage of fewer parameters, simple structure, and easy implementation. It has attracted the attention of many researchers. With the deepening of the research, improved methods and practical applications of the SCA have been proposed. The literature shows that this algorithm very much uses to study the real engineering problems by examples: In [28], SCA is used to optimize the position and size of the capacitor bank to solve voltage instability and line loss problems in the radial distribution system. In [28], [29], SCA is used to optimize the connecting rods of car engine. In [40], the SCA is used for frequency band selection and allows to obtain a subset of frequency bands with higher classification accuracy.

Many recent studies that are presented improvements on the SCA algorithm to overcome the drawbacks of traditional SCA: Fernández et al [30] proposed a binary percentile SCA (BPSCOA) to solve the set coverage problem (SCP), Tuncer [31] proposed a new weighted chaotic SCA (LDW-SCSA) based on the integration of chaos in the SCA for numerical function optimization, Liang et al [32] presented an oppositional chaotic SCA (COSCA) for solving global optimization problems, Zamli et al [33] proposed an adaptive SCA for solving the combinatorial testing problem and Abdel-Fatah et al [34] introduced a modified SCA (MSCA) based on Lévy Flight distribution.

In this paper, we will use an improved version of SCA called ISCA to deal with the clustering problem whose objective is to increase the quality of the segmented images. The ISCA process works as follows, the marked individuals update their position using a combination between two mechanisms which takes into account the first is the given random position of the search space and the second is the position of the best solution found so far. The experimental results of the ISCA algorithm were compared with those of well-known metaheuristics such as GA, PSO and original SCA.

III. METHODOLOGY

A. ORIGINAL SCA

The basic SCA proposed by [21] could be a modern optimization strategy to solve optimization problems. The SCA makes some irregular candidate arrangements at the beginning and forces them to fluctuate outwards or towards the finest arrangement using a numerical demonstration based on sine and cosine capabilities. In addition, some arbitrary and versatile factors are coordinates in this algorithm to focus on investigation and misuse of space at distinct stages of optimization.

The SCA is a stochastic algorithm [21] that contains several iterations. We update, for each iteration, the solutions according to the sine or cosine function by the following expressions:

$$X(i,j)_{t+1} = X(i,j)_t + r_1 * \sin(r_2) * |r_3 * P(j)_t - X(i,j)_t|, \quad (1)$$

$$X(i,j)_{t+1} = X(i,j)_t + r_1 * \cos(r_2) * |r_3 * P(j)_t - X(i,j)_t|, \quad (2)$$

Such as,

$X(i,j)_t$: the position of the current solution (i) in dimension (j) at the iteration (t).

$P(j)_t$: the position of the best individual in dimension (j) at the iteration (t),

$|e|$: the absolute value of (e).

r_1, r_2 and r_3 are three random variables.

The two expressions (1) and (2) presented blow give the following formula:

$$X(i,j)_{t+1} = \begin{cases} X(i,j)_t + r_1 * \sin(r_2) * |r_3 * P(j)_t - X(i,j)_t| & \text{if } r_4 < 0{,}5 \\ X(i,j)_t + r_1 * \cos(r_2) * |r_3 * P(j)_t - X(i,j)_t| & \text{if } r_4 \geq 0{,}5 \end{cases}, \quad (3)$$

Such as $r_4 \epsilon [0 \;\; 1]$.

Expression (3) represents a circular search model [21]. In this model, the best solution is located in the center of a circle and the search area surrounds this solution. This area is divided into sub-areas representing possible exploration areas for Xi. The parameters r_1,r_2,r_3 and r_4 are defined as follows:

r_1 controls how Xi varies in these areas. Indeed, if $r_1 > 1$ then Xi moves towards P. Otherwise, this point moves away from P. In addition, r_1 is used in the balancing of the exploration and exploitation phases, since it allows to be reduced by the expression (4).

r_2 control how far Xi moves along its direction according to r_1.

r_3 brings a random weight to the destination in order to stochastically emphasize ($r_3 > 1$) or de-emphasize ($r_3 < 1$) the destination effect in the distance definition.

r_4 to switches between the sine and cosine components of equation (3).

The random number r_1 determine the exploration phase (when $r_1 > 1$) or exploitation phase (when $r_1 < 1$). The value of r_1 which is adopted in the SCA is given by the formula (4):

$$r_1 = a - a * \frac{t}{T}, \quad (4)$$

where T is the maximum number of iterations to be given as a stopping criterion of the algorithm. t is the current iteration and a is a constant.

The SCA algorithm is given in Algorithm 1.

Algorithm1: Original SCA.
1. Initialize a set of search agents (solutions) (X)
2. ***repeat***
3. Evaluate each of the search agents by the objective Function
4. Update the best solution obtained so far (P = X)
5. Update the parameters r_1,r_2 and r_3
6. Update the poison of search agents using Eq. (3)
7. ***until*** (t < *T*)
8. Return the best solution obtained so far as the global optimum

B. IMPROVED SINE COSINE ALGORITHM

Although the sine-cosine algorithm effectively balances the ability to explore and produce the solution space by four parameters (r_1,r_2,r_3 and r_4), but it suffers from some limitations, such as slow convergence, slow excision time and also trapping in local optima due to premature convergence in optimization problems. In order to increase the performance of the algorithm and avoid local optima. We present a new ISCA algorithm that takes into account two solution positions: the first is the position of the best solution found so far, and the second is a random position taken in the search space. The proposed idea is to make the generation of the potential solution render not only the best solution, but also another randomly generated solution position. With this method, the global search capability and local search capability of the algorithm during the exploration period are improved, and the probability of falling into the local minimum is reduced. To translate this idea concretely, we propose that the new position of the solution (i) is determined by the following relation:

$$X(i,j)_{t+1} = (1-r_1) * X(i,j)_t + r_1 * R, \quad (5)$$

where R is a random position in the interval of the search space.

The weights (1-r_1) and (r_1) represent the weights relative to the current position $X(i,j)t$ and the random position R respectively. These parameters affect the position of the new solution. Indeed, the latter is influenced by the random position R at the beginning of the search. This influence is degraded iteratively by the first parameter. Consequently, at the end of the search, the position of the new solution tends towards its current position. This operation of attraction between the position R and the current position expresses the passage from divergence to convergence. In addition, we introduce some improvements to the system of equations (3). First, we propose to merge the two sine and cosine formulas into a single formula: instead of switching between two formulas, we want to limit the position update to only one formula among the two previous ones. Second, to obtain a good compromise between intensification and diversification, we present a new equation combined by the two random variables c and r1. This allows to ensure the efficient transition between exploration and exploitation.

The updates of the new agent positions and the diversification strategy are shown in the following equation (6):

$$X(i,j)_{t+1} = \begin{cases} (1-r_1) * X(i,j)_t + r_1 * R & \text{if } c < r_1 \\ P(j)_t + r_1 * \sin(r_2) * |r_3 * P(j)_t - X(i,j)_t| & \text{if } c \geq r_1 \end{cases}, \quad (6)$$

where:

c is given by the following formula $c = b * r_1$.

b is an integer generally greater than 1 which is used to amplify the parameter c.

The larger the parameter b, the more likely it is that we will have a diversification, otherwise we intensify the search for the best solution P.

Algorithm 2 below illustrates the pseudocode of our ISCA algorithm:

Algorithm 2: ISCA Algorithm

1 Initialize the positions of the agents (X)
2 Compute the (fitness) for any agent
3 Select the best position ($best_{pos}$)
4 Set the max numbers of iteration T
5 ***While*** (t<*T*)
  $r_1 = a - a * \frac{t}{T}$
  for each agent Xi in the population ***do***
    $r_2 = (2 * pi) * rand()$
    $r_3 = rand()$
    $r_4 = b * rand()$
    if ($r_4 < r_1$) **then**
      $m = get\_rand\_position();$
      $X(j) = (1 - r_1) * X(j) + r_1 * m$
    else
      $X(j) = best_{pos(j)} + r_1 * (\sin(r_2)) * abs(r_3 * best_{pos(j)} - X(j))$
    end
6 ***end***
7 Recalculate the (fitness) for each agent (X)
8 Get the best position ($best_{pos}$).
9 $t = t + 1$
10 ***end of while***

IV. PROPOSED CLUSTERING METHOD

In this section, we will address the image segmentation problem using the clustering approach. This proposed method uses a clustering technique based on the ISCA algorithm to find combinations of features that maximize the inter-class distance and minimize the intra-class distance. The goal of our approach is to find an optimal point in the search space that minimizes the fitness function that we formulated in expression (7) in the section below. The search space for each feature, represented by an individual dimension, and the range of each dimension from 0 to 1 are very large and thus require a smart search method.

The approach proposed in this paper is based on the ISCA algorithm to find a better classification of image pixels that have the same similarity. The ISCA optimizer allows a better exploration of the search space in order to improve the image segmentation results, especially the classification performance. We proved that this exploration capability is guaranteed thanks to a strategy based on two mechanisms: first, two solution positions are considered, the first position is always the best solution already found, and the second is another random position chosen in the search space. With this combination, we expect to obtain a better exploration of the latter, which leads to cut the space of individuals (pixels) into homogeneous areas according to a similarity criterion (criterion of proximity of their attribute vectors in the space of representation between the individuals). The most commonly used criterion in the literature is the Euclidean distance, which finds the minimum distance between the points with each of the available clusters and assigns the point to the cluster.

Image segmentation by clustering has been the subject of various studies that have led to a multitude of approaches. These approaches can be differentiated according to the similarity measure used during the classification of pixels. Any function that evaluates the degree of similarity between pixels can be used for this purpose. The measures used by similarity-based clustering methods are considered local measures (e.g., the similarity between two pixels or the similarity between pixels and the center of a region). In this paper, we used the nonlinear cost function to solve the clustering problem as the one treated in classical clustering algorithms like k-means, this function is the Euclidean distance. The minimization of this objective function (fitness), which is one of the main objectives of clustering, leads to a higher similarity in each cluster and to the difference with other clusters. This function is formulated by the following expression:

$$fitness(x_i, c_j) = \sum_{i=1}^{N} \sum_{j=1}^{K} \left\| x_i - c_{ij} \right\|^2, \qquad (7)$$

where c_{ij} refers to the cluster center vector of the ith member of the jth cluster (c_{ij}).

N and *K* are the number of picture elements and the number of clusters, respectively.

The process of our approach is realized by the following step: in the first step, the positions of the solutions are randomly initialized in the search space. Then, a loop of T (maximum iteration) steps is executed for the consistent clusters. At each iteration of this loop, the score (fitness) of each solution is calculated using the objective function defined in formula (7). After this evaluation phase, the best solution ($best_{pos}$) in the population is determined based on the best score. The updating of the solution positions is done using the system of equations (6) in which each agent swaps between its two formulas, serving the diversification and intensification strategies, respectively. This alternation is controlled using the two parameters r_1 and r_4, where r_1 decreases steadily with each iteration while the variable r_4 receives random values between 0 and 1. The diversification phase is performed if the condition ($r_4 < r_1$) is verified, otherwise intensification is performed. Furthermore, the update of the parameters (r_1, r_4) is adjusted to obtain an appropriate transaction of the diversification strategy to the intensification, which ensures a good balance between them.

The main steps of the algorithm proposed by our approach are presented in algorithm 2 below:

Algorithm 2: Algorithm of the proposed method

Step 1: Initialize the problem and the parameters of the algorithm.

The performance of the clustering algorithm is evaluated using the fitness function described in equation (7). In this article, each cluster center is considered as a decision variable.

Step 2: Initialize the search agents (solutions).

Search agents are initialized with image points chosen at random from the given set of points.

Step 3: Evaluation of the research agents.

The performance of search agents is calculated using the fitness function, which is mentioned in equation (7). The position of the best search agent is updated using the fitness values.

Step 4: Update the position of the search agents.

The parameter values are updated. The position of the search agents is updated using equations (6) or (5) depending on the value of r_4 relative to r_1.

Step 5: stoppage criteria.

The calculation is complete when the stoppage criterion is met. Otherwise, steps 3 and 4 are repeated. Finally, the best search agent is selected and is considered the best cluster center of the set of points in the image.

The advantages of the ISCA optimizer over other metaheuristics because it is characterized by its fast convergence speed and also simple to implement, the application of this algorithm in the field of clustering shows a very important efficiency since it is able to produce optimal cluster centers and gathers the data of the same cluster with greater similarities, while the data of different clusters have greater dissimilarities [35].

The proposed method is illustrated in the form of the flowchart in Fig. 1.

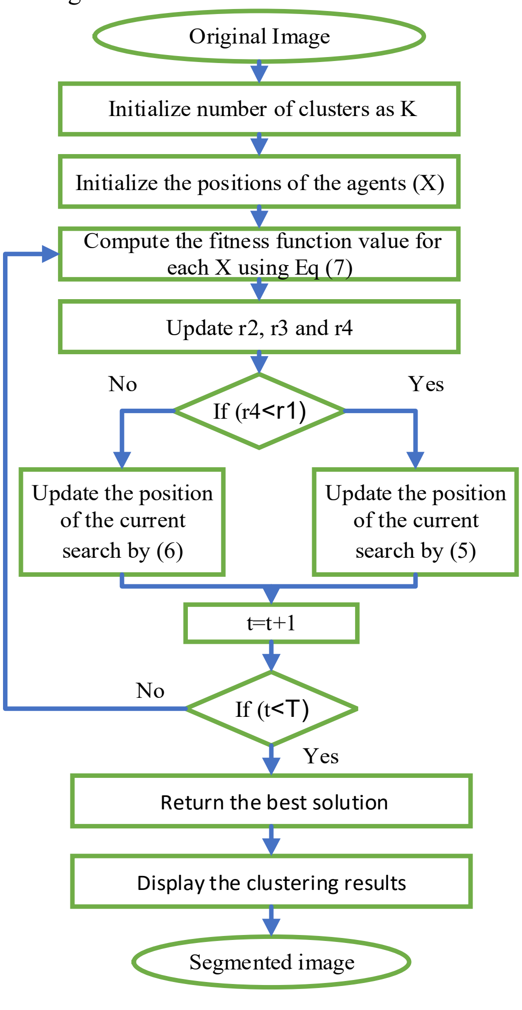


Figure 1. The flowchart of the proposed approach

V. EXPERIMENTAL RESULTS

In order to study the performance of the proposed method with the ISCA algorithm in the field of image segmentation by clustering, a series of experiments has been performed using natural reference images. In addition, to evaluate and validate our proposal we have made a comparison with other metaheuristic algorithms under the same conditions, namely: the original SCA algorithm [36], genetic algorithms (GA) [37] and the particle swarm optimization algorithm (PSO) [38]. The experiment was performed on a 64-bit Windows 10 platform on a 7th generation Intel Core (TM) i5 processor with 8 GB of RAM using MATLAB 2014b. The results obtained by the different algorithms were analyzed on a set of quantitative measures to evaluate the effectiveness of the proposed method: on the one hand the performance of the ISCA optimization algorithm and on the other hand, the quality of the segmented image. We performed several tests by changing the number of clusters k at each run to verify the results of the experiments. In this section, we will only present the results obtained with k=4 to avoid putting all these different tests. In order to get a fair comparison, the results were calculated after running each algorithm 20 times and the population size for all algorithms is set to 30 and each algorithm runs the same maximum number of iterations which was 200. The parameters for each algorithm are taken from the original reference.

Fig. 2 shows original images selected from the database (BSD300) [37], which are used in the experiments. Four images are used to evaluate the performance of each algorithm in the comparisons. The names of these images are Image1, Image2, Image3 and Image4.

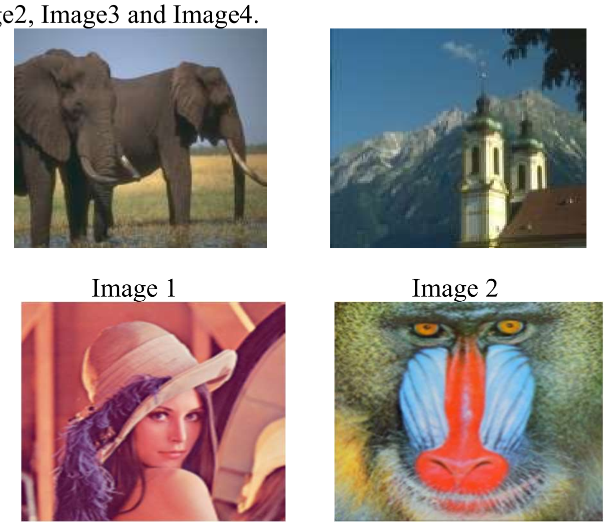
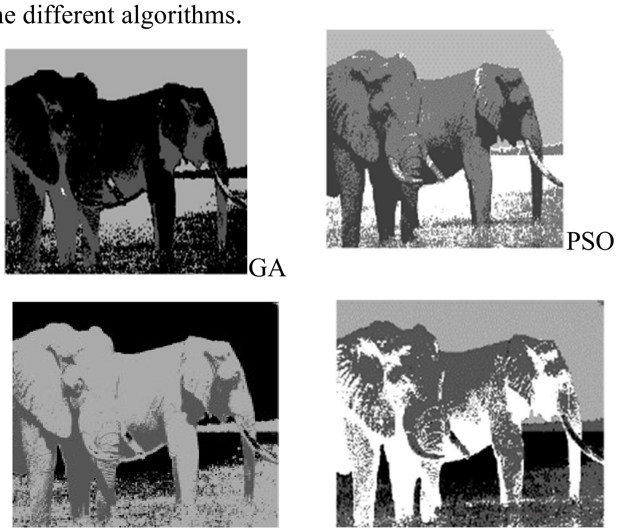


Image 1 Image 2

Image 3 Image 4

Figure 2. The four original images selected in the database BSD300.

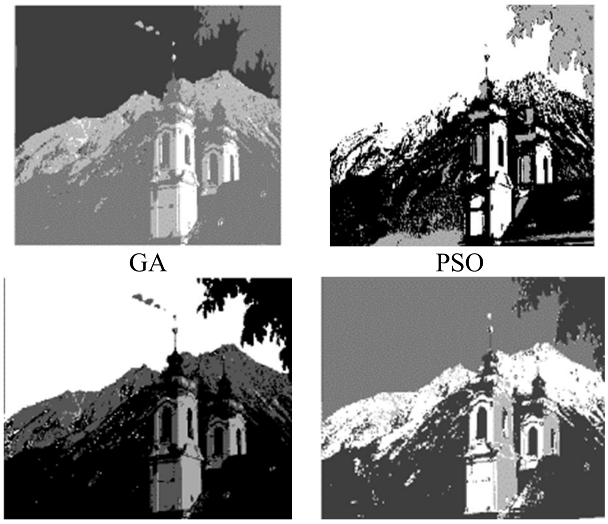
Figs. 3, 4, 5 and 6 show the image segmentation results of the different algorithms.



GA PSO

SCA Proposed

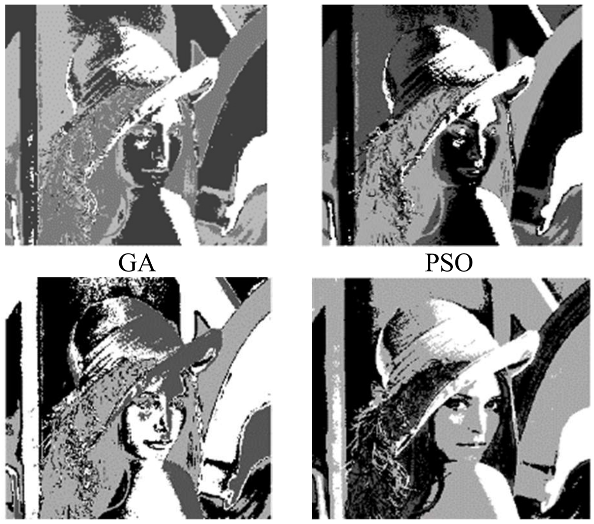
Figure 3. The results of image Image1



GA PSO

SCA Proposed

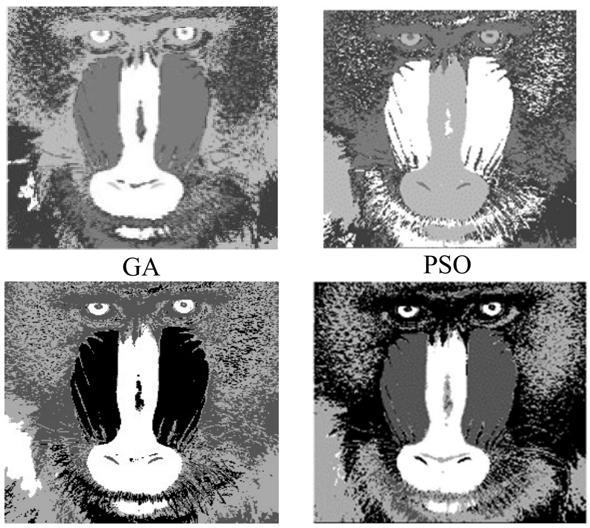
Figure 4. The results of image Image2



GA PSO

SCA Proposed

Figure 5. The results of image Image3



GA PSO

SCA Proposed

Figure 6. The results of image Image4

A. PERFORMANCE MEASURE

In order to evaluate the effectiveness of each algorithm in determining the best clusters to improve the image segmentation process, a set of metrics is used [39]: structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), as well as root mean square error (RMSE), which are given in Equations 8,9,10 respectively.

The SSIM is developed to measure the visual quality of a segmented image, compared to the original image.

$$SSIM(I, I_s) = \frac{(2\mu_I \mu_{I_s} + c_1)(2\sigma_{I,I_s} + c_2)}{({\mu_I}^2 + {\mu_{I_s}}^2 + c_1)({\sigma_I}^2 + {\sigma_{I_s}}^2 + c_2)}, \qquad (8)$$

where:

μ_I and μ_{I_s} represent the mean intensity of the images I and I_s, respectively;

σ_I and σ_{I_s} represent the standard deviation of I and I_s, respectively;

σ_{I,I_s} is the covariance between I and I_s (where the value of two constants are set to $c_1 = 6.50$ and $c_2 = 58.52$)

The BSNR is used to measure the quality of the output image. It is represented by the proportion between the maximum powers that can be achieved and the spurious noise that influences the likeness of the image.

$$BSNR = 20 log_{10}(\frac{255}{RMSE}) \qquad (9)$$

$$RMSE = \sqrt{\frac{\sum_1^M \sum_1^Q [I(i,j) - I_S(i,j))]^2}{M \times Q}} \qquad (10)$$

where I and I_s represent the original image and the segmented image of size $M \times Q$, respectively.

In addition, additional measures are described to evaluate the performance of the proposed method for image segmentation. These metrics quantitatively evaluate the quality of the cluster. The internal quality compares different sets of clusters without any reference to external knowledge. A good clustering method is one with high similarity within the cluster and low similarity between clusters. To show the performance of our clustering approach, we studied the cluster quality of the segmented image using three measures[40] XB, CE and PC, given in formulas (11), (12) and (13) respectively.

The Xie and Beni index (XB) determines the ratio of the total variation of the interior of a cluster to the separation of the clusters:

$$XB = \frac{\sum_{i=1}^{c} \sum_{k=1}^{n} (u_{i,k})^m \|x_k - v_i\|^2}{n \min_{ik} \|x_k - v_i\|^2}. \qquad (11)$$

The classification entropy (CE) calculates the blurring of the cluster partition:

$$CE = -\frac{1}{n} \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik} \log u_{ik}. \qquad (12)$$

The partition coefficient (PC) determines the amount of overlap between clusters:

$$PC = \frac{1}{n} \sum_{i=1}^{c} \sum_{k=1}^{n} (u_{ik})^2 \qquad (13)$$

B. RESULTS AND DISCUSSIONS

The results of the comparison of the proposed ISCA clustering method with the other three algorithms are presented in Tables 1-4 and Figs. 7-9.

Figs. 7, 8 and 9 below, shows the values of RMSE, BSNR and SSIM respectively obtained by the four methods depending on the test images used in the experiment.

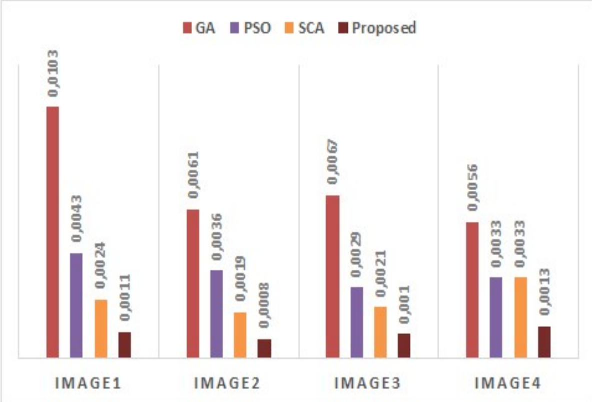


Figure 7. RMSE values obtained by the four algorithms.

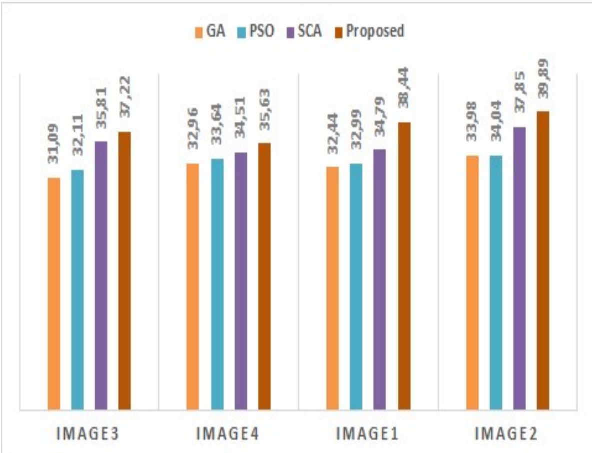


Figure 8. PSNR values obtained by the four algorithms.

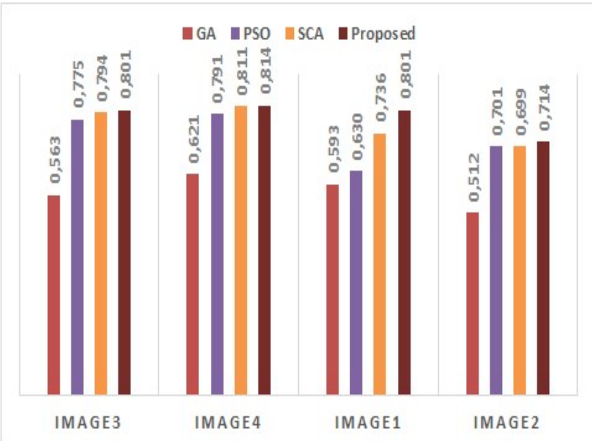


Figure 9. SSIM values obtained by the four algorithms.

To discuss the quality of the segmented image we note that the low value of RMSE means the image is of good quality as well as the low value of PSNR and SSIM means the image is of poor quality. In terms of RMSE, PSNR and SSIM measurements, the results of the comparison are given in Figs. 7-9. Based on these results, we can see that note approach has the best PSNR and SSIM values, followed by SCA and PSO; the worst is GA. Moreover, according to the RMSE measure, the values obtained by our approach are very low compared to other algorithms of comparison clustering. These results indicate that the ISCA algorithm is very efficient and able to escape local optima. So, we can say that our clustering approach gives good results, which shows that the segmented image is of better quality compared to other algorithms.

The metric values of the test images are shown in Tables 1, 2, 3 and 4 below. To facilitate the comparison of XB, CE and PC values for each image, we note that a clustering method is better and performs well if the XB and CE values are low while the PC value is high.

Table 1. Results of the four clustering methods on Image1

	XB	CE	PC
GA	3,7862e+03	0,4933	0,6494
PSO	2,8192e+03	0,4314	0,7764
SCA	2,0143e+03	**0,3920**	0,8244
Proposed	**2,0001e+03**	**0,3920**	**0,8414**

Table 2. Results of the four clustering methods on Image2

	XB	CE	PC
GA	527,5608	0,5391	0,7384
PSO	550,7512	**0,4297**	0,7586
SCA	516,3101	0,4335	0,7745
Proposed	**511,2811**	0,4301	**0,7911**

Table 3. Results of the four clustering methods on Image3

	XB	CE	PC
GA	543,9865	0,4221	0,7643
PSO	2,1982e+03	0,3923	0,7132
SCA	2,0982e+03	0,3819	**0,7710**
Proposed	**423,9804**	**0,3655**	0,7664

Table 4. Results of the four clustering methods on Image4

	XB	CE	PC
GA	1,3154e+03	0,4221	0,7643
PSO	1,3128e+03	0,3923	0,7132
SCA	1,2682e+03	0,3819	0,7710
Proposed	**1,2312e+03**	**0,2971**	**0,7876**

Tables 1-4 show the results of the comparison between the proposed ISCA-based clustering approach and the other techniques mentioned above in terms of cluster quality measures for each image used in the experiment where the bold value in each column indicates the best result obtained among the four algorithms. We found that our approach is approximately better in terms of cluster quality than the other techniques. From these tables, we notice that the values of XB, CE and PC obtained are varied from one image to another, but overall will guide to find good clusters which explains that our approach is slightly better than other methods for all test images.

To validate the performance of the proposed method, we will discuss the results obtained for each image separately. From the results of image1, we can see that the segmented image obtained by our method is better than the others because it is sharper and displays all the areas that introduce the original image. Moreover, according to Table 1, the XB and PC values of our proposed approach are slightly better than the other methods. But in terms of EC, clustering by SCA obtains the same value as the proposed ISCA clustering. From the results of image2 displayed in Table 2, we can see the obvious edges in the image of our method, and these XB, CE values are much smaller than those of other methods. In terms of PC, clustering by SCA achieves the best result. Table 3 shows the results of image3, from these results, the segmentation effect of our approach is quite obvious. Moreover, the PC and XB values of our proposed approach are better than those of other approaches. In terms of EC, PSO clustering obtains the best result. Finally, Table 4 shows the results of image4, these results clearly show that our approach performs better than other clustering versions, the segmented image has sharp edges and significant features and the obtained XB, CE and PC values are better compared to other approaches.

From the previous results, it can be concluded that the proposed ISCA-based clustering algorithm avoids the limitations of the original SCA, PSO and GA algorithms. Moreover, it provides better results in image segmentation, based on performance metrics (SSIM, PSNR and RMSE) than the other three algorithms, and also the quality of clusters is discussed by three parameters (XB, CE and PC), the results obtained validate that our technique is better. In particular, the CE values of the proposed approach are close to zero, which indicates that the cohesion within the clusters is very high.

In summary, the experimental results used that the proposed ISCA-based clustering provides better values in terms of cluster quality measures that allow to obtain good qualities of the segmented image.

VI. CONCLUSIONS

Image segmentation is one of the most important tasks in image analysis and computer vision. Clustering is one of the most popular and powerful methods to perform image segmentation. In this technique, this problem is formalized as a combinatorial optimization problem. Hence the use of metaheuristics. In this paper, we have introduced a new and improved version of the SCA algorithm (named ISCA) to handle the clustering problem. This algorithm is based on a new strategy of moving individuals in the search space. More precisely, the update of the positions of the solutions is performed by taking into account two positions of the solutions in their generations: the best solution plus a second randomly generated position. Thanks to this strategy, the proposed approach ensures a wide exploration of the search process and converges to the optimal solution. The use of this ISCA algorithm has shown several advantages, including significant efficiency and simplicity of its design which is based on the sine and cosine mathematical functions to perform the exploitation and exploration of the search space. Our proposed approach is compared with other methods based on several evaluation criteria and the experimental results on several test images show the performance of the proposed approach. Based on the proposed ISCA-based clustering results, it can be used in other applications including feature selection.

References

[1] M. Antony and J. Anitha, "A survey of moving object segmentation methods," *International Journalof Advanced Research in Electronics and Communication Engineering (IJARECE)*, vol. 1, no. 4, pp. 591-595, 2012.

[2] X. Chen, B. M. Williams, S. R. Vallabhaneni, G. Czanner, R. Williams, and Y. Zheng, "Learning active contour models for medical image segmentation," *Proceedings of the IEEE/CVF conference on computer vision and pattern* recognition, 2019, pp. 11632–11640. https://doi.org/10.1109/CVPR.2019.01190.

[3] H. Moussaoui, M. Benslimane, and N. el Akkad, "A novel brain tumor detection approach based on fuzzy c-means and marker watershed algorithm," *Proceedings of the International Conference on Digital Technologies and Applications*, 2021, pp. 871–879. https://doi.org/10.1007/978-3-030-73882-2_79.

[4] M. Es-Sabry, N. el Akkad, M. Merras, A. Saaidi, and K. Satori, "A new color image encryption algorithm using multiple chaotic maps with the intersecting planes method," *Sci Afr*, p. e01217, 2022. https://doi.org/10.1016/j.sciaf.2022.e01217,

[5] O. Berezsky, O. Pitsun, N. Batryn, K. Berezska, N. Savka, and T. Dolynyuk, "Image segmentation metric-based adaptive method," *Proceedings of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, 2018, pp. 554–557. https://doi.org/10.1109/DSMP.2018.8478579.

[6] A. Javed, Y.-C. Kim, M. C. K. Khoo, S. L. D. Ward, and K. S. Nayak, "Dynamic 3-D MR visualization and detection of upper airway obstruction during sleep using region-growing segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 2, pp. 431–437, 2015. https://doi.org/10.1109/TBME.2015.2462750.

[7] Z. Faska, L. Khrissi, K. Haddouch, and N. el Akkad, "A powerful and efficient method of image segmentation based on random forest algorithm," *Digital Technologies and Applications*, 2021, pp. 893–903. https://doi.org/10.1007/978-3-030-73882-2_81.

[8] M. Gong, H. Li, X. Zhang, Q. Zhao, and B. Wang, "Nonparametric statistical active contour based on inclusion degree of fuzzy sets," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 5, pp. 1176–1192, 2015. https://doi.org/10.1109/TFUZZ.2015.2505328.

[9] L. Khrissi, N. E. Akkad, H. Satori, and K. Satori, "Color image segmentation based on hybridization between Canny and k-means," *Proceedings of the 2019 7th Mediterranean Congress of Telecommunications (CMT)*, 2019, pp. 1-4, https://doi.org/10.1109/CMT.2019.8931358.

[10] L. Khrissi, H. Satori, K. Satori, and N. el Akkad, "An efficient image clustering technique based on fuzzy c-means and cuckoo search algorithm," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 423–432, 2021, https://doi.org/10.14569/IJACSA.2021.0120647.

[11] D. Mahapatra, "Semi-supervised learning and graph cuts for consensus based medical image segmentation," *Pattern Recognition*, vol. 63, pp. 700–709, 2017. https://doi.org/10.1016/j.patcog.2016.09.030.

[12] D. Kaushik, U. Singh, P. Singhal, and V. Singh, "Medical image segmentation using genetic algorithm," *International Journal of Computer Applications*, vol. 81, no. 18, pp. 10–15, 2013. https://doi.org/10.5120/14222-2220.

[13] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using K-means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, 2015. https://doi.org/10.1016/j.procs.2015.06.090.

[14] P. Contreras and F. Murtagh, "Fast hierarchical clustering from the Baire distance," in *Classification as a Tool for Research*, Springer, 2010, pp. 235–243. https://doi.org/10.1007/978-3-642-10745-0_25.

[15] J. Kaur, S. Agrawal, and R. Vig, "A methodology for the performance analysis of cluster-based image segmentation," *International Journal of Engineering Research and Applications*, vol. 2, no. 2, pp. 664–667, 2012.

[16] J. Lee and D. Perkins, "A simulated annealing algorithm with a dual perturbation method for clustering," *Pattern Recognition*, vol. 112, p. 107713, 2021. https://doi.org/10.1016/j.patcog.2020.107713.

[17] B. Khorram and M. Yazdi, "A new optimized thresholding method using ant colony algorithm for MR brain image segmentation," *J Digit Imaging*, vol. 32, no. 1, pp. 162–174, 2019. https://doi.org/10.1007/s10278-018-0111-x.

[18] T. Y. Tan, L. Zhang, C. P. Lim, B. Fielding, Y. Yu, and E. Anderson, "Evolving ensemble models for image segmentation using enhanced particle swarm optimization," *IEEE Access*, vol. 7, pp. 34004–34019, 2019. https://doi.org/10.1109/ACCESS.2019.2903015.

[19] L. Khrissi, N. el Akkad, H. Satori, and K. Satori, "Image segmentation based on K-means and genetic algorithms," *Advances in Intelligent Systems and Computing*, 2020, vol. 1076, pp. 489–497. https://doi.org/10.1007/978-981-15-0947-6_46.

[20] X. Yue and H. Zhang, "Improved hybrid bat algorithm with invasive weed and its application in image segmentation," *Arabian Journal for Science and Engineering*, vol. 44, no. 11, pp. 9221–9234, 2019. https://doi.org/10.1007/s13369-019-03874-y.

[21] S. Karthikeyan and T. Christopher, "A hybrid clustering approach using artificial bee colony (ABC) and particle swarm optimization," *International Journal of Computer Applications*, vol. 100, no. 15, pp. 1-6, 2014. https://doi.org/10.5120/17598-8057.

[22] S. Mirjalili, "SCA: a sine cosine algorithm for solving optimization problems," *Knowl Based Syst*, vol. 96, pp. 120–133, 2016. https://doi.org/10.1016/j.knosys.2015.12.022.

[23] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*, Prentice-Hall, Inc., 1988.

[24] V. Singh, "Sunflower leaf diseases detection using image segmentation based on particle swarm optimization," *Artificial Intelligence in Agriculture*, vol. 3, pp. 62–68, 2019. https://doi.org/10.1016/j.aiia.2019.09.002.

[25] S. Tongbram, B. A. Shimray, L. S. Singh, and N. Dhanachandra, "A novel image segmentation approach using fcm and whale optimization algorithm," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–15, 2021. https://doi.org/10.1007/s12652-020-02762-w.

[26] H. Jia, C. Lang, D. Oliva, W. Song, and X. Peng, "Hybrid grasshopper optimization algorithm and differential evolution for multilevel satellite image segmentation," *Remote Sensing*, vol. 11, no. 9, pp. 1-26, 2019, https://doi.org/10.3390/rs11091134.

[27] M. E. Farmer and D. Shugars, "Application of Genetic Algorithms for Wrapper-based Image Segmentation and Classification," *Proceedings of the 2006 IEEE International Conference on Evolutionary Computation*, 2006, pp. 1300–1307. https://doi.org/10.1109/CEC.2006.1688459..

[28] S. R. Biswal and G. Shankar, "Optimal sizing and allocation of capacitors in radial distribution system using sine cosine algorithm," *Proceedings of the 2018 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, 2018, pp. 1–4. https://doi.org/10.1109/PEDES.2018.8707739.

[29] B. S. Yıldız and A. R. Yıldız, "Comparison of grey wolf, whale, water cycle, ant lion and sine-cosine algorithms for the optimization of a vehicle engine connecting rod," *Materials Testing*, vol. 60, no. 3, pp. 311–315, 2018. https://doi.org/10.3139/120.111153.

[30] A. Fernández, A. Peña, M. Valenzuela, and H. Pinto, "A binary percentile sin-cosine optimisation algorithm applied to the set covering problem," *Proceedings of the Computational Methods in Systems and Software*, 2018, pp. 285–295. https://doi.org/10.1007/978-3-030-00211-4_25.

[31] T. Tuncer, "Ldw-scsa: Logistic dynamic weight based sine cosine search algorithm for numerical functions optimization," arXiv preprint arXiv:1809.03055, 2018. https://doi.org/10.21533/pen.v6i2.159.

[32] X. Liang, Z. Cai, M. Wang, X. Zhao, H. Chen, and C. Li, "Chaotic oppositional sine–cosine method for solving global optimization problems," *Engineering with Computers*, pp. 1–17, 2020. https://doi.org/10.1007/s00366-020-01083-y.

[33] K. Z. Zamli, F. Din, A. B. Nasser, and A. Alsewari, "Combinatorial test suite generation strategy using enhanced sine cosine algorithm," *Proceedings of the ECCE2019*, Springer, 2020, pp. 127–137. https://doi.org/10.1007/978-981-15-2317-5_12.

[34] S. Abdel-Fatah, M. Ebeed, and S. Kamel, "Optimal reactive power dispatch using modified sine cosine algorithm," *Proceedings of the 2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*, 2019, pp. 510–514. https://doi.org/10.1109/ITCE.2019.8646460.

[35] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*, Elsevier, 2011.

[36] M. Abd Elaziz, N. Nabil, A. A. Ewees, and S. Lu, "Automatic data clustering based on hybrid atom search optimization and sine-cosine algorithm," *Proceedings of the 2019 IEEE congress on evolutionary computation (CEC)*, 2019, pp. 2315–2322. https://doi.org/10.1109/CEC.2019.8790361.

[37] V. Jaiswal, V. Sharma, and S. Varma, "An implementation of novel genetic based clustering algorithm for color image segmentation," *Telkomnika*, vol. 17, no. 2, pp. 1461–1467, 2019. https://doi.org/10.12928/telkomnika.v17i3.10072.

[38] M. Omran, A. P. Engelbrecht, and A. Salman, "Particle swarm optimization method for image clustering," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 19, no. 03, pp. 297–321, 2005. https://doi.org/10.1142/S0218001405004083.

[39] Z. Liu and R. Laganière, "Phase congruence measurement for image similarity assessment," *Pattern Recognition Letters*, vol. 28, no. 1, pp. 166–172, 2007. https://doi.org/10.1016/j.patrec.2006.06.019.

[40] B. Balasko, J. Abonyi, and B. Feil, *Fuzzy Clustering and Data Analysis Toolbox*, Department of Process Engineering, University of Veszprem, Veszprem, 2005.

***LAHBIB KHRISSI** received the bachelor's and master's degrees from SMBA-Fez University in 2000 and 2007 respectively. He is currently working toward the PhD degree in the LISAC Laboratory at faculty of Sciences, Sidi Mohammed Ben Abdellah University, Fez, Morocco. His current research interests include Data mining, Image processing and Artificial intelligence.*

***NABIL EL AKKAD** received the PhD degree from SMBA-Fez University in 2014. He is currently a Professor of Computer Science at National School of Applied Sciences (ENSA) of Fez, Sidi Mohammed Ben Abdellah University, Fez, Morocco. He is a member of the LISA and LIIAN Laboratories.*

His research interests include Artificial intelligence, Image processing, Data mining, Camera self calibration, 3D reconstruction, Machine learning, Real-time rendering and Cryptography.

***HASSAN SATORI** is an Associate Professor, Department of Computer Science, Faculty of Sciences, Dhar El Mahraz, Sidi Mohammed Ben Abdellah University -Fez- Morocco. He is Ph.D in Speech Recognition and Signal Processing in 2009. He received a MSc in Physics from the Mohamed Premier University, and a Ph.D in Nanotechnology and computer simulation, also from the same University, in 2001. From 2001-2002 he was a postdoctoral fellow in Physics, mathematical modeling and computer simulation at the Chemnitz University of Technology, Germany, and was then a postdoctoral fellow in computer simulation at Department of Interface Chemistry and Surface Engineering, Max-Planck-Institut, Düsseldorf, Germany, (2003-2005). Dr. Hassan SATORI has large academic teaching and research experience in the fields.*

***KHALID SATORI** received the PhD degree from the National Institute for the Applied Sciences INSA at Lyon in 1993. He is currently a Professor of Computer Science at SMBA-Fez University. He is the director of the LIIAN Laboratory.*

His research interests include real-time rendering, Image-based rendering, virtual reality, biomedical signal, camera self calibration, genetic algorithms and 3D reconstruction.