

Automatic Data Clustering Based Mean Best Artificial Bee Colony Algorithm

Ayat Alrosan¹, Waleed Alomoush², Mohammed Alswaitti^{3,*}, Khalid Alissa⁴, Shahnorbanun Sahran⁵, Sharif Naser Makhadmeh⁶ and Kamal Alieyan⁷

¹Deanship of Information and Communication Technology, Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia

²Computer Department, Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia

³School of Electrical and Computer Engineering, Department of Information and Communication Technology, Xiamen University Malaysia, Sepang, 43900, Malaysia

⁴Department of Computer Science, College of Computer Science and Information Technology, Imam Abdulrahman bin Faisal University, Dammam, Saudi Arabia

⁵Center for Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia

⁶School of Computer Sciences, Universiti Sains Malaysia, 11800, Penang, Malaysia

⁷Faculty of Computer Sciences and Informatics, Amman Arab University, Amman, Jordan

*Corresponding Author: Mohammed Alswaitti. Email: alswaitti.mohammed@xmu.edu.my

Received: 14 December 2020; Accepted: 28 January 2021

Abstract: Fuzzy C-means (FCM) is a clustering method that falls under unsupervised machine learning. The main issues plaguing this clustering algorithm are the number of the unknown clusters within a particular dataset and initialization sensitivity of cluster centres. Artificial Bee Colony (ABC) is a type of swarm algorithm that strives to improve the members' solution quality as an iterative process with the utilization of particular kinds of randomness. However, ABC has some weaknesses, such as balancing exploration and exploitation. To improve the exploration process within the ABC algorithm, the mean artificial bee colony (MeanABC) by its modified search equation that depends on solutions of mean previous and global best is used. Furthermore, to solve the main issues of FCM, Automatic clustering algorithm was proposed based on the mean artificial bee colony called (AC-MeanABC). It uses the MeanABC capability of balancing between exploration and exploitation and its capacity to explore the positive and negative directions in search space to find the best value of clusters number and centroids value. A few benchmark datasets and a set of natural images were used to evaluate the effectiveness of AC-MeanABC. The experimental findings are encouraging and indicate considerable improvements compared to other state-of-the-art approaches in the same domain.

Keywords: Artificial bee colony; automatic clustering; natural images; validity index; number of clusters



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Clustering of data is a statistical approach used for managing large data volume. It is a multivariate analytical approach, which identifies patterns and relationships that exist amongst data. By data clustering, a user could divide data into relatively homogeneous groups. By reorganizing these groups, the user may be able to utilize the original data volume efficiently. Clustering accuracy is crucial because, according to [1–3], clustered data that does not accurately denote the original data will have adverse consequences. In other words, data clustering is a set of patterns or points represented as vectors of attributes, measurement methods, or objects that exist in multidimensional feature space [4,5].

Clustering algorithms consist of two types, namely, the partitional clustering, which generates several partitions, and hierarchical clustering, which generates only one partition [4,5]. As highlighted by [5,6], there are two types of partitional clustering, namely, hard (crisp) and fuzzy (soft). Every data point in crisp clustering belongs to one cluster only. In the latter clustering of fuzzy, at the same time, data points might relate to much more than single cluster based on a certain degree of fuzzy membership. Between these two types (crisp and fuzzy), fuzzy clustering appears to be more acceptable for datasets that exhibit undefined bounds between regions or clusters.

Over the last few years, clustering methods have been demonstrated to be effective, particularly in tasks of categorization, requiring semi or full automation [7,8]. Clustering is employed to classify identical data points according to similarity and distance, based on the standard approach. In inter-cluster, the similarity increases as the distances decrease while in intra-cluster, simultaneously increasing distances will reduce similarity.

Clustering has many useful features, making it one of the most well-known in many fields of image segmentation, pattern recognition, machine learning, data mining, etc. [3,9]. One of its features is the partitional method of clustering [2]. It is free from issues such as static-behaviour, where the elements of data of a cluster are unmovable to another cluster. Additionally, this type of clustering also does not have the problem of overlapping clusters' separation inability, which is common in hierarchical clustering. Amongst these fuzzy clustering techniques, FCM algorithm [10] is a useful approach that has been applied in many domains because it has robust ambiguity characteristics and can maintain rather more information than hard segmentation techniques [11,12].

However, the Fuzzy clustering-based approach still has significant weaknesses; for instance, its ability to obtain an automatic approach without prior knowledge of the number of clusters and centroid locations. Furthermore, determining the number of clusters within a particular dataset is a significant challenge and the unavailability of experts, operators, and any prior knowledge has contributed to this challenge as well. Accordingly, many researchers have worked on the successful implementation of clustering methods in the last few years to find the number of the appropriate clusters within a particular dataset without experts and operators.

Metaheuristic optimization search algorithms such as artificial bee colony (ABC) [13], bees algorithm (BA) [14,15], harmony search (HS) [16,17], firefly algorithm (FA) [18,19], and cuckoo search (CS) [20,21], have shown success in many fields such as clustering [11,22], image processing [23], scheduling [24] and others real-world applications [25]. Moreover, a fuzzy clustering approach based on metaheuristic optimization search algorithms is considered a suitable choice for solving the problems related to fuzzy clustering [3,26]. The following is a discussion of some of the methods.

A fuzzy clustering FCM was proposed in Ouadfel et al. [27] based on a modified ABC called MoABC-FCM. Additionally, Hancer et al. [28] presented an image segmentation approach using ABC algorithm to detect tumours from MRI images. In Alrosan et al. [29], ABC algorithms and fuzzy c-means were combined and called ABC-FCM. The combination used two types of MRI brain images namely simulated brain data, and actual MRI images, for its performance. Automatic Clustering-based Improved Binary ABC (IDisABC) which was proposed by [30]. In Kuo et al. [3] an automatic customer segmentation method based on an improved artificial bee colony and fuzzy clustering was proposed. Further, Kuo et al. [3], Alrosan et al. [23] proposed a new image clustering method based on improved ABC called (MeanABC-FCM). This method used MRI images, and the outcomes show promising results in this field.

A method, fuzzy automatic clustering named AFDE for image segmentation problem was proposed by [31]. Relevantly, Alia et al. [32] proposed a novel automatic algorithm of fuzzy clustering (DC) by the hybrid harmony search HS with FCM to produce an automatic segment method called DCHS. Further, an automatic hard clustering algorithm called DCPSO was proposed by [33–35]. While an algorithm called Fuzzy based genetic algorithm with variable string length FVGA to automatically find a fitting clusters number with the matching fuzzy clustering outcomes was proposed by [36]. A fuzzy clustering algorithm was proposed in Alomoush et al. [37] using a combination of a firefly optimizer with an FCM algorithm named FFCM [38]. In Agbaje et al. [26] an automatic data clustering method was proposed based on improved FA algorithm combined with the Particle Swarm Optimization (PSO) algorithm.

This paper proposes an automatic fuzzy clustering approach based on mean artificial bee colony called (AC-MeanABC). The proposed method uses the (MeanABC) algorithm to determine the appreciate cluster number and initial location of centroids. The method modifies the search equation depending on mean previous and global best solutions to reach the best balance between exploration and exploitation. The remaining portions of the paper are as follows. The type 2 fuzzy set is described in Section 2, related background on the employed techniques are presented in Sections 4 and 5. The proposed method is presented in Section 5, and Section 6 presents the experiments and results, wherein Section 7 concludes the paper.

2 Type 2 Fuzzy Set

FCM is an unsupervised method that can identically partite objectives based on the similarity attribute, which tends to increase the similarities of entities within the same set and decrease the similarities of entities within different sets [2,39]. The objective function defined by FCM is reformulated using type 2 fuzzy set as in Eq. (1) below.

$$J(c_1, c_2, \dots, c_n) = \sum_{k=1}^c \sum_{j=1}^n h_{kj}^2 \|x_{kj} - v_{kj}\| \quad (1)$$

$\{v_j\}_{j=1}^c$. The location of centroids c refers to the standard in the sign of product $\|\cdot\|$ for the element x_i to the j th centre's location; furthermore, the new value of membership h_{kj} , and type 2 fuzzy set includes five steps whereby the first step involves the selection of fuzzy clusters number c , followed by the second step which involves the selection of location centres v_1, v_2, \dots, v_c . The third step is the computation of elements of fuzzy membership value as in Eq. (2) where μ_{kj} is the membership degree of pixel k to the j th cluster, while the fourth step involves calculating the

cluster centres as in Eq. (3). As the final step, the aforementioned four steps are repeated until the limit of iterations (t) is reached or until a stopping condition is met as in Eq. (4).

$$h_{kj} = \mu_{kj} - \frac{1 - \mu_{kj}}{2} \quad (2)$$

$$V_{kj} = \frac{\sum_{j=1}^n h_{kj}^m \cdot x_{kj}}{\sum_{j=1}^n h_{kj}^m} \quad (3)$$

$$\|v_{new} - v_{old}\| < \varepsilon \text{ where } \varepsilon < 0.001 \quad (4)$$

3 The Artificial Bee Colony (ABC)

A standard ABC algorithm was proposed in Karaboga et al. [13] and this algorithm was inspired by the life of bee swarm. The bees in ABC are divided into three kinds: the employed bees focus on seeking sources of food and transmitting their search information to the onlookers. Meanwhile, the onlooker bees retain more searches to achieve better solutions depending on the quality of the discovered food sources. A source of food in the flock is a possible solution, and the amount of nectar determines their fitness values. The ABC algorithm randomly initializes the source of food as in Eq. (5).

$$x_{i,j} = x_{min,j} + rand(0, 1)(x_{max,j} - x_{min,j}) \quad (5)$$

where ($i = 1, \dots, SN$) and SN are the solutions ($j = 1, 2, \dots, D$), D is the parameters value, while $x_j \text{ min}$ and $x_j \text{ max}$ are the highest and lowest bound of the dimension j , respectively. The computation of the updated source of food is as in Eq. (6).

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) \quad (6)$$

where: $K \in (1, 2, \dots, SN)$ and $j \in (1, 2, \dots, D)$ indexes which are randomly chosen and $K \neq i$. ϕ_{ij} is a value from $[-1, 1]$. The balancing of the candidate food exploration and exploitation fitness of food source (solutions) $fit(x_i)$ will be found from the value of fitness function $f(x_i)$ that is computed by Eq. (7).

$$fit(x_i) = \begin{cases} \frac{1}{(1 + f(x_i))} & iF f(x_i) \geq 0 \\ 1 + abs(f(x_i)) & iF f(x_i) < 0 \end{cases} \quad (7)$$

That onlooker selects solutions with certainty. This certainty means that correlation exists between the fitness value of a source of food and the bees employed. The probability of fitness is as in Eq. (8).

$$p^i = \frac{fit_i}{\sum_{n=1}^{sn} fit_i} \quad (8)$$

In this regard, when several trials of the solution are not changing, the “limit,” is reached, and employee bees are converted to scout bees to abandon their solutions. Accordingly, Eq. (5) shows how the scout bees start new searches and solutions randomly until they reached the criterion of termination.

4 The Mean Artificial Bee Colony (MeanABC)

The ABC has some disadvantages in terms of unbalanced search behaviour, and this has become a common limitation with other optimization algorithms. The authors used an improved ABC based mean global best (MeanABC) [23] to solve these disadvantages. In Alrosan et al. [23] an improved ABC algorithm called MeanABC was proposed, and this algorithm presents a unique modified search formula for the bee phase employed based on the knowledge about the best candidates to produce a new neighbourhood source of food as in Eq. (9)

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) + \Psi_{ij} \left[\left(\frac{pbest_i + gbest}{2} - x_{ij} \right) + \left(\frac{pbest_i - gbest}{2} - x_{ij} \right) \right] \quad (9)$$

where $x_{i,j}$ would be the first expression for the present position, $\phi_{ij}(x_{ij} - x_{kj})$ is the variation among the current position and the new position, and ϕ_{ij} is a random value from $[-1,1]$. Both expressions are identical to traditional ABC algorithms. The third expression is the modified search equation in MeanABC. Here, the first side of the third expression is essential for switching the bee's present position towards the mean value of the positive way of the best global position and the positive way that is its own best position,

$$\left(\frac{pbest_i + gbest}{2} - x_{ij} \right).$$

The second concerns with switching the present mean position value of ($pbest$), which is the positive way of own best position and the ($-gbest$) negative way of the best global position.

$$\left(\frac{pbest_i - gbest}{2} - x_{ij} \right).$$

Based on the above: ψ is a random value between $[0, C]$, and C is a positive number. The constant value of C plays a critical role in balancing the candidate food exploration and exploitation [38]. Fig. 1 shows the MeanABC pseudo-code.

```

1: Initialize the population of solution  $x_i$ 
   Initialize the control parameters: colony size  $SN$ ,
   Maximum Cycle Number ( $MCN$ ),  $limit=L$ 
   Initialize the nonnegative constant number parameter  $C$ 
   Initialize the positions of the food sources by Eq5.
2: Evaluate the fitness of the population.
3: Cycle=1
4: Repeat
5: FOR each employee bee {
   Generate new solutions  $v_{ij}$  in the neighbourhood of  $x_{ij}$  by the Eq9.
   
$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) + \Psi_{ij} \left[ \left( \frac{pbest_i + gbest}{2} - x_{ij} \right) + \left( \frac{pbest_i - gbest}{2} - x_{ij} \right) \right]$$

   Calculate the fitness value  $fit_i$  for the new solution by Eq7.
   A greedy selection process is applied between  $v_{ij}$  and  $x_{ij}$ ;
   the value of  $x_i$  and  $v_i$  is selected as the new  $x_i$ .}
6: Calculate the probability value  $p_i$  for the solutions by Eq8.
7: FOR each onlooker bee {
   Select a solution depending on  $p_i$ .
   Generate new solution  $v_{ij}$  for the onlooker bee from  $x_i$  solution
   by Eq9.
   Calculate the fitness value  $fit_i$  for the new solution by using
   Eq7.
   Apply a greedy selection process between  $v_{ij}$  and  $x_{ij}$ ;
   the value of  $x_i$  and  $v_i$  is selected as the new  $x_i$ .}
8: IF a food source cannot be improved under a predefined number of
   generations (the food source is abandoned).
9: The scouts' phase:
   For an abandoned food source  $x_i$ , the scout bee
   randomly generates a new solution to replace it by Eq9.
10: Memorize the best food position (solution) achieved so far.
11: Cycle=Cycle+1
12: Until Cycle=Maximum Cycle Number ( $MCN$ ).

```

Figure 1: The pseudo-code of MeanABC algorithm

5 The Proposed Method (AC-MeanABC)

AC-MeanABC is an unsupervised data clustering method that employs the ability of Mean-ABC to seek the best solution that comprises a set number of clusters and centroids' location. The pseudo-code for AC-MeanABC algorithm is available in Fig. 2. The steps of AC-MeanABC are described as follows.

1:	Initialize the population size N_p , $ClusMaxNum$, $ClusMinNum$, Maximum Cycle Number (MCN), $limit=L$, the real number m . Initialize the nonnegative constant number parameter C . The number of clusters $ClusNo$ is initialized by Eq10.
2:	Calculate the membership matrix h_{kj} using Eq2.
3:	Update the centres' values by Eq3.
4:	Compute dissimilarity between centroids by objective function J_i as in Eq 1 :
	Cycle=1
	Repeat
5:	FOR each employee bee { Generate new solutions v_{ij} in the neighbourhood of x_{ij} by the Eq9. Calculate the membership matrix h_{kj} using Eq2. Update the centres' values by Eq3. Calculate the fitness value fit_i for the new solution by Eq7. IF $fit_i < fit_j$; THEN , update the food source vector encodes, ELSE keep it as it is }
6:	Calculate the probability value p_i for the solutions by Eq8.
7:	FOR each onlooker bee { Select a solution (vector encodes) depending on p_i . Generate new solution v_{ij} for the onlooker bee from x_i solution by Eq9. Calculate the membership matrix h_{kj} using Eq2. Update the centres' values by Eq3. Calculate the fitness value fit_i for the new solution by using Eq7. Apply a greedy selection process between v_{ij} and x_{ij} , the value of x_i and v_i is selected as the new x_i .}
8:	If a food source cannot be improved under a predefined number of generations, the food source is abandoned.
9:	The scouts' phase: For an abandoned food source x_i , the scout bee Randomly generates new solutions to replace it by Eq.10.
10:	Memorize the best food position (solution) achieved so far.
11:	Cycle=Cycle+1
12:	Until Cycle=Maximum Cycle Number (MCN).
13:	Do the Clustering by setting the optimal cluster centres for FCM.

Figure 2: The pseudo-code of AC-MeanABC algorithm

5.1 Initialization

The AC-MeanABC algorithm begins by initializing the N_p food sources. Each food source N_p represents a vector that encodes the cluster number and centroid. Each N_p food sources' vector length can change based on the randomly generated number of clusters for each vector. Also, the N_p size range (number of clusters) depends on the dataset. Each N_p vector encodes several clusters, denoted by ' $ClusNo$,' which is the number of N_p vector encode between the maximum number of clusters ($ClusMaxNum$) and the minimum number of clusters ($ClusMinNum$). The

food source initialization initializes the number of clusters and centroids value in MeanABC as in Eq. (10).

$$ClustNo = ClustMinNum + ClusMaxNum + rand() \cdot (ClusMaxNum - ClustMinNum) \quad (10)$$

The values of $ClusMaxNum$ and $ClustMinNum$ are chosen based on the images used for segmentation. In this regard, the length of the vector is allowed to change if $ClustNo$ is changed. Additionally, each length of a vector in Np must reach to highest clusters number ($ClusMaxNum$). Thus, for a matrix representation, it consists of d -dimensional space and the Np vector length ($ClusMaxNum, d$), where d represents the centroid locations. In case the vector number of clusters is less than $ClusMaxNum$, the vector is taken by centroids in random locations, while the rest of unused vector elements are represented as don't care '#' sign [32,40]. To clarify the 'don't care' elements, let $d = 2$ and $ClusMaxNum = 6$, which means that the feature space is 2-dimensional ($d = 2$) and the maximum number of clusters is $ClusMaxNum = 6$. So, suppose that one of the Np vectors has only 3-candidate cluster centres such as $Np = \{(113, 9), (80, 4), (42, 5)\}$, then, these three centres will be set in the Np vector while the rest of the elements are labelled as don't care '#' sign $Np = \{113 \ 9 \ \# \ \# \ \# \ \# \ 80 \ 4 \ \# \ \# \ 425\}$. Also, to determine the best solutions, the fitness function $J(c_1, c_2, \dots, c_n)$ in Eq. (1) is used. The maximum value of $J(c_1, c_2, \dots, c_n)$ is considered as the best solution.

5.2 Employed Bees Phase

This phase of AC-MeanABC clustering algorithm includes finding the best Np vector with the optimal number of clusters and centroid locations. To update the Np vector value, the new food source position is calculated using previously mentioned Eq. (9) in section (4).

$$v_{ij} = x_{ij} + \phi_{ij} (x_{ij} - x_{kj}) + \Psi_{ij} \left[\left(\frac{pbest_i + gbest}{2} - x_{ij} \right) + \left(\frac{pbest_i - gbest}{2} - x_{ij} \right) \right] \quad (9)$$

Here, $i = [1, 2, \dots, Np]$, $j = [1, 2, \dots, Cn]$, $x_{j,i}$ is a randomly chosen j th parameter of the k th individual, and i is one of the Np food sources with $i \neq k$. If any parameter of v_i exceeds its predetermined boundaries, it should be adjusted to fit the appropriate range. A scale factor ϕ_{ij} is a random number between $[-1, 1]$. Meanwhile, ψ is a uniform random number from $[0, C]$, where C is a nonnegative constant number. Here, the $pbest$ represents the previous best value of each vector encodes namely $NP = (C_1, C_2, C_3, C_4, \dots, C_n)$, and the $gbest$ represents the previous best value of all vector encodes $i = [1, 2, \dots, Np]$. After generating the new candidate vector encodes, the new candidate solution's fitness value is calculated as in Eq. (7), say, fit_j . If $fit_i < fit_j$; then, update the food source vector encodes, else, keep it as it is.

5.3 Onlooker Bees Phase

In the onlooker bee stage, each onlooker selects a source of food with a probability determined by the amount of nectar (fit) from a source of food exchanged by employed bees. As in Eq. (8), the probability is determined. When the probability for fit_i of the vector encodes is high, the food source is selected. Else, the vector will be considered as don't care '#' solution. After the employed bee phase and a food source have been chosen, a neighbour source v_i is calculated by Eq. (9), and its $f(x)$ is calculated. Then, the process of greedy is used between v_i and x_i .

5.4 Scout Bees Phase

When trials of solutions are not changing, and it reached to “limit,” employee bees become scout bees and abandon their solutions. Here, is where, the scout bees begin new searches and solutions, randomly via Eq. (10). These aspects are iterated until the criterion of termination has been reached.

5.5 Solutions Evaluation and Clustering

The degree of appropriateness or goodness is measured by its fitness value of each Mean-ABC solution. Each data is allocated to one or more clusters by using the fuzzy membership categories. These fuzzy membership values are determined using the fuzzy membership equation. Consequently, a cluster validity index is used as an indicator for the appropriateness of the clustering. This index is typically exercised to establish the quality of different solutions obtained using different settings in a particular clustering algorithm (or for solutions given by different algorithms).

In this paper, the cluster validity index of VI is used for the appropriateness of the clustering [35]. The VI index is presented as follows:

$$VI = (s \times N(\mu, \sigma) + 1) \times \frac{\text{intra}}{\text{inter}} \quad (11)$$

where s is a constant number, and $N(\mu, \sigma)$ is a Gaussian distribution as in Eq. (12) with mean and standard deviation, where c_n is the number of clusters

$$N(\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{\left[-\frac{(c_n - \mu)^2}{2\sigma^2}\right]} \quad (12)$$

Eqs. (13) and (14) represent these intra clusters and the inter-cluster.

$$\text{intra} = \frac{1}{N_p} \sum_{c_n=1}^{c_n} \sum_{z \in c_n} \|z_p - m_k\|^2 \quad (13)$$

$$\text{inter} = \min \left\{ \|m_k - m_{kk}\|^2 \right\} \text{ where } k \neq kk \quad (14)$$

where N_p is the number of patterns in a dataset, Z_p is a pattern in the cluster, while m_k , and c_n is the k th centroids of the cluster and the number of clusters.

6 Experiments and Results

The AC-MeanABC fuzzy clustering algorithm is conducted as a fully automatic data clustering to find the number of clusters in datasets. The experiments and results are represented in three parts. Firstly, the AC-MeanABC parameters setting and the most appropriate values are chosen. Secondly, AC-MeanABC is conducted on 11 benchmark clustering datasets selected from the UCI databases such as Iris, Ecoli, Wisconsin, Wine, Dermatology, Glass, Aggregation, R15, D31, Libras movement and Wholesaler customers. Some details about these datasets are given in Tab. 1.

The third part includes AC-MeanABC being conducted with five natural images obtained from the Berkeley1 dataset [30,41] such as Lena, Jet, MorroBay, Mandril, and Pepper. All-natural images have the size is 512×512 except the Pepper image, which is of 256×256 .

Table 1: Clustering datasets details

Datasets	Number of instances	Number of attributes	Number of clusters	Datasets	Number of instances	Number of attributes	Number of clusters
Iris	150	4	3	Aggregation	788	2	7
Ecoli	327	8	5	R15	600	2	15
Wisconsin	683	9	2	D31	3100	2	31
Wine	178	13	3	Libras movement	360	28	15
Dermatology	358	34	6	Wholesaler customers	440	8	3/4/5
Glass	214	9	6				

6.1 AC-MeanABC Parameters Setting

To obtain the best outcomes from any clustering optimization algorithm, the suitable selection of parameters is critical because these parameters are essential in the algorithm's performance and accuracy. To determine the values of parameters of AC-MeanABC (i.e., Population size SN, limit, maximum cycle number (MCN), nonnegative constant parameter (C), and Termination criterion TC) have to be determined. The values of these parameters were set in [Tab. 2](#) based on related works [[3,30](#)]. In [Tab. 3](#), other relevant parameters to the dataset, such as the maximum number of clusters N_p , μ and σ [[3,30](#)].

Table 2: AC-MeanABC parameters settings

Approach	Population size SN (Np)	Limit	Maximum cycle number (MCN)	Nonnegative constant parameter C	Termination criterion TC
Benchmark data clustering (part1)	20	50	500	1.43	2
Benchmark data clustering (part2)	30	100	9000	1.43	2
Image clustering	50	100	10000	1.43	2

6.2 Benchmark Clustering Datasets

In this experiment, AC-MeanABC was performed with 11 benchmark clustering datasets, which were chosen from the UCI database. The benchmark datasets were divided into two parts, whereby part1 of the benchmark datasets included Glass, Aggregation, R15, D31, Libras movement and wholesaler customers, while part2 of the benchmark datasets included Iris, Ecoli, Wisconsin, Wine and Dermatology.

6.2.1 Part1

In this part, the AC-MeanABC outcomes were compared with the standard ABC algorithm and other related works such as iABC, AKCBCO, AKC-MEPSO, DCPG, DCPSO and DCGA [[3,42,43](#)]. The parameter settings for AC-MeanABC algorithm were selected as in [Tabs. 2](#) and [3](#) based on the same parameter setting in [[3](#)]. Accordingly, [Tabs. 4](#) and [5](#) illustrate these

benchmark data clustering, the results of VI index, and their optimal cluster numbers (# OC). [Tab. 4](#) shows that AC-MeanABC outperforms other related works in minimizing the VI index in most of the outcomes, where iABC outperforms AC-MeanABC in R15 and D31. Additionally, [Tab. 5](#) indicates that AC-MeanABC exceeds other state-of-the-art methods in finding the optimal number of clusters for Glass, Aggregation, R15, Libras movement and wholesaler customers, while iABC outperforms AC-MeanABC in D31.

Table 3: Other parameters settings

Datasets	Np	μ	σ	Datasets	Np	μ	σ
Iris	5	0	1	D31	56	1	2
Ecoli	20	2	1	Libras movement	19	1	2.5
Wisconsin	13	0	1	Wholesaler customers	1	0	0.8
Wine	15	0	1	Lena	20	2	1
Dermatology	20	2	1	Jet	20	2	1
Glass	15	1	1.5	MorroBay	15	1	1
Aggregation	28	1	2	Mandril	20	2	1
R15	24	1	2.5	Peppe	20	3	1

Table 4: Benchmark data clustering part1 outcomes of VI index

Datasets	Measure	ACMeanABC	iABC	ABC	AKC-BCO	AKC-MEPSO	DCPG	DCPSO	DCGA
Glass	Average	0.3118	0.3644	0.3752	2.6411	2.1352	0.4700	0.5400	0.60
	SD	0.0020	0.0017	0.0053	1.1412	1.0154	0.0500	0.0800	0.13
Aggregation	Average	0.2952	0.3628	0.3706	2.3986	0.7543	0.6431	0.6311	0.52
	SD	0.0121	0.0117	0.0159	0.7359	0.2752	0.0368	0.0445	0.00
R15	Average	0.2753	0.2162	0.2164	0.9885	0.3271	0.2800	0.2700	0.40
	SD	0.0011	0.0013	0.0013	1.5429	0.0549	0.0500	0.0400	0.06
D31	Average	0.3188	0.3056	0.3056	3.8074	0.5769	0.4700	0.8670	0.42
	SD	0.0030	0.0004	0.0004	2.6032	0.2293	0.1300	0.1900	0.01
Libras movement	Average	0.5810	0.6270	0.6610	2.0972	1.8171	0.6262	0.6380	0.79
	SD	0.0024	0.0079	0.0136	0.3840	0.1956	0.0311	0.0357	0.03
Wholesaler customers	Average	0.2019	0.2040	0.2041	0.5104	0.4622	0.3010	0.2419	0.29
	SD	0.0005	0.0008	0.0002	0.0535	0.0894	0.0665	0.0324	0.05

In [Tab. 6](#), the Friedman test was carried out and the results were used in comparing the performance of AC-MeanABC and other related works. As shown, the proposed algorithm outperforms all comparable methods in [Tab. 6](#). The box whisker for the results of clustering algorithms in [Tab. 4](#) is displayed in [Fig. 3](#).

6.2.2 Part2

In this part, the AC-MeanABC outcomes were compared with other related works such as discrete binary artificial bee colony DisABC, GA based clustering algorithms, improved discrete binary artificial bee colony IDisABC [28] and dynamic clustering based particle swarm optimization DCPSO [33]. The parameter settings for DC-MeanABC algorithm were selected as in [Tabs. 2](#) and [3](#) based on the same parameter setting in [28].

Table 5: Benchmark data clustering part1 outcomes of clusters number

Datasets	#OC	Measure	AC-MeanABC	iABC	ABC	AKC-BCO	AKC-MEPSO	DCPG	DCPSO	DCGA
Glass	6	Average	6.22	7.03	6.53	3	3.98	5.93	5.43	4.97
		SD	0.14	0.18	0.51	0	1.24	0.77	0.61	0.57
Aggregation	7	Average	7.21	7.23	7.9	11.73	8.37	6.43	6.07	6.00
		SD	0.00	1.28	2.44	2.6	4.63	0.90	0.74	0.00
R15	15	Average	15	15	15	9.57	11.13	8.03	7.13	8.00
		SD	0	0	0	1.31	1.25	1.37	1.11	1.46
D31	31	Average	31.20	31.17	31.4	28.13	22.63	28.06	21.00	10.41
		SD	0.44	0.38	0.5	4.15	2.72	6.24	3.16	0.91
Libras movement	15	Average	15	15.73	15.37	4.2	4.13	6.87	6.33	8.55
		SD	0.12	0.87	1.1	0.61	0.35	0.97	1.24	0.78
Wholesaler customers	3	Average	3.28	6.87	6	3.37	3.43	7.73	6.30	3.66
		SD	0.001	0.57	0	0.49	0.57	2.21	2.02	0.72

Table 6: Mean ranks obtained by Friedman test for the clustering algorithms outcomes in [Tab. 4](#)

Variable	Mean rank
AC-MeanABC	1
iABC	2
ABC	2.8
DCPG	3.62
DCGA	3.8
DCPSO	4
AKC-MEPSO	5.92
AKC-BSO	6.97

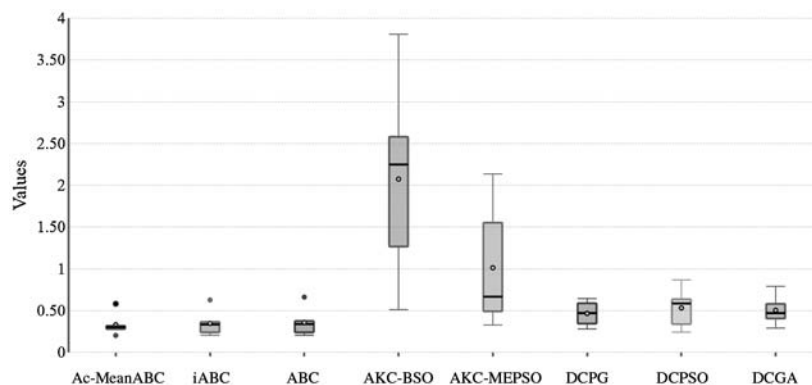


Figure 3: Box-whisker to present the outcomes of five clustering algorithms in [Tab. 4](#)

[Tabs. 7](#) and [8](#) illustrate these benchmark data clustering, the results of VI index, and their optimal cluster numbers (# OC). [Tab. 7](#) shows that AC-MeanABC outperforms other related works in minimizing the VI index in all outcomes. Additionally, [Tab. 8](#) indicates that AC-MeanABC exceeds other state-of-the-art methods in finding the optimal number of clusters for

Ecoli and Wine datasets. In the rest of the three cases, all approaches reached an optimal number of clusters in the Iris dataset. Further, AC-MeanABC, IDisABC, and DisABC obtained the same outcomes in Wisconsin dataset with a preference for AC-MeanABC in standard deviation value. In the Dermatology dataset, DCPSO and AC-MeanABC are very close in terms of result to find the best number of clusters. Still, the DCPSO standard deviation value is much higher than that of AC-MeanABC.

Table 7: Benchmark data clustering part2 outcomes of VI index

Datasets	AC-MeanABC	IDisABC	DisABC	GA	DCPSO
	Average (SD)	Average (SD)	Average (SD)	Average (SD)	Average (SD)
Iris	0.070 (0.035)	0.0974 (0.010)	0.0982 (0.010)	0.1182 (0.020)	0.1042 (0.013)
Ecoli	0.229 (0.041)	0.3073 (0.057)	0.3841 (0.069)	0.5351 (0.176)	0.3691 (0.086)
Wisconsin	0.091 (0.017)	0.135 (0.022)	0.135 (0.022)	0.1424 (0.039)	0.1368 (0.018)
Wine	0.192 (0.068)	0.3251 (0.035)	0.3365 (0.029)	0.4426 (0.145)	0.3518 (0.055)
Dermatology	0.257 (0.051)	0.3968 (0.043)	0.4328 (0.048)	0.5717 (0.138)	0.4600 (0.045)

Table 8: Benchmark data clustering part2 outcomes of clusters number

Datasets	#OC	AC-MeanABC	IDisABC	DisABC	GA	DCPSO
		Average (SD)	Average (SD)	Average (SD)	Average (SD)	Average (SD)
Iris	3	3 (0)	3 (0)	3 (0)	3 (0)	3 (0)
Ecoli	5	5.024 (0.388)	5.166 (0.379)	5.333 (0.606)	7.7 (1.600)	5.866 (0.776)
Wisconsin	2	2.033 (0.011)	2.033 (0.018)	2.033 (0.018)	2.133 (0.681)	2.066 (0.253)
Wine	3	3.10 (0.251)	3.3 (0.534)	3.4 (0.498)	4.3 (1.557)	3.9 (0.712)
Dermatology	6	6.043 (0.390)	5.56 (0.504)	5.533 (0.571)	6.966 (1.790)	6.033 (0.999)

To show the significance of improvement of the AC-MeanABC, Friedman test was performed and the results are shown in Tab. 9. As shown clearly, the mean ranking value of AC-MeanABC is higher than that of IDisABC, DisABC, DCPSO and GA. where the best value is the minimum value and it's represented in bold. Also, Fig. 4 represents the box whisker for the results of five clustering algorithms as in Tab. 7.

Table 9: Mean ranks obtained by Friedman test for the five clustering algorithms outcomes in Tab. 7

Variable	Mean rank
(1) AC-MeanABC	1
(2) IDisABC	2
(3) DisABC	3
(4) DCPSO	4
(5) GA	5

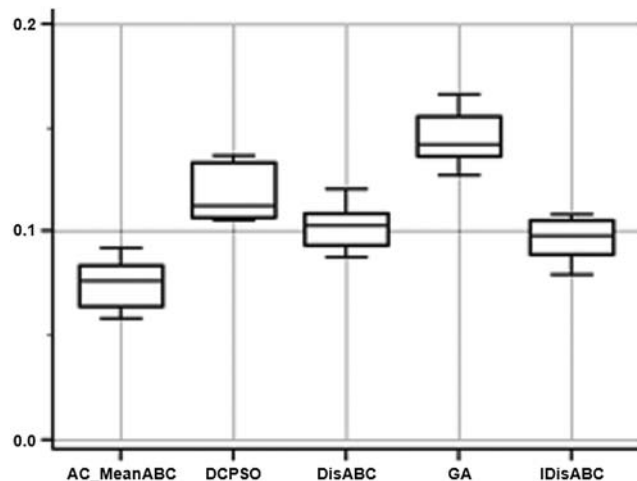


Figure 4: Box-whisker to present the outcomes of five clustering algorithms in [Tab. 7](#)

6.2.3 Natural Images

In this experiment, five natural images were obtained from the Berkeley1 segmentation dataset [30,41] such as Lena, Jet plane, MorroBay, Mandril, and Pepper images. In this experiment, the AC-MeanABC outcomes were compared with other related works such as discrete binary artificial bee colony DisABC, GA based clustering algorithms, and improved discrete binary of ABC called IDisABC [30] and dynamic clustering based PSO that is called DCPSO [35]. The parameter settings for AC-MeanABC algorithm were selected based on the same parameters settings in [30], as in [Tabs. 2](#) and [3](#).

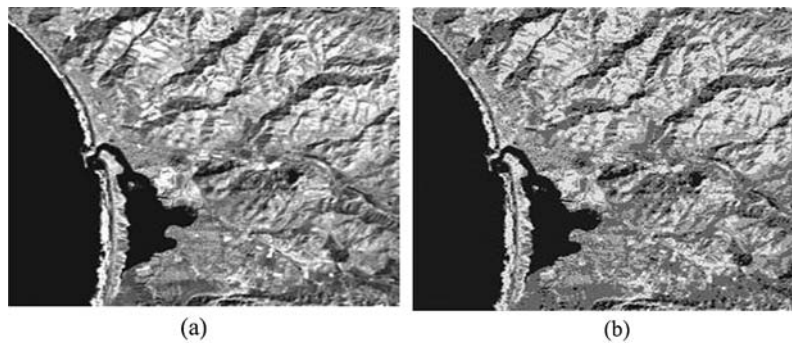
[Tabs. 10](#) and [11](#) illustrate these natural images, the VI index results, and the corresponding cluster numbers from ground truth image (# AC). [Tab. 10](#) shows that AC-MeanABC outperforms other related works in minimizing the VI index in all cases. [Tab. 11](#) shows the number of clusters that are determined automatically using algorithm AC-MeanABC and other state-of-the-art methods. [Tab. 8](#) shows that AC-MeanABC-FCM outperforms other related works in finding the optimal number of clusters for Lena, Jet plane, MorroBay, and Pepper images. In contrast, the IDisABC approach outperforms AC-MeanABC and other approaches in finding the optimal number of clusters for the Mandrill image, but AC-MeanABC-FCM and IDisABC are very close in terms of the result of Mandrill image. [Figs. 5a, 6a, 7a, 8a](#) and [9a](#) show the original image of MorroBay, Lena, Mandrill, Jet plane and Pepper, where [Figs. 5b, 6b, 7b, 8b](#) and [9b](#) show the clustering images by AC-MeanABC of MorroBay, Lena, Mandrill, Jet plane and Pepper images.

Table 10: Natural image clustering outcomes of VI index

Images	AC-MeanABC	IDisABC	DisABC	GA	DCPSO
	Average (SD)	Average (SD)	Average (SD)	Average (SD)	Average (SD)
Lena	0.0662 (0.0107)	0.0982 (0.0118)	0.1032 (0.0141)	0.1395 (0.0259)	0.1126 (0.0130)
Jet	0.0810 (0.0098)	0.0922 (0.0209)	0.0959 (0.0305)	0.1517 (0.0595)	0.1366 (0.1183)
MorroBay	0.0577 (0.0090)	0.0789 (0.0093)	0.0875 (0.0125)	0.1268 (0.0395)	0.1053 (0.0339)
Mandrill	0.0920 (0.0095)	0.1043 (0.0109)	0.1045 (0.0111)	0.1419 (0.0324)	0.1077 (0.0115)
Pepper	0.0764 (0.0151)	0.1081 (0.0142)	0.1201 (0.0150)	0.1662 (0.0595)	0.1323 (0.0460)

Table 11: Obtained number of clusters for natural image by clustering approaches

Images	#AC	AC-MeanABC	IDisABC	DisABC	GA	DCPSO
		Average (SD)	Average (SD)	Average (SD)	Average (SD)	Average (SD)
Lena	6	5.989 (0.018)	5.9 (0.922)	5.666 (0.546)	7.1 (1.516)	6.695 (1.063)
Jet	6	6.000 (0.002)	5.733 (0.868)	5.5 (0.861)	6.733 (1.680)	6.033 (0.964)
MorroBay	4	4.158 (0.232)	4.333 (0.479)	4.333 (0.479)	4.621 (1.146)	4.333 (0.546)
Mandrill	6	5.950 (0.520)	5.966 (0.905)	6.166 (1.0199)	7.4 (1.652)	6.9 (1.471)
Pepper	7	6.979 (0.026)	6.733 (0.691)	6.566 (0.568)	7.466 (1.105)	7.333 (1.212)

**Figure 5:** (a) Original MorroBay and (b) segmented MorroBay by DC-MeanAB**Figure 6:** (a) Original Lena and (b) segmented Lena by (AC-MeanABC)

In [Tab. 12](#), the Friedman test was carried out to compare the performance between the proposed algorithm and other related works. As shown, AC-MeanABC outperforms IDisABC, DisABC, DCPSO and GA and the best value is presented in bold; the best value is considered the minimum value of all clustering algorithms in [Tab. 9](#). The box whisker for the results of five clustering algorithms in [Tab. 10](#) is displayed in [Fig. 10](#).



Figure 7: (a) Original mandrill and (b) segmented mandrill by (AC-MeanABC)

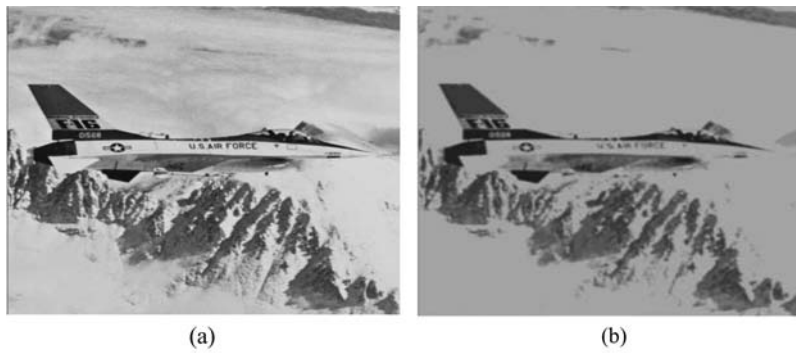


Figure 8: (a) Original jet and (b) segmented jet by (AC-MeanABC)

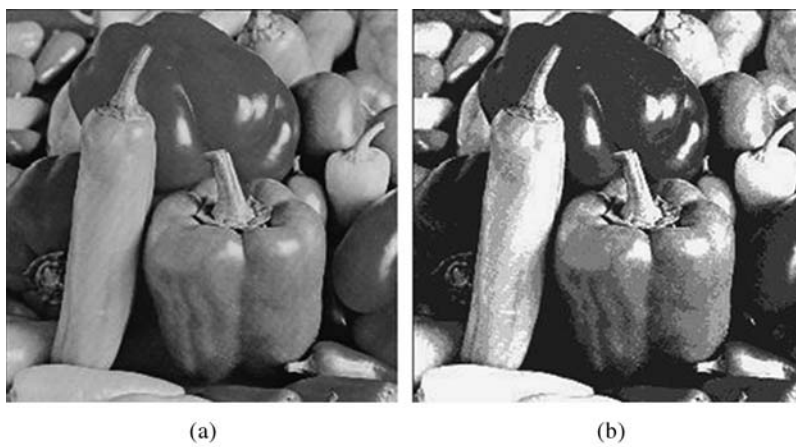


Figure 9: (a) Original pepper and (b) segmented pepper by (AC-MeanABC)

Table 12: Mean ranks obtained by Friedman test for the five clustering algorithms outcomes in [Tab. 10](#)

Variable	Mean rank
(1) AC-MeanABC	1
(2) IDisABC	2
(3) DisABC	3
(4) DCPSO	4
(5) GA	5

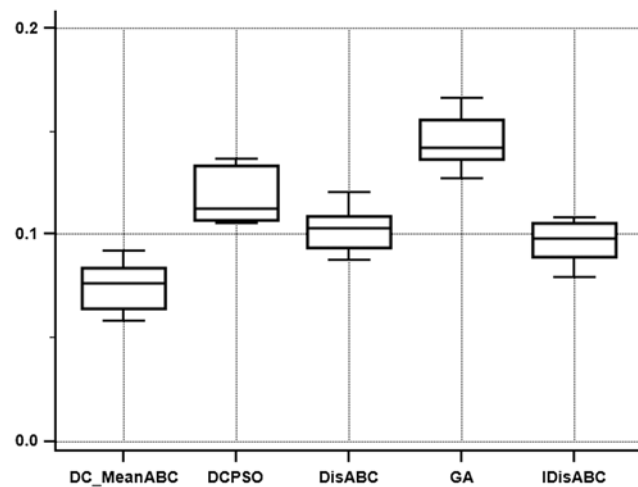


Figure 10: Box-whisker to present the outcomes of five clustering algorithms in [Tab. 10](#)

7 Conclusion

In this paper, the automatic fuzzy clustering based on the MeanABC search method called AC-MeanABC was proposed to solve the challenges of determining the number of clusters (region) and cluster centroids. AC-MeanABC clustering method used the capability of the MeanABC algorithm to explore the search space in positive and negative directions to search for the near-optimal number of clusters and centroids values. The experiments and results were obtained using 11 benchmark datasets and 5 natural images. These experiments compared the AC-MeanABC with other clustering methods such as iABC, ABC, AKCBCO, AKC-MEPSO, DCPG, DCGA IDisABC, DisABC, DCPSO, and GA. In conclusion, the clustering results of the AC-MeanABC are better than those of the state-of-the-art techniques in determining the optimal number of clusters and the value of validity index VI.

Acknowledgement: Thanks to our families & colleagues who supported us morally.

Funding Statement: This research was supported by the Research Management Center, Xiamen University Malaysia under XMUM Research Program Cycle 4 (Grant No: XMUMRF/2019-C4/IECE/0012).

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

References

- [1] A. Bose and K. Mali, "Fuzzy-based artificial bee colony optimization for gray image segmentation," *Signal Image and Video Processing*, vol. 10, no. 6, pp. 1–8, 2016.
- [2] J. Nayak, B. Naik and H. Behera, "Fuzzy c-means (FCM) clustering algorithm: A decade review from 2000 to 2014," *Computational Intelligence in Data Mining*, vol. 2, pp. 133–149, 2015.
- [3] R. Kuo and F. E. Zulvia, "Automatic clustering using an improved artificial bee colony optimization for customer segmentation," *Knowledge and Information Systems*, vol. 57, no. 2, pp. 331–357, 2018.
- [4] A. K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.
- [5] A. E. Ezugwu, "Nature-inspired metaheuristic techniques for automatic clustering: A survey and performance study," *SN Applied Sciences*, vol. 2, no. 2, pp. 273, 2020.
- [6] N. Jain and V. Srivastava, "Data mining techniques: A survey paper," *IJRET International Journal of Research in Engineering and Technology*, vol. 2, no. 11, pp. 2319–1163, 2013.
- [7] H. Zhou and G. Schaefer, "An overview of fuzzy c-means based image clustering algorithms," *Foundations of Computational Intelligence*, vol. 2, pp. 295–310, 2009.
- [8] R. Kuo and F. E. Zulvia, "An improved differential evolution with cluster decomposition algorithm for automatic clustering," *Soft Computing*, vol. 23, no. 18, pp. 8957–8973, 2019.
- [9] S. S. Nikam, "A comparative study of classification techniques in data mining algorithms," *Oriental Journal of Computer Science & Technology*, vol. 8, no. 1, pp. 13–19, 2015.
- [10] J. C. Bezdek, "Pattern recognition with fuzzy objective function algorithms," in *Advanced Applications in Pattern Recognition*, 1st ed., USA: Springer Press, pp. 1–272, 1981.
- [11] W. Alomoush, A. Alrosan, N. Norwawi, Y. M. Alomari, D. Albashish *et al.*, "A survey: Challenges of image segmentation based fuzzy C-means clustering algorithm," *Journal of Theoretical & Applied Information Technology*, vol. 96, no. 16, pp. 5153–5170, 2018.
- [12] W. Alomoush and A. Alrosan, "Metaheuristic search-based fuzzy clustering algorithms," arXiv preprint arXiv: 1802.08729, 2018.
- [13] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm," *Journal of Global Optimization*, vol. 39, no. 3, pp. 459–471, 2007.
- [14] W. A. Hussein, S. Sahran and S. N. H. Sheikh Abdullah, "The variants of the bees algorithm (BA): A survey," *Artificial Intelligence Review*, vol. 47, no. 3, pp. 67–121, 2017.
- [15] W. A. Hussein, S. Sahran and S. N. H. S. Abdullah, "A fast scheme for multilevel thresholding based on a modified bees algorithm," *Knowledge-Based Systems*, vol. 101, pp. 114–134, 2016.
- [16] Z. W. Geem, J. H. Kim and G. V. Loganathan, "A new heuristic optimization algorithm: Harmony search," *Simulation*, vol. 76, no. 2, pp. 60–68, 2001.
- [17] A. A. Alomoush, A. A. Alsewari, H. S. Alamri, K. Z. Zamli, W. Alomoush *et al.*, "Modified opposition based learning to improve harmony search variants exploration," in *Int. Conf. of Reliable Information and Communication Technology*, Johor, Malaysia, pp. 279–287, 2019.
- [18] X. S. Yang, "Firefly algorithm, stochastic test functions and design optimisation," arXiv preprint arXiv: 1003.1409, 2010.
- [19] W. Alomoush, K. Omar, A. Alrosan, Y. M. Alomari, D. Albashish *et al.*, "Firefly photinus search algorithm," *Journal of King Saud University—Computer and Information Sciences*, vol. 32, no. 5, pp. 599–607, 2018.

- [20] X. S. Yang and S. Deb, "Cuckoo search via lévy flights," in *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*, Coimbatore, India, 210–214, 2009.
- [21] W. Alomoush, "Cuckoo search algorithm based dynamic parameter adjustment mechanism for solving global optimization problems," *International Journal of Applied Engineering Research*, vol. 14, no. 24, pp. 4434–4440, 2019.
- [22] D. Karaboga and C. Ozturk, "A novel clustering approach: Artificial bee colony (ABC) algorithm," *Applied Soft Computing*, vol. 11, no. 1, pp. 652–657, 2011.
- [23] A. Alrosan, W. Alomoush, N. Norwawi, M. Alswaitti and S. N. Makhadmeh, "An improved artificial bee colony algorithm based on mean best-guided approach for continuous optimization problems and real brain MRI images segmentation," *Neural Computing and Applications*, vol. 62, no. 3, pp. 736, 2020.
- [24] B. A. Aldeeb, M. A. Al-Betar, A. O. Abdelmajeed, M. J. Younes, M. AlKenani *et al.*, "A comprehensive review of uncapacitated university examination timetabling problem," *International Journal of Applied Engineering Research*, vol. 14, no. 24, pp. 4524–4547, 2019.
- [25] M. Alauthman, A. Almomani, M. Alweshah, W. Omoushd and K. Alieyane, *Machine Learning for Phishing Detection and Mitigation, Machine Learning for Computer and Cyber Security: Principle, Algorithms, and Practices*, 1st ed., vol. 1. Boca Raton, Florida, USA: CRC Press, pp. 26, 2019.
- [26] M. B. Agbaje, A. E. Ezugwu and R. Els, "Automatic data clustering using hybrid firefly particle swarm optimization algorithm," *IEEE Access*, vol. 7, pp. 184963–184984, 2019.
- [27] S. Ouadfel and S. Meshoul, "Handling fuzzy image clustering with a modified ABC algorithm," *International Journal of Intelligent Systems and Applications*, vol. 4, no. 2, pp. 65, 2012.
- [28] E. Hancer, C. Ozturk and D. Karaboga, "Extraction of brain tumors from MRI images with artificial bee colony based segmentation methodology," in *2013 8th Int. Conf. on Electrical and Electronics Engineering (ELECO)*, Bursa, Turkey, pp. 516–520, 2013.
- [29] A. Alrosan, N. Norwawi, W. Ismail and W. Alomoush, "Artificial bee colony based fuzzy clustering algorithms for MRI image segmentation," in *Int. Conf. on Advances in Computer Science and Electronics Engineering—CSEE*, Kuala Lumpur, Malaysia, pp. 225–228, 2014.
- [30] C. Ozturk, E. Hancer and D. Karaboga, "Dynamic clustering with improved binary artificial bee colony algorithm," *Applied Soft Computing*, vol. 28, no. 3, pp. 69–80, 2015.
- [31] S. Das and A. Konar, "Automatic image pixel clustering with an improved differential evolution," *Applied Soft Computing*, vol. 9, no. 1, pp. 226–236, 2009.
- [32] O. Alia, R. Mandava and M. E. Aziz, "A hybrid harmony search algorithm for MRI brain segmentation," *Evolutionary Intelligence*, vol. 4, no. 1, pp. 31–49, 2011.
- [33] M. G. Omran, "Particle swarm optimization methods for pattern recognition and image processing," Ph.D. dissertation. University of Pretoria, South Africa, 2006.
- [34] M. G. Omran, A. P. Engelbrecht and A. Salman, "Dynamic clustering using particle swarm optimization with application in unsupervised image classification," in *Fifth World Enformatika Conf. (ICCI)*, Prague, Czech Republic, pp. 199–204, 2005.
- [35] M. G. Omran, A. Salman and A. P. Engelbrecht, "Dynamic clustering using particle swarm optimization with application in image segmentation," *Pattern Analysis and Applications*, vol. 8, no. 4, pp. 332–344, 2006.
- [36] U. Maulik and S. Bandyopadhyay, "Fuzzy partitioning using a real-coded variable-length genetic algorithm for pixel classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 5, pp. 1075–1081, 2003.
- [37] W. Alomoush, S. N. H. S. Abdullah, S. Sahran and R. I. Hussain, "MRI brain segmentation via hybrid firefly search algorithm," *Journal of Theoretical & Applied Information Technology*, vol. 61, no. 1, pp. 73–90, 2014.
- [38] W. K. Alomoush, S. Abdullah, S. Sahran and R. I. Hussain, "Segmentation of MRI brain images using FCM improved by firefly algorithms," *Journal of Applied Sciences*, vol. 14, no. 1, pp. 66–71, 2014.
- [39] J. C. Bezdek, C. Coray, R. Gunderson and J. Watson, "Detection and characterization of cluster substructure i. linear structure: Fuzzy c-lines," *SIAM Journal on Applied Mathematics*, vol. 40, no. 2, pp. 339–357, 1981.

- [40] S. Saha and S. Bandyopadhyay, "A new point symmetry based fuzzy genetic clustering technique for automatic evolution of clusters," *Information Sciences*, vol. 179, no. 19, pp. 3230–3246, 2009.
- [41] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Computer Vision, 2001, ICCV*, Vancouver, British Columbia, Canada, pp. 416–423, 2001.
- [42] R. Kuo, Y. Huang, C. C. Lin, Y. H. Wu and F. E. Zulvia, "Automatic kernel clustering with bee colony optimization algorithm," *Information Sciences*, vol. 283, pp. 107–122, 2014.
- [43] R. Kuo, Y. Syu, Z. Y. Chen and F. C. Tien, "Integration of particle swarm optimization and genetic algorithm for dynamic clustering," *Information Sciences*, vol. 195, no. 1, pp. 124–140, 2012.