

Research Article

Automatic Segmentation of Medical Images Using Fuzzy c-Means and the Genetic Algorithm

Omid Jamshidi and Abdol Hamid Pilevar

Medical Intelligence Lab, Department of Computer Engineering, Bu Ali Sina University, Hamedan, Iran

Correspondence should be addressed to Omid Jamshidi; o.jamshidi@gmail.com

Received 14 December 2012; Revised 17 January 2013; Accepted 17 January 2013

Academic Editor: Hiroshi Watabe

Copyright © 2013 O. Jamshidi and A. H. Pilevar. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Magnetic resonance imaging (MRI) segmentation is a complex issue. This paper proposes a new method for estimating the right number of segments and automatic segmentation of human normal and abnormal MR brain images. The purpose of automatic diagnosis of the segments is to find the number of divided image areas of an image according to its entropy and with correctly diagnose of the segment of an image also increased the precision of segmentation. Regarding the fact that guessing the number of image segments and the center of segments automatically requires algorithm test many states in order to solve this problem and to have a high accuracy, we used a combination of the genetic algorithm and the fuzzy c-means (FCM) method. In this method, it has been tried to change the FCM method as a fitness function for combination of it in genetic algorithm to do the image segmentation more accurately. Our experiment shows that the proposed method has a significant improvement in the accuracy of image segmentation in comparison to similar methods.

1. Introduction

Image segmentation is one of the difficult issues in the field of image processing. Image segmentation is the process of assigning a label to every pixel in an image so that pixels with the same label share certain visual characteristics. Many applications such as object identification, feature extraction, and object position identifications and classification require accurate image segmentation. Several methods of medical image segmentation have been proposed, such as edge based, region based, or a combination of both. The purpose of medical image segmentation is to provide a more meaningful image which can be more easily understood and analyzed.

The edge-based methods use edge information in an image to determine the boundaries of objects and, hence, to form closed regions that determine different objects in an image. In some image segmentation methods, this method has been consistently used with the edge of the area for segmenting magnetic resonance imaging (MRI).

Chun and Yang performed image segmentation according to the edge information [1]. In addition to edge information, they made use of a similarity measure which was

obtained as the median pixel variance parameters. The method used a fuzzy validity function as well as genetic algorithms and tried to find limit and suitable search space for image segmentation. Finding the main edges and removing redundant edges were the main issues in this method. Moreover, atlas-based segmentation methods were successfully employed for different applications. For instance, Heckemann segmented 67 brain images using 29 marked images [2].

On the other hand, k-means and the fuzzy c-means (FCM) are two successful region-based approaches. FCM can be obtained by a little modification in the k-means algorithm. FCM has been successfully used for MR image segmentation.

Neural network method needs a large number of training data and long time required for network training. For many experts, manual image segmentation is difficult and time consuming, and this leads to the use of automatic method.

Among the image segmentation methods, the FCM is more popular. Vast usage of it is due to its simplicity and accuracy. However, FCM has weaknesses in noise detection; many attempts are done for covering this weakness. In [3],

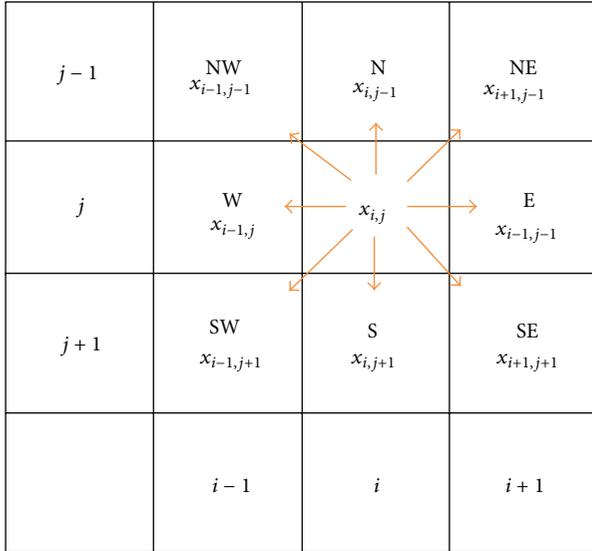


FIGURE 1: Definition of eight directions.

with using the objective function FCM and the use of neighbor pixels in addition to the pixel and also pixel division have been used. In [4–6], FCM method is used, and for improving the accuracy in image segmentation, membership function is changed. The defined object function in FCM is used, the local and nonlocal data segmentation are done, and the effect of each neighbor in the objective function is tested [7]. According to the total biases and gradient of sparse matrix, the FCM objective function has been changed [8]. The weight of local points in the objective function is calculated, and fuzzy clustering algorithm and continuity of local areas are used to make a better classification [9–11]. In [12], it is tried to change the FCM method to improve image segmentation. In order to reduce noise impact, local and nonlocal information have been used for image segmentation. In this paper, a new method is used to determine the dissimilarity with respect to local and nonlocal distances of each pixel. An algorithm based on histogram and using predefined window class centers was introduced. Then, the modified FCM method that in its neighboring areas are also used, and ultimately a new technique that is called a neighborhood-based membership ambiguity correction is used to eliminate small noises [13]. In [14], a hierarchical genetic method is used. In this method, chromosome is divided into two parts; that is, randomly placed one section of chromosomes with center point and the other with a random value of zero. The centers are determined by enabling or disabling. In [15], the genetic method and LVQ network for hierarchical classification of MR images are used. In this method, each chromosome is divided into two parts: one part is LVQ network weight and the other is determined by enabling or disabling the neurons of the LVQ network. In [16], a combination of statistical expectation maximization (EM) and pulse coupled neural network (PCNN) for segmentation of MR images has been used.

In most of the articles in the area of FCM segmentation, they believe that it is an effective method.

2. The Presented Method

Noise is inevitable in medical images. Therefore, this is essential to reduce the effect of noise prior to image segmentation. Since the method presented in this paper is based on FCM, we can perform noise reduction after noise detection and then perform image segmentation. Alternatively, we can estimate the probability of presence of noise in each pixel and change the effect of that pixel in FCM method accordingly. In this paper, we apply noise reduction to improve the image and, hence, to obtain a more reliable image segmentation.

2.1. Method to Reduce Noise on the Image. In this paper, we distinguish noise with respect to local neighborhood of each pixel. In addition, we consider the neighboring pixels while detection of noise. Considering the fact that we aim at a fast image segmentation method, we need to make use of a fast noise reduction method.

2.1.1. Detection of the Noisy Pixels. To detect a noisy pixel with the help of neighboring pixels, we followed the method proposed by [17] and defined the neighborhood of a pixel as shown in Figure 1.

Distance of central pixel with one of 8 first-order neighbors is calculated using

$$\text{Distance with NW} = |x_{i-1,j-1} - x_{i,j}|. \quad (1)$$

The distance of the corresponding pixel from all its neighbors was obtained. These distances were sorted in an ascending order, and 5 neighbors with the longest distances are selected. If the average of these 5 values was greater than a certain threshold, it means that at least five of the neighboring pixels have been very different so that the pixel is not considered as an edge position, and it is probably a noisy pixel. As we will discuss later, even if a pixel was wrongly detected as noisy, it would not affect much the algorithm as the noise correction method we use can usually handle it.

2.1.2. Correction of the Noisy Pixels. When a pixel was identified as noisy, we use its neighboring pixels to correct it. The reason we use neighboring pixels lies in the fact that a pixel is expected to be similar to its surrounding pixel. To do this, we select the pixel where total distance of its neighbor pixels with neighboring pixels of noisy pixel is less; this means that Euclidean distance between each neighboring pixel selected from the neighboring pixel in the same position of noisy pixel is calculated. Pixels to produce the lowest value means that their neighbors are more like together. Two pixels with similar neighbor values are expected to be close to each other, and, hence, the pixel can be replaced by the noisy pixel.

2.2. Image Segmentation. Given the fact that FCM method is simple and accurate and also due to the fuzzy nature of the image segmentation, many works on segmentation have used this method. Different works have tried to tackle some of the weaknesses associated with FCM. An important problem in image segmentation is to guess the number of

parts of an image. Some of the methods in the literature have ignored this factor and performed image segmentation with a predefined number of segments. However, in some specific cases such as MR images, the number of segments of an image can be guessed. But even in medical images, sometimes when image quality is low or variant or there exist tumors in the image, it is not possible to estimate the number of segments. Having an accurate estimate of the number of segments in an image is crucial as parameters that determine accuracy as boundary overlaps highly depend on that.

In this work, we try segment to increase the accuracy in image segmentation in an automatic manner and without expert aid. In [18], objective function is based on the distance of each point from determined centers, and the extent of membership of each point to these centers is used to determine the centers in the next steps.

By considering the predefined parameters and number of segments and by using the principle of maximum entropy and genetic algorithm, we try to improve the existing parameters. In this method, the value of the membership of each pixel is clearly observable. Differences in the number of segments of an image directly affect the membership of each pixel. The method proposed in this paper utilizes the genetic algorithm and the FCM objective function and automatically estimates the number of segments for each given image.

2.2.1. Chromosome Representation. The values corresponding to segment centers are placed in each chromosome. Hence, the length of a chromosome shows the number of different groups of data or, in other words, the number of segments in an image. We can control the maximum number of segments by measuring the length of each chromosome. Each cell of chromosomes can be initialized according to pixel values in the range of 0 to 255. For the first initialization, these numbers will be randomly generated. Table 1 is an example of a chromosome whose values are randomly generated.

If the chromosomes are initialized in the range 0 to 255, it means that the number of segments is equal to the length of the chromosome. But we want to set the number of image segments automatically. Some cells of chromosomes should have values that can show that they are not valid centers. To represent this, we can use negative values. The negative values of a cell correspond to the fact that this cell should not be considered as a data center. Hence, the number of nonnegative values in a chromosome determines the number of data centers. Some examples of chromosomes with length 6 are given in Table 2.

The first chromosome in Table 2 shows three centers with the values of 150.7, 121.1, and 94.3. The second chromosome shows six valid centers, whereas the third one contains four valid centers.

2.2.2. Initial Population. We construct the initial population using the method described earlier. To do this, we create arbitrary number of randomly initialized chromosomes. Also, we want to assign negative value to some allele of chromosomes. As mentioned, the negative values are not considered as centers of data. The number of valid data can

be different in each chromosome. According to a fixed and predetermined probability, the values of chromosome alleles will change to negative. The number of valid data in each chromosome can be at least 2 and at most equal to its total length.

2.2.3. Fitness Computation. With respect to the accuracy of FCM method, we leverage this method to calculate the compatibility. For each chromosome, using the estimated data centers and the pixel values of the image, we construct a matrix u that consists of membership values for each pixel. The elements of this matrix are calculated according to

$$u_{ij} = \frac{1}{\sum_{k=1}^c (\|x_i - c_j\| / \|x_i - c_k\|)^{2/(m-1)}}. \quad (2)$$

The objective function of Formula (2) is defined as

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, \quad 1 \leq m < \infty, \quad (3)$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the data center j , x_i is the i th dimension of the d -dimensional measured data, c_j is the d -dimension center of the data center, and $\|*\|$ is any norm expressing the similarity between any measured data and the center, while c_j are set of centers that are stored in chromosomes.

With respect to J_m value that is considered as fitness, the best value of the J_m occurs when each data center is placed in the center of data sets, and this causes J_m be small, therefore as much J_m be smaller, then the approach is better, and we are try to get the smallest amount for J_m .

If all of the data for the chromosome are a valid set of data centers or that the entire chromosome is equal to the number of valid data, J_m is calculated according to the aforementioned method that is acceptable, but with respect to the fact that in each chromosome, the number of valid data as class centers is different, the value of J_m obtained for comparing chromosomes fitness value is not proper and proportional to the number of the centers, and the fitness should be corrected such that the comparison between chromosomes becomes possible.

It should be noted that with increasing the number of centers, the distance of each x_i from the nearest center decreases, and it is expected that by increasing centers and decreasing the distance of each x_i with nearest center, the J_m value is reduced, and because the J_m is generally less in chromosome with more valid numbers, usually chromosomes with high number of centers are selected, but our favorite state is when data centers are being guessed in the centers of data sets and automatically determine the minimum required number of data centers with respect to data entropy. Based on the importance of the J_m value, we need to add another parameter to correct fitness value according to the changes in the number of the centers. For improvement of the presented method, we assign a penalty factor for recognizing the increment of the number of data centers. We add y to show the factor of penalty. Parameter y is obtained from the total

TABLE 1: A sample chromosome with length of 6.

Value in each chromosome, are the data centers. Thus, the length of each chromosome shows the number of different categories or in other words shows the number of segment of image. We can control the maximum number of the data sets with controlling the length of chromosomes. Each cell of chromosomes initialized can be in the range of 0 to 255. In the first place, the numbers will be randomly generated. An example of randomly generated chromosome is shown in here.	20.7	42	92.8	120	198
---	------	----	------	-----	-----

TABLE 2: Three chromosomes of the initial population generated randomly.

-170	150.7	-170	121.1	-170	94.3
83.1	73.5	159.7	196.7	63.3	224.9
111.7	-170	-170	38.1	187.1	65.2

TABLE 3: Different modes of pixels.

Detected class	True class	
	Abnormal	Normal
Abnormal	N_{tp}	N_{fp}
Normal	N_{fn}	N_{tn}

length of the valid data of each chromosome to total length of chromosome according to

$$y = \frac{\text{length of valid data of chromosome}}{\text{total length of chromosome}}. \quad (4)$$

Then, with multiplying, J_m in y fitness is calculated for each chromosome according to

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2 \times y, \quad 1 \leq m < \infty. \quad (5)$$

2.2.4. Selection. We use the roulette wheel technique to produce the mating pool of chromosomes. The main idea of the roulette wheel technique is to associate more chance to better chromosomes.

2.2.5. Crossover. Crossover is the next step after the selection of parent chromosomes. In this step, a new offspring is generated as a result of combining two parents.

2.2.6. Mutation. Each allele of the chromosome changes according to the probability P_m . Mutation is used to perform a search over the entire range of answers. Figure 2 shows an overview of our approach to segmentation.

3. Results

To evaluate the proposed method, we use an MR image of brain of size 217×181 pixels, each of whose pixels range from zero to 255 (256 gray level). The number of initial population is set to 20 and the maximum number of generations to 100. The number of individuals to be replaced in each generation

TABLE 4: Parameters for genetic operations.

Representation	Bit representation
Crossover type	One point crossover
Crossover rate	0.7
Mutation type	Bit mutation
Mutation rate	0.1

is set to 40% of the total population. As mentioned earlier, we utilize roulette wheel selection technique. Images obtained by the proposed method determine the appropriate number of segments and the estimated center for each.

In the resulting segments, in order to verify the correctness of segments which are in the same group, we use different colors which are random.

Figure 3 shows the output of the proposed method for automatic segmentation of the original MR image.

In order to verify the performance of our algorithm, we use sensitivity, specificity, Jaccard, and k index parameters measures. If A and B are the automatic and manual segmentations of an image, respectively, then $T_p = B \cap A$ will be the true positive, and $F_p = A - B$, $F_n = B - A$ will be the false positive and false negative, respectively.

According to Table 3, sensitivity is defined as [19]

$$\text{Sensitivity} = \frac{N_{tp}}{N_{tp} + N_{fn}}. \quad (6)$$

Specificity as

$$\text{Specificity} = \frac{N_{tn}}{N_{fp} + N_{tn}}. \quad (7)$$

Similarity is defined as [20]

$$k(A, B) = \frac{2|A \cap B|}{|A + B|} \times 100\% = \frac{2T_p}{|B| + |A|} \times 100\%. \quad (8)$$

Jaccard index is defined as

$$J(A, B) = \frac{|A \cap B|}{A \cup B} \times 100\% = \frac{|T_p|}{|T_p| + |F_n| + |F_p|} \times 100\%. \quad (9)$$

The parameters settings for the genetic operations in this experiment are determined as shown in Table 4. Table 5 shows the performance of the proposed method when the initial number of chromosomes is set to 20 and 100 generations.

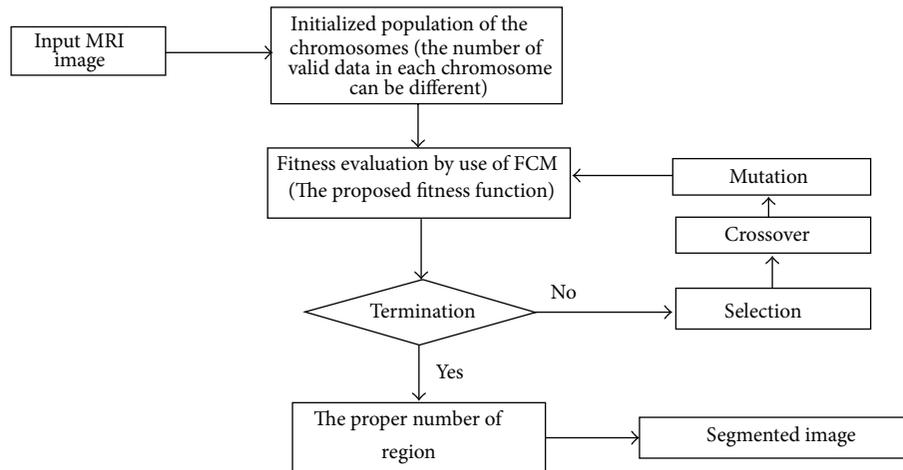


FIGURE 2: The schematic flow chart of segmentation.

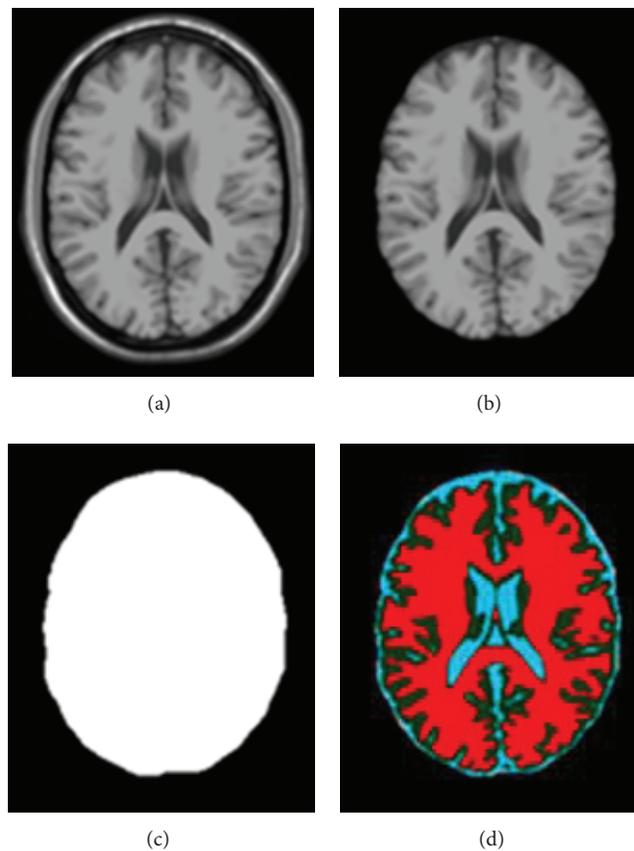


FIGURE 3: Automatic segmentation results of our algorithm. (a) is the original image, (b) is the brain for segmentation, (c) is the background of the image removed from brain, and (d) is MR brain image which is divided into 3 segments (shown with 3 different colors).

As shown in Table 5, every time we increase the length of the chromosomes, the output results become more accurate. The answers are different in each run of the algorithm because the chromosomes are initialized randomly in each run of algorithm. Also, maximum number of iteration of the genetic algorithm is 100.

In addition, in Figure 4, we show the performance of our algorithm on a noisy image.

The overall results of our method on a noisy image are given in Table 6.

4. Suggestions

Given that the proposed method is based on the inner class distance, it has been attempted to act according to

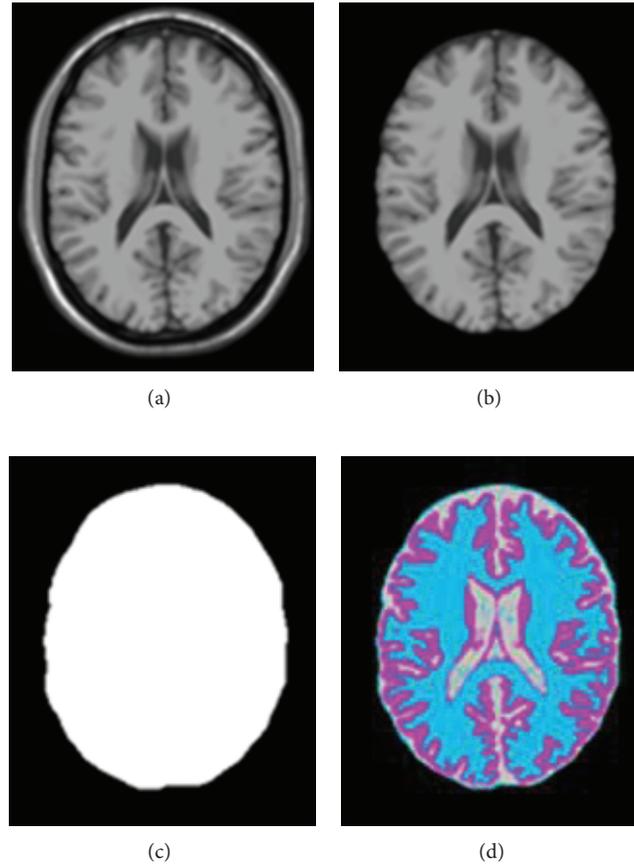


FIGURE 4: Performance of our algorithm on an image with 3% noise. Image (a) is the original image, (b) is the brain for segmentation, (c) is the background of the image removed from brain, and (d) is MR brain image which is divided into 3 segments (shown with 3 different colors).

TABLE 5: Results obtained from the proposed method.

Initial number of chromosome length	Sensitivity	Specificity	Jaccard similarity	k index	CPU time (s)	Number of final center
5	0.967	0.974	0.913	0.919	12.5	3
5	0.961	0.964	0.912	0.921	12.9	3
5	0.971	0.978	0.891	0.895	12.7	3
5	0.966	0.959	0.921	0.931	10.5	3
6	0.963	0.941	0.928	0.933	11.8	3
6	0.968	0.959	0.933	0.945	12.9	3
6	0.973	0.976	0.935	0.937	14.4	3
6	0.970	0.967	0.938	0.827	14.7	3
7	0.979	0.977	0.944	0.958	15.8	3

the distance with the nearest center and correction center. However, one can think of using outer class distance to obtain increased classification accuracy. In this case, in addition to each pixel belonging to the nearest center, the parameter of being distant from the other points is also taken into account. With respect to these two parameters, the precision of medical image segmentation methods can be expected to be increased.

5. Conclusion

FCM is a popular clustering method and has been widely applied for medical image segmentation. However, traditional FCM always suffers from noise in the images. Although many researchers have developed various extended algorithms based on FCM, none of them are flawless. A method based on genetic algorithm with use of FCM is proposed in

TABLE 6: Results obtained from the proposed method.

Initial number of chromosome length	Sensitivity	Specificity	Jaccard similarity	k index	CPU time (s)	Number of final centers
5	0.960	0.964	0.911	0.913	12.4	3
5	0.954	0.962	0.912	0.919	12.8	3
5	0.959	0.965	0.892	0.897	12.9	3
5	0.961	0.955	0.918	0.927	11.4	3
6	0.958	0.943	0.925	0.931	11.8	3
6	0.962	0.954	0.931	0.940	13.1	3
6	0.965	0.967	0.932	0.935	14.2	3
6	0.961	0.965	0.929	0.932	14.9	3
7	0.968	0.972	0.938	0.949	16.3	3

this paper. In our algorithm, local neighbor pixels are used. We tested our algorithm on simulated MR images. In this method, we obtained the right number of the segments fully automatically. We reduced the number of iterations of genetic algorithm and increased the convergence speed by applying proposed technique.

References

- [1] D. N. Chun and H. S. Yang, "Robust image segmentation using genetic algorithm with a fuzzy measure," *Pattern Recognition*, vol. 29, no. 7, pp. 1195–1211, 1996.
- [2] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," *NeuroImage*, vol. 33, no. 1, pp. 115–126, 2006.
- [3] Z. M. Wang, Y. C. Soh, Q. Song, and K. Sim, "Adaptive spatial information-theoretic clustering for image segmentation," *Pattern Recognition*, vol. 42, no. 9, pp. 2029–2044, 2009.
- [4] S. R. Kannan, S. Ramathilagam, R. Devi, and E. Hines, "Strong fuzzy c -means in medical image data analysis," *Journal of Systems and Software*, vol. 85, no. 11, pp. 2425–2438, 2012.
- [5] S. R. Kannan, S. Ramathilagam, R. Devi, and A. Sathya, "Robust kernel FCM in segmentation of breast medical images," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4382–4389, 2011.
- [6] S. R. Kannan, A. Sathya, S. Ramathilagam, and R. Devi, "Novel segmentation algorithm in segmenting medical images," *Journal of Systems and Software*, vol. 83, no. 12, pp. 2487–2495, 2010.
- [7] B. Caldairou, N. Passat, P. A. Habas, C. Studholme, and F. Rousseau, "A non-local fuzzy segmentation method: application to brain MRI," *Pattern Recognition*, vol. 44, no. 9, pp. 1916–1927, 2011.
- [8] Y. He, M. Y. Hussaini, J. Ma, B. Shafei, and G. Steidl, "A new fuzzy c -means method with total variation regularization for segmentation of images with noisy and incomplete data," *Pattern Recognition*, vol. 45, no. 9, pp. 3463–3471, 2012.
- [9] Z.-X. Ji, Q.-S. Sun, and D. S. Xia, "A framework with modified fast FCM for brain MR images segmentation," *Pattern Recognition*, vol. 44, no. 5, pp. 999–1013, 2011.
- [10] Z.-X. Ji, Q.-S. Sun, and D. S. Xia, "A modified possibilistic fuzzy c -means clustering algorithm for bias field estimation and segmentation of brain MR image," *Computerized Medical Imaging and Graphics*, vol. 35, no. 5, pp. 383–397, 2011.
- [11] Z. Ji, Q. Sun, Y. Xia, Q. Chen, D. Xia, and D. Feng, "Generalized rough fuzzy c -means algorithm for brain MR image segmentation," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 2, pp. 644–655, 2011.
- [12] J. Wang, J. Kong, Y. Lu, M. Qi, and B. Zhang, "A modified FCM algorithm for MRI brain image segmentation using both local and non-local spatial constraints," *Computerized Medical Imaging and Graphics*, vol. 32, no. 8, pp. 685–698, 2008.
- [13] K. Sikka, N. Sinha, P. K. Singh, and A. K. Mishra, "A fully automated algorithm under modified FCM framework for improved brain MR image segmentation," *Magnetic Resonance Imaging*, vol. 27, no. 7, pp. 994–1004, 2009.
- [14] C.-C. Lai and C.-Y. Chang, "A hierarchical evolutionary algorithm for automatic medical image segmentation," *Expert Systems with Applications*, vol. 36, no. 1, pp. 248–259, 2009.
- [15] J. Y. Yeh and J. C. Fu, "A hierarchical genetic algorithm for segmentation of multi-spectral human-brain MRI," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1285–1295, 2008.
- [16] J. C. Fu, C. C. Chen, J. W. Chai, S. T. C. Wong, and I. C. Li, "Image segmentation by EM-based adaptive pulse coupled neural networks in brain magnetic resonance imaging," *Computerized Medical Imaging and Graphics*, vol. 34, no. 4, pp. 308–320, 2010.
- [17] H.-C. Chen and W.-J. Wang, "Efficient impulse noise reduction via local directional gradients and fuzzy logic," *Fuzzy Sets and Systems*, vol. 160, no. 13, pp. 1841–1857, 2009.
- [18] W.-B. Tao, J.-W. Tian, and J. Liu, "Image segmentation by three-level thresholding based on maximum fuzzy entropy and genetic algorithm," *Pattern Recognition Letters*, vol. 24, no. 16, pp. 3069–3078, 2003.
- [19] A. R. Van Erkel and P. M. T. Pattynama, "Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology," *European Journal of Radiology*, vol. 27, no. 2, pp. 88–94, 1998.
- [20] J. K. Udupa, V. R. LeBlanc, Y. Zhuge et al., "A framework for evaluating image segmentation algorithms," *Computerized Medical Imaging and Graphics*, vol. 30, no. 2, pp. 75–87, 2006.



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

