

Image Clustering based on Artificial Intelligence Techniques

Baydaa Ibraheem Khaleel

Baydaa_ibraheem@yahoo.com

College of Computer Science and Mathematics

University of Mosul

Received on: 15/8/2013

Accepted on: 11/11/2013

ABSTRACT

Clustering has been widely used in data analysis and pattern recognition and classification. The objective of cluster analysis is the classification of objects according to similarities among them, and organizing of data into groups. There are several algorithms for clustering large data sets or streaming data sets, Their aims are to organize a collection of data items into clusters. These such items are more similar to each other within a cluster, and different than they are in the other clusters. We have take the advantage of classification abilities of Artificial Intelligence Techniques (AITs) to classify images data set into a number of clusters. The Gath-Geva (GG) fuzzy clustering algorithm, Artificial Bee Colony algorithm(ABC), Radial Basis Function Network(RBF), and then combined Gath-Geva algorithm with (RBF) algorithm to produce Fuzzy RBF (FRBF) method were applied using images data set to classify this data set into a number of clusters (classes). Each cluster will contain data set with most similarity in the same cluster and most dissimilarity with the different clusters. We compute the classification rate, and false rate on this data set. Finally we make comparisons between results obtained after applying these algorithms on this images data set. The FRBF is better than the other three methods that applied in this research such as G-G, ABC, RBF, because the FRBF was obtained higher classification rate in testing state equal (96.8571) and low false alarm equal(3.1429).

Keywords: *Gath-Geva(GG) fuzzy clustering algorithm; Artificial Bee Colony algorithm (ABC); Radial Basis Function (RBF) Neural Network; Fuzzy RBF (FRBF).*

عنقدة الصور المعتمدة على التقنيات الذكائية الاصطناعية

بيداء ابراهيم خليل

كلية علوم الحاسوب والرياضيات

جامعة الموصل

تاريخ القبول: 11/11/2013

تاريخ الاستلام : 15/8/2013

المخلص

العنقدة تستخدم بشكل واسع في تحليل البيانات وتمييز الأنماط والتصنيف. أهداف تحليل العنقود هو تصنيف الاشياء طبقا للتشابه فيما بينها، وتنظيم البيانات على شكل مجاميع. هناك خوارزميات عدة لعنقدة مجاميع البيانات الكبيرة أو سيل من مجاميع البيانات، أهدافها تنظيم مجاميع البيانات في عناقيد ، وعناصر البيانات هذه الموجودة في العنقود الواحد تشبه بعضها البعض الآخر وتختلف عن عناصر البيانات في عناقيد أخرى. لقد تم اخذ الفائدة

لقابليات التصنيف للتقنيات الذكائية الاصطناعية لتصنيف مجموعة بيانات الصور الى عدد من العناقيد. فقد تم تطبيق خوارزمية العنقدة المضببة G-G وخوارزمية مستعمرة النحل الاصطناعية ABC وشبكة دالة القاعدة الشعاعية RBF وبعد ذلك تم دمج خوارزمية G-G مع شبكة RBF لتنتج شبكة دالة القاعدة الشعاعية المضببة FRBF. تصنف مجموعة بيانات الصور الى عدد من العناقيد او الاصناف، حيث كل عنقود يحوي على مجموعة بيانات متشابه في نفس العنقود و غير متشابه مع العناقيد الاخرى. تم حساب نسبة التصنيف ونسبة الخطا لهذه البيانات. وأخيرا تم مقارنة النتائج التي تم الحصول عليها بعد تطبيق هذه الخوارزميات على مجموعة بيانات الصور. FRBF هي افضل من الطرق الثلاثة الاخرى التي تم تطبيقها في هذا البحث وهي G-G, ABC, RBF، وذلك لان FRBF حصلت على أعلى نسبة تصنيف في مرحلة الاختبار والتي تساوي (96.8571) واقل نسبة خطأ وهي (3.1429).

الكلمات المفتاحية : خوارزمية العنقدة المضببة G - G ، خوارزمية مستعمرة النحل الاصطناعية ABC، وشبكة دالة القاعدة الاصطناعية RBF، شبكة القاعدة الاصطناعية المضببة FRBF.

1. Introduction

We are living in a world full of data. Every day, people encounter a large amount of information and store or represent it as data, for further analysis and management. One of the vital means in dealing with these large data is to classify or group them into a set of categories or clusters[1]. Clustering is the process of grouping a data set in such a way that the similarity between data within a cluster is maximized while the similarity between data of different clusters is minimized[2]. Clustering or classification systems are either supervised or unsupervised[1], unsupervised clustering takes an unlabelled set of data and partition it into groups of examples, without additional knowledge. Supervised clustering , on the other hand, assumes that the class structure is already known. It takes a set of examples with class labels[3].

Fuzzy clustering is a branch in clustering analysis and it is widely used in the pattern recognition field, such as Gath-Geva algorithm that takes the size and density of clusters for classification[4].

Artificial Bee Colony(ABC) algorithm, an optimization strategy based on the intelligent behavior of honey bee swarm, is one of the swarm intelligence system, it is a population-based algorithm that can provide the best possible solution for different mathematical problems by using inspiration techniques from nature.

Artificial Neural Networks(ANNs) are the most novel and powerful artificial tool suitable for solving combinatorial problems such as predication and classification. ANNs are being used extensively for solving universal problems intelligently like continuous, discrete, and clustering[5].

The rest of the research is organized as follows: Section 2 previous work, Section 3 Gath-Geva Fuzzy Clustering Algorithm, section 4 describes artificial Bee Colony (ABC) algorithm, section 5 discusses Radial Basis Function (RBF), section 6 presents the proposed method which combined Gath-Geva fuzzy clustering with RBF to produce a proposed method called Fuzzy Radial Basis Function (FRBF), section 7 performance measures, section 8 describes the experimental and results obtained, section 9 conclusion .

2. Previous Work

In particular several classification or clustering algorithms and artificial intelligence techniques were used for images clustering or classification. In 2012 Thai, Hai, and Thuy[6] used the combination between feedforward artificial neural networks(FANN) with support vector machine (SVM) to classify Roman numeral images, the average classification rate was 86%. Akgundogdu et al.[7], used support vector machine and genetic algorithm (GA) for bone sample images, with 12 training and 6 testing data, the success rate for SVM was equal 83.33%. Banerjee, et al.[8], used artificial bee colony to classify band satellite images for different features, the accuracy rate for barren was equal 93.65%. Arias et al. [9] used multi layer perceptron MLP, and Self Organized Maps SOM neural network to classify satellite images, error for MLP is equal 29.65%, and error for SOM equal 18.00%. Rajendran, Madheswaran[10], used Hybird Association Rule with Decision Tree Classification (HARC) to classify brain images the accuracy rate equal 95%. Zhang and Qiu[11] used unsupervised neuro-fuzzy system and some methods to classify Hyperspectral imaging, the overall accuracy for SOM is 73.1%, fuzzy c-means FCM is equal 80.9%, DFLVQ is equal 83.5%, GFSOM is equal 88.6%, ISODATA is equal 53.4%. Luping and Zhang[12], used SVM to classify fingerprint, the accuracy rate is equal 89.3%. Raja, Shanmugam[13], used wavelet features based on backpropagation network to classify images such as, "opencountry" category, and "War Tank" category, the classification rate is equal 82.5%. Arpita and Mahua[14] proposed a computer assisted treatment planning system implementing genetic algorithm based on Neuro-fuzzy approach. The boundary based features of the tumour lesions appearing in breast have been extracted for classification. The shape features represented by fourier descriptors, introduce a large number of feature vectors with classification rate at 87%.

3. Gath-Geva Fuzzy Clustering Algorithm

Gath-Geva (GG) fuzzy clustering algorithm takes the size and density of clusters for classification[15]. The objective function based on the minimization of the sum of weighted squared distances between the data points and cluster centers is described in the following[16]:

$$J(Z, U, V) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m D_{ik}^2 \quad \dots\dots\dots (1)$$

Where Z is the set of data, $U = [\mu_{ik}]$ is the fuzzy partition matrix, $V = [V_1, V_2, \dots, V_c]^T$ is the set of centers of the clusters, c is the number of clusters, N is the number of the data, m is the fuzzy coefficient, D_{ik} is the difference between X_k and the cluster center, where equation (5) calculates D_{ik} , μ_{ik} is the membership degree between the i -th cluster and k -th data, which satisfies conditions:

$$\mu_{ik} \in [0,1]; \sum_{i=1}^c \mu_{ik} = 1 \quad \dots\dots\dots (2)$$

The minimum of (U, V) is calculated as follows:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c (D_{ik} / D_{jk})^{2/(m-1)}} \quad \dots\dots\dots (3)$$

$$V_i = \frac{\sum_{k=1}^N (\mu_{ik})^m X_k}{\sum_{k=1}^N (\mu_{ik})^m} \dots\dots\dots (4)$$

The norm of distance between i-th cluster and k-th data is :

$$D_{ik}(X_k, V_i) = \frac{\sqrt{\det(F_{mi})}}{P_i} * \exp\left(\frac{1}{2}(X_k - V_i)^T F_{mi}^{-1} (X_k - V_i)\right) \dots\dots\dots (5)$$

$$F_{mi} = \frac{\sum_{k=1}^N (\mu_{ik})^m (X_k - V_i)(X_k - V_i)^T}{\sum_{k=1}^N (\mu_{ik})^m} \dots\dots\dots (6)$$

where the X_k denote the input, F_{mi} is the fuzzy covariance matrix of the i-th cluster, μ_{ik} is the fuzzy partitioning matrix, m is the weighting exponent controls the 'fuzziness' of the resulting cluster and P_i is aprior probability of selecting the i-th cluster. The distance in Eq. (5) is used in the calculation of P_i , the probability of selecting the i-th cluster given the k-th data point, is given by:

$$P_i = \frac{1}{N} \sum_{k=1}^N \mu_{ik} \dots\dots\dots (7)$$

The step of the GG algorithm are listed as follows[15][16][17]:

- Step1: determining the number of clusters; c and m value(let $m = 2$), given the converge error, $\varepsilon > 0$.
- Step2: generate the matrix U randomly, U must satisfy the condition of equation (2).
- Step3: compute the parameters of the model using equations(4), (6), (7).
- Step4: calculate the norm of distance utilizing equation(5).
- Step5: update the partition matrix U using equation(3).
- Step6: stop if $\|U^{(l)} - U^{(l-1)}\| \leq \varepsilon$, else go to step 3.

4. Artificial Bee Colony Algorithm

Artificial Bee Colony(ABC) algorithm is a swarm intelligence method which simulates intelligence forging behavior of honey bees[18]. It was first proposed by Karaboga and then further developed by Karaboga, Basturk and Akay et al. [19]. In the model of ABC algorithm, there are three groups of bees; employed bees, onlooker bees and scout bees in the colony of artificial bees. Firstly, half of the colony consists of the employed bees and the second half consists the onlookers. Employed bees go to the food sources, and then they share the nectar and the position information of the food sources with the onlooker bees which are waiting on the dance area determine to choose a food source. The employed bee whose food source has been abandoned by the bees becomes a scout bee that carries out random search in the simulating model. The goal of bees in the ABC model is to find the best solution, the position of a food source represents a possible solution to the optimization problem and the nectar amount of a food source corresponds to the quality (fitness) of the associated solution. In the ABC algorithm, the number of employed bees is equal to the number of food sources which is also equal to the number of onlooker bees. There is only one employed bee for each food source whose first position is randomly generated. At each iteration of the

algorithm, each employed bee determines a new neighboring food source of its currently associated food source by Equation (8), and computes the nectar amount of this new food source:

$$v_{ij} = X_{ij} + \varphi(X_{ij} - X_{kj}) \quad \dots\dots\dots (8)$$

Where k is a randomly selected food source different from i , j is a randomly selected dimension. φ is a random number between $[-1,1]$. If the nectar amount of this new food source is higher than that of its currently associated food source, then this employed bee moves to this new food source, otherwise it continues with the old one. After all employed bees complete the search process, they share the information about their food sources with onlooker bees. An onlooker bee evaluates the nectar information taken from all employed bees and chooses a food source with a probability related to its nectar amount by Equation (9).

$$P_i = \frac{fitness_i}{\sum_{j=1}^{SN} fitness_j} \quad \dots\dots\dots (9)$$

$$fitness_i = \begin{cases} \frac{1}{1 + f_i} & \text{if } f_i \geq 0 \\ 1 + abc(f_i) & \text{if } f_i < 0 \end{cases} \quad \dots\dots\dots (10)$$

where $fitness_i$ is the fitness value of the solution i which is proportional to the nectar amount of the food source in the position i and SN is the number of food sources which is equal to the number of employed bees. Once all onlookers have selected their food sources, each of them determines a new neighboring food source of its selected food source and computes its nectar amount. Providing that this amount is higher than that of the previous one, and then the bee memorizes the new position and forgets the old one. The employed bee becomes a scout bee when the food source which is exhausted by the employed and onlooker bees is assigned as abandoned. In other words, if any solution cannot be improved further through a predetermined number of cycles which is called limit parameter, the food source is assigned as an abandoned source and employed bee of that source becomes a scout bee. In that position, scout generates randomly a new solution by Equation (11).

$$X_{ij} = X_j^{\min} + rand(0,1)(X_j^{\max} - X_j^{\min}) \quad \dots\dots\dots (11)$$

Where $i=1,2,\dots, SN$, $j=1,2,\dots,D$. SN the number of food sources and equals to half of the colony size. D is the dimension of the problem, representing the number of parameters to be optimized. X_j^{\min} and X_j^{\max} are lower and upper bounds of the j th parameter. The fitness of food sources will be evaluated. Additional, counters which store the numbers of trials of each bee are set to 0 in this phase[18][19].

The main steps can be described as follows[20]:

Step1: Initialize the bee colony $X = \{x_i | i=1,2,\dots,n\}$, Where n denotes the population size, x_i is the i th bee.

Step2: According to the fitness function, calculate the fitness f_i of each employed bee x_i , and record the maximum nectar amount as well as

the corresponding food source.

Step3: Each employed bee produces a new solution v_i in the neighborhood of the solution in its memory by Eq(8).

Step4: Use the greedy criterion to update x_i . Compute the fitness of v_i . If v_i is superior to x_i , x_i is replaced with v_i ; otherwise x_i is remained.

Step5: According to the fitness f_i of x_i , get the probability value P_i via formulas (9) and (10).

Step6: Depending on the probability P_i , onlookers choose food sources, search the neighborhood to generate candidate solutions, and calculate their fitness.

Step7: Use the greedy criteria to update the food sources.

Step8: Memorize the best food source and nectar amount achieved.

Step9: Check whether there are some abandoned solutions or not. If true, replace them with some new randomly-generated solutions by Eq(11).

Step10: Repeat steps (3)-(9), until the maximum number of iterations is researched or stop conditions are satisfied.

As mentioned above, the fitness function is a key component of ABC algorithm, which evaluates the foraging quality of the colony, i.e. the accuracy of possible solutions. Besides, some control parameters, such as the number of employ bees or onlooker bees, the limit times for abandonment, the maximum number of iterations or stop conditions, need to be assigned. They would have a direct influence on the speed and stability of convergence [21].

5. Radial Basis Function (RBF).

An RBF network is a three-layer feed forward neural network which consists of an input layer, a hidden layer and an output layer[22]. The hidden layer transforms the input vector from the input layer to a higher dimension, by applying a nonlinear transformation from the input space to the hidden layer vector space. Then, from the hidden layer, there is a linear transformation to the output layer. This Artificial Neural Network (ANN) would be suitable for application to a pattern classification. The dimension of the hidden layer is defined by the number of hidden neurons[23][24].

The structure of a traditional RBF network is shown in figure 1.

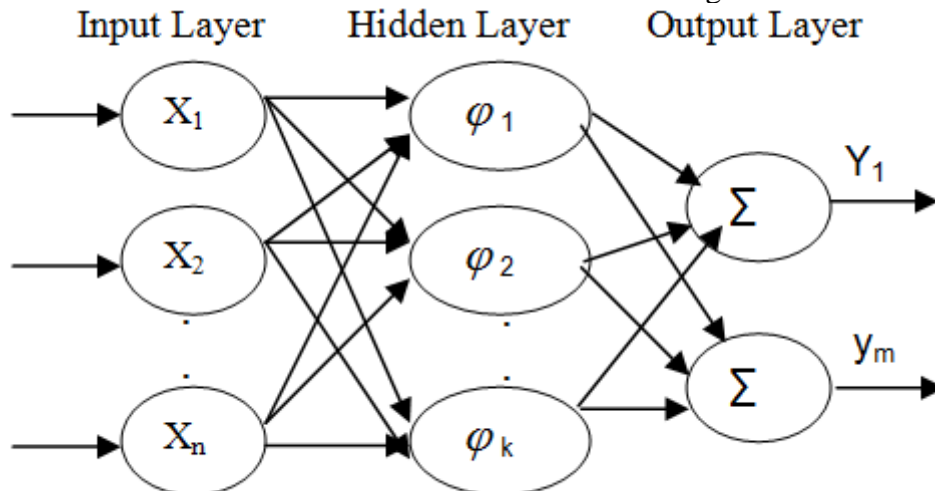


Figure 1. The RBF network structure

The number of neurons in input layer depends on number of input attributes of

training samples. The number of neurons in the output layer depends on number of target class labels. Assume that x is an n -dimensional input vector, there are m neurons in output layer and k neurons in hidden layer. It can be formulated as :

$$\varphi(x) = \exp\left(-\frac{\|x - c_j\|^2}{2\sigma^2}\right) \dots\dots\dots (12)$$

$$y_i = \sum_{j=1}^k w_{ij} \varphi_j \dots\dots\dots (13)$$

where $i=\{1,2,\dots,m\}$, $j=\{1,2,\dots,k\}$, $\varphi(x)$ is the output of the j th hidden node, C_j is the center vector of Gaussian kernel function which has the same number of dimension with x , and σ is the width of Gaussian kernel function at the j th hidden node. Besides, y_i is the output of the i th output neurons and W_{ij} is the connect weight from the j th hidden node to the i th output node.

And then apply the sigmoid function for each neuron in the output layer by this equation:

$$\psi_k(x) = f\left[\sum_j^L w_{jk} \varphi_j\right] \dots\dots\dots (14)$$

where L is the number of hidden nodes. Update the weight between output and hidden layers by using equations (15) – (16).

$$w_{jk}^{new} = w_{jk}^{old} + \Delta w_{jk} \dots\dots\dots (15)$$

$$\Delta w_{jk} = \eta(t_k - \Psi_k) \varphi_j \dots\dots\dots (16)$$

Where η is the learning rate, it is a small number takes value between (0,1)[25][26].

6. Fuzzy Radial Basis Function Network

Radial basis functions neural network has its origins in techniques for performing exact interpolation of a set of data point. We can generalized to design a Fuzzy Radial Basis Function by adding Gath-Geva Fuzzy clustering algorithm to weight equation (16) to produce a novel method called Fuzzy Radial Basis Function (FRBF), and the algorithm of new method FRBF as follows:

Step1: A vector pair (x,y) of the training set, is selected randomly.

Step2: compute the Gaussian Basis Function by the equation(12).

Step3: update the weight of each neuron, the weight updation is performed in accordance to the following

$$w_{jk}^{new} = w_{jk}^{old} + R_i * \eta(t_k - \Psi_k) \varphi_j \dots\dots\dots (17)$$

$$\Delta w_{jk} = R_i * \eta(t_k - \Psi_k) \varphi_j \dots\dots\dots (18)$$

Where R_i is the fuzzy scaling function given by:

$$R_i = (\mu_{ik})^m \text{ where}$$

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c (D_{ik} / D_{jk})^{2/(m-1)}} \dots\dots\dots (19)$$

The scaling function R_i depends on the fuzzy generator m which is a real number greater than 1.

Step4: Calculate the output of output layer by equation (14).

7. Performance Measures

To evaluate the accuracy of a method that was used in this research, we use some indicators: Classification Rate (CR), is obtained by dividing the number of correctly classified items in each category by the number of training set used for that category, False Rate(FR); FR is the ratio between the number of false positive and the total number of testing dataset. The value of (CR), is expected to be as large as possible, while the value of FR is expected to be as small as possible [27][[28][29].

8. Experimental And Results

MATLAB language was used to implemented for all algorithms and equations in this research. To train Gath-Geva(G-G)fuzzy clustering algorithm, Artificial Bee Colony(ABC), and Radial Basis Function(RBF), Fuzzy Radial Basis Function (FRBF), a larger number of training samples are needed. Training and testing dataset obtained from <http://www.flowers.vg/flowers>[30]. Dataset consist of (239) images with different category of flowers images. We used (116) images for training state, and (123) images for testing state. Each image is normalized using equation (20):

$$X = \frac{x - \min}{\max - \min} \dots\dots\dots (20)$$

Where, x is the numerical value, min is the lower bound of the attribute values, and max is the upper bound of the attribute values[31][32].

A number of experiments were performed to evaluate the system as well as to investigate the effect of G-G, ABC parameters and RBF, FRBF network parameters. Training and testing dataset consist of 10 categories of images. Each category consists of number of images samples as shown in Table (1).

Table (1) The Number Of Samples Of Flower Image Dataset, With The Number Of Flowers In Each Category

Type of samples	Training dataset		Testing dataset	
	No. of samples	samples rate	No. of samples	samples rate
White lilies flower single	15	12.93	10	8.13
White lilies flowers	16	13.80	12	9.76
Yellow black lily	10	8.62	14	11.38
Red white lily flower	10	8.62	6	4.88
Bright orange lily flower	10	8.62	10	8.13
Orange lily flower	10	8.62	7	5.69
White light red carnation flower	15	12.93	12	9.76
Red carnation	12	10.34	32	26.01
White gardenia	8	6.90	8	6.50
White gerbera daisy	10	8.62	12	9.76
Total	116	100%	123	100%

we applied (G-G) , ABC, RBF, and FRBF on flowers images dataset. In the training state we applied these algorithms on training data set that contain (116) flowers images to classify this dataset into (10) classes or clusters. In training state the results of classification obtained is 100% for all these algorithms. In the testing state we used testing dataset that contain (123) flowers images, and applied these algorithm on this dataset. Table(2) and table (3) show the results of the testing by G-G, ABC algorithms with the classification rate for each category of the flowers and over all classification rate for it. ABC was classified a greater number than the G-G algorithm. The ABC is better than G-G method, because ABC give higher classification rate equal (95.0237) than G-G which give classification rate equal(93.3451), and also ABC give lower false rate equal (4.9763) than G-G method which give false rate equal (6.6549).

Table (2) Results Of Testing State Using G-G Fuzzy Clustering Algorithm

Type of samples	Testing dataset	Classification rate (CR)%	False rate (FR)%
	No. of samples		
White lilies flower single	10	100	0
White lilies flowers	12	91.666	8.334
Yellow black lily	14	78.571	21.429
Red white lily flower	6	100	0
Bright orange lily flower	10	90	10
Orange lily flower	7	85.714	14.286
White light red carnation flower	12	100	0
Red carnation	32	100	0
White gardenia	8	87.5	12.5
White gerbera daisy	12	100	0
Total	123	93.3451	6.6549

Table (3) Results Of Testing State Using ABC Algorithm

Type of samples	Testing dataset	Classification rate (CR)%	False rate (FR)%
	No. of samples		
White lilies flower single	10	90	10
White lilies flowers	12	91.666	8.334
Yellow black lily	14	92.857	7.143
Red white lily flower	6	100	0
Bright orange lily flower	10	90	10
Orange lily flower	7	85.714	14.286
White light red carnation flower	12	100	0
Red carnation	32	100	0
White gardenia	8	100	0
White gerbera daisy	12	100	0
Total	123	95.0237	4.9763

The input data in training state is 116 images and in testing state is 123 images. And the resolution of each image is (114*86). Different parameter setting for ABC as follows: population size=200; and Maximum Cycle Number (MCN) = 10. In G-G method, the row number of membership function U is (10 rows) according to the number of output classes, and the column number of U is (9804) according to the number of pixels in each image in dataset, and the fuzzy coefficient value (m) equal 2. The output of these algorithms are 10 classes each class contain different number of samples.

Table(4) and table (5) show the results of the testing by RBF, FRBF algorithms with the classification rate for each category of the flowers. FRBF was classified a greater number than the RBF algorithm. The FRBF is better than RBF method, because FRBF give higher classification rate equal (96.8571) than RBF which give classification rate equal(89.6189), and also FRBF give lower false rate equal (3.1429) than RBF method which give false rate equal(10.3811).

Table (4) Results Of Testing State Using RBF Algorithm

Type of samples	Testing dataset	Classification rate (CR)	False rate (FR)
	No. of		

	samples		
White lilies flower single	10	90	10
White lilies flowers	12	83.333	16.667
Yellow black lily	14	71.428	28.572
Red white lily flower	6	100	0
Bright orange lily flower	10	80	20
Orange lily flower	7	71.428	28.572
White light red carnation flower	12	100	0
Red carnation	32	100	0
White gardenia	8	100	0
White gerbera daisy	12	100	0
Total	123	89.6189	10.3811

Table (5) Results Of Testing State Using FRBF Algorithm

Type of samples	Testing dataset	Classification rate (CR)	False rate (FR)
	No. of samples		
White lilies flower single	10	100	0
White lilies flowers	12	100	0
Yellow black lily	14	92.857	7.143
Red white lily flower	6	100	0
Bright orange lily flower	10	90	10
Orange lily flower	7	85.714	14.286
White light red carnation flower	12	100	0
Red carnation	32	100	0
White gardenia	8	100	0
White gerbera daisy	12	100	0
Total	123	96.8571	3.1429

RBF and FRBF used the input dataset(flower images dataset). And the resolution of each image in training and testing state is (114*86). The number of the input nodes in RBF and FRBF are (114 nodes) in input layer, and 10 nodes in hidden layer, the output is 10 classes which are equal to number of nodes in output layer.

Table (6) shows the comparison between G-G, ABC, RBF, FRBF for 10 classes with over all classification rate that obtained for G-G is equal(93.3451). and for ABC is equal to (95.0237). the classification rate that was obtained for RBF is equal (89.6189) and for FRBF is the higher classification rate which is equal to (96.8571).

Table (6) Comparision Between (G-G, ABC, RBF, FRBF) Algorithms

Measures	G-G	ABC	RBF	FRBF
Classification rate (CR)%	93.3451	95.0237	89.6189	96.8571
False rate (FR)%	6.6549	4.9763	10.3811	3.1429

Figure (2) shows a sample of images, there are 10 categories, for a total of 116 training dataset and 123 testing dataset of images, different categories contain very different number of flowers

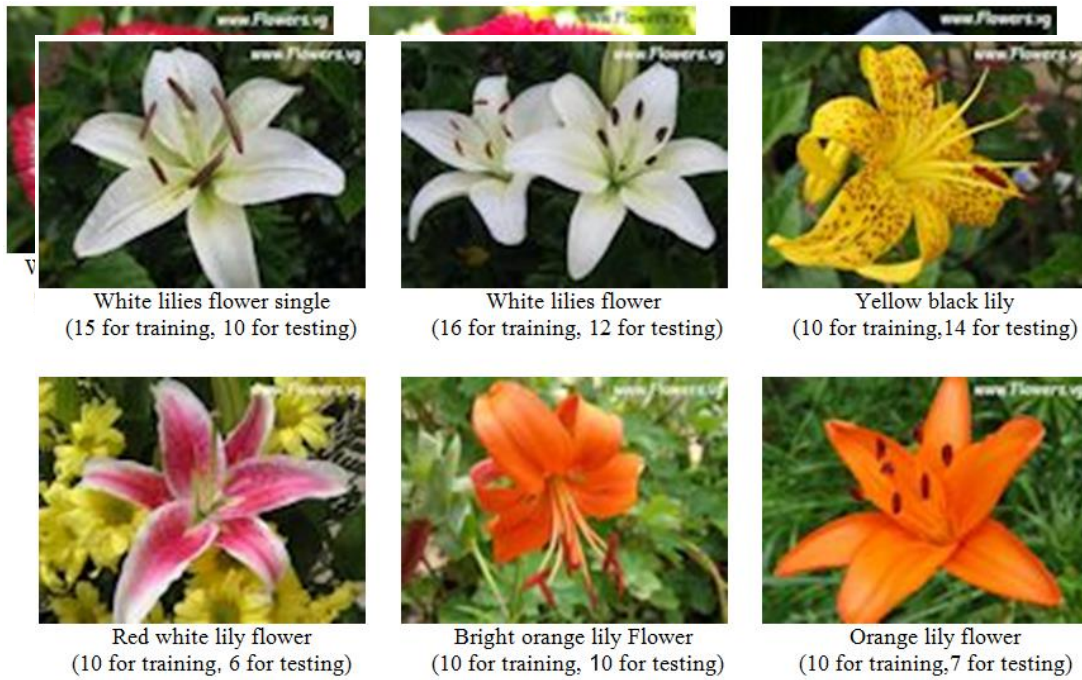


Figure 2. A sample of the flower image database, with the number of flowers in each category [30].

9- Conclusion

In this research, we presented some of artificial intelligence techniques such as G-G clustering algorithm, ABC algorithm, RBF neural network, and then we combined G-G method to RBF neural network to produce the proposed method called Fuzzy RBF (FRBF) neural network. In training state all these algorithms obtained high classification rate equal 100%. But in testing state, these algorithms obtained different classification and false rate, The proposed method FRBF has achieved high classification rate equal (96.8571), and low false rate equal (3.1429) on testing large number of flower images.

REFERENCES

- [1] Xu R. , (2005), "Survey Of Clustering Algorithms", IEEE Transaction On Neural Networks, Vol. 16, No. 3.
- [2] Kwok T., Smith K., Lozano S., And Taniar D., (2002), "Parallel Fuzzy C-Means Clustering For Large Data Sets", © Springer- Verlag Berlin Heidelberg, Pp. 365-374, Euro-Par (2002).
- [3] Grira N., Crucianu M., And Boujemaa N.,(2005), "Semi-Supervised Fuzzy Clustering With Pair Wise-Constrained Competitive Agglomeration", IEEE International Conference On Fuzzy Systems (Fuzzy IEEE).
- [4] Liu H., Jen B., Yih J., And Yu Y., ,(2009), " Fuzzy C-Means Algorithm Based On Standard Mahalanobis Distances" International Symposium On Information Processing (ISIP '09), Pp.422-427, 21-23.
- [5] Line J., And Wu S. , (2012), "Fuzzy Artificial Bee Colony System With Cooling Schedule For The Segmentation Of Medical Images By Using Of Spatial Information", Journal Of Applied Sciences, Engineering And Technology, 4(17): 2973-2980.
- [6] Thai L., Hai T., And Thuy N. ,(2012), " Image Classification Using Support Vector Machine And Artificial Neural Network", I. J. Information Technology And Computer Science, Vol. 5, Pp 32-38.
- [7] Akgundogdu A., Jennane R., Et Al.,(2010), " 3D Image Analysis And Artificial Intelligence For Bone Disease Classification", Journal Of Medical System, Vol. 34, No. 5, Pp 815-828.
- [8] Banerjee S., Bharadwaj A., Et Al. ,(2012), "Remote Sensing Image Classification Using Artificial BEE COLONY Algorithm", International Journal Of Computer Science And Informatics, Vol. 2, Issue 3, ISSN(PRINT) 2231-5292, Pp 67-72.
- [9] Arias S., Gomez H, Et Al.,(2009), " Satellite Image Classification By Self-Organized Maps On GRID Computing Infrastructures", Proceeding Of The Second EELA-2 Conference.
- [10] Rajendran P., And Madheswaran M.,(2010), " Hybrid Medical Image Classification Using Association Rule Mining With Decision Tree Algorithm", Journal Of Computing, Vol. 2, Issue 1, Pp 127- 136.
- [11] Zhang C., And Qiu F., (2012), " Hyperspectral Image Classification Using An Unsupervised Neuro-Fuzzy System",Journal Of Applied Remote Sensing, Vol.6.
- [12] Luping J., And Zhang Y.,(2007)," SVM- Based Fingerprint Classification Using Orientation Field", 3rd International Conference On Natural Computation, Vol. 2, Pp 724-727.
- [13] Raja S., Shanmugam A., (2010), " Wavelet Features Based War Scene Classification Using Artificial Neural Network", IJCSE International Journal On Computer Science And Engineering, Vol. 02, NO. 09, Pp 3033-3037.
- [14] Arpita D., And Mahua B., (2008), " GA Based Neuro Fuzzy Techniques For Breast Cancer Identification", Intl. Machine Vision & Image Processing Conf.,

Pp 136-141.

- [15] Sivaraman E., Arulselvi S., Natarajan S.,(2010),“ Modeling Of An Inverted Pendulum Based On Fuzzy Clustering Techniques”, International Journal Of Computer Applications, Volume 9, No. 4.
- [16] Xing Z., Zhang Y., Hou Y., And Jia L.,(2007), “ On Generating Fuzzy Systems Based On Pareto Multi-Objective Cooperative Coevolutionary Algorithm”, International Journal Of Control, Automation, And Systems, Vol. 5, No. 4, Pp. 444-455.
- [17] Parmar G., And Mitra S.,(2011), “ Effective Analysis Of Fuzzy Unsupervised Clustering Algorithms For Brain Tissue Segmentation In Single Channel MR Image ”, International Journal Of Bio-Science And Bio-Technology, Vol. 3, No. 2.
- [18] Karaboga D., And Ozturk C., (2010), “ Fuzzy Clustering With Artificial Bee Colony Algorithm”, Scientific Research And Essays Vol 5(14), Pp. 1899-1902.
- [19] Yan X., Zhu Y., Zou W., And Wang L.,(2012), “A New Approach For Data Clustering Using Hybrid Artificial Bee Colony Algorithm”, Neurocomputing 97, Pp. 241-250.
- [20] Ma M., Liang J., Guo M., Fan Y., And Yin Y.,(2011), “ SAR Image Segmentation Based On Artificial Bee Colony Algorithm”, Applied Soft Computing 11, Pp. 5205-5214.
- [21] Abdelhamid N., Halim M., And Fakhr M.,(2013), "Bees Algorithm-Based Document Clustering“, International Conference On Information Technology ICIT.
- [22] Castro L., And Zuben F.,(2001), “An Immunological Approach To Initialize Centers Of Radial Basis Function Neural Networks”, Proceeding Of V Brazilian Conference On Neural Networks, Pp. 79-84.
- [23] Sriram P.,(2005), “Developing Neural Network Applications Using Labview”, Thesis, University Of Missouri-Columbia.
- [24] Yang P., Zhu Q., And Zhong X.,(2009), “Subtractive Clustering Based RBF Neural Network Model For Outlier Detection”, Journal Of Computers, Vol. 4, No. 8.
- [25] Liu D. , Zhang H. , Polycarpou M.,(2011), " Advances In Neural Networks " , ISBN 978-3-642-21110-2, Springer Heidelberg Dordrecht, Newyork.
- [26] Olivas E. , And Guerrero J. ,(2010), " Handbook Of Research On Machine Learning Applications And Trends", ISBN 978-1-60566-766-9, IGI Global, Hershey , New York.
- [27] Chimphee W., Abdullah A., Sap M., Chimphee S., And Srinoy S., (2007), "A Rough-Fuzzy Hybrid Algorithm For Computer Intrusion Detection ", The International Arab Journal Of Information Technology, Vol.4, No.3.
- [28] Guan Y., Gorbani A., And Belacel N. , (2003), “Y-Means: A Clustering Method For Intrusion Detection”, IEEE.
- [29] Banerjee S., Bharadwaj A., Gupta D., And Panchal V., (2012), “ Remote Sensing Image Classification Using Artificial Bee Colony Algorithm”,

International Journal Of Computer Science And Informatics, Volume 2, Issue 3.

- [30] [Http//Www.Flowers.Vg/Flowers](http://www.Flowers.Vg/Flowers).
- [31] Samadzadegan F., And Ferdosi E., (2012), " Classification Of Polarimetric SAR Images Based On Optimum Svms Classifier Using Bees Algorithm", International Conference On Intelligent Computational Systems Pp. 106-111, Dubai.
- [32] Gomez J., And Dasgupta D. , (2002), "Evolving Fuzzy Classification For Intrusion Detection", IEEE, Proceeding Of The 2002 IEEE, Workshop On Information Assurnace, United State Military Academy, Pp 1-6.