Research Article

Segmentation of Brain Tissues from Magnetic Resonance Images Using Adaptively Regularized Kernel-Based Fuzzy C-Means Clustering

Ahmed Elazab,[1,2,3] Changmiao Wang,[1,2] Fucang Jia,[1,2] Jianhuang Wu,[1,2] Guanglin Li,[1,2] and Qingmao Hu[1,2]

[1]*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Boulevard, Shenzhen 518055, China*
[2]*University of Chinese Academy of Sciences, 52 Sanlihe Road, Beijing 100864, China*
[3]*Faculty of Computers and Information, Mansoura University, Elgomhouria Street, Mansoura 35516, Egypt*

Correspondence should be addressed to Qingmao Hu; qm.hu@siat.ac.cn

An adaptively regularized kernel-based fuzzy C-means clustering framework is proposed for segmentation of brain magnetic resonance images. The framework can be in the form of three algorithms for the local average grayscale being replaced by the grayscale of the average filter, median filter, and devised weighted images, respectively. The algorithms employ the heterogeneity of grayscales in the neighborhood and exploit this measure for local contextual information and replace the standard Euclidean distance with Gaussian radial basis kernel functions. The main advantages are adaptiveness to local context, enhanced robustness to preserve image details, independence of clustering parameters, and decreased computational costs. The algorithms have been validated against both synthetic and clinical magnetic resonance images with different types and levels of noises and compared with 6 recent soft clustering algorithms. Experimental results show that the proposed algorithms are superior in preserving image details and segmentation accuracy while maintaining a low computational complexity.

1. Introduction

Image segmentation is to partition an image into meaningful nonoverlapping regions with similar features. Segmentation of brain magnetic resonance (MR) images is necessary to differentiate white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). Such segmentation is essential for studying anatomical structure changes and brain quantification [1]. It is also a prerequisite for tumor growth modeling as tumors diffuse at different rates according to the surrounding tissues [2]. Due to potential existence of noise, bias field, and partial volume effect, segmentation of brain images remains challenging.

Image segmentation techniques can be roughly categorized into [3] thresholding, region growing, clustering, edge detection, and model-based methods. Clustering is an unsupervised learning strategy that groups similar patterns into clusters and can be hard or soft. Soft clustering is preferred as

every pixel can be assigned to all clusters with different membership values [4, 5]. The most popular soft clustering methods applied to MR images are [4] fuzzy C-means (FCM) clustering [6, 7], mixture modeling, and hybrid methods of both.

Although the FCM algorithm comes with good accuracy in the absence of noise, it is sensitive to noise and other imaging artifacts. Therefore, enhancements have been tried to improve its performance by including local spatial and grayscale information [8–14] which will be briefly elaborated in Section 2.

A mixture model composed of a finite number of Gaussians has been employed for brain MR image segmentation. The main strategy to incorporate the local information into the mixture model is to use hidden Markov random fields for more accurate segmentation [15]. Nikou et al. [16] proposed a hierarchical and spatially constrained mixture model that takes into account spatial information by imposing distinct smoothness priors on the probabilities of each cluster and

pixel neighborhoods. In [17], a nonparametric Bayesian model for tissue classification of brain MR images known as Dirichlet process mixture model was explored. Nguyen and Wu [18] introduced a way to incorporate spatial information between neighboring pixels into the Gaussian mixture model (GMM).

To be more robust to noise and attain fast convergence, FCM and GMM can be combined. Chatzis and Varvarigou [19] embedded the hidden Markov random field model into the FCM objective function to explore the spatial information. Chatzis [20] introduced a methodology for training finite mixture models under fuzzy clustering principle with a dissimilarity function to incorporate the explicit information into the fuzzy clustering procedure. Recently, Ji et al. [21] employed robust spatially constrained FCM (RSCFCM) algorithm for brain MR image segmentation by introducing a factor for the spatial direction based on the posterior probabilities and prior probabilities.

In [22], Li et al. presented an algorithm for brain tissue classification and bias estimation using a coherent local intensity clustering. Later they explored multiplicative intrinsic component optimization (MICO) [23] to improve the robustness and accuracy of tissue segmentation in the presence of high level bias field.

Generally, the current brain MR image segmentation algorithms suffer from one or more of the following shortcomings: lack of robustness to outliers [8, 9, 13], high computational cost [8, 13, 14, 16, 21], prior adjusting of crucial or many parameters [8–11, 21], limited segmentation accuracy in the presence of high level noise [8, 11, 19, 22, 23], and loss of such image details like CSF [9, 13, 14, 21]. In this paper, a new soft clustering framework is to be explored for better handling of the aforementioned segmentation problems.

The rest of this paper is organized as follows. Related work of FCM algorithm is presented in Section 2. The proposed framework is then elaborated in Section 3. Experiments on synthetic and clinical MR images are presented in Section 4. Sections 5 and 6 are devoted to discussion and conclusion, respectively.

2. Related Work

The FCM algorithm in its original form assigns a membership value to each pixel for all clusters in the image space. For an image I with set of grayscales x_i at pixel i ($i = 1, 2, \ldots, N$), $X = \{x_1, x_2, \ldots, x_N\} \subset R^k$ in k-dimensional space and cluster centers $v = \{v_1, v_2, \ldots, v_c\}$ with c being a positive integer ($2 < c \ll N$), there is a membership value u_{ij} for each pixel i in the jth cluster ($j = 1, 2, \ldots, c$). The objective function of the FCM algorithm is [7]

$$J_{\text{FCM}} = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \left\| x_i - v_j \right\|^2, \tag{1}$$

where m is a weighting exponent to the degree of fuzziness, that is, $m > 1$, and $\left\| x_i - v_j \right\|^2$ is the grayscale Euclidean

distance between pixel i and center v_j. The membership u_{ij} should be constrained to the following:

$$\forall i \in [1, N], \quad j \in [1, c]:$$
$$\sum_{j=1}^{c} u_{ij} = 1, \quad u_{ij} \in [0, 1], \quad 0 \leq \sum_{i=1}^{N} u_{ij} \leq N. \tag{2}$$

The membership function and cluster centers are updated iteratively in an alternating process known as alternate optimization. The membership function and cluster centers are

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\left\| x_i - v_j \right\|^2 / \left\| x_i - v_k \right\|^2 \right)^{1/(m-1)}},$$
$$v_j = \frac{\sum_{i=1}^{N} u_{ij}^m x_i}{\sum_{i=1}^{N} u_{ij}^m}. \tag{3}$$

As the objective function in (1) does not include any local information, the original FCM is very sensitive to noise and the accuracy of clustering in the presence of noise and image artifacts will decrease. To overcome this problem, Ahmed et al. [8] modified the objective function by adding a term for the spatial information of neighboring pixels. This algorithm is denoted as FCM_S with the following objective function:

$$J_{\text{FCM\_S}} = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \left\| x_i - v_j \right\|^2$$
$$+ \frac{\alpha}{N_R} \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \left(\sum_{r \in N_i} \left\| x_r - v_j \right\|^2 \right), \tag{4}$$

where α is a parameter to control the spatial information of neighbors with $0 < \alpha \leq 1$, N_i is the set of pixels around pixel i, and N_R is the cardinality of N_i.

The FCM_S algorithm is computationally expensive as the local neighborhood term has to be calculated in each iteration step. To overcome this drawback, Chen and Zhang [10] replaced the term $(1/N_R) \sum_{r \in N_i} \left\| x_r - v_j \right\|^2$ with $\left\| \overline{x}_i - v_j \right\|^2$, where $\overline{x}$ is the grayscale of a filtered image that could be calculated once in advance, and used kernel function to replace the Euclidean distance. The enhancement could be in two forms, that is, FCM_S1 by using the average filter and FCM_S2 by adopting the median filter. Their objective function is as follows:

$$J_{\text{FCM\_S1,2}} = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \left\| x_i - v_j \right\|^2 + \alpha \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \left\| \overline{x}_i - v_j \right\|^2. \tag{5}$$

Although the accuracy has been improved, it is sensitive to high level noises and different types of noises. In addition, the parameter α, which has a great impact on the performance, is set manually with care and requires prior information about noise.

Yang and Tsai [12] proposed a Gaussian kernel-based FCM method with the parameter η_j calculated in every

iteration to replace α for every cluster. Similar to FCM_S1 and FCM_S2, this method has two forms: GKFCM1 and GKFCM2 for average and median filters, respectively. The parameter η_j is estimated using kernel functions:

$$\eta_j = \frac{\min_{j' \neq j} \left(1 - K\left(v_j, v_j\right)\right)}{\max_k \left(1 - K\left(v_k, \overline{x}\right)\right)}, \tag{6}$$

where K is the kernel function. The replacement of α with η_j could yield better results than FCM_S1 and FCM_S2. However, for good estimation of η_j, cluster centers should be well separated which might not be always true; hence the algorithm has to iterate many times to converge. Moreover, the learning scheme requires a large number of patterns and many cluster centers to find the optimal value for η_j.

To tackle the problem of parameter adjustment, Krinidis and Chatzis [13] proposed the FLICM algorithm with a fuzzy factor that combined both spatial and grayscale information of the neighboring pixels. The fuzzy factor $G_{ij} = \sum_{k \in N_i, i \neq k} (1/(1 + d_{ik}))(1 - u_{ij})^m \|x_i - v_j\|^2$ was embedded into (1) as follows:

$$J_{\text{FLICM}} = \sum_{i=1}^{N} \sum_{j=1}^{c} \left[u_{ij}^m \left\|x_i - v_j\right\|^2 + G_{ij} \right], \tag{7}$$

where pixel i is the center of the local window, pixel j is in the neighborhood, and d_{ik} is the spatial Euclidean distance between pixels i and k.

Although FLICM algorithm enhances robustness to noise and artifacts, it is slow since the fuzzy factor (G_{ij}) has to be calculated in each iteration. Moreover, G_{ij} is heavily affected by spatial Euclidean distance from the central pixel to its neighboring pixels to lose small image details due to the smoothing effect.

To enhance the FLICM algorithm, Gong et al. [14] developed KWFLICM algorithm with a trade-off weighted fuzzy factor to control the local neighbor relationship and replaced the Euclidean distance with kernel function. The weighted fuzzy factor $\acute{G}_{ij}$ of KWFLICM is

$$\acute{G}_{ij} = \sum_{k \in N_i, i \neq k} \omega_{ik} \left(1 - u_{ij}\right)^m \left(1 - K\left(x_i, v_j\right)\right), \tag{8}$$

where ω_{ik} is the trade-off weighted fuzzy factor of pixel k in the local window around the central pixel i and $1 - K(x_i, v_j)$ is the kernel metric function. The trade-off weighted fuzzy factor combines both the local spatial and grayscale information [14]. Because of the trade-off weighted fuzzy factor, its computational cost increases substantially. In addition, the algorithm is unable to preserve small image details.

In addition to the abovementioned shortcomings, Szilágyi [24] pointed out serious theoretical mistakes in FLICM and KWFLICM. It was shown that the iterative optimization nature of FLICM and KWFLICM did not minimize their objective functions; instead, they iterated until the partition matrices converged. Furthermore, their objective functions intended to employ local contextual information but theoretically failed and were not even suitable for creating a valid partition [24].

To this end, a new way to modify the existing FCM clustering is explored with adaptive regularization for contextual information. The proposed framework utilizes a new parameter to control the effect of pixel neighbors based on the heterogeneity of local grayscale distribution. A weighted image is devised that combines the local contextual information with respect to the heterogeneity of local grayscale distribution and the original grayscale that is calculated once in advance to reduce the computational cost. To improve segmentation accuracy and robustness to outliers, a kernel function is employed to replace the Euclidean distance metric. Validation against both synthetic and clinical MR data has been carried out to compare the proposed algorithms with 6 recent soft clustering algorithms in terms of segmentation accuracy and computational costs.

3. Proposed Algorithms

We introduce a regularizing parameter to enhance segmentation robustness and preserve image details, devise a weighted image, and adopt the Gaussian radial basis function (GRBF) for better accuracy.

3.1. The Introduced Regularization Term. The parameter α used in [8–10] is usually set in advance to control the desirable amount of contextual information. Indeed, using a fixed α for every pixel is not appropriate since noise level differs from one window to another. In addition, setting such parameter needs prior knowledge about noise which is not always available in reality. Hence, adaptive calculation of α is necessary according to the pixel being processed.

To be adaptive to noise level of the pixel being processed, we first calculate the local variation coefficient (LVC) to estimate the discrepancy of grayscales in the local window to be normalized with respect to the local average grayscale. In the presence of noise to have high heterogeneity between the central pixel and its neighbors, LVC will increase. Consider

$$\text{LVC}_i = \frac{\sum_{k \in N_i} \left(x_k - \overline{x}_i\right)^2}{N_R * \left(\overline{x}_i\right)^2}, \tag{9}$$

where x_k is the grayscale of any pixel falling in the local window N_i around the pixel i, N_R is the cardinality of N_i, and $\overline{x}_i$ is its mean grayscale. Next, LVC_i is applied to an exponential function to derive the weights within the local window:

$$\zeta_i = \exp\left(\sum_{k \in N_i, i \neq k} \text{LVC}_k\right), \tag{10}$$

$$\omega_i = \frac{\zeta_i}{\sum_{k \in N_i} \zeta_k}.$$

The ultimate weight assigned to every pixel is associated with the average grayscale of the local window:

$$\varphi_i = \begin{cases} 2 + \omega_i, & \overline{x}_i < x_i \\ 2 - \omega_i, & \overline{x}_i > x_i \\ 0, & \overline{x}_i = x_i. \end{cases} \tag{11}$$

(a)

30	30	30	30 A	30	30
30	30	30	30	30 B	30
30	30	30	165	30	30
30	30	123	30	30	30 C
30	30	30	30	30	30
30	30	30	30	30	30

(b)

0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	1.989	1.957	1.992	0.000
0.000	1.960	1.896	2.623	1.957	0.000
0.000	1.952	2.085	1.896	1.989	0.000
0.000	1.978	1.952	1.960	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000

(c)

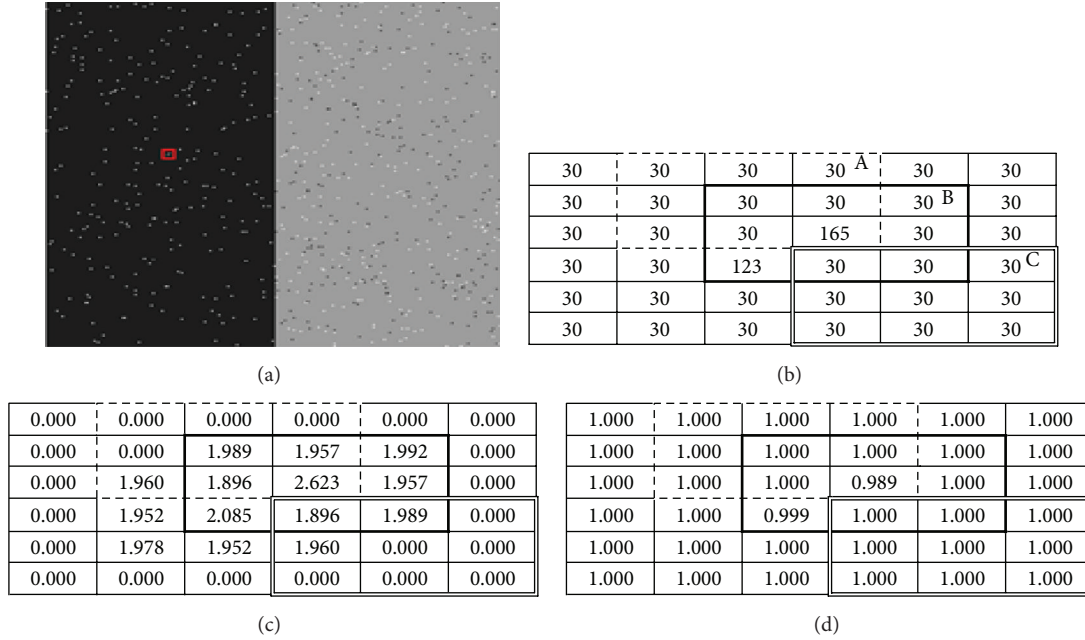| 1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	0.989	1.000	1.000
1.000	1.000	0.999	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000

(d)

FIGURE 1: Calculation and effect of the regularization parameter in different cases. (a) Noisy image. (b) 6×6 subimage, red rectangle from (a), with 3 different windows A, B, and C. (c) Weights associated with each pixel using the proposed method. (d) Membership values after three iterations.

The parameter φ_i assigns higher values for those pixels with high LVC (for pixel i being brighter than the average grayscale of its neighbors, φ_i will be $2 + \omega_i$, and ω_i will be large when the sum of LVC within its neighborhood is large) and lower values otherwise. When the local average grayscale is equal to the grayscale of the central pixel, φ_i will be zero and the algorithm will behave as the standard FCM algorithm. The value 2 in (11) is set through experiments to balance between the convergence rate and the capability to preserve details. The proposed parameter φ_i is embedded into (5) to replace α. Figure 1 shows the calculation of φ_i with different cases of noise.

Here are some remarks on the parameter φ_i.

The first point to emphasize is that φ_i is only relevant to the grayscales within a specified neighborhood, which is very different from FCM_S, FLICM, and KWFLICM, where the contextual information is expressed, respectively, by $\sum_{x_r \in N_i} \|x_r - v_j\|^2$, G_{ij}, and $\acute{G}_{ij}$, all containing a loop on the neighborhood and the cluster center v_j. Due to its irrelevance to clustering parameters, φ_i could be calculated in advance before the clustering process which can greatly reduce the computational cost. On the contrary, FCM_S, FLICM, and KWFLICM will need to update the contextual weights at each iteration, which is the main reason why they have higher computational cost.

The second point is that the contextual information provided by φ_i is based on the heterogeneity of grayscale distribution within the local neighborhood, which is completely different from existing enhanced versions of FCM to base the contextual information on the difference between the grayscales of neighboring pixels and cluster centers. As a result, the proposed φ_i tends to yield a homogeneous clustering according to local grayscale distribution while existing enhanced FCM algorithms tend to make the clustering to have more homogeneous labels by incorporating the contextual information.

3.2. Devising a Weighted Image. In addition to making $\overline{x}$, respectively, the grayscale of average/median filter of the original image, $\overline{x}$ can also be replaced with the grayscale of the newly formed weighted image $\overline{\xi}$:

$$\overline{\xi}_i = \frac{1}{2 + \max(\varphi_i)} \left(x_i + \frac{1 + \max(\varphi_i)}{N_R - 1} \sum_{r \in N_i} x_r \right), \quad (12)$$

where x_r and N_i are, respectively, the grayscale and neighborhood of pixel i and N_R is the cardinality of N_i. Formula (12) is inspired by the weighted image in [9] but utilizes φ_i explicitly to make the weighted image free from parameters that are difficult to adjust.

3.3. Measuring Distance Using Kernel Function. The Euclidean distance metric is generally simple and computationally inexpensive, but it is sensitive to perturbations and outliers. Recently, with popular usage of support vector machine, a new direction appears to use kernel functions. The kernel functions are able to project the data into higher dimensional space where the data could be more easily separated [25]. To do this, a so-called kernel trick has been adopted that can transform linear algorithm to nonlinear one using dot product [26]. Using the kernel trick, the Euclidean distance

term $\|x_i - v_j\|^2$ can be replaced with $\|\phi(x_i) - \phi(v_j)\|^2$ that is defined as

$$\left\|\phi(x_i) - \phi(v_j)\right\|^2 = K(x_i, x_i) + K(v_j, v_j) \\ - 2K(x_i, v_j), \tag{13}$$

where K is the kernel function.

In this paper, we use GRBF kernel [27]:

$$K(x_i, v_j) = \exp\left(-\frac{\|x_i - v_j\|^2}{2\sigma^2}\right), \tag{14}$$

where σ is the kernel width.

Using GRBF, the kernel function in (13) will be

$$\left\|\phi(x_i) - \phi(v_j)\right\|^2 = 2\left(1 - K(x_i, v_j)\right). \tag{15}$$

The choice of the kernel width σ is still a problem and has to be selected carefully. If it is large the exponential effect will be almost linear. On the contrary, if it is small, the cluster boundaries will be sensitive to outliers [27]. In [10], σ was set to a fixed value of 150 while in [12] the authors used sample variance to estimate σ. Similar to [14], we calculate σ based on the distance variances among all pixels:

$$\sigma = \left[\frac{\sum_{i=1}^{N}\left(d_i - \bar{d}\right)^2}{N-1}\right]^{1/2}, \tag{16}$$

where $d_i = \|x_i - \bar{x}\|$ is distance from the grayscale of pixel i to the grayscale average of all pixels and $\bar{d}$ is the average of all distances d_i.

3.4. The Proposed Framework. The proposed adaptively regularized kernel-based FCM framework is denoted as ARKFCM. First, we calculate the adaptive regularization parameter φ_i associated with every pixel to control the contextual information using (11). The objective function is defined as

$$J_{\text{ARKFCM}} = 2\left[\sum_{i=1}^{N}\sum_{j=1}^{c} u_{ij}^m \left(1 - K(x_i, v_j)\right)\right. \\ \left. + \sum_{i=1}^{N}\sum_{j=1}^{c} \varphi_i u_{ij}^m \left(1 - K(\bar{x}_i, v_j)\right)\right]. \tag{17}$$

Under the conditions specified in (2), the minimization of $J_{\text{ARKFCM}}(u, v)$ can be calculated through an alternate optimization procedure using (derivation is given in the Appendix)

$$u_{ij}$$

$$= \frac{\left(\left(1 - K(x_i, v_j)\right) + \varphi_i\left(1 - K(\bar{x}_i, v_j)\right)\right)^{-1/(m-1)}}{\sum_{k=1}^{c}\left(1 - K(x_i, v_k) + \varphi_i\left(1 - K(\bar{x}_i, v_k)\right)\right)^{-1/(m-1)}} \tag{18}$$

$$v_j = \frac{\sum_{i=1}^{N} u_{ij}^m \left(K(x_i, v_j) x_i + \varphi_i K(\bar{x}_i, v_j) \bar{x}_i\right)}{\sum_{i=1}^{N} u_{ij}^m \left(K(x_i, v_j) + \varphi_i K(\bar{x}_i, v_j)\right)}. \tag{19}$$

When $\bar{x}$ is replaced with the grayscale of the average/median filter of the original image, the algorithm is denoted as $\text{ARKFCM}_1/\text{ARKFCM}_2$. When $\bar{x}_i$ is replaced with the weighted image $\bar{\bar{\xi}}_i$ defined in (12), the algorithm is denoted as ARKFCM_w. The main steps for the proposed algorithms are as follows:

(1) Initialize threshold $\varepsilon = 0.001$, $m = 2$, loop counter $t = 0$, v, and $u^{(0)}$.

(2) Calculate the adaptive regularization parameter φ_i.

(3) Calculate $\bar{x}_i$ for ARKFCM_1 and ARKFCM_2 or $\bar{\xi}$ for ARKFCM_w.

(4) Calculate cluster centers $v_j^{(t)}$ using $u^{(t)}$ as in (19).

(5) Calculate the membership function $u^{(t+1)}$ with (18).

(6) If max $\|u^{(t+1)} - u^{(t)}\| < \varepsilon$ or $t > 100$ then stop; otherwise, update $t = t + 1$ and go to step (4).

4. Experiments

In this section, we present the experiments on both synthetic and clinical MR images. For validation, the proposed algorithms (ARKFCM_1, ARKFCM_2, and ARKFCM_w) are compared with 6 recent soft clustering algorithms, namely, GKFCM1 [12], GKFCM2 [12], FLICM [13], KWFLICM [14], MICO [21], and RSCFCM [23]. As we are unable to have faithful implementation of RSCFCM [23], RSCFCM is compared only against the common data with results reported from [23]. The algorithms are implemented using MATLAB software package (a demo version is freely available online (http://www.mathworks.com/matlabcentral/fileexchange/54141-arkfcm-algorithm)). All the experiments are conducted with window size of 3×3 pixels, maximum number of iterations $t = 100$, and $\varepsilon = 0.001$. The accuracy of segmentation is measured using the Jaccard Similarity (JS) metric [28] which is defined as the ratio between the intersection and union of segmented volume S_1 and ground truth volume S_2:

$$\text{JS}(S_1, S_2) = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|}. \tag{20}$$

4.1. Experiments on Synthetic Brain MR Images. The following experiments are carried out using Simulated Brain Database (SBD) [29] which contains a set of realistic MR volumes produced by an MR imaging simulator with ground truths of CSF, GM, and WM available.

The first experiment is to segment a T1-weighted axial slice (number 100) with 217×181 pixels corrupted with 7% noise and 20% grayscale nonuniformity into WM, GM, and CSF. Figure 2 shows the segmentation results while Table 1 summarizes the JS and average running times.

The second experiment is to segment a T1-weighted sagittal slice (number 100) with 181×217 pixels corrupted with 7% noise and 20% grayscale nonuniformity. This image is chosen to show the capability of preserving details.

TABLE 1: JS and running time on the T1-weighted axial slice (number 100) from SBD with 7% noise and 20% grayscale nonuniformity.

Algorithm	GKFCM1	GKFCM2	FLICM	KWFLICM	MICO	RSCFCM	ARKFCM$_1$	ARKFCM$_2$	ARKFCM$_w$
WM	0.930	0.933	0.941	0.937	0.894	0.898	0.940	0.941	0.940
GM	0.822	0.855	0.860	0.852	0.782	0.842	0.864	0.868	0.865
CSF	0.781	0.847	0.829	0.805	0.860	0.882	0.863	0.867	0.867
Average	0.844	0.879	0.876	0.865	0.845	0.874	**0.889**	**0.892**	**0.891**
Time (s)	0.911	0.586	3.403	139.470	0.673	2.530	**0.356**	**0.282**	**0.329**

TABLE 2: JS and running time on the T1-weighted sagittal slice (number 100) from SBD with 7% noise and 20% grayscale nonuniformity.

Algorithm	GKFCM1	GKFCM2	FLICM	KWFLICM	MICO	ARKFCM$_1$	ARKFCM$_2$	ARKFCM$_w$
WM	0.773	0.775	0.771	0.765	0.672	0.785	0.788	0.786
GM	0.796	0.806	0.794	0.791	0.669	0.815	0.816	0.816
CSF	0.834	0.852	0.835	0.824	0.869	0.871	0.872	0.875
Average	0.801	0.811	0.800	0.794	0.736	**0.824**	**0.825**	**0.825**
Time (s)	1.377	1.797	5.392	166.717	0.751	**0.348**	**0.329**	**0.322**

| (a) | (b) | (c) | (d) | (e) | (f) |

| (g) | (h) | (i) | (j) | (k) |
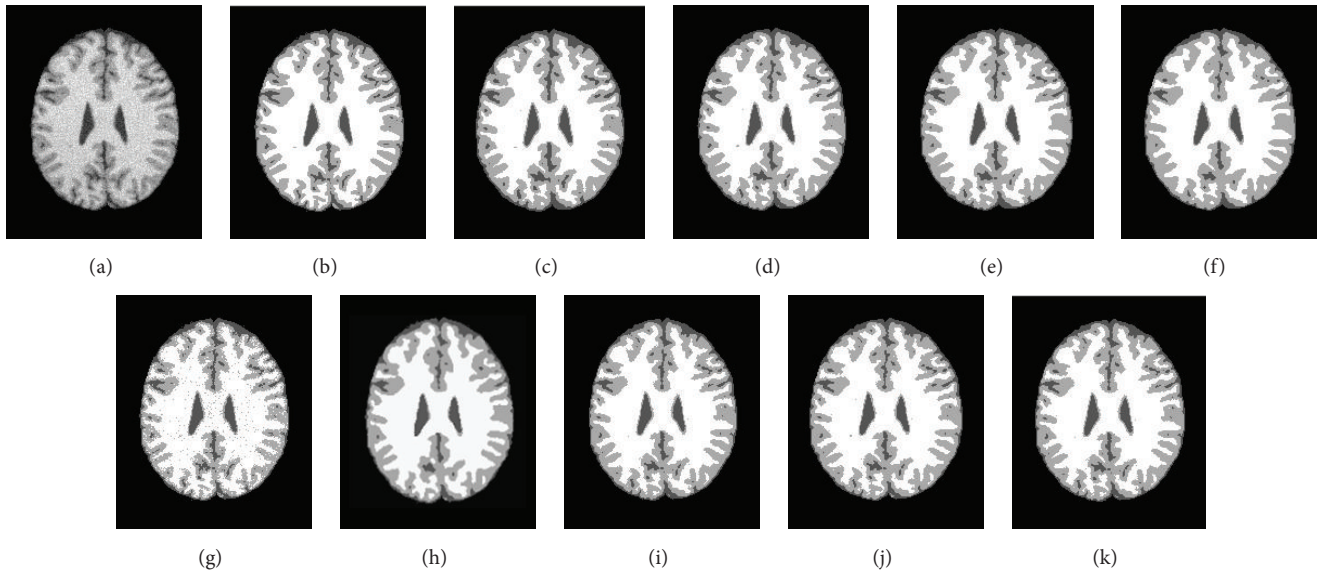
FIGURE 2: Segmentation results on a T1-weighted axial slice (number 100) from SBD with 7% noise and 20% grayscale nonuniformity. (a) Original image. (b) Ground truth. (c) GKFCM1 results. (d) GKFCM2 results. (e) FLICM results. (f) KWFLICM results. (g) MICO results. (h) RSCFCM results. (i) ARKFCM$_1$ results. (j) ARKFCM$_2$ results. (k) ARKFCM$_w$ results.

The segmentation results and JS are presented in Figure 3 and Table 2, respectively.

The third experiment is to check the robustness to Rician noise that commonly affects MR images [30]. The segmentation results and the JS with average running times of a T1-weighted axial slice (number 91) with 217×181 pixels corrupted with 10% Rician noise are given in Figure 4 and Table 3, respectively.

4.2. Experiments on Clinical Brain MR Images with Tumors. We experimented two T1-weighted axial slices (slices numbers 80 and 86, denoted, resp., as Brats1 and Brats2) with 240×240 pixels, respectively, from files pat266_1 and pat192_1 (available from MICCAI BRATS 2014 challenge, https://www.virtualskeleton.ch/BRATS/Start2014) (Figures 5 and 6).

From black to white are, respectively, background, CSF, GM, and WM. It should be noted that clustering is carried out only for CSF, GM, and WM, with the pathology region being considered as background.

Since images Brats1 and Brats2 come only with ground truth for the pathology and have no ground truth for normal tissues, a quantification measure other than JS should be used. To this end, an entropy-based metric proposed by Zhang et al. [31] is adopted. This is to maximize the uniformity of pixels within each segmented region and to minimize the uniformity across the regions. First, the entropy of every region j is calculated:

$$H\left(R_j\right) = -\sum_{m \in V_j} \frac{L_j(m)}{S_j} \log \frac{L_j(m)}{S_j}, \tag{21}$$

TABLE 3: JS and running time on the T1-weighted axial slice (number 91) from SBD corrupted with 10% Rician noise.

Algorithm	GKFCM1	GKFCM2	FLICM	KWFLICM	MICO	ARKFCM$_1$	ARKFCM$_2$	ARKFCM$_w$
WM	0.931	0.921	0.929	0.927	0.885	0.926	0.925	0.922
GM	0.804	0.827	0.831	0.821	0.777	0.838	0.840	0.834
CSF	0.826	0.872	0.861	0.846	0.889	0.887	0.892	0.887
Average	0.850	0.874	0.874	0.865	0.850	**0.884**	**0.886**	**0.881**
Time (s)	1.836	1.869	2.953	124.922	0.649	**0.218**	**0.220**	**0.218**
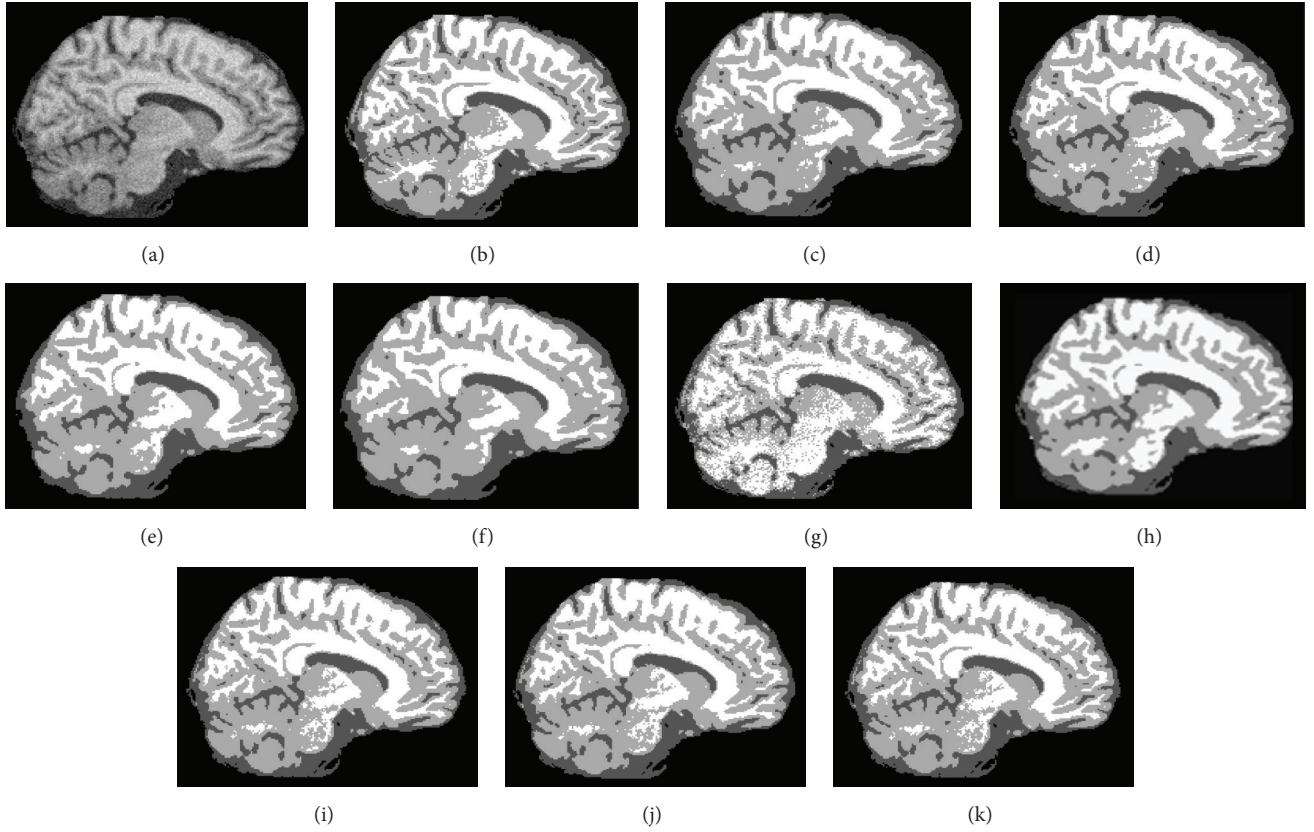
FIGURE 3: Segmentation results on a T1-weighted sagittal slice (number 100) from SBD with 7% noise and 20% grayscale nonuniformity. (a) Original image. (b) Ground truth. (c) GKFCM1 results. (d) GKFCM2 results. (e) FLICM results. (f) KWFLICM results. (g) MICO results. (h) RSCFCM results. (i) ARKFCM$_1$ results. (j) ARKFCM$_2$ results. (k) ARKFCM$_w$ results.

where V_j is the set of all possible grayscales in region j, $L_j(m)$ is the number of pixels belonging to region j with grayscale m, and S_j is the area of region j. For the grayscale image I, the E measure is calculated as

$$E = \sum_{j=1}^{c} \left(\frac{S_j}{S_I}\right) H\left(R_j\right) + \left(-\sum_{j=1}^{c} \left(\frac{S_j}{S_I}\right) \log \frac{S_j}{S_I}\right), \quad (22)$$

where the first term represents the expected region entropy of the segmentation and the second term is the layout entropy. The measure E yields smaller value for better segmentation and higher otherwise. Table 4 summarizes the segmentation accuracy in terms of E and running times for Brats1 and Brats2. It should be noted that the E metric is calculated only for WM, GM, and CSF regions without considering the background.

5. Discussion

FCM clustering is a well-known soft clustering method that assigns a membership degree for each pixel to every cluster. As the segmentation accuracy of FCM algorithm decays in the presence of noise, artifacts, and increased number of clusters, many investigations have been carried out on using contextual information to enhance the quality of segmentation. But how the contextual information shall be used effectively is still a challenge.

We propose an adaptively regularized kernel-based FCM clustering framework, with new parameter φ_i that adaptively controls the contextual information according to the heterogeneity of grayscale distribution within the local neighborhood. The new parameter is estimated using local variation coefficient among pixels within a specified neighborhood. A weighted image is devised that combines the original

TABLE 4: Segmentation performance measure E and running time on the Brats1 and Brats2 images.

Image	Measure	GKFCM1	GKFCM2	FLICM	KWFLICM	MICO	ARKFCM$_1$	ARKFCM$_2$	ARKFCM$_w$
Brats1	E	1.311	1.288	1.392	1.424	1.309	**1.274**	**1.270**	**1.279**
	Time (s)	2.591	1.516	9.510	635.413	1.747	**1.185**	**1.111**	**1.272**
Brats2	E	1.307	1.297	1.336	1.340	1.298	**1.273**	**1.271**	**1.279**
	Time (s)	1.576	1.115	5.170	400.295	1.546	**1.091**	**0.827**	**0.975**
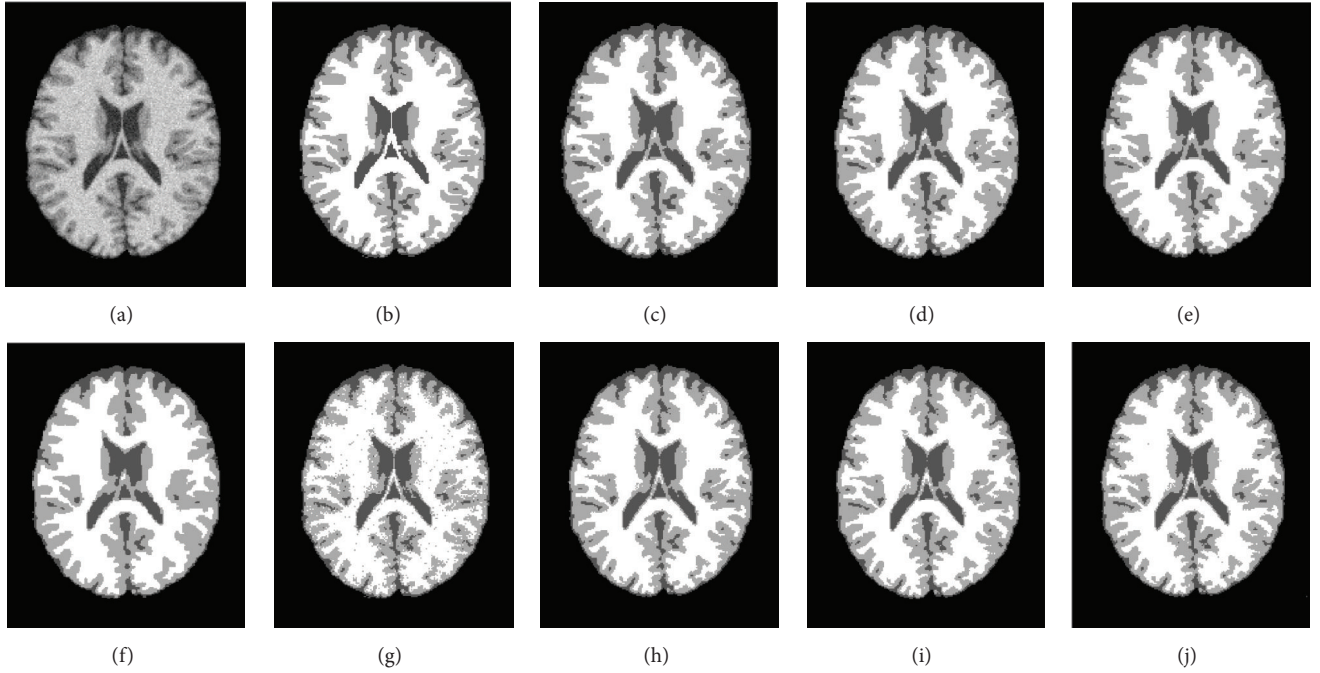
FIGURE 4: Segmentation results on a T1-weighted axial slice (number 91) from SBD with 10% Rician noise. (a) Original image. (b) Ground truth. (c) GKFCM1 results. (d) GKFCM2 results. (e) FLICM results. (f) KWFLICM results. (g) MICO results. (h) ARKFCM$_1$ results. (i) ARKFCM$_2$ results. (j) ARKFCM$_w$ results.
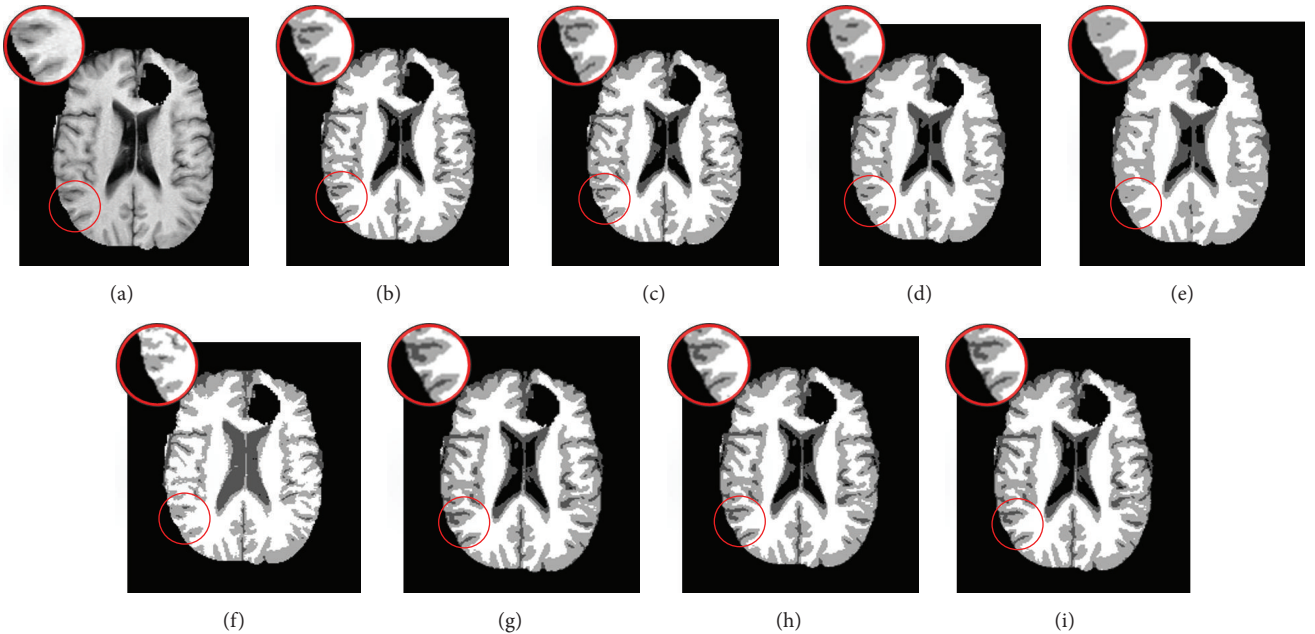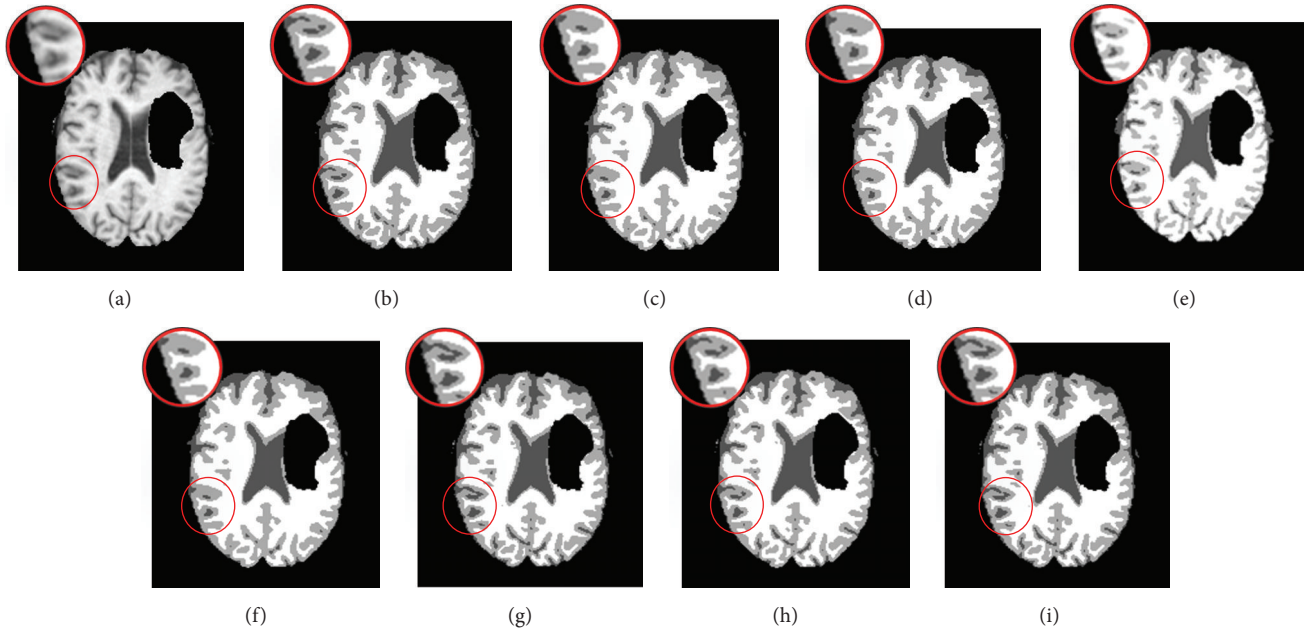

FIGURE 5: Segmentation results on the Brats1 image. (a) Original image. (b) GKFCM1 results. (c) GKFCM2 results. (d) FLICM results. (e) KWFLICM results. (f) MICO results. (g) ARKFCM$_1$ results. (h) ARKFCM$_2$ results. (i) ARKFCM$_w$ results.

FIGURE 6: Segmentation results on the Brats2 image. (a) Original image. (b) GKFCM1 results. (c) GKFCM2 results. (d) FLICM results. (e) KWFLICM results. (f) MICO results. (g) ARKFCM$_1$ results. (h) ARKFCM$_2$ results. (i) ARKFCM$_w$ results.

image and the parameter φ_i to represent the image contextual information embedded through the weighting procedure. Furthermore, a GRBF is adopted to replace Euclidean distance for better partitioning and to be less sensitive to outliers. The proposed framework can be in the form of 3 algorithms: ARKFCM$_1$, ARKFCM$_2$, and ARKFCM$_w$ for the local average grayscale to be replaced with the grayscale of the average filter, median filter, and the devised weighted image, respectively. Experiments have been carefully carried out to show the superiority of the proposed algorithms in comparison with 6 recent soft clustering algorithms to be discussed further.

5.1. Segmentation Accuracy. The differences in JS between the proposed algorithms and the other 6 algorithms are clear for the axial slice with 7% noise and 20% grayscale nonuniformity (Table 1 and Figure 2) and become more distinctive in preserving details in the sagittal slice corrupted with the same noise (Table 2 and Figure 3). The average JSs of ARKFCM$_1$, ARKFCM$_2$, and ARKFCM$_w$ are, respectively, 0.889, 0.892, and 0.891 for the axial slice as shown in Figure 2, which are better than the other 6 algorithms with range 1.3–4.8% (Table 1). For the sagittal slice with 7% noise and 20% grayscale nonuniformity shown in Figure 3, the average JSs of ARKFCM$_1$, ARKFCM$_2$, and ARKFCM$_w$ are, respectively, 0.824, 0.825, and 0.825, which are better than the other algorithms with range 1.4–8.9% (Table 2). For the axial slice corrupted with 10% Rician noise shown in Figure 4, the average JSs of ARKFCM$_1$, ARKFCM$_2$, and ARKFCM$_w$ are, respectively, 0.884, 0.886, and 0.881, which are better than the other 6 algorithms with range 1.20–3.6% (Table 3). The higher JSs may imply that the proposed algorithms achieve a better balance in preserving image details in the presence of noise and grayscale inhomogeneity.

For the clinical brain MR images with tumors (Brats1 and Brats2) shown in Figures 5 and 6, the proposed algorithms achieve smaller E than other algorithms (Table 4). For Brats1/Brats2 image, ARKFCM$_2$ attains the lowest E of 1.270/1.271, which is smaller than the other 5 algorithms with ranges 0.018–0.154 and 0.025–0.109, respectively. Visibly, GKFCM1 (Figures 5(b) and 6(b)) and GKFCM2 (Figures 5(c) and 6(c)) come with good results but they are unable to preserve small details such as CSF being wrongly broken possibly due to the difficulty in estimating the parameter η_j. On the other hand, FLICM (Figures 5(d) and 6(d)) and KWFLICM (Figures 5(e) and 6(e)) yield smooth results and do not preserve many details which affects the segmentation accuracy of CSF (CSF is small as compared with surrounding GM and WM) due to the effects of G_{ij} and $\acute{G}_{ij}$, respectively. For MICO algorithm (Figures 5(f) and 6(f)), the edges are not smooth enough; hence the CSF is not well preserved. Finally, results of ARKFCM$_1$, ARKFCM$_2$, and ARKFCM$_w$ (Figures 5(g) and 6(g), Figures 5(h) and 6(h), and Figures 5(i) and 6(i)) show good balance between smooth borders and preserving image details due to the introduction of adaptive local contextual information measure φ_i to replace the fixed value of α or the oversmoothing factors G_{ij} or $\acute{G}_{ij}$.

5.2. Computational Cost. In terms of computational cost, the objective function of FCM algorithm in its original form [6] contains only the difference between the grayscale of the current pixel i and the cluster centers v_j. This is basically to cluster grayscales as there is no spatial information, so it has the smallest computational cost and can be implemented based on grayscale histogram to reduce the computational cost further [9, 11]. The enhancement of original FCM is to add the local contextual information to make it robust

to noise and image artifacts at the expense of increased computational cost [7–14].

The GKFCM1 and GKFCM2 algorithms [12] use the parameter η_j to replace α and need an additional loop on the number of clusters to be calculated at each pixel to update the local contextual information. So they have higher computational cost than the original FCM, FCM_S1, and FCM_S2 at each iteration.

The FLICM algorithm [13] introduces G_{ij} which needs an additional loop on the neighborhood of the current pixel to calculate the local information in every iteration; thus it has a high computational cost. As an extension of FLICM, KWFLICM [14] introduces $\acute{G}_{ij}$ that requires two additional loops on the neighborhood, so it has the highest computational cost at each iteration.

The RSCFCM algorithm [21] uses a spatial fuzzy factor that is constructed based on the posterior and prior probabilities and takes the spatial direction into account. That increases the complexity as many parameters have to be optimized and consequently the computational complexity.

MICO algorithm [23] comes with fast calculations due to the convexity of its energy function particularly in the presence of less noise but tends to have many iterations in the presence of high level noise.

The proposed algorithms incorporate the local contextual information by introducing LVC, which is a measure of grayscale heterogeneity and has nothing to do with the cluster centers. So it can be calculated once in advance and hence reduce the complexity of the clustering procedure.

The eventual computational cost will be the multiplication of the computational cost for each iteration and the number of iterations for convergence. The number of iterations will be dependent on the initialization as well as the objective function. To make fair comparisons, initializations are all set randomly and the average number of iterations for convergence is then recorded for 10 converged times. From Figure 7, it can be seen that the average iteration times will be data dependent, with $ARKFCM_2$, $ARKFCM_w$, and $ARKFCM_1$ having the minimum number of iterations followed by FLICM, GKFCM2, and GKFCM1, respectively.

The running times for the tested images agree well with the above analysis as illustrated in Figure 8. For SBD images given in Figures 2, 3, and 4, the KWFLICM algorithm takes the longest time (124–166 seconds), followed by FLICM (3–5.4 seconds), RSCFCM (about 2.5 seconds), GKFCM1/GKFCM2 (0.6–1.9 seconds), and the proposed algorithms (0.22 to 0.36 seconds). For the clinical MR images in Figures 5 and 6, it takes more time (due to more iterations) as the images are more complicated, but the trend remains the same with the proposed algorithms having the lowest computational cost (Table 4 and Figure 8).

5.3. Neighborhood Size.
The local neighborhood window size is a crucial factor to determine the smoothness of clustering and details to be preserved. We have experimented with different window sizes and found that a window size of 3×3 pixels achieves the best balance between segmentation accuracy and computational cost. Increasing window size to 5×5 pixels has very small impact on the JS but 7×7 pixels
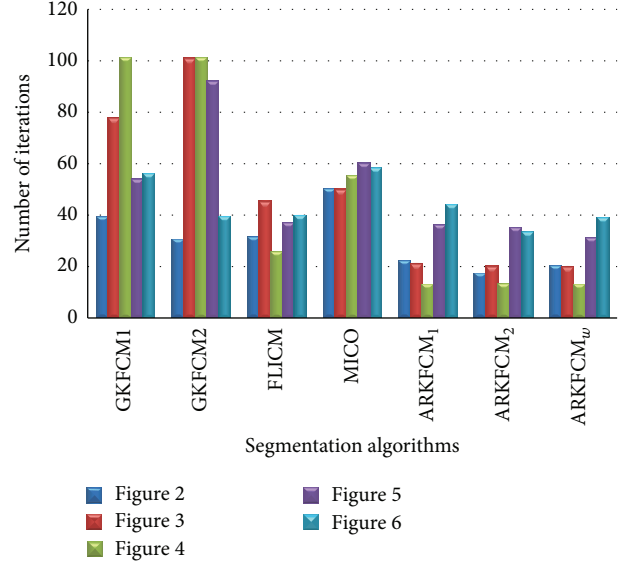

FIGURE 7: Average number of iterations of all the algorithms (except RSCFCM and KWFLICM) on Figures 2(a), 3(a), 4(a), 5(a), and 6(a).
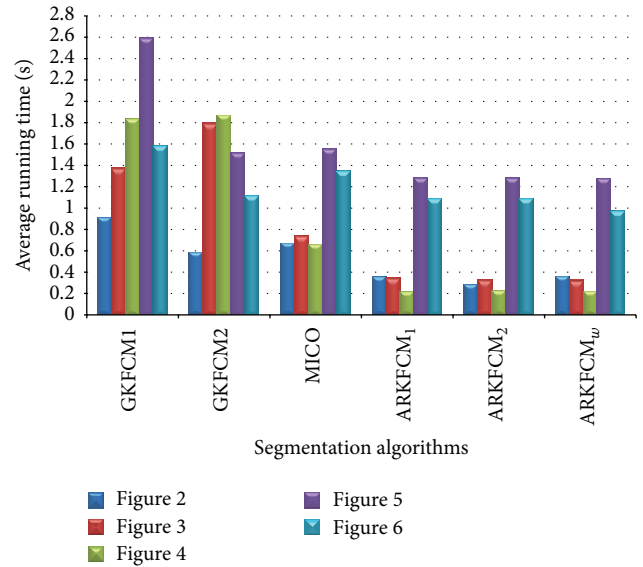

FIGURE 8: Average running times of all the algorithms (except FLICM, RSCFCM, and KWFLICM) on Figures 2(a), 3(a), 4(a), 5(a), and 6(a).

or more will, significantly, decrease the accuracy, such as losing image details like CSF. Therefore, it is recommended to use a local window of size 3×3 pixels for constructing the neighborhood.

5.4. Limitation.
From the experiments, it was found that the proposed algorithms could be sensitive to severe noise in small areas between edges that have width 1 or 2 pixels (e.g., area between ventricles, Figures 4(h), 4(i), and 4(j)). Potentially, this may be solved by embedding edge detection technique into the clustering process which is yet to be explored.

6. Conclusion

An adaptively regularized kernel-based FCM framework has been proposed to enhance the original FCM for higher segmentation accuracy at low computational cost. The framework can be in the form of three algorithms that employ the heterogeneity of grayscales in the neighborhood employed for local contextual information. The main advantages are adaptiveness to local context, enhanced robustness, and independence of clustering parameters to decrease computational cost. GRBF kernel has been adopted as a distance metric. We validated the proposed algorithms on synthetic and clinical MR images with tumors. The proposed algorithms attain a higher JS (Tables 1, 2, and 3) and lower entropy measure E (Table 4) than the 6 recent soft clustering algorithms and could preserve small image details (Figures 2, 3, 4, 5, and 6). In addition, the proposed algorithms have a low computational cost and are, to the best of our knowledge, the only algorithms that are adaptive to local context and do not include cluster centers. Therefore, they attain a trade-off between high segmentation accuracy and low computational cost. The proposed algorithms can be a potential tool for segmenting brain MR images for further processing and other images as well.

Appendix

The objective function J_{ARKFCM} given in equation (17) with conditions in (2) is a constrained minimization problem which can be solved using Lagrange multiplier method. Let the minimization problem of J_{ARKFCM} be written as follows:

$$
\begin{aligned}
L_{\text{ARKFCM}} &\left(u_{ij}, v_j, \lambda_i, \varphi_i\right) \\
&= 2\sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \left(1 - K\left(x_i, v_j\right)\right) \\
&\quad + 2\sum_{i=1}^{N} \sum_{j=1}^{c} \varphi_i u_{ij}^m \left(1 - K\left(\overline{x}_i, v_j\right)\right) \\
&\quad + \sum_{i=1}^{N} \lambda_i \left(1 - \sum_{j=1}^{c} u_{ij}\right).
\end{aligned}
\tag{A.1}
$$

By taking the derivative of L_{ARKFCM} with respect to u_{ij} and setting the result to zero, for $m > 1$, we have

$$
\begin{aligned}
\frac{\partial L_{\text{ARKFCM}}}{\partial u_{ij}} &= 2mu_{ij}^{m-1}\left(1 - K\left(x_i, v_j\right)\right) \\
&\quad + 2m\varphi_i u_{ij}^{m-1}\left(1 - K\left(\overline{x}_i, v_j\right)\right) - \lambda_i \\
&= 0.
\end{aligned}
\tag{A.2}
$$

By using (A.2) we can calculate u_{ij} as follows:

$$
\begin{aligned}
u_{ij} \\
= \left[\frac{\lambda_i}{2m\left(1 - K\left(x_i, v_j\right) + \varphi_i\left(1 - K\left(\overline{x}_i, v_j\right)\right)\right)}\right]^{1/(m-1)}.
\end{aligned}
\tag{A.3}
$$

Since $\sum_{k=1}^{c} u_{ik} = 1 \ \forall v_k$, from (A.3) we can get

$$
\sum_{k=1}^{c} \left[\frac{\lambda_i}{2m\left(1 - K\left(x_i, v_k\right) + \varphi_i\left(1 - K\left(\overline{x}_i, v_k\right)\right)\right)}\right]^{1/(m-1)} = 1,
\tag{A.4}
$$

$$
\begin{aligned}
\lambda_i \\
= \frac{1}{\left[\sum_{k=1}^{c}\left(2m\left(1 - K\left(x_i, v_k\right) + \varphi_i\left(1 - K\left(\overline{x}_i, v_k\right)\right)\right)\right)^{-1/(m-1)}\right]^{m-1}}.
\end{aligned}
\tag{A.5}
$$

By substituting (A.5) into (A.3) with necessary calculations, the final u_{ij} is

$$
\begin{aligned}
u_{ij} \\
= \frac{\left(\left(1 - K\left(x_i, v_j\right)\right) + \varphi_i\left(1 - K\left(\overline{x}_i, v_j\right)\right)\right)^{-1/(m-1)}}{\sum_{k=1}^{c}\left(1 - K\left(x_i, v_k\right) + \varphi_i\left(1 - K\left(\overline{x}_i, v_k\right)\right)\right)^{-1/(m-1)}}.
\end{aligned}
\tag{A.6}
$$

Similarly, by taking the derivative of L_{ARKFCM} with respect to v_j and setting the result to zero, we have

$$
\begin{aligned}
\frac{\partial L_{\text{ARKFCM}}}{\partial v_j} &= \sum_{i=1}^{N} u_{ij}^m K\left(x_i, v_j\right)\left(x_i - v_j\right) \\
&\quad + \sum_{i=1}^{N} \varphi_i u_{ij}^m K\left(\overline{x}_i, v_j\right)\left(\overline{x}_i - v_j\right) = 0.
\end{aligned}
\tag{A.7}
$$

After necessary calculations, the final v_j is

$$
v_j = \frac{\sum_{i=1}^{N} u_{ij}^m \left(K\left(x_i, v_j\right)x_i + \varphi_i K\left(\overline{x}_i, v_j\right)\overline{x}_i\right)}{\sum_{i=1}^{N} u_{ij}^m \left(K\left(x_i, v_j\right) + \varphi_i K\left(\overline{x}_i, v_j\right)\right)}.
\tag{A.8}
$$

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] N. Sharma and L. M. Aggarwal, "Automated medical image segmentation techniques," *Journal of Medical Physics*, vol. 35, no. 1, pp. 3–14, 2010.

[2] A. Elazab, Q. Hu, F. Jia, and X. Zhang, "Content based modified reaction-diffusion equation for modeling tumor growth of low grade glioma," in *Proceedings of the 7th Cairo International Biomedical Engineering Conference (CIBEC '14)*, pp. 107–110, Giza, Egypt, December 2014.

[3] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annual Review of Biomedical Engineering*, vol. 2, pp. 315–337, 2000.

[4] I. Despotović, B. Goossens, and W. Philips, "MRI segmentation of the human brain: challenges, methods, and applications," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 450341, 23 pages, 2015.

[5] S. Ruan, C. Jaggi, J. Xue, J. Fadili, and D. Bloyet, "Brain tissue classification of magnetic resonance images using partial volume modeling," *IEEE Transactions on Medical Imaging*, vol. 19, no. 12, pp. 1172–1186, 2000.

[6] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, vol. 3, no. 3, pp. 32–57, 1973.

[7] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Advanced Applications in Pattern Recognition, Springer, Boston, Mass, USA, 1981.

[8] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 21, no. 3, pp. 193–199, 2002.

[9] L. Szilagyi, Z. Benyo, S. M. Szilagyi, and H. S. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 724–726, Cancun, Mexico, September 2003.

[10] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 34, no. 4, pp. 1907–1916, 2004.

[11] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognition*, vol. 40, no. 3, pp. 825–838, 2007.

[12] M.-S. Yang and H.-S. Tsai, "A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction," *Pattern Recognition Letters*, vol. 29, no. 12, pp. 1713–1725, 2008.

[13] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1328–1337, 2010.

[14] M. Gong, Y. Liang, J. Shi, W. Ma, and J. Ma, "Fuzzy C-means clustering with local information and kernel metric for image segmentation," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 573–584, 2013.

[15] M. A. Balafar, "Gaussian mixture model based segmentation methods for brain MRI images," *Artificial Intelligence Review*, vol. 41, no. 3, pp. 429–439, 2014.

[16] C. Nikou, N. P. Galatsanos, and A. C. Likas, "A class-adaptive spatially variant mixture model for image segmentation," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 1121–1130, 2007.

[17] A. R. Ferreira da Silva, "A Dirichlet process mixture model for brain MRI tissue classification," *Medical Image Analysis*, vol. 11, no. 2, pp. 169–182, 2007.

[18] T. M. Nguyen and Q. M. J. Wu, "Fast and robust spatially constrained gaussian mixture model for image segmentation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 4, pp. 621–635, 2013.

[19] S. P. Chatzis and T. A. Varvarigou, "A fuzzy clustering approach toward hidden Markov random field models for enhanced spatially constrained image segmentation," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 5, pp. 1351–1361, 2008.

[20] S. Chatzis, "A method for training finite mixture models under a fuzzy clustering principle," *Fuzzy Sets and Systems*, vol. 161, no. 23, pp. 3000–3013, 2010.

[21] Z. Ji, J. Liu, G. Cao, Q. Sun, and Q. Chen, "Robust spatially constrained fuzzy c-means algorithm for brain MR image segmentation," *Pattern Recognition*, vol. 47, no. 7, pp. 2454–2466, 2014.

[22] C. Li, C. Xu, A. Anderson, and J. C. Gore, "MRI tissue classification and bias field estimation based on coherent local intensity clustering: a unified energy minimization framework," in *Information Processing in Medical Imaging: 21st International Conference, IPMI 2009, Williamsburg, VA, USA, July 5–10, 2009. Proceedings*, vol. 5636 of *Lecture Notes in Computer Science*, pp. 288–299, Springer, Berlin, Germany, 2009.

[23] C. Li, J. C. Gore, and C. Davatzikos, "Multiplicative intrinsic component optimization (MICO) for MRI bias field estimation and tissue segmentation," *Magnetic Resonance Imaging*, vol. 32, no. 7, pp. 913–923, 2014.

[24] L. Szilágyi, "Lessons to learn from a mistaken optimization," *Pattern Recognition Letters*, vol. 36, no. 1, pp. 29–35, 2014.

[25] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *The Annals of Statistics*, vol. 36, no. 3, pp. 1171–1220, 2008.

[26] K.-R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, "An introduction to kernel-based learning algorithms," *IEEE Transactions on Neural Networks*, vol. 12, no. 2, pp. 181–201, 2001.

[27] C. R. Souza, *Kernel Functions for Machine Learning Applications*, 2010, http://crsouza.com/2010/03/kernel-functions-for-machine-learning-applications/.

[28] U. Vovk, F. Pernuš, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI," *IEEE Transactions on Medical Imaging*, vol. 26, no. 3, pp. 405–421, 2007.

[29] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans, "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database," 2011, http://brainweb.bic.mni.mcgill.ca/brainweb/.

[30] L. He and I. R. Greenshields, "A non-local maximum likelihood estimation method for Rician noise reduction in MR images," *IEEE Transactions on Medical Imaging*, vol. 28, no. 2, pp. 165–172, 2009.

[31] H. Zhang, J. E. Fritts, and S. A. Goldman, "An entropy-based objective evaluation method for image segmentation," in *Proceedings of the Storage and Retrieval Methods and Applications for Multimedia*, vol. 5307 of *Proceedings of SPIE*, pp. 38–49, January 2004.