EXTRACTION OF BRAIN TUMORS FROM MRI IMAGES WITH ARTIFICIAL BEE COLONY BASED SEGMENTATION METHODOLOGY

Emrah Hancer[1], Celal Ozturk[2] and Dervis Karaboga[3]

Department of Computer Engineering, Erciyes University, Turkey
[1]emrahhancer@erciyes.edu.tr, [2]celal@erciyes.edu.tr and [3]karaboga@erciyes.edu.tr

Abstract

Image segmentation plays significant role in medical applications to extract or detect suspicious regions. In this paper, a new image segmentation methodology based on artificial bee colony algorithm (ABC) is proposed to extract brain tumors from magnetic reasoning imaging (MRI), one of the most useful tools used for diagnosing and treating medical cases. The proposed methodology comprises three phases: enhancement of the original MRI image (pre-processing), segmentation with the ABC based image clustering method (processing), and extraction of brain tumors (post-processing). The proposed methodology is compared and analyzed on totally 9 MRI images shooting in different positions from a patient with the methodologies based on K-means, Fuzzy C-means and genetic algorithms. It is observed from the experimental studies that the segmentation process with the ABC algorithm obtains both visually and numerically best results.

1. Introduction

Segmentation is one of the most significant requirements in the analysis of medical images. However, the complicated structures of the inside of the human body cause problems in segmentation i.e. prescribing appropriate therapy. Magnetic resonance imaging (MRI) is a technique primarily used in medical conditions to get high quality images of organs, soft tissues, bone and virtually all other internal human body structures. In other words, "MRI possesses good contrast resolution for different tissues and has advantages over computerized tomography (CT) for brain tissues due to its superior contrast properties [1]". On account of those advantages, MRI images have become a basic source of medical image segmentation, especially brain segmentation. Brain MRI segmentation is mainly applied to the following fields [2]: 1) automatic or semiautomatic diagnosis of regions to be treated prior to the surgery, 2) diagnosis of tumors before and after surgical intervention for response assessment, and 3) tissue classification. This paper particularly concerns with the second field.

Traditionally, segmentation of brain MRI images is processed manually by radiologists. However, manual segmentation is high time consumption and cause unavoidable mistakes. To sort out these problems, researchers proposed various segmentation methods based on thresholding [3], boundary detection [4], region growing [5], and clustering [6]. Thresholding, the process of separating the regions regarding to pixel intensity and color, is one of the simplest methods using for segmentation. Although thresholding methods can be efficiently applied to distinguish the objects from background when the histogram of objects and background is apparently distributed, these methods ignore all the spatial information of an image and cannot deal with the noise. Therefore they cannot reach adequate performance in segmentation of brain MRI images. Boundary detection methods try to find rapid pixel change between the regions at the boundary. That is achieved by gradient operator, e.g., Sobel and Prewitt operators. These methods are very dependent to the noise and describing edge pixel members as boundaries of regions is a challenging process. Region growing methods rely on the neighboring pixels having similar values within the region. To achieve the satisfactory results in region growing methods, the regions must be homogenously appeared in an image and it requires a user intervention to select the seed candidates and to determine the criteria of region growth [5]. Image clustering which is the process of partitioning patterns such as pixels into groups or clusters using similarity criterion is mostly used method in brain MRI segmentation. The most well-known clustering methods are K-means [7]. Fuzzy C-Means (FCM) [8]. K-means, one of the simplest clustering methods, is based on the similarity criterion such that Euclidean distance between the patterns and the cluster centroids and FCM, the fuzzy version of K-means, is based on updating the memberships over the cluster centroids. Although, they are very popular in brain MRI segmentation, they are very dependent to initial conditions i.e. the centroids or memberships should be determined in an efficient way to improve the performance of clustering quality. In addition, they can adversely affect from noise i.e. that is not convenient for FCM since brain MRI images include considerable unknown and uncertain noise [9]. To minimize the drawbacks of FCM, Shen et.al. [1] proposed an improved FCM considering the difference between the neighboring pixels in the image and the relative locations of neighboring pixels. That method therefore does not only consider pixel intensities, but also consider neighboring the pixel intensities and locations. The degree of neighboring attractions is defined by the parameters determined by simple artificial neural networks (ANN). Hence, the performance of the proposed method relies on these parameters.

In this paper, artificial bee colony based image segmentation methodology proposed to extract brain tumors from MRI images. The proposed methodology is constructed on three stages. In first stage, original image is enhanced and eliminated from noise by 2D 3x3 median filter. In second stage, artificial bee colony based image clustering method is applied to the enhanced image. In the last stage, the segmented image is converted into binary image using thresholding and then connected component labeling method is employed to extract brain tumor.

The organization of the paper is as follows. In Section 2, the artificial bee colony algorithm is presented. In Section 3, the proposed segmentation methodology is introduced and the algorithms applied for analysis are expressed. The experimental studies are demonstrated to compare the proposed algorithm with the others in Section 4 and finally, the paper is concluded.

2. Artificial Bee Colony Algorithm

The artificial bee colony (ABC) algorithm [10] proposed by Karaboga in 2005 is one of the most popular swarm intelligence algorithm and has been used in diverse real world problems such as clustering [6], neural network training [11], numerical problems [12] and etc. The model of the ABC algorithm is established on the honey bee swarm in such a way that the ABC algorithm simulates the foraging behaviours of honey bees such as employed bees, onlooker bees and scout bees in the hive. In this model, a food source is represented by a solution, the nectar amount of a food source is demonstrated by the quality of solution i.e. objective function value of the solution, and exploring new food sources refers with getting more qualified solutions.

Employed and onlooker bees explore new food sources in the neighborhood of current food sources. Onlooker bees start exploration process after the employed bees completed their process. The difference between the employed and onlooker bees on the way of applying the exploration process is that onlooker bees tend to search in the neighborhood of the food sources owning more nectar, but employed bees search in the neighborhood of their associated food sources. After the completion of the exploration process, scout bees control whether there is an exploited (abandoned) food source. If so, scout bees find new food sources with any nectar amount and replace them with the abandoned food sources. A general framework of the ABC algorithm is demonstrated in Fig. 1.
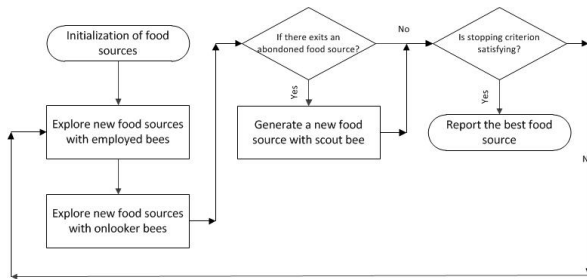

Fig. 1. General framework of the ABC algorithm

The new food source in the neighborhood of a current food source is explored by;

$$v_{ij} = x_{ij} + rand(-1,1)(x_{ij} - x_{kj}) \quad (1)$$

where j and k are randomly chosen parameter and neighborhood, and $rand(-1,1)$ is a random number between -1 and 1.

The food source selected for the exploration by onlooker bee is described by;

$$p_i = \frac{fit_i}{\sum_{i=1}^{SN} fit_i} \quad (2)$$

where fit_i is the fitness value of food source $i=1,...,SN$ and SN is the number of food sources.

The new food source produced by a scout bee is as follows;

$$x_{ij} = x_j^{min} + rand(x_j^{max} - x_j^{min}) \quad (3)$$

where $j=1,...,D$, and D is the number of parameters; x_j^{min} is the minimum and x_j^{max} is the maximum values of parameter j.

3. The Proposed Segmentation Methodology

The extraction of brain tumors from MRI images is handled in three stages such as pre-processing, processing and post-processing presented in Fig. 2. The detailed description of the proposed methodology is as follows;
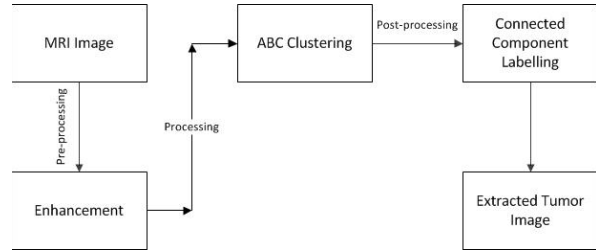

Fig. 2. A general framework of the proposed segmentation methodology

a) Pre-processing: In this stage, the main purpose is to enhance the input MRI image or eliminate noise in order to prepare it to the segmentation process since noise appeared on image might ruin segmentation quality. Various filters such as average, median and adaptive Gaussian filter etc. have been proposed to clear the image from unfavorable candidates. From all those filters, 2D 3x3 median filter is experimentally selected for the pre-processing stage.

b) Processing: In this stage, segmentation is applied to the enhanced image. For the segmentation, artificial bee colony based clustering method [13] is selected since the performance of that method is satisfactory. The objective function defined by Eq. 4 is employed as in [14];

$$fit(x_i, Z) = w_1 d_{max}(Z, x_i) + w_2(z_{max} - d_{min}(Z, x_i)) + w_3 J_e \quad (4)$$

where $x_i=\{m_{i,1}, m_{i,2},...,m_{i,k}...,m_{i,d}\}r$ is the solution represented by centroids, Z is the pixel intensities of image, z_{max} is the maximum pixel intensity, $d_{max}(Z, x_i)$ is the intra cluster separation defined by Eq. 5, $d_{min}(Z, x_i)$ is the inter cluster separation defined by Eq. 6, J_e is the quantization error described by Eq. 7, and w's are the degree of parameters such as $d_{max}(Z, x_i)$, $d_{min}(Z, x_i)$ and J_e, respectively.

$$d_{max}(Z, x_i) = max\left\{\sum_{\forall z_p \in C_{i,k}} \frac{d(z_p, m_{i,k})}{n_{i,k}}\right\} \quad (5)$$

where $n_{i,k}$ is the total number of pixels in cluster $C_{i,k}$, $d(z_p, m_{i,k})$ is the Euclidean distance between the pattern z_p and centroid $m_{i,k}$ inside the cluster $C_{i,k}$.

$$d_{min}(Z, x_i) = min\{d(m_{i,j}, m_{i,k})\}, j \neq k \quad (6)$$

$$J_e = \frac{\sum_{k=1}^{K} \sum_{\forall z_p \in C_k} d(z_p, m_k)/n_k}{K} \quad (7)$$

where K is the total number of clusters and n_k is the number of patterns in cluster C_k. The objective function evaluation of a solution is expressed in Fig. 3.

c) Post-processing: After segmentation, tumor extraction process is carried out. Firstly, segmented image is filtered if it includes a great amount of noise. After the elimination of noise, grey level thresholding is employed to convert that image to the binary image. Lastly, connected component labeling based on pixel connectivity is applied to extract the tumor region.

Fig. 3. The fitness evaluation of a solution

3. The Algorithms Used In the Study

To compare and analyze the proposed methodology, three algorithms such as K-means, Fuzzy C-means (FCM) and Genetic Algorithm (GA) are employed. This is because K-means and FCM is most widely used clustering algorithms and GA is the most well-known and the earliest population based algorithm.

3.1. K-means

K-means [7] proposed by MacQueen in 1967 iteratively partitions the dataset into K clusters using similarity measure (Euclidean Distance, cosine, Hamming and etc.). The algorithm tries to minimize the following objective function;

$$J = \sum_{k=1}^{K} \sum_{\forall z_p \in C_k} d(z_p, m_k) \qquad (8)$$

In K-means, the number of clusters K in which the data set will be partitioned is firstly predetermined. Secondly, cluster centroids are randomly initialized from dataset. Thirdly, each pattern in dataset is assigned to the cluster which has the closest centroid and then the positions of K centroids are recomputed. The second and third steps are repeated until there is no improvement in the positions of centroids.

3.2. Fuzzy C-Means

Fuzzy C-Means (FCM) developed by Dunn [15] in 1973 and improved by Bezdek [16] in 1981 allows patterns of dataset belong to two or more clusters. The FCM algorithm is based on minimizing the following objective function;

$$J_m = \sum_{p=1}^{N} \sum_{k=1}^{K} u_{pk}{}^{\mu} d(z_p, m_k) \qquad (9)$$

where N is the total number of patterns, u_{pk} is the membership degree of pattern z_p and μ is a number greater than 1.

FCM consists of four steps. In the first step, the membership matrix U_{NxK} is initialized. In the second step, the cluster centroids are calculated by Eq. 10. In the last step, the membership matrix is updated by Eq. 11. Second and last steps are repeated until the improvement between the updated and old membership matrix is smaller than € determined by a user.

$$m_j = \frac{\sum_{p=1}^{N} u_{pj}{}^{\mu} z_p}{\sum_{p=1}^{N} u_{pj}{}^{\mu}} \qquad (10)$$

$$u_{pj} = \frac{1}{\sum_{k=1}^{K} \left(\frac{d(z_p, m_j)}{d(z_p, m_k)}\right)^{\frac{2}{1-\mu}}} \qquad (11)$$

3.3. Genetic Algorithm

Genetic algorithm (GA) [17] proposed by Holland in 1975 is based on natural selection and evolution. In GA, each individual known as chromosome represents a solution of the handled problem. A chromosome comprises bit strings called as genes. The main purpose is to evolve chromosomes through crossover and mutation operators. To evolve chromosomes, new generation, called offspring, are generated by merging two chromosomes using a crossover operator and by modifying bit string of a chromosome using a mutation operator. In this way, the crossover operator satisfies cooperation and the mutation operator satisfies diversity in population. Then, new population is selected from the current and generated chromosomes according to their fitness values.

4. Experimental Studies

To verify the effectiveness of the proposed methodology, nine MRI images called as mri70-mri150 (256x256 pixels) from a patient are chosen, and the K-means, FCM and GA methods are employed. The methods are evaluated in terms of both numerical and visual results. Numerically, the methods are analyzed by calculating CS measurement [18] and the Turi's VI index [19] which are the clustering validity indexes to measure the clustering quality, defined by Eq. 12 and 13, respectively. Visually, the methods are compared by using the output images of extracted tumors, illustrated in Fig. 4. The visual results of GA is not presented in Fig 4 since it is suggested from the visual experiments that it cannot be possibly to illustrate apparent differences between the output images of the GA and ABC algorithms.

$$CS = \frac{\sum_{k=1}^{K} \frac{1}{n_k} \sum_{z_p \in C_k} \max_{z_q \in C_k} d(z_p, z_q)}{\sum_{k=1}^{K} \min_{k \neq kk} d(m_k, m_{kk})} \qquad (11)$$

$$VI = (c \times N(2,1) + 1) \times \frac{intra}{inter} \qquad (12)$$

where c is a constant, N(2,1) is the Gaussian distribution, and the other components of Eq. (12) are given as follows;

$$intra = \frac{1}{N} \sum_{k=1}^{K} \sum_{z_p \in C_k} d(z_p, m_k)^2 \qquad (13)$$

$$inter = \min\{d(m_k, m_{kk})^2\} \, where \, k \neq kk \qquad (14)$$

The following parameters are used for the ABC, GA, FCM and K-means methods: the number of food sources is taken 50, the value of limit is set to 100 for ABC; and the number of chromosomes is taken 50 for GA. The other parameters of GA such as crossover rate and mutation rate are chosen as 0.8 and 0.2, respectively. The maximum number of iterations (cycles) for ABC and GA is set to 250. Thus, the number of iterations of the K-means and FCM methods is set to 12500. The μ parameter of FCM which determines the level of cluster fuzziness is set to 2. The number of clusters is selected as 3 for the image mri70,

mri80 and mri90, and for the other images, the number of clusters is selected as 4.

The results of VI and CS indexes are demonstrated in Table 1 from which that can be inferred the ABC based segmentation method outperforms the others almost in all cases. In fact, FCM and K-means generally cannot obtain the appropriate results against the ABC and GA based segmentation methods. When particularly comparing the performances of the ABC and GA based segmentation methods, the ABC based segmentation method is superior to the GA based segmentation method in terms of obtained VI values, although the ABC based segmentation method cannot obtain the best results against GA in four problems in terms of obtained CS values. Moreover, K-means has the worst performances when comparing to the others. Therefore, the numerical performance of the methods such as ABC, GA, FCM and K-means can be ranked as first, second, third and last positions, respectively.

Table 1. The results of CS and VI indexes

	ABC		GA		FCM		K-means	
	CS	VI	CS	VI	CS	VI	CS	VI
mri 70	**0.61** (0.01)	**0.16** (0)	0.62 (0.06)	0.18 (0.02)	0.81 (0)	0.22 (0)	0.81 (0.01)	0.22 (0)
mri 80	**0.58** (0)	**0.16** (0)	0.60 (0.05)	0.17 (0.02)	0.81 (0)	0.20 (0)	0.83 (0.01)	0.21 (0)
mri 90	**0.51** (0)	**0.16** (0)	0.52 (0.02)	0.17 (0.02)	0.67 (0)	0.19 (0)	0.70 (0)	0.20 (0)
mri 100	0.60 (0.02)	**0.08** (0.03)	**0.59** (0.04)	0.10 (0.02)	0.91 (0)	0.13 (0)	0.97 (0.03)	0.12 (0.01)
mri 110	0.70 (0.01)	**0.08** (0.01)	**0.68** (0.05)	0.11 (0.03)	0.89 (0)	0.10 (0)	1.02 (0.03)	0.13 (0)
mri 120	0.64 (0.01)	**0.07** (0)	**0.59** (0.04)	0.10 (0.02)	0.69 (0)	0.11 (0)	0.71 (0.02)	0.12 (0)
mri 130	**0.56** (0)	**0.10** (0)	0.61 (0.06)	0.11 (0.02)	0.76 (0)	0.11 (0)	0.80 (0.01)	0.14 (0.01)
mri 140	0.67 (0)	**0.07** (0)	**0.59** (0.05)	0.11 (0.03)	0.74 (0)	0.10 (0)	0.77 (0)	0.11 (0)
mri 150	**0.53** (0.02)	**0.09** (0.01)	0.56 (0.03)	0.11 (0.03)	0.69 (0)	0.12 (0)	0.73 (0.01)	0.13 (0)

In addition to the results of the VI and CS indexes, the success of ABC and GA can be also clearly illustrated in Fig. 4 and 5 in which the results of the mri70, mri80, mri110 and mri120 images are appeared. Fig. 4 and 5 demonstrate that the ABC based image segmentation method obtains the most meaningful visual results where the K-means and FCM based segmentation methods cannot separate the tumor regions from the other parts.

5. Conclusions

In this paper, the artificial bee colony based image segmentation methodology is proposed to extract the region of tumors from MRI images. The proposed methodology is analyzed and compared with the K-means, FCM, and GA based image segmentation methodologies. The both visual and numerical results indicate that the ABC based image segmentation methodology outperforms the others and can be efficiently used in tumor segmentation of MRI images. The future work is to improve the performance of the proposed clustering methodology.

6. References

[1] S. Shen, W. Sandham, M. Granat, and A. Sterr, "MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization," *Trans. Info. Tech. Biomed.,* vol. 9, pp. 459-467, 2005.

[2] M. Forouzanfar, N. Forghani, and M. Teshnehlab, "Parameter optimization of improved fuzzy c-means clustering algorithm for brain MR image segmentation," *Engineering Applications of Artificial Intelligence,* vol. 23, pp. 160-168, 2010.

[3] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron. Imaging.,* vol. 13, pp. 146-168, 2004.

[4] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*: Prentice-Hall, 2002.

[5] R. Adams and L. Bischof, "Seeded region growing," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 16, pp. 641-647, 1994.

[6] D. Karaboga and C. Ozturk, "A novel clustering approach: Artificial Bee Colony (ABC) algorithm," *Applied Soft Computing,* vol. 11(1), pp. 652-657, 2011.

[7] J. MacQueen, " Some methods for classification and analysis of multivariate observations," presented at the Proc. 5th Berkeley Symp. Math. Stat. Probability, 1967.

[8] J. C. Bezdek, "Cluster validity with fuzzy sets," *Journal of Cybernetics,* vol. 3, pp. 58–72, 1974.

[9] M. G. H. Omran, A. Salman, and A. P. Engelbrecht, "Dynamic clustering using particle swarm optimization with application in image segmentation," *Pattern Analysis and Applications,* vol. 8, pp. 332-344, Feb 2006.

[10] 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, 2007.

[11] D. Karaboga and C. Ozturk, "Neural Networks Training By Artificial Bee Colony Algorithm On Pattern Classification," *Neural Network World,* vol. 19, pp. 279-292, 2009.

[12] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," *Applied Soft Computing,* vol. 8, pp. 687-697, 2008.

[13] E. Hancer, C. Ozturk, and D. Karaboga, "Artificial Bee Colony Based Image Clustering," presented at the IEEE Congress on Evolutionary Computation, CEC 2012, Brisbane, Australia, 2012.

[14] M. Omran, "Particle Swarm Optimization Methods for Pattern Recognition and Image Processing," University of Pretoria, Environment and Information Technology, 2004.

[15] J. C. Dunn, "Well separated clusters and optimal fuzzy partitions," *J. Cybern.,* vol. 4, pp. 95–104, 1974.

[16] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algoritms*. New York: Plenum Press, 1981.

[17]   J. Holland, "Genetic algorithms," *Scholarpedia,* vol. 7, p. 1482, 2012.

[18] C.-H. Chou, M.-C. Su, and E. Lai, "A new cluster validity measure and its application to image compression," *Pattern Analysis and Applications,* vol. 7, pp. 205-220, 2004.

[19]R. H. Turi, "Clustering-Based Colour Image Segmentation," PhD Thesis, Monash University, Australia, 2001.

Fig 4. First column is the original images; second, third and last columns are the segmentation results of ABC, K-means and FCM, respectively.
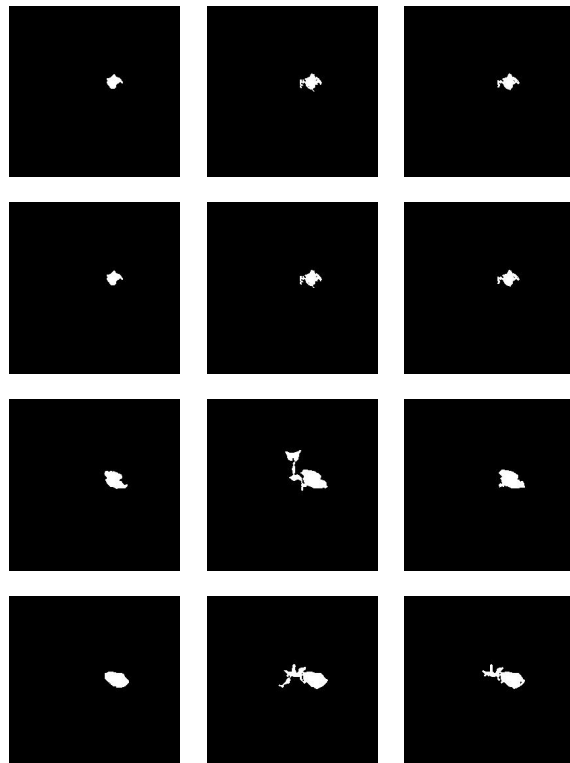

Fig 5. First, second and last columns are the extraction results after segmentation process of ABC, K-means and FCM, respectively.