Feature selection using cuckoo search algorithm for object classification

Nawal Murad Khatur[1] and Abdul Ameer Abdullah Karim[2]

[1]Department of computers,University of Technology,Baghdad,Iraq

[2]Department of computers,University of Technology,Baghdad,Iraq

nwalmurad@yahoo.com , 110004@uotechnology.edu.iq

Abstract

In any object classification system, there is a need to extract features and use them to classify the objects, most of the extracted features have succeed to classify some object but failed to classify others. Feature selection is a general problem used for dimensionality reduction purposes. Feature selection aims to select important features. The major objective of the this paper is to use the Binary Cuckoo Search(BCS) to select important features from the set of the extracted features. The feature extracted were grey level features (texture features), the size features, shape features and the optimal feature has been selected by BCS. Support vector machine (SVM) classifier used in BCS as a fitness function. The cuckoo search algorithm select only seven feature from (25) features where performed average accuracy 92% and improve the classification time from (1.55) second at average to (57) millisecond at average.

Keyword: object classification , Binary Cuckoo Search(BCS)**,** Feature Selection (FS)**,** K Fold Cross Validation

1.Introduction

The main objective of image analysis is to extract information that can be used for the purpose of solving application-based problems. This is achieved through intelligently decreasing the amount of image data by extracting features which representing the image under consideration . If we desire a system to distinguish objects of different types , we must first decide which features of that objects, will be measured.

Feature Selection (FS) for classification can be defined as a significant process used for the purpose of finding minimal subset of features from the original data through aiming to remove irrelevant and redundant data. This

process aims for improving the accurateness of classification, shorten computational time of classification algorithms, and reducing the complexity of classification model [1].

New meta-heuristic approach for the continuous optimization (Cuckoo Search (CS)), that depends on attractive reproduction approach of cuckoo birds has been proposed to be used in this thesis in order to perform the task of feature selection. It has been identified that the CS had improved results than results achieved via certain familiar nature-inspired optimization methods, like Genetic Algorithms GA and Particle Swarm Optimization (PSO). Today, CS is utilized in nearly all domains and area of FS, image processing, function optimization and additional applications used in real world [2]

2. Related work

1. **In 2013, Yildiz A.R.** has introduced the algorithm of CS as novel optimization approach that has used to solve manufacturing optimization problems.The efficiency of Cuckoo Search, milling optimization problem has been solved and the results have been put to comparison with the results acquired through the use of other known optimization approaches. The acquired results has demonstrated that Cuckoo Search was extremely robust and efficient method for the optimization [3].

2. In 2015, Medjahed S.A et.al, have shown that spectral band selection was significant and active research field in hyper spectral image classification. Hyper spectral images consist of irrelevant information and redundant measurements that decrease the accuracy rate of classification. The researchers used binary form of CS algorithm being more efficiency that PSO and GA. The suggested method used k-nearest neighbor classifier. The results indicated that CS algorithm has offered better accuracy rate when compared to other methods on 3 experiments [4].

3. **In 2017, Alia A. F. and Taweel A.,** have state that Rough set theory (RST) was effective approaches to feature selection, but it uses complete search to search for all subsets of features. The complete search is expensive and high cost for large data. Thus, CS algorithm was utilized for replacing reduction part in RST. This paper has developed a novel algorithm for FS

depending on hybrid Binary CS and rough set theory for classification on nominal datasets. The results showed that this algorithm has achieved better FS compared to Genetic and PSO algorithm[5].

4. In 2016, Abd El Aziz and Hassanien et.al, have presented modified CS algorithm with rough sets for handling high dimensionality data via FS. The new CS utilizes rough sets theory for building fitness function which is taken into consideration the classification quality and the number of features in reduce set. The suggested algorithm was validated and examined benchmark on various bench mark datasets. The learning algorithms (SVMs and k-NN) were used as classifier method. The results indicated that the suggested algorithm might considerably enhance the performance of classification[6].

3.Optimization Algorithms

Optimization can be defined as the process of searching for the best solutions to a certain task of concern, and this searching process may be performed with the use of multiple factors that fundamentally make up a system of developing factors. This system may be developed via iterations based on a group of mathematical equations or rules. As a result, this type of systems shows some developing properties, which leads to self-organizing states corresponding to some search space optima. As soon as self-organized states are obtained, it can be said that the system converges. Which is why, designing a sufficient algorithm of optimization corresponds to mimicking the progress of self-organizing systems [7].

3.1 Cuckoo Search (CS)

CS is a novel nature-inspired meta-heuristic algorithm. It has been proposed in the year of 2009 by Suash Deb and Xin-She Yang, according to the brood parasitism of some species of cuckoos as they lay their eggs in other birds' nests.

Cuckoos mimic pattern and color regarding their eggs for the purpose of matching that of their hosts, after that the host eggs will be evicted by the cuckoo chick. In the case when host bird notices that these eggs do not belong to them, it is going to leave the nest and creates a new one somewhere else or just through the foreign eggs [3,8,9].

Figure1: Cuckoo eggs mimic pattern and color of hosts eggs

3.2The rules for CS

For uncomplicatedness in explaining standard CS, the next 3 idealized rules could be applied [3,8,9]:

- Single egg at a time will be laid by each cuckoo, and then dumped in a nest that has been chosen at a random way.
- The optimum nests that have eggs of good-quality are going to be passed on to next generations.
- The number of existing host nests will be fixed, eggs that have been by cuckoo will be found via host bird with probability. In this event, host bird could leave the nest and create another one somewhere else or just throws away the eggs.

For the purpose of carrying out this view point, the next uncomplicated representations could be used, that a solution is represented via each egg in the nest. Clearly, the algorithm could be extended to extra difficult situation in which all nests have multiple eggs which represent set of solutions. For the current introduction, the easiest method will be used in which all nests have just one egg. In such situation, no distinction exists between eggs, cuckoo or nest, since all nests corresponds to single egg that is also representing single cuckoo [7,10].

3.3 Lévy Flights and Cuckoo Search

Animals explore environment in a random way and at every time they are looking for food. There is a possibility of using a random walk for approximating the path tracked via the animal, for the next step is based on the present location and the possibility of moving in a certain direction. The direction that is going to be followed could be fixed depending on the probability which could be modeled in a mathematical way. Researches have

indicated that the flight's behavior of certain insects and animals has certain correspondence with Lévy probability distribution [9].

Cuckoos behavior is combined in the CS with Lévy flights for the purpose of effectively search new nests. Pauly Lévy (French mathematician) has introduced Lévy flights, that is representing model of random walks specified via by their step lengths that follow a power-law distribution [11].

Let $x_i(t)$ represents i-th probable solution at t-th iteration. Cuckoo i generate new solution $x_i^{(t+1)}$ through Lévy flights, based on Eq. (1):

$$\boldsymbol{x_i^{(t+1)} = x_i^{(t)} + \alpha \oplus levy(s.\lambda)} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

Where α can be defined as the size of the step which follows Lévy distribution, *t* can be defined as the step size, and α > 0 is specified as the step size scaling factor/parameter. The entry wise product (multiplication) $\oplus$ is considered to be similar to those applied in particle swarm optimization, and $x_i^{(t+1)}$ represents (*t+1*) *th* egg (feature) at the nest i (solution), i=1,2..,m and t= 1,2,….,d and **s** can be defined as the step size from Lévy distribution [9, 10,11].

Lévy flights are mainly random walks with aleatory steps from Lévy distribution as indicated in Eq. (2):

$$\boldsymbol{levy(s.\lambda) \sim s^{\lambda}} \qquad \boldsymbol{(1 < \lambda <= 3)} \ \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

This being the case, cuckoo flight essentially signifies a process of random walk that follows distribution depending on power-law with heavy tail. Thus, cuckoo search is extra effective in discovering the search space since the steps length of the algorithm is considered to be considerably longer in the long run . Lastly, the nest that has eggs with least quality will be swapped to new ones depending on probability Pa in the range of [0,1]. [9, 10,11,12]

The parameters *Pa,* λ and α presented in cuckoo search are utilized for finding the locally and globally better solutions, respectively. Parameters pa and α are considered to be extremely significant in fine-tuning the solution vectors and could be utilized in correcting the algorithm's convergence rate. The standard cuckoo search algorithm utilizes fixed value for *α and Pa* [10].

In the present study, the parameters regarding cuckoo search are n (number of cuckoos, size of solutions that are implemented by the algorithm operators at each iteration), *Pa* (probability of a certain cuckoo abandoning the nest to create one somewhere else) and α (the step size of Lévy flight).

The Classical Cuckoo Search Algorithm

In this section, the main steps of the cuckoo search algorithm is presented in details as shown in Algorithm 1.

Algorithm 1 : Cuckoo search algorithm adapted from [5]
Start 1: initiate the value of the size of the host nest N, probability pa ∈ [0, 1] and maximal number of iterations Max_itr. 2: Set t = 0 {Counter initializing}. 3: **for** (i = 1 : i ≤ N) do Produce the initial population of N host x^t_i {N is size population}. 5: Perform an evaluation of the fitness function f(x^t_i). 6: **end for** 7: **repeat** 8: Produce a new solution (Cuckoo) x_i^{t+1} in a random manner via L´evy flight. 9: Perform an evaluation of the solution's fitness function) x_i^{t+1} f(x_i^{t+1}) 10: Randomly select a nest *xj* amongst N solutions. 11: **if** (f(x_i^{t+1}) > f(x^t_j)) 12: Replace solution x_j^{t+1} with solution x_i^{t+1} 13: **end if** 14: Abandon a fraction *Pa* of worse nests. 15: Generate new nests at new locations with the use of the L´evy flight a fraction *Pa* of worse nests 16: Save the best solutions (nests that have the quality solutions) 17: Give Ranking to solutions and obtain the current best one 18: Set *t = t + 1*. {Increasing the counter of the iteration }. 19: **until** ($t < Max_{itr}$). {Satisfying the criteria of termination }. 20: Produce the optimal solution.

4. The proposed method

the proposed method contains three stage which are preprocessing and features extraction and features selection using cuckoo search.

4.1 preprocessing

The preprocessing step contains three steps the first is converting the color image to a grey level image, the second is performing median filter on the image and the third is performing the histogram equalization on the grey level image.

4.1.1 Converting Color Image to Gray Level Image

In this step, the color image will be converted into gray level image where the reason beyond that is to discard the color component and preserve only the lightness component , so that dealing with one band (Lightness) is easier than dealing with three bands (R, G, B) and dealing with grey level image is much faster than dealing with color images.

4.1.2 Noise removal

Median filter has been used in this paper to remove noise. Median filter is employed because it works efficiently with salt-and-pepper noise , besides, this filter is preserve the image details (edge) which is required in the next stages of the proposed system.

4.1.3 Histogram Equalization

Histogram equalization is employed in this paper in order to enhance the quality of the image. Histogram equalization perform on the processed image. In this paper, it has been found that performing Histogram Equalization after Median filter gives better image quality.

4.2 Feature Extraction

The main aim of extract features is to generate features vector from each image in dataset. First, image must be segmented into object and background. Second, use number of measures to extract particular features from the object. The extracted features has been categorized in three categories includes the grey level features (texture features), the size features, shape features. We obtain (25) features from each images as show in Table1.

Table(1) : Summary of the extracted image features

Feature type	number of Feature	Description
size measurement	6	Object area ,object perimeter , object high and width , major axis, minor axis.
shape measurement	12	Rectangularity, Circularity, Compactness, Aspect Ratio, Eccentricity, Convex Area, Convex Perimeter, Roundness, Convexity ,Solidity, Sphericity, Elongation
gray level features	7	Mean ,Standard Deviation, Variance, Energy, Entropy, Skewness, Kurtosis

4.3 Feature Selection Using Cuckoo Search Algorithm.

From the feature extraction stage, ***twenty five*** features were extracted where this number of features is considered as a large number which has a negative impact on the classification speed, hence this large number of feature must be reduced (dimensionality reduction) in order to improve the classification performance. Cuckoo Search Algorithm is used to select the most important features which have a powerful effect on the classification operation. The input to cuckoo search algorithm is 25 extracted features while the output of cuckoo search algorithm is 7 selected features.

5. K Fold Cross Validation

The efficiency of the proposed system is evaluate by using K fold cross validation. The features vector has been randomly divided into K parts which have equal size approximately. Classifier is trained with K-1 parts and just one part for test. The average accuracy of support vector machine(SVM) classifier is calculated from the average of K fold. Table(2) illustrates Parameter used for BCS.

Table (2) illustrates Parameter used for BCS.

Parameter	Values
Number of nests(population)	426
Number of eggs	25
Number of iteration	100
P_a	0.3
K (fold cross validation)	5
A	1
λ	1.3

6. The Proposed System Specifications

In order to implement the proposed system, it is required both software and hardware. Below the hardware and software specification associated with the proposed system.

6.1 Software Specification

The proposed system has been built using VB.net programming languages, while the system interface is built using jave programming languages under Microsoft windows 7 ultimate.

6.2 Hardware Specification

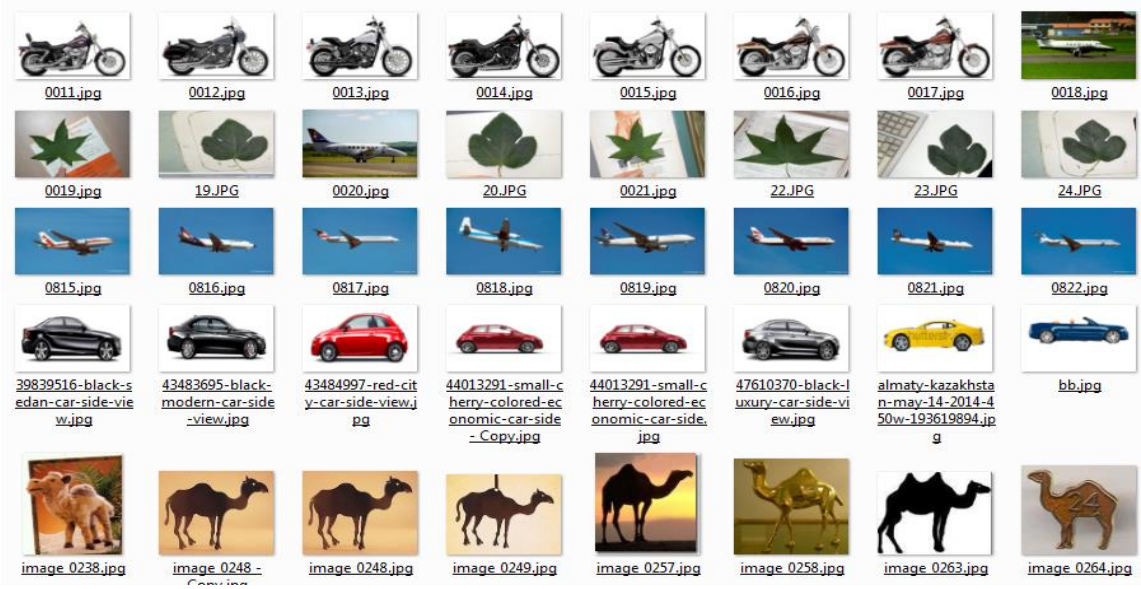
The computer which has been used to implement the algorithms of the proposed system has properties listed in table (3)

Table(3) computer properties

Processor	Pentium(R) Dual-Core CPU
Installed memory (RAM)	3.00 GB
System type	32-bit
Pan and touch	No pen or touch is available

7. Dataset

The dataset has been collected from various web sites. all the collected images has been resized to a unified size which is 896 × 592. The constructed Database contains 426 images of five classes: camel, leaves, cars, airplane and motorcycle, it contains 53 camel, 135 leaves, 120 motorcycle, 45 cars and 80 airplane. Figure(2) shows examples for this dataset.



Figure(2) shows examples for this dataset.

8. Results

The proposed feature selection method independent on support vector machine classifier select (7) features using cuckoo search algorithm. The classification time improved from (1.55) second at average to (57) millisecond at average and the average values of recall, precession and accuracy was (0.90 , 0.95, 0.92) respectively in case of using the reduce number of features (7) feature**.**

9. Conclusion

Cuckoo search used for select import features for each object, Binary cuckoo search algorithm (BCS) has been used to select (7) features from(25) features vector. K-fold cross validation can be considered as a powerful approach used in binary cuckoo search Network especially when the number of patterns in the training data is relatively small. In this paper the binary cuckoo search give a highest accuracy ratio of 92%.

References

[1] Alia A. F. , **2015,**"Modified Binary Cuckoo Search using rough set theory for Feature Selection", Birzeit University , MSc thesis.

[2] Yang X. S. , **2014**, "Cuckoo Search and Firefly algorithms, Theory and Applications", **Springer** International Publishing Switzerland.

[3] Yildiz A. R. , **2013**, " Cuckoo search algorithm for the selection of optimal machining parameters in milling operations ", International Journal of Advanced Manufacturing Technology , **Springer**.

[4] Medjahed S. A. , Saadi T. A., Benyettou A., and Ouali M. , **2015,** " Binary Cuckoo Search Algorithm for Band Selection in Hyperspectral Image Classification ", International Journal of Computer Science IAENG.

[5] Alia A. F. and Taweel A., **2017**, "Feature Selection based on Hybrid Binary Cuckoo Search and Rough Set Theory in Classification for Nominal Datasets " , Information Technology and Computer Science.

[6] Abd El Aziz M. and Hassanien A. E. , **2016**," Modified cuckoo search algorithm with rough sets for feature selection", **Springer**.

[7] Yang X. S. and Deb S. , **2014**, " Cuckoo Search: Recent Advances and Applications",Neural Computing and Applications, vol. 24, No. 1, pp. 169-174.

[8] Rodrigues D., Pereira L. A. M., Almeida T. N. S. and Papa J. P., **2013**, " BCS: A Binary Cuckoo Search Algorithm for Feature Selection", **IEEE**.

[9] Krishnammal P. M. and Raja S.S., **2015,** "Automated Brain Image classification using Neural Network Approach and Abnormality Analysis" , International Journal of Engineering and Technology (IJET).

[10] Gunavathi C. and Premalatha K., **2015,** "Cuckoo search optimization for feature selection in cancer classification: a new approach", International journal of data mining and bioinformatics, Vol. 13, No. 3.

[11] Pereira L. A. M., Rodrigues D., Almeida T. N. S., Ramos C. C. O. , Souza A. N., Yang X.-S. and Papa J. P., **2014**, " A Binary Cuckoo Search and its Application for Feature Selection ", **Springer** International Publishing Switzerland**.**

[12] Moghadasian M. and Hosseini S. P., **2014** "Binary cuckoo optimization algorithm for feature selection in high-dimensional datasets" , International conference on innovative engineering technologies (ICIET), pp. 18-21.