New Method to Optimize Initial Point Values of Spatial Fuzzy c-means Algorithm

Iman Omidvar Tehrani*, Subariah Ibrahim*, Habib Haron*
* Departement of Computing, Universiti Teknologi Malaysia, Johor, Malaysia

Article Info	ABSTRACT
	Fuzzy based segmentation algorithms are known to be performing well on medical images. Spatial fuzzy C-means (SFCM) is broadly used for medical image segmentation but it suffers from optimum selection of seed point initialization which is done either manually or randomly. In this paper, an enhanced SFCM algorithm is proposed by optimizing the SFCM initial point values. In this method in order to increasing the algorithm speed first the approximate initial values are determined by calculating the histogram of the original image. Then by utilizing the GWO algorithm the optimum initial values could be achieved. Finally By using the achieved initial values, the proposed method shows the significant improvement in segmentation results. Also the proposed method performs faster than previous algorithm i.e. SFCM and has better convergence. Moreover, it has noticeably improved the clustering effect.

Corresponding Author:

Habib Haron,
Departement of Computing,
Universiti Teknologi Malaysia,
UTM Skudai, Johor, Malaysia.
Email:habib@utm.my

1. INTRODUCTION

The The process of dividing the original image into multiple meaningful regions in such a way that there is no intersection among them is called image segmentation. Segmentation is usually difficult to be done because of region inhomogeneity, blurred region boundaries and noise. Many image segmentation algorithms have been published over the past decade [1-5]. These algorithms have been used in applications such as medical imaging analysis, object classification and image retrieval [6]. Generally these algorithms are categorized into four groups including edge detection, clustering, region based and thresholding. clustering based algorithms play an important role in many applications [7] such as medical-image processing [8]. These algorithms divide the objects or patterns into different groups in such a way that the samples in a group are more similar to each other than the samples of other group.

K-means also known as hard c-means is one of the good examples of clustering based algorithms [9, 10]. The image pixels are grouped into segments based on their intensity level in such a way that each pixel can be only in one segment. Fuzzy c-means (FCM) which was first presented by Dunn [11] is considered as another approach for image clustering. It was used by Bezdak [12] as the universal classification based algorithm. The FCM algorithm classifies image pixels into groups however each pixel can be associated into multiple groups based on a degree of membership. FCM has proven to have outstanding performance on various medical applications [13, 14].

In spite of good performance of FCM, there are still some weaknesses in the algorithm such as the sensitivity to noise and the lack of a good strategy for the initial seed point placement [15]. Chuang et al. [14] proposed Spatial FCM (SFCM) algorithm to overcome the FCM noise issue by modifying the FCM objective function and taking the advantage of the neighborhood pixels. Regardless of the FCM noise issue which is

enhanced by SFCM technique, initialization step of the algorithm also plays a principal role on the quality of segmented images and standard SFCM algorithm fails to have a proper strategy for this case [15]. In order to overcome the initialization issue of FCM, recently a hybridization algorithm of particle swarm optimization (PSO) which is a technique of population based clustering with FCM namely (PSO+FCM) was proposed by researchers [7, 15-19]. Zhang et al. [16] used PSO as an initialization step in possibilistic c-means clustering (PCM) [17] to locate the best initial positions of cluster centers.

In order to enhance the spectral characteristics of features for clustering, Liu Hanli et al [19] used the PSO-FCM on the image data to enhance the accuracy of wetland extraction. Farhad et al [7] used PSO-FCM with four iterations to the particles in the swarm for every eight generations such that the fitness value of each particle was improved. The result of using PSO-FCM on hyperspectral data, in two spaces data and feature showed its higher ability in segmentation than fuzzy clustering [7].

A hybrid fuzzy clustering algorithm that incorporates FCM into Quantum-behaved PSO namely QPSO+FCM algorithm was proposed by Wang et.al [18]. The QPSO has less parameters and higher convergent capability of the global optimizing than PSO algorithm. So the iteration algorithm was replaced by the QPSO based on the gradient descent of FCM, which makes the algorithm have a strong global searching capacity and avoids the local minimum problems of FCM and in a large degree avoid depending on the initialization values.

Although it is possible to find the optimum initial position values by PSO and QPSO but the resulted values are not always optimized because both of them are trapped into local optimal solution and fail to find the global best value [20, 21]. The problem of trapping in local solution was fixed by gray wolf algorithm (GWO) [22]. The GWO algorithm is able to provide very competitive results compared to PSO, gravitational search algorithm (GSA), differential evolution (DE), evolutionary programming (EP), and evolution strategy (ES) meta-heuristics.

In this research, GWO is utilized instead of PSO to find the global best initial seed positions of SFCM algorithm for MRI image segmentation. Moreover, the histogram of image is used to increase the convergence speed. The rest of the paper is organized as follows. In Section 2, clustering based algorithms including FCM and SFCM are introduced. The GWO as an optimization algorithm is presented in section 3. In Section 4, the proposed algorithm which is based on improving the SFCM using GWO is presented. The results of the proposed method are presented in Section 5. Finally, we have drawn the conclusion in Section 6.

2. CLUSTERING BASED ALGORITHMS

FCM is known as the most common partitioning method [23]. Assuming that $X=\{i_1,\ldots,i_N\}$ be a set of image pixels, FCM divides these pixels based on a degree of membership (μ) by calculating the cluster centers $\{v_1,\ldots,v_C\}$ and minimizing the following sum of squared objective function:

$$J_{FCM} = \sum_{m=1}^{C} \sum_{n=1}^{N} \mu_{mn}{}^{l} ||i_n - v_m||^2 \qquad (1)$$

Where l is a variable which is greater than 1 and it controls the level of fuzziness of the segmentation result; C and N are the total number of regions and image pixels respectively. The following conditions must be satisfied in the FCM objective function.

$$\sum_{m=1}^{C} \mu_{mn} = 1 ; \qquad 0 \leq \mu_{mn} \leq 1 ; \qquad \sum_{n=1}^{N} \mu_{mn} > 0 . \qquad (2)$$

The membership functions μ_{mn} and the centroids v_m are updated iteratively by the following two equations:

$$\mu_{mn} = \frac{||i_n - v_m||^{-2/(l-1)}}{\sum_{k=1}^{C} ||i_n - v_k||^{-2/(l-1)}} \qquad (3)$$

$$V_i = \frac{\sum_{n=1}^{N} \mu_{mn}^{l} i_n}{\sum_{n=1}^{N} \mu_{mn}^{l}} \qquad (4)$$

When the pixels that are close to their centroids are having high membership values and those that are far away are having low values, it can be concluded that the FCM algorithm is optimized.

Although FCM algorithm performs well however the lack of spatial information is one of the weaknesses of FCM [13, 14, 24]. By taking into account spatial information, it is possible to make FCM robust against image artifacts and noise.

Spatial FCM (SFCM) was proposed by Chuang *et al.* [14]. This algorithm has modified FCM objective function to include spatial information by following equations:

$$\mu'_{mn} = \frac{\mu^p_{mn} h^q_{mn}}{\sum^C_{k=1} \mu^p_{kn} h^q_{kn}} \tag{5}$$

Where, *p* and *q* are controlling the level of fuzziness and the level of effect of the spatial information respectively. The variable h_{mn} which includes spatial information can be calculated by the following equation:

$$h_{mn} = \sum_{k \in N_n} \mu_{nk} \tag{6}$$

where N_n indicates a local window which is placed around the image pixel *n*. The two other variables μ_{mn} and v_m are updated iteratively based on Eqs. (3) and (4).

3. GREY WOLF OPTIMIZATION (GWO)

GWO is a meta-heuristic optimization algorithm inspired by grey wolves (Canis lupus) [22]. The GWO algorithm mimics the leadership hierarchy and hunting mechanism of grey wolves in nature. In this algorithm the population is divided into four groups: alpha (α), beta (β), delta (δ), and omega (ω). The first three fittest wolves are considered as α, β, and δ who guide other wolves (ω) toward promising areas of the search space. During optimization, the wolves update their positions around α, β, or δ as follows:

$$\vec{D} = |\vec{C} \cdot \overrightarrow{X_p}(t) - \vec{X}(t)| \tag{7}$$

$$\vec{X}(t+1) = \overrightarrow{X_p}(t) - \vec{A} \cdot \vec{D} \tag{8}$$

Where *t* indicates the current iteration, $\vec{A} = 2a.r_1 a$, $\vec{C} = 2.r_2$, $\overrightarrow{X_p}$ is the position vector of the prey, $\vec{X}$ indicates the position vector of a grey wolf, *a* is linearly decreased from 2 to 0, and r_1, r_2 are random vectors in [0,1]. The concepts of position updating using Eqs. (7) and (8) are illustrated in Figure1.
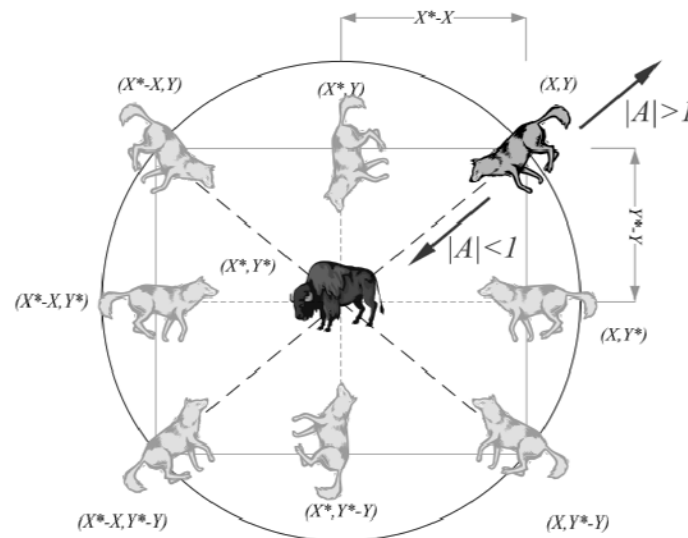

Figure 1. Position updating mechanism of search agents and effects of *A* on it [25]

It may be seen in this figure that a wolf in position (*X*, *Y*) is able to relocate itself around the prey with the proposed equations. Although seven of possible locations have been shown in Figure 1, the random parameters *A* and *C* allow the wolves to relocate to any position in the continuous space around the prey.

In the GWO algorithm, it is always assumed that *α*, *β*, and *δ* are likely to be the position of the prey (optimum). During optimization, the first three best solutions obtained so far are assumed as *α*, *β*, and *δ* respectively. Then, other wolves are considered as *ω* and able to re-position with respect to *α*, *β*, and *δ*. The mathematical model proposed to re-adjust the position of *ω* wolves are as follows:

$$\overrightarrow{D_\alpha} = |\vec{C_1} \cdot \overrightarrow{X_\alpha} - \vec{X}| \tag{9}$$

$$\overrightarrow{D_\beta} = |\vec{C_2} \cdot \overrightarrow{X_\beta} - \vec{X}| \tag{10}$$

$$\overrightarrow{D_\delta} = |\overrightarrow{C_3} \cdot \overrightarrow{X_\delta} - \vec{X}| \tag{11}$$

where X_α shows the position of the *α*, X_β shows the position of the *β*, X_δ is the position of *δ*, C_1, C_2, C_3 are random vectors and *X* indicates the position of the current solution.

Equations (9-11) calculate approximate distance between the current solution and alpha, beta, and delta respectively. After defining the distances, the final position of the current solution is calculated as Equations (12-15):

$$\overrightarrow{X_1} = \overrightarrow{X_\alpha} - \vec{A_1} \cdot \left(\overrightarrow{D_\alpha}\right) \tag{12}$$

$$\overrightarrow{X_2} = \overrightarrow{X_\beta} - \overrightarrow{A_2} \cdot \left(\overrightarrow{D_\beta}\right) \tag{13}$$

$$\overrightarrow{X_3} = \overrightarrow{X_\delta} - \overrightarrow{A_3} \cdot \left(\overrightarrow{D_\delta}\right) \tag{14}$$

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{15}$$

Where X_α shows the position of the alpha, X_β shows the position of the beta, X_δ is the position of delta, A1, A2, A3 are random vectors, and t indicates the number of iterations. As may be seen in these equations, the equations (9-11) define the step size of the ω wolf toward α, β, and δ respectively. The equations (12-15) then define the final position of the ω wolves. It may also be observed that there are two vectors: A and C. These two vectors are random and adaptive vectors that provide exploration and exploitation for the GWO algorithm.

As shown in Figure 1, the exploration occurs when A is greater than 1 or less than -1. The vector C also promotes exploration when it is greater than 1. In contrary, the exploitation is emphasized when |A|<1 and C<1. It should be noted here that A is decreased linearly during optimization in order to emphasize exploitation as the iteration counter increases. However, C is generated randomly throughout optimization to emphasize exploration/exploitation at any stage, a very helpful mechanism for resolving local optima entrapment.After all, the pseudo code of the GWO algorithm is presented in Figure 2:

Initialize the grey wolf population $X_i=(i=1,2,…,n)$
Initialize a, A and C
Calculate the fitness of each search agent
X_α= the best search agent
X_β = the second best search agent
X_δ = the third best search agent
While (t<Max number of iterations)
   For each search agent
       Update the position of the current search agent by equation (15)
   End For
   Update a, A and C
   Calculate the fitness of all search agents
   Update X_α, X_β, X_δ
   t=t+1
End While
Return X_α

Figure 2. Pseudo code of GWO algorithm

The exploration of this algorithm is very high and requires it to avoid local optima. Moreover, the balance of exploration and exploitation is very simple and effective in solving challenging problems as per the results of the real problems in [22].

4.   Proposed Algorithm
The goal of the proposed algorithm is to segment the given MRI brain image into three regions including White Matter (WM), Gray Matter (GM) and Cerebrospinal Fluid (CSF). In this algorithm, first the original image as shown in Figure 3 is provided to the algorithm then the number of search agents and the maximum number of iterations are initialized.
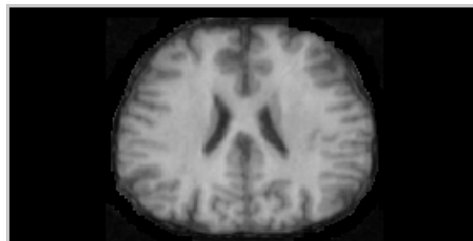

Figure 3. The original MRI image

The flowchart of proposed algorithm is depicted in Figure 5. In this method, in order to speed up the optimization, the search agents are initialized by approximating the initial seed points which can be calculated by our proposed histogram based method. The histogram of the original image is shown in Figure 4.
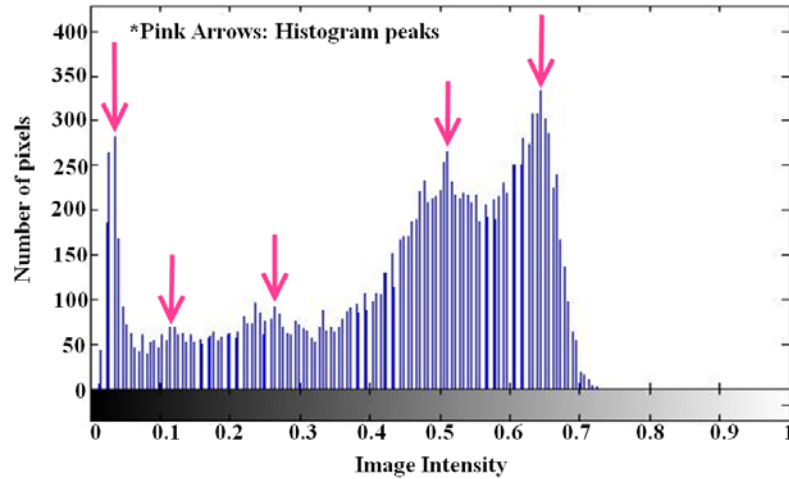
Figure 4. Histogram of the original image and its peaks

The horizontal axis shows the intensity of the image and the vertical axis shows the total number of pixels having the intensity presented in the horizontal vector in Figure 4. Finally as shown by pink arrows in Figure 4, the peaks of the histogram are located and the intensity of the peaks can be considered as potential seed point values and the combination of these values is used as the initial value for each search agent.Now, the fitness function that is going to be used by GWO is achieved by calculating the Euclidean distance between centroid points and all pixels for each region. The result is a decimal number that is the lesser the better.

$$F = \sum_{n=1}^{N} \sum_{m=1}^{C} ||i_n - v_m||^2 \tag{16}$$

In Equation 16, N is the total number of pixels, C is the number of regions in the original image, i_n is the intensity of nth pixel and v_m is the intensity of initial point of the mth region.

After calculating the fitness function, the positions of search agents are updated by GWO algorithm. The process of updating search agent positions and fitness function calculation are repeated iteratively until the maximum number of iterations is reached or there is no significant improvement that can be calculated based on the following equation:

$$|F_{t+1} - F_t| < \varepsilon \tag{17}$$

Where, F_t is the fitness of the best achieved solution in t^{th} iteration and ε less than 0.01. Finally, the optimized initial points are achieved and they are used directly in SFCM algorithm to perform the segmentation procedure.
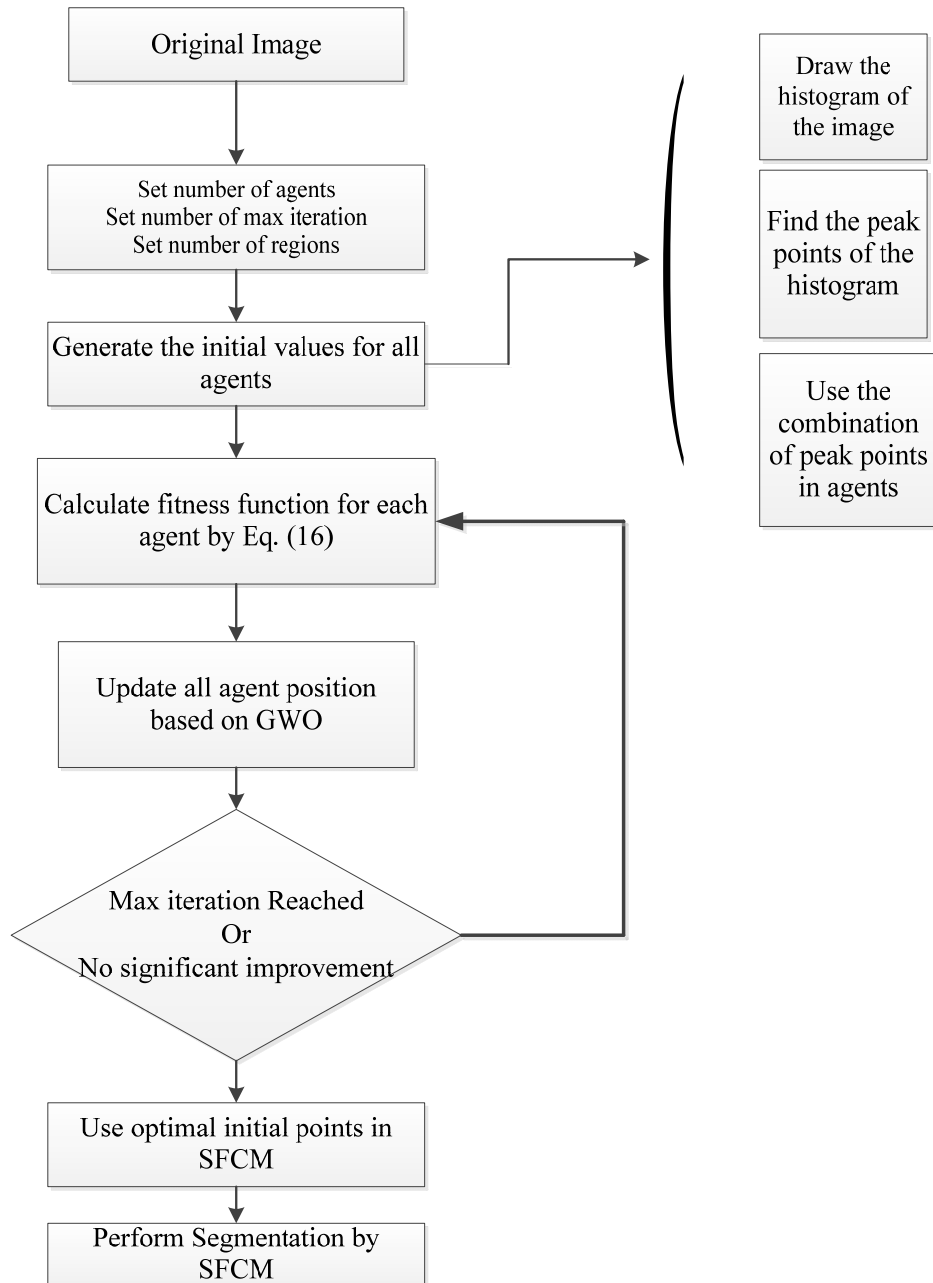
Figure 5. The flowchart of the proposed algorithm

5.   RESULTS AND ANALYSIS

In In order to verify the proposed algorithm, thirty search agents are utilized to find the optimum result with maximum of fifteen iterations. The input to the algorithm is a brain MRI image as shown in Figure 2 and the output is three intensity values where each of them represents one region of the brain. Finally in order to evaluate the proposed method, three datasets from ISBR [26] have been used. T1 weighted images are selected for the evaluation purpose because they have better white matter / gray matter (WM/GM) intensities [27].

After performing the proposed method on the original image in Figure 2, the following optimum initial points (V_1, V_2 and V_3) were obtained where each represents one region of the brain:

v_1=0.49227, v_2=0.24708, v_3=0.6635

Those optimum initial values are then used in the standard SFCM algorithm and the segmentation results are shown in Figure 6. The segmented region is shown as white color in the segmentation results.

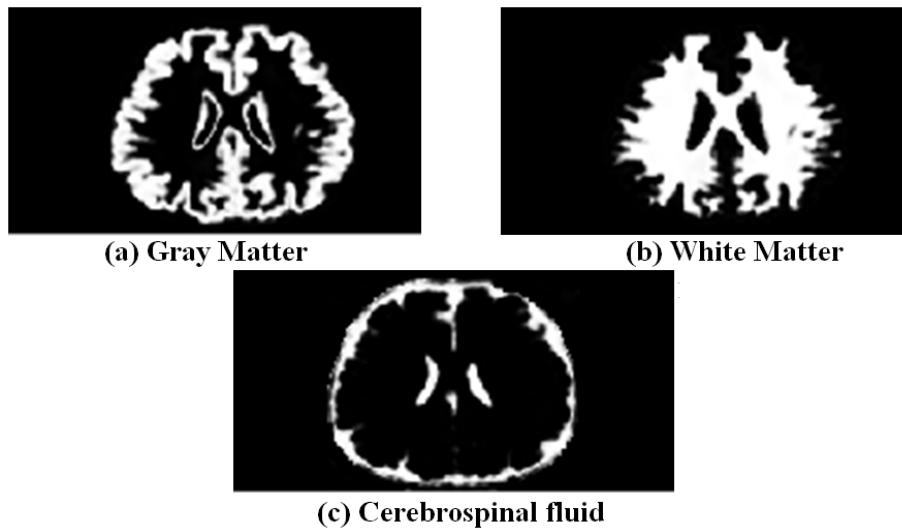**(a) Gray Matter**   **(b) White Matter**

(c) Cerebrospinal fluid

Figure 6. Segmentation result of the proposed SFCM-GWO algorithm

The convergence curve of the proposed method is shown in Figure 7. According to the diagram, it can be observed that the error rate which is calculated by the fitness function has reached its minimum value after fifteen iterations.
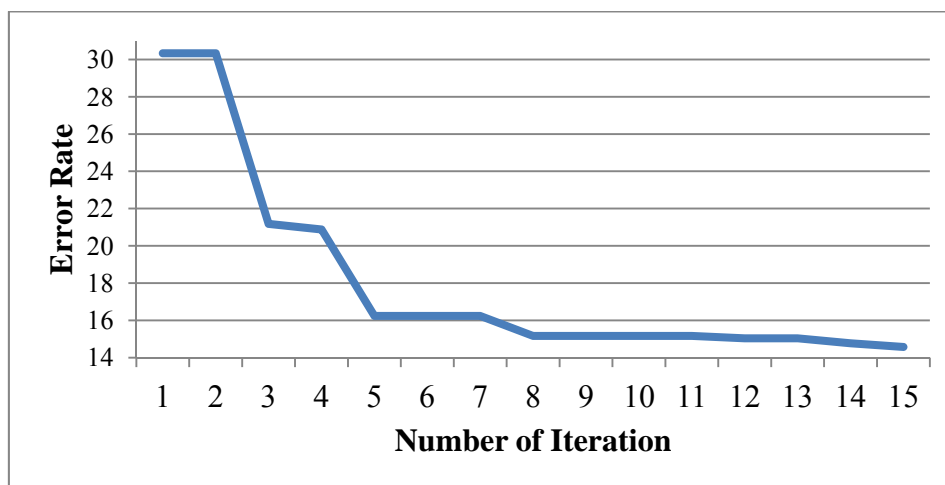

Figure 7. The convergence curve of the proposed method

In order to achieve the quantitative analysis of the proposed method, the experiment was performed on three MRI brain datasets and manual segmentation for all of the datasets are provided by medical experts. The results that are obtained from the proposed method is compared with manual ones based on similarity index (SI), true positive (TP), false positive (FP) and false negative (FN). The results that are achieved in Table 1 indicate that the similarity between the proposed method and manual segmentation are very close and also it can be observed that SI and TP are above 0.7 and 0.8 respectively which show high similarity on the other hand FP and FN are below 0.19 which means a few number of pixels are not match with manual segmentation.

Table 1. Numerical analysis of the proposed method with automatic point initialization

	Brain Tissue	SI	TP	FP	FN
Dataset 1	WM	0.8571	0.9492	0.1074	0.0505
	GM	0.7611	0.8202	0.0776	0.1794
Dataset 2	WM	0.8431	0.9079	0.0769	0.0917
	GM	0.7394	0.8327	0.1262	0.1670
Dataset 3	WM	0.8578	0.8928	0.0408	0.1068
	GM	0.7002	0.8006	0.1434	0.1990

The original SFCM algorithm with random initialization is performed on the same datasets and the results are compared with manual segmentation (Table 2). It can be observed that good results are achieved in dataset 1 however the algorithm fails to obtain good SI on other datasets.

Table 2. Original SFCM algorithm with random point initialization

	Brain Tissue	SI	TP	FP	FN
Dataset 1	WM	0.8508	0.9363	0.1005	0.0633
	GM	0.7721	0.9341	0.2098	0.0656
Dataset 2	WM	0.5793	0.9841	0.6986	0.0156
	GM	0.4935	0.6068	0.2294	0.3929
Dataset 3	WM	0.5718	0.9942	0.7387	0.0055
	GM	0.7538	0.9785	0.2981	0.0212

By comparing the proposed method and SFCM, it is concluded that the achieved results by the proposed method have higher similarity index values with less errors. In table 3, a comparison between the proposed method and SFCM is performed based on the average number of iterations and average number of running the algorithm.

Table 3. A comparison between SFCM and GWO-SFCM

	Average Algorithm Execution Number	Average number of iterations
SFCM	5	27
GWO-SFCM	1	4

According to Table 4.3, the proposed methods can achieve the result in the first attempts and with 4 iterations however SFCM algorithm needs to be executed five times in average and it will perform the segmentation after twenty seven iterations. Also the proposed method does not require any adjustments prior the segmentation process.

6.   CONCLUSION

In this paper, an enhanced SFCM algorithm was presented. In order to improve the initialization step of SFCM, we have used GWO algorithm to find the optimum initial point values. Three datasets from ISBR were used to evaluate the proposed algorithm and it was observed that good results were achieved as compared with standard SFCM algorithm. Moreover we have seen that the segmentation results were obtained in smaller number of iterations when using the proposed algorithm.

REFERENCES

[1]   Tran, T.T., V.T. Pham, and K.K. Shyu, *Image segmentation using fuzzy energy-based active contour with shape prior.* Journal of Visual Communication and Image Representation, 2014. 25(7): 1732-1745.
[2]   Zhang, X., et al., *Hybrid region merging method for segmentation of high-resolution remote sensing images.* ISPRS Journal of Photogrammetry and Remote Sensing, 2014. 98(0): 19-28.
[3]   Patino-Correa, L.J., et al., *White matter hyper-intensities automatic identification and segmentation in magnetic resonance images.* Expert Systems with Applications, 2014. 41(16): 7114-7123.
[4]   Aja-Fernández, S., A.H. Curiale, and G. Vegas-Sánchez-Ferrero, *A local fuzzy thresholding methodology for multiregion image segmentation.* Knowledge-Based Systems, (0).
[5]   Li, B.N., et al., *Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation.* Computers in Biology and Medicine, 2011. 41(1): 1-10.
[6]   Ugarriza, L.G., et al., *Automatic Image Segmentation by Dynamic Region Growth and Multiresolution Merging.* Ieee Transactions on Image Processing, 2009. 18(10): 2275-2288.

[7]   Samadzadegan, F. and A.A. Naeini. *Fuzzy clustering of hyperspectral data based on particle swarm optimization.* in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011 3rd Workshop on.* 2011. IEEE.

[8]   Maxwell, J.C. and J.J. Thompson, *A treatise on electricity and magnetism.* Vol. 1. 1892: Clarendon.

[9]   Hartigan, J.A., *Clustering algorithms.* 1975.

[10]   MacQueen, J. *Some methods for classification and analysis of multivariate observations.* in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability.* 1967. Oakland, CA, USA.

[11]   Dunn, J.C., *A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters.* 1973.

[12]   Bezdek, J., *Pattern recognition with fuzzy objective function algorithm.* New York: Plenum Press, 1981.

[13]   Cai, W.L., S.C. Chen, and D.Q. Zhang, *Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation.* Pattern Recognition, 2007. 40(3): 825-838.

[14]   Chuang, K.S., et al., *Fuzzy c-means clustering with spatial information for image segmentation.* Computerized Medical Imaging and Graphics, 2006. 30(1): 9-15.

[15]   Benaichouche, A., H. Oulhadj, and P. Siarry, *Improved spatial fuzzy c-means clustering for image segmentation using PSO initialization, Mahalanobis distance and post-segmentation correction.* Digital Signal Processing, 2013. 23(5): 1390-1400.

[16]   Zhang, Y., et al., *Image segmentation using PSO and PCM with Mahalanobis distance.* Expert Systems with Applications, 2011. 38(7): 9036-9040.

[17]   Krishnapuram, R. and J.M. Keller, *A possibilistic approach to clustering.* Fuzzy Systems, IEEE Transactions on, 1993. 1(2): 98-110.

[18]   Wang, H., et al. *Scalability of hybrid fuzzy c-means algorithm based on quantum-behaved PSO.* in *Fuzzy systems and knowledge discovery, 2007. fskd 2007. fourth international conference on.* 2007. IEEE.

[19]   Liu, H., et al. *Multi-temporal MODIS-data-based PSO-FCM clustering applied to wetland extraction in the Sanjiang Plain.* in *International Conference on Earth Observation Data Processing and Analysis.* 2008. International Society for Optics and Photonics.

[20]   Liu, B., et al., *Improved particle swarm optimization combined with chaos.* Chaos, Solitons & Fractals, 2005. 25(5): 1261-1271.

[21]   Noel, M.M. and T.C. Jannett. *Simulation of a new hybrid particle swarm optimization algorithm.* in *System Theory, 2004. Proceedings of the Thirty-Sixth Southeastern Symposium on.* 2004. IEEE.

[22]   Mirjalili, S., S.M. Mirjalili, and A. Lewis, *Grey wolf optimizer.* Advances in Engineering Software, 2014. 69: 46-61.

[23]   Coleman, G.B. and H.C. Andrews, *Image segmentation by clustering.* Proceedings of the IEEE, 1979. 67(5): 773-785.

[24]   Li, B.N., et al., *Integrating FCM and Level Sets for Liver Tumor Segmentation.* 13th International Conference on Biomedical Engineering, Vols 1-3, 2009. 23(1-3): 202-205.

[25]   Mirjalili, S., *How effective is the Grey Wolf optimizer in training multi-layer perceptrons.* Applied Intelligence, 2015: 1-12.

[26]   *The datsets are provided from ISBR website.* Available from: http://www.cma.mgh.harvard.edu/ibsr/.

[27]   Schultz, R.T. and A. Chakraborty, *Magnetic resonance image analysis*, in *Neuroimaging I.* 1996, Springer. 9-51.