Hindawi
Mathematical Problems in Engineering
Volume 2019, Article ID 5984649, 17 pages
https://doi.org/10.1155/2019/5984649

Research Article

A Robust Bias-Correction Fuzzy Weighted C-Ordered-Means Clustering Algorithm

Wenyuan Zhang,[1,2] Tianyu Huang ,[1] and Jun Chen[1]

[1]*College of Information Science and Engineering Yanshan University, Qinhuangdao, Hebei, China*
[2]*College of Mathematics and Information & Science Technology Hebei Normal University of Science & Technology, Qinhuangdao, Hebei, China*

Correspondence should be addressed to Tianyu Huang; htyu3000@163.com

Received 19 January 2019; Revised 12 April 2019; Accepted 15 May 2019; Published 18 June 2019

Academic Editor: Bonifacio Llamazares

This paper proposes a modified fuzzy C-means (FCM) algorithm, which combines the local spatial information and the typicality of pixel data in a new fuzzy way. This new algorithm is called bias-correction fuzzy weighted C-ordered-means (BFWCOM) clustering algorithm. It can overcome the shortcomings of the existing FCM algorithm and improve clustering performance. The primary task of BFWCOM is the use of fuzzy local similarity measures (space and grayscale). Meanwhile, this new algorithm adds a typical analysis of data attributes to membership, in order to ensure noise insensitivity and the preservation of image details. Secondly, the local convergence of the proposed algorithm is mathematically proved, providing a theoretical preparation for fuzzy classification. Finally, data classification and real image experiments show the effectiveness of BFWCOM clustering algorithm, having a strong denoising and robust effect on noise images.

1. Introduction

Clustering analysis is an important technique in data analysis, having a wide range of applications in data mining technique [1], image processing [2, 3], computer vision [4], and artificial intelligence [5]. Image segmentation is a type of image processing, dividing an image into specific nonoverlapping regions with unique characteristics, and putting forward goals of interest. It is a key step from image processing to image analysis. The existing image segmentation methods are mainly divided into the following categories: threshold-based segmentation method, region-based segmentation method, edge-based segmentation method, specific theory-based segmentation method, etc. Specifically, these methods include clustering [6, 7], region growing [8], Watershed Transformation [9], active contour model [10], MeanShift [11], Graph Cut [12], spectral clustering [13], Markov random field [14], Neural network [15], etc. The process of image segmentation is also a marking process, in which pixels belonging to the same area are given the same number. Image segmentation has always been one of the most challenging tasks in image processing and computer vision [16, 17]. Nowadays, there have been many methods for image segmentation [18–21]; however, they are not robust and effective enough for a large number of different images.

How to segment a given image? And how to segment it in a rational way? These are the major problems which need to be solved by the clustering algorithm. A solution to such problems: under a certain measure, the data in classes is made to have as short distance within classes as possible and as large distance between classes as possible and meanwhile the summary of whole distance is minimum. FCM clustering algorithm was first proposed by Dunn [22, 23] and later developed by Bezdek [24–30]. It is a classical clustering algorithm, a combination of clustering analysis and fuzzy theory. It is an unsupervised machine learning algorithm, which can be divided by constantly modifying the clustering center and membership matrix. FCM algorithm is sensitive to noise due to its restricted condition that total membership of elements to all centers is 1. Thus, local spatial information is necessarily introduced into the objective function to improve the robustness of the FCM algorithm for image segmentation.

This paper proposes a new algorithm to make the algorithm more robust by a detailed study of FCM clustering algorithm. Its formula is as (∗).

1.1. Introduction to FCM Clustering with Local Spatial Information. The traditional Hard C-means (HCM) clustering algorithm is very simple (the membership degree is only 0 or 1) [31], resulting in poor image segmentation, while FCM is more tolerant of ambiguity and retain more original image information. In addition, it is effective for images with simple texture and background. Therefore, FCM is better than HCM. The shortcoming of FCM is that it only takes the gray information into consideration while ignoring the spatial information. This algorithm cannot segment images with complex texture and background, because it cannot avoid the noise interference. In order to solve the problems mentioned above, many scholars have proposed to introduce local spatial information into the objective function. For example, Tolias [32] proposed to add spatial information to membership degree. Ahmed [33] (1999) and Pham [34] (2001) proposed a bias-correction FCM algorithm. Considering that the membership degree of the central pixel is affected by the pixels in the neighborhood, this algorithm adds the neighborhood pixel information to the objective function of the FCM algorithm as a penalty term and attempts to make neighborhood pixels and central pixels produce more consistent clustering results in the segmentation process. Liew [35] used a new neighborhood weighted dissimilarity measure for the Euclidean distance in the FCM objective function. Compared with FCM, this algorithm uses more spatial information, which can better realize the clustering and segmentation of image data. In order to make the FCM clustering algorithm perform better, Szilćgyi [36] introduced a new factor, starting from the standard FCM and its bias-corrected version. An enhanced fuzzy C-means clustering (EnFCM) algorithm was proposed which can reduce the amount of computation required, and provides a fast way to segment high quality brain images. Wu [37] replaced the Euclidean norm with an exponential function in C-means clustering, greatly improving the antinoise ability of the algorithm. Chen and Zhang [38] proposed a FCM clustering algorithm based on median filtering or mean filtering and a kernel method for improving the algorithm. A class of robust non-Euclidean distance measures is introduced to derive new objective functions, enhancing the robustness of the original clustering algorithm to noise and outliers. Yang et al. [39] proposed a Gaussian kernel-based FCM (GKFCM) algorithm, a generalization of FCM, BCFCM, KFCMS1, and KFCMS2 algorithms, having higher efficiency and robustness. Cai, Chen, and Zhang [40] proposed a fast generalized FCM clustering algorithm. It uses new factors as a measure for local similarity (space and grayscale), thus ensuring the noise immunity and details of images and removing the global parameter α. However, there is still a parameter that needs to be experimentally determined in the algorithm. Wang et al. [41] proposed local and nonlocal FCM clustering algorithms to reduce the effects of noise in the image segmentation process. By using a new dissimilarity index to replace the usual distance measures, this method incorporates local spatial context and nonlocal information into the standard FCM clustering algorithm, where the image local information takes into account the weighted distance between the pixel and the class center.

Table 1: Noise immunity fuzzy clustering algorithm table.

Fuzzy clustering	The objective function
FCM	$J_m = \sum_{i=1}^{C}\sum_{k=1}^{N} u_{ik}^m \left\|x_k - v_i\right\|^2 . (*)$
FCMS FLICM FCMS1 FCMS2	$J_m = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^m \left\|x_k - v_i\right\|^2 + \sum_{i=1}^{c}\sum_{k=1}^{n} G_{ik}. (**)$
FCMS	$G_{ik} = \frac{\alpha}{N_R} u_{ik}^m \sum_{r\in N_k} \left\|x_r - v_i\right\|^2$ Among them: α is a parameter used to control the effect of the neighborhood terms. N_R is the cardinality of u_{ik}, x_r is the neighbor of, N_i is the neighbor set in the window around.
FLICM	$G_{ik} = \sum_{\substack{r\in N_i \\ i\neq r}} \frac{1}{d_{ir}+1} \left(1-u_{rk}\right)^m \left\|x_r - v_k\right\|^2 . (***)$ Among them: d_{ir} represents the spatial Euclidean distance between pixels v_i and x_r. Obviously, G_{ik} is more complicated than in FCMS, so FLICM has a higher computational complexity than FCMS.
FCMS1 FCMS2	$G_{ik} = \alpha {u_{ik}}^m \left\|\overline{x}_k - v_i\right\|^2 . (****)$ Among them: $\overline{x}_k$ is the average value or median value of adjacent pixels within the window surrounding. G_{ik} have a more simplified form in FCMS1 and FCMS2 than the FCMS, and the clustering time can be reduced because $\alpha \sum_{r\in N_i} \left\|x_r - v_k\right\| / N_R$ is replaced by $\alpha \left\|\overline{x}_k - v_i\right\|^2$.

The target criteria and differences of FCM, FCMS, FLICM, FCMS1, and FCMS2 are presented in Table 1.

Spatial neighborhood information is of great significance for image segmentation. FCMS and EnFCM [39] are good algorithms due to their short calculation time, because it performs a Gray Histogram-Based clustering rather than summing the pixels of images. The number of gray levels in an image is usually much smaller than that of pixels, and the calculation time is very short. However, only the neighborhood gray value information is added to the clustering process, while the distance information of the spatial neighborhood points is not taken into account. The effect of all neighborhood points on the central point is the same, which is obviously not in line with the actual situation. The neighborhood points that are closer to the central point should have more effect on the central point, while the neighborhood points that are farther away from the central point should have less effect on it.

However, when EnFCM algorithm constructs linear weighted sum images, the lack of spatial information makes

the image segmentation results sensitive to salt and pepper noise. In order to improve the segmentation results obtained by EnFCM, Cai et al. [40] proposed a fast generalized FCM algorithm (FGFCM). Inspired by the EnFCM algorithm, the FGFCM algorithm first redefines the linear weighted sum images by utilizing grayscale and spatial information of each pixel's neighborhood window of the original image and then performs clustering on its gray histogram. This algorithm introduces a new factor as a local similarity measure, which is aimed to ensure the noise immunity and detail retention of image segmentation. Meanwhile, the empirical adjustment parameter α required in EnFCM is removed, and finally a gray histogram is clustered. The FGFCM algorithm can further improve the balance between noise and detail, enhance the robustness of the algorithm to noise, and improve the computational efficiency of FCM to some extent. However, it requires more parameters than EnFCM. The contribution strength of the spatial neighborhood information of FGFCM algorithm is hard to control and needs to be set manually. In addition, the strength value is a global variable, and the distribution of noise cannot be fully considered.

In order to develop a new FCM algorithm, Krinidis and Chatzis [42] proposed a robust fuzzy local information C-mean (FLICM) clustering algorithm. The FLICM algorithm has the following advantages: first, G_{ij} (fuzzy factors) not containing parameters other than necessary m and C is used. Therefore, there is no need to adjust parameters (such as parameters in the EnFCM algorithm) to balance image noise and image detail. In addition, compared with the traditional FCM algorithm, it can greatly improve the segmentation effect for images with noise by using spatial and grayscale information. It replaces parameters used in the EnFCM with a new fuzzy factor, which can be translated into an objective function that ensure the noise immunity and image detail preservation. FLICM overcomes the problem of parameter selection and improves image segmentation performance, but it is not robust enough for local information of different images and fixed spatial distance. Meanwhile, FLICM uses variance information in the pixel neighborhood to modify the fuzzy weight factors, thus improving the robustness and noise immunity of this algorithm. However, the image denoising effect is still poor for weak contrast images.

Gong et al. [43] used variable local coefficients instead of fixed spatial distances and proposed a variant of FLICM (the formula as shown in Table 1 $(***)$). A modified robust FLICM clustering algorithm is obtained by replacing G_{ij} with a newly constructed fuzzy factor G'_{ij}, which can utilize more local context information in the image. In order to further improve the segmentation performance of the FLICM algorithm, Gong et al. [43] proposed a fuzzy factor that links local spatial information with local gray information to obtain a new algorithm called kernel weighted fuzzy local information C-means (KWFLICM). In addition, the fuzzy weighting factor is modified by variance information in the pixel neighborhood to adaptively control the local spatial relationship, thus enhancing the robustness of FLICM to noise and outliers. The KWFLICM algorithm replaces the Euclidean distance measures in the FLICM algorithm with the nuclear distance measures and proposes a new local information weighted fuzzy factor to improve the antinoise performance of this algorithm. However, KWFLICM has a higher computational complexity than FLICM.

Ahmed et al. [33] proposed an adaptively regularized kernel-based FCM framework (ARKFCM), aiming at the problem of the segmentation of brain magnetic resonance image. This algorithm utilizes the heterogeneity of grayscales in the neighborhood to obtain the local context information and replaces the Standard Euclidean distance with the Gaussian radial basis kernel functions. In this way, the robustness of preserving the image details is improved, the independence of cluster parameters is enhanced, and the computing costs are reduced. Lei T et al. [44] proposed an improved FCM algorithm based on morphological reconstruction and membership filtering (FRFCM), which is faster and more robust than FCM algorithm. By introducing morphological reconstruction operation, the local spatial information is incorporated into the FRFCM algorithm to ensure the antinoise ability and image detail preservation.

Saranathan [45] introduced the nonlocal spatial information into the objective function by using variant parameters adaptive to the noise level of each pixel of the image, which has been extended to FCM to overcome the parameter selection to improve the robustness to noise. The FGFCM algorithm and the FLICM algorithm both use coordinate information to add spatial distance information to the clustering process. Zhao et al. [46] believe that the Euclidean distance cannot reflect the spatial information of the image well. Therefore, they proposed a neighborhood weighted fuzzy C-means (NWFCM) clustering algorithm, namely, Neighborhood Weight (NW) distance to reduce the running time of FLICM and KWFLICM and strengthen the robustness of the algorithm to noise. Although NWFCM uses more parameters, NWFCM is faster than FLICM and KWFLICM because image filtering is performed before the iteration begins.

However, the NWFCM algorithm is still time-consuming due to the calculation of the NW distance and the parameter selection. Guo et al. [47] proposed an adaptive Noise Detection-based FCM (NDFCM). This algorithm overcomes the above-mentioned shortcoming, in which the trade-off parameters are automatically adjusted by measuring the local variance of the gray scale.

1.2. Introduction to Fuzzy C-Ordered-Means Clustering Algorithm. Fuzzy C-ordered-means (FCOM) [48] clustering algorithm combines FCM clustering with robust ordered statistics, in which Huber's M-estimation [49] and Yager's ordered weighted averaging (OWA) operators [50] are used together for fuzzy clustering to significantly improve its robustness. The typicality of each data item is calculated instead of attribute empowerment. The data items are sorted, and their typicality is updated in each iteration of the clustering process. The data closer to the cluster center are of high typicality, while those farther from the cluster center low typicality. In this way, outliers and noisy data items are classified as low typicality without damaging the clustering

process. Thus this method has better performance in terms of noise and outlier data. Its formula is as shown in (1).

Fuzzy weighted C-ordered-means (FWCOM) clustering algorithm [51] proposed a new fuzzy weighted C-ordered-means clustering algorithm based on FCOM. This algorithm looks for clusters within the fuzzy subspace of the original task space and assigns weights to dimensions (attributes) in each cluster. The weight is a number from the unit interval $[0, 1]$. In order to enhance the robustness of the algorithm to outliers and noise, it is combined with sorting techniques.

The formula of FWCOM [51] $\beta_{ik} \in [0, 1]$ is as follows:

$$J(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} \beta_{ik} (u_{ik})^m \sum_{d=1}^{p} (x_{kd} - v_{id})^2 \tag{1}$$

The restricted condition is as follows:

$$\forall k \in X$$
$$\sum_{i=1}^{c} \beta_{ik} u_{ik} = f_k \tag{2}$$

Its membership degree formula is as follows:

$$u_{ik} = \frac{f_i \left(\sum_{d=1}^{p} (x_{id} - v_{kd})^2 \right)^{1/(1-m)}}{\sum_{i=1}^{c} \beta_{ik} \left(\sum_{d=1}^{p} (x_{id} - v_{kd})^2 \right)^{1/(1-m)}} \tag{3}$$

The c-th clustering center of the d-th attribute is calculated using the following formula:

$$v_{cd} = \frac{\sum_{k=1}^{n} \beta_{ck} u_{ik}^m h(e_{ckd}) x_{kd}}{\sum_{k=1}^{n} \beta_{ck} u_{ik}^m h(e_{ckd})} \tag{4}$$

For the above-mentioned algorithm, this paper proposes an adaptive filtering based fuzzy clustering algorithm by combining the FWCOM clustering algorithm with the FCM clustering algorithm of spatial information. This algorithm first determines the nonlocal parameter balance factor according to the strength of the local spatial information, meanwhile calculating the typicality of each data item, sorting the data items, and updating their typicalities in each iteration of the clustering process, with an aim to more accurately reflect the spatial structure information contained in the image. Then it uses this balance factor to make an effective combination of the filtered image and the median filtered image of the original image. The obtained adaptive filtered image adaptively determines the degree of filtering according to the noise intensity, thus improving the algorithm's ability to suppress noise and the robustness of the algorithm.

Table 2 lists symbols used in the paper.

2. A New Algorithm BFWCOM with Incorporating Spatial Information and Fuzzy Weights

A robust algorithm should have the following properties: (1) It should have fairly good precision in the assumed model. (2) Small deviations from the model assumption should only cause damage to a small amount of performance. (3) Large deviations from the model assumption should not cause disaster.

The standard FCM objective function for partitioning $X = \{x_i\}_{i=1}^{n}$ into C cluster: this method is also used for clustering in the paper.

2.1. Criterion Function for BFWCOM Algorithm. The FCM clustering algorithm uses a quadratic loss function as a different measure between the data and the cluster center. The reason for using this method is for simplicity and low computational burden. However, this method is sensitive to noise and outliers. In many literatures, there are many suggestions for robust loss functions.

For example, Huber proposed a mathematical formula [49, 50]:

$$\tau_{HUB}(e) = \begin{cases} \dfrac{e^2}{\delta^2}, & |e| \le \delta \\ \dfrac{|e|}{\delta^2}, & |e| > \delta \end{cases} \tag{5}$$

where $\delta > 0$ represents a parameter. Another well-known robust loss function is a logarithmic function.

$$\tau_{HUB}(e) = \begin{cases} 0, & e = 0 \\ \log\left(1 + e^2\right), & e \ne 0 \end{cases} \tag{6}$$

Use $D(x_k, v_i) = \tau(x_k - v_i)$ as a discrepancy measure between the k-th data and the i-th prototype center and an additional weighting. The criterion function for BFWCOM is as follows:

$$J(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} \beta_{ik} u_{ik}^m D(x_k, v_i) + \sum_{i=1}^{c} \sum_{k=1}^{n} \frac{\alpha}{N_R} u_{ik}^m \sum_{r \in N_i} \left\| x_r - v_i \right\|^2 \tag{7}$$

N_R is the cardinality of u_{ki}, x_r is the neighbor of x_i, N_i is the neighbor set in the window around x_i. α is a parameter used to control the effect of the neighborhood terms. The relative importance of the adjustment terms is inversely proportional to the signal-to-noise ratio (SNR). A lower signal-to-noise ratio requires a higher parameter value.

The area of application to which the criterion function faces is as follows:

$$M_{fnc} = \left\{ U \in \Re_{cN} \mid \forall u_{ik} \in [0, 1];\ \sum_{i=1}^{c} \beta_{ik} u_{ik} = f_k;\ 0 < \sum_{k=1}^{N} u_{ik} < N \right\} \tag{8}$$

Table 2: Symbols used in the paper.

Symbol	Meaning
c	Number of clusters
X	Set of data items, $X = \{x_1, x_2, \cdots, x_n\}$
n	Number of data items, $n = \|X\|$
x_i	Data item, $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^T$
p	Number of attributes
x_r	the neighbor of, x_i
N_R	the cardinality of, u_{ki}
N_i	the neighbor set in the window around x_i
α	a parameter used to control the effect of the neighborhood terms
U	Membership matrix $c \times n$
u_{ik}	Membership of the i-th item to the k-th cluster
V	Matric of cluster centers, $V = (v_1, v_2, \cdots, v_c)^T$
v_i	Centre of i-th cluster, $v_i = (v_{i1}, v_{i2}, \ldots, v_{ip})^T$
β_{ik}	Typicality of the k-th data item with respect to i-th cluster
f_k	Global typicality of the k data item, cf. Eq.(42)
m	Weighting exponent for memberships
e_{idk}	Residual of d-th of k-th datum from the centre of i-th cluster
$\diamondsuit$	S-norm, fc. Eq.(42)
η_k	the bias field at the k-th pixel

2.2. The Laplace Algorithm Used to Determine the Membership Degree and the Central Formula

2.2.1. The Membership Degree Formula. The Laplace algorithm used to give the specific formula is as follows:

$$G(U,V) = \sum_{i=1}^{c}\sum_{k=1}^{n}\beta_{ik}u_{ik}^{m}D(x_k, v_i) + \sum_{i=1}^{c}\sum_{k=1}^{n}\frac{\alpha}{N_R}u_{ik}^{m}SD(x_r, v_i) - \lambda\left(\sum_{i=1}^{c}\beta_{ik}u_{ik} - f_k\right) \tag{9}$$

where $SD(x_r, v_i) = \sum_{r\in N_i}\|x_r - v_i\|$, x_r is the neighbor of x_i, and $D(x_k, v_i)$ is the distance from the neighbor point x_k to the center v_i.

Take the partial derivative of formula (9) with respect to u_{ik}:

$$\frac{\partial G(U,V)}{\partial u_{ik}} = m\beta_{ik}u_{ik}^{m-1}D(x_k, v_i) + m\frac{\alpha}{N_R}u_{ik}^{m-1}SD(x_r, v_i) - \lambda\beta_{ik} = 0 \tag{10}$$

$$\frac{\lambda}{m} = u_{ik}^{m-1}\left(D(x_k, v_i) + \frac{(\alpha/N_R)\,SD(x_r, v_i)}{\beta_{ik}}\right) \tag{11}$$

The membership degree u_{ik} is derived:

$$u_{ik} = \left(\frac{\lambda}{m}\right)^{1/(m-1)} \cdot \left(D(x_k, v_i) + \frac{(\alpha/N_R)\,SD(x_r, v_i)}{\beta_{ik}}\right)^{1/(1-m)} \tag{12}$$

$$D(x_k, v_i) = \sum_{d=1}^{p}(x_{kd} - v_{id})^2 \tag{13}$$

Available from formula (8),

$$\sum_{i=1}^{c}\beta_{ik}u_{ik} = f_k \tag{14}$$

Substitute formula (12) into formula (14):

$$f_k = \sum_{i=1}^{c}\beta_{ik}u_{ik} = \sum_{i=1}^{c}\beta_{ik}\left(\frac{\lambda}{m}\right)^{1/(m-1)} \cdot \left(D(x_k, v_i) + \frac{(\alpha/N_R)\,SD(x_r, v_i)}{\beta_{ik}}\right)^{1/(1-m)} \tag{15}$$

The following formula is derived from formula (15):

$$\left(\frac{\lambda}{m}\right)^{1/(m-1)} = f_k\left(\sum_{i=1}^{c}\beta_{ik}\left(D(x_k, v_i) + \frac{(\alpha/N_R)\,SD(x_r, v_i)}{\beta_{ik}}\right)^{1/(1-m)}\right)^{-1} \tag{16}$$

The obtained membership degree formula is as follows:

$$u_{ik} = \frac{f_k\left(D\left(x_k, v_i\right) + \left(\alpha/N_R\right) SD\left(x_r, v_i\right) \beta_{ik}^{-1}\right)^{1/(1-m)}}{\sum_{i=1}^{c} \beta_{ik} \left(D\left(x_k, v_i\right) + \left(\alpha/N_R\right) SD\left(x_r, v_i\right) \beta_{ik}^{-1}\right)^{1/(1-m)}} \tag{17}$$

2.2.2. The Clustering Center Formula. The distance formula $D(x_k, v_i)$ can also be written as

$$D\left(x_k, v_i\right) = \left(x_k - v_i\right)^T A \left(x_k - v_i\right) \tag{18}$$

Formula (7) can be written as follows:

$$\begin{aligned} J\left(U, V\right) = &\sum_{i=1}^{c} \sum_{k=1}^{n} \beta_{ik} u_{ik}^m \left(x_k - v_i\right)^T A \left(x_k - v_i\right) \\ &+ \sum_{i=1}^{c} \sum_{k=1}^{n} \frac{\alpha}{N_R} u_{ik}^m \sum_{r \in N_i} \left(x_r - v_i\right)^T A \left(x_r - v_i\right) \end{aligned} \tag{19}$$

Take the derivative of formula (19) with respect to v_i:

$$\begin{aligned} \frac{\partial J}{\partial v_i} = &-2A \sum_{k=1}^{n} \beta_{ik} u_{ik}^m \left(x_k - v_i\right) \\ &- 2A \sum_{k=1}^{n} \frac{\alpha}{N_R} u_{ik}^m \sum_{r \in N_i} \left(x_r - v_i\right) = 0 \end{aligned} \tag{20}$$

The clustering center formula as follows can be obtained by deriving formula (20):

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^m \left(\beta_{ik} x_k + \left(\alpha/N_R\right) \sum_{r \in N_i} x_r\right)}{\sum_{k=1}^{n} u_{ik}^m \left(\beta_{ik} + \alpha\right)} \tag{21}$$

2.3. Clustering Prototype Updating. The objective function $J(U, V)$ can be minimized in a manner similar to the standard FCM algorithm. The first-order derivatives of the objective function $J(U, V)$ with respect to u_{ki}, v_i, and η_k are set 0, resulting in three necessary but insufficient conditions for local extrema. In the following subsections, we will obtain these three conditions.

According to literature [33], the additive bias field formula $y_k = x_k + \eta_k$, $\forall k \in \{1, 2, \ldots, N\}$ is given, where x_k and y_k are the true and observed log-transformed intensities at the kth pixel, respectively, and η_k is the bias field at the kth pixel.

$$J_m\left(U, V\right) = \sum_{i=1}^{c} \sum_{k=1}^{n} \left(\beta_{ik} u_{ik}^m D_{ik} + \frac{\alpha}{N_R} u_{ik}^m \gamma_i\right) \tag{22}$$

where

$$D_{ik} = \left\| y_k - \eta_k - v_i \right\|^2 \tag{23}$$

$$\gamma_i = \sum_{r \in N_i} \left\| y_r - \eta_r - v_i \right\|^2 \tag{24}$$

Use Laplace's theorem:

$$\begin{aligned} G_m = &\sum_{i=1}^{c} \sum_{k=1}^{n} \left(\beta_{ik} u_{ik}^m D_{ik} + \frac{\alpha}{N_R} u_{ik}^m \gamma_i\right) \\ &- \lambda \left(\sum_{i=1}^{c} \beta_{ik} u_{ik} - f_k\right) \end{aligned} \tag{25}$$

Take the derivative of u_{ik}, $1 \le i \le n$; $1 \le k \le c$:

$$\frac{\partial G_m}{\partial u_{ik}} = m \beta_{ik} u_{ik}^{m-1} D_{ik} + m \frac{\alpha}{N_R} u_{ik}^{m-1} \gamma_i - \lambda \beta_{ik} = 0 \tag{26}$$

Sort out the formula:

$$u_{ik} = \left(\frac{\lambda}{m}\right)^{1/(m-1)} \left(D_{ik} + \frac{\alpha \gamma_i}{\left(N_R \beta_{ik}\right)}\right)^{1/(1-m)} \tag{27}$$

Sort out the formula by substituting $\sum_{i=1}^{c} \beta_{ik} u_{ik} = f_k$:

$$\sum_{i=1}^{c} \beta_{ik} \left(\frac{\lambda}{m}\right)^{1/(m-1)} \left(D_{ik} + \frac{\alpha \gamma_i}{\left(N_R \beta_{ik}\right)}\right)^{1/(1-m)} = f_k \tag{28}$$

$$\begin{aligned} &\left(\frac{\lambda}{m}\right)^{1/(m-1)} \\ &= f_k \left(\sum_{i=1}^{c} \beta_{ik} \left(D_{ik} + \frac{\alpha \gamma_i}{\left(N_R \beta_{ik}\right)}\right)^{1/(1-m)}\right)^{-1} \end{aligned} \tag{29}$$

The following formula is available from formula (27) and by substituting formula (29):

$$u_{ik} = \frac{f_k \left(D_{ik} + \alpha \gamma_i / \left(N_R \beta_{ik}\right)\right)^{1/(1-m)}}{\sum_{i=1}^{c} \beta_{ik} \left(D_{ik} + \alpha \gamma_i / \left(N_R \beta_{ik}\right)\right)^{1/(1-m)}} \tag{30}$$

Available from formulas (23) and (24), formula (19) can also be written as follows:

$$\begin{aligned} J_m = &\sum_{i=1}^{c} \sum_{k=1}^{n} \left(\beta_{ik} u_{ik}^m D\left(y_k - \eta_k, v_i\right)\right. \\ &\left. + \frac{\alpha}{N_R} u_{ik}^m \sum_{r \in N_i} D\left(y_r - \eta_r, v_i\right)\right) \end{aligned} \tag{31}$$

among

$$\begin{aligned} D\left(y_k - \eta_k, v_i\right) &= \left(y_k - \eta_k - v_i\right)^T A \left(y_k - \eta_k - v_i\right) \\ D\left(y_r - \eta_r, v_i\right) &= \left(y_r - \eta_r - v_i\right)^T A \left(y_r - \eta_r - v_i\right) \end{aligned} \tag{32}$$

Use the Laplace's method:

$$\begin{aligned} G_m = &\sum_{i=1}^{c} \sum_{k=1}^{n} \left[\beta_{ik} u_{ik}^m D\left(y_k - \eta_k, v_i\right)\right. \\ &\left. + \frac{\alpha}{N_R} u_{ik}^m \sum_{r \in N_i} D\left(y_r - \eta_r, v_i\right)\right] + \lambda \left(\sum_{i=1}^{c} \beta_{ik} u_{ik} - f_k\right) \end{aligned} \tag{33}$$

Take the derivative of

$$\frac{\partial G_m}{\partial v_i} = -2A\sum_{k=1}^{n}\beta_{ik}u_{ik}^{m}\left(y_k - \eta_k - v_i\right) - 2A\sum_{k=1}^{n}\frac{\alpha}{N_R}u_{ik}^{m}\sum_{r\in N_i}\left(y_r - \eta_r - v_i\right) = 0 \tag{34}$$

The clustering center formula as follows can be obtained by deriving formula (34):

$$v_i = \frac{\sum_{k=1}^{n}u_{ik}^{m}\left(\beta_{ik}\left(y_k - \eta_k\right) + \left(\alpha/N_R\right)\sum_{r\in N_i}\left(y_r - \eta_r\right)\right)}{\sum_{k=1}^{n}u_{ik}^{m}\left(\beta_{ik} + \alpha\right)} \tag{35}$$

2.4. Bias Field Estimation. In a similar manner, take the derivative of J_m with respect to η_k, and the result is set to 0.

$$\sum_{i=1}^{c}\frac{\partial}{\partial\eta_k}\sum_{k=1}^{n}\left(\beta_{ik}\left(u_{ik}\right)^{m}D_{ik} + \frac{\alpha}{N_R}u_{ik}^{m}\gamma_i\right) = 0 \tag{36}$$

Because only the k-th term in the second sum depends on, then we can have

$$\sum_{i=1}^{c}\frac{\partial}{\partial\eta_k}\beta_{ik}\left(u_{ik}\right)^{m}\left(y_k - \eta_k - v_i\right)^2 = 0 \tag{37}$$

Differentiate the distance function:

$$\sum_{i=1}^{c}\beta_{ik}\left(u_{ik}\right)^{m}\left(y_k - \eta_k - v_i\right) = 0 \tag{38}$$

Therefore, the zero gradient condition of the bias field estimator is expressed as

$$\eta_k = y_k - \frac{\sum_{i=1}^{c}\beta_{ik}\left(u_{ik}\right)^{m}v_i}{\sum_{i=1}^{c}\beta_{ik}\left(u_{ik}\right)^{m}} \tag{39}$$

Here we should first discuss the ordering of data items. For each data item in each cluster, the properties of each attribute are computed separately. To do this, the data items are sorted by their distance from the cluster center. The d-th attribute of the closest data item is marked with the ordinal number 1, while that of the farthest data item n (n stands for the number of data items). The value β_{ckd} represents the characteristic of the d-th attribute relative to the k-th data item of the c-th cluster and is calculated with the function:

$$\beta_{ckd} = \left[\left(\frac{p_c n - \chi_{ckd}}{2p_l n} + \frac{1}{2}\right)\wedge 1\right]\vee 0 \tag{40}$$

It is called piecewise linear ordered weighted averaging (PLOWA) or

$$\beta_{ckd} = \frac{1}{1 + \exp\left[\left(2.944/p_a n\right)\left(\chi_{ckd} - p_c n\right)\right]} \tag{41}$$

Among them, $\wedge$ and $\vee$, respectively, represent the minimum and maximum operations. These two functions, which can be called weighting functions, are not increased relative to the parameter $\chi_{ckd} \in \{1, 2, \ldots, n\}$. For $\chi_{ckd} = p_c n$, both functions are equal to 0.5. Parameters $p_l > 0$ and $p_a > 0$ affect the slope. In the case of a piecewise linear function, for $\chi_{ckd} \in [p_c n - p_l n, p_c n + p_l n]$, its value is linearly reduced from 1 to 0 (see Figure 2 for details). In the case of the S-function, for $\chi_{ckd} \in [p_c n - p_a n, p_c n + p_a n]$, its value dropped from 0.95 to 0.05. The functions defined by formula (40) and formula (41) are, respectively, called S-ordered weighted average (SOWA) and piecewise linear ordered weighted averaging (PLOWA) in the rest of the work. If residual ordering is not applied, this is equivalent to using a uniform weighting function Based on OWA-UOWA for all $\chi_{ckd}, \beta_{ckd} = 1$. Then we call this case an unordered cluster (or no weighting function).

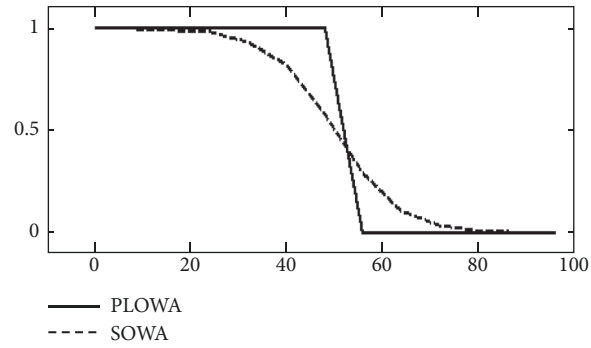

Figure 1: The PLOWA and SOWA weighting functions for $n = 100$, $p_a = 0.2$, $p_c = 0.5$, and $p_l = 0.2$.

It is called S-ordered weighted average (SOWA) [50]. In these two functions, χ_{ckd} represents the k-th data item index after reordering the distance relative to the c-th cluster of the d-th attribute. The two functions are shown in Figure 1, n = 100 and, $p_a = 0.2$, $p_c = 0.5$, and $p_l = 0.2$,

The importance of this data item affects the central location of the cluster. The global characteristic of the k-th data item is calculated using s-norm ($\diamondsuit$) of all cluster data item characteristics:

$$f_i = \beta_{1i}\diamondsuit\beta_{2i}\diamondsuit\cdots\diamondsuit\beta_{ci} \tag{42}$$

We use the largest operator as the s-norm operator.

3. Pseudocode, Complexity Calculation, and Convergence of BFWCOM Algorithm

3.1. Pseudocode and Complexity Calculation of BFWCOM Algorithm. By writing the pseudocode mentioned in Algorithm 1, it can be seen that it is a 4-fold loop with a complexity of $O(N^4)$.

3.2. Proof of the Convergence of BFWCOM Algorithm. Algorithm 1 starts from initializing the membership degree matrix $U^{(0)}$. The local minimum point or saddle point of the fuzzy C line (FCL) is its solution point. The BFWCOM algorithm can converge to a local minimum point.

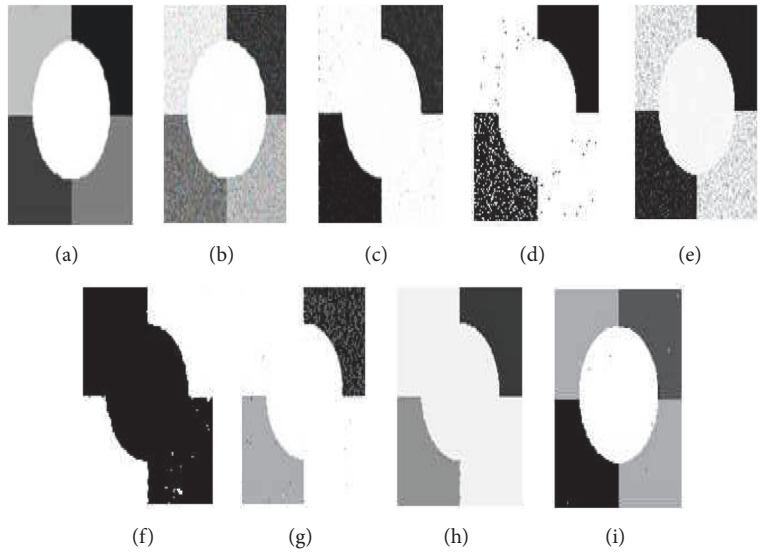

Figure 2: Comparison of segmentation results on the synthetic image from denoising base on salt & pepper. (a) Original image; (b) the image added salt & pepper (zero mean and 20% variance); (c) FCMS result; (d) FCMS1 result; (e) FCMS2 result; (f) FLICM result; (g) ARKFCM result; (h) FRFCM result; (i) BFWCOM result.

Definition 1. In formula (22), $\{J_m(U,V), m \in [1,\infty]\}$. Let BFWCOM algorithm obtain the best clustering result (U^*, V^*) by minimizing

$$\forall U \in M_{fnc}, \quad J_m(U^*, V^*) \le J_m(U, V^*) \tag{43}$$

$$\forall V \in R^{sc}, \quad J_m(U^*, V^*) \le J_m(U^*, V) \tag{44}$$

Ω is called the solution set of the problem FCL.

Definition 2. Let $J : M_{fnc} \longrightarrow \mathfrak{R}^{sc}; J(U) = V = (V_1, V_2, \cdots, V_c)$. Calculate V_i using formula (35).

Let $F : \mathfrak{R}^{sc} \longrightarrow P(M_{fnc}); F(V) = (V_1, V_2, \cdots, V_c) = \{U \in M_{fnc} \mid (U,V)$ *subject to* (17) *and* (35)$\}$.

BFWCOM operator (point-to-set mapping) $T_m : M_{fnc} \times \mathfrak{R}^{sc} \longrightarrow P(M_{fnc} \times \mathfrak{R}^{sc})$ is defined as

$$T_m = M_2 \circ M_1 \tag{45}$$

where

$$\begin{aligned} M_1 &: M_{fnc} \times \mathfrak{R}^{sc} \longrightarrow P\left(M_{fnc}\right); \\ M_1(U,V) &= \{(U,V) \mid U \in F(V)\} \\ M_2 &: P\left(M_{fnc}\right) \longrightarrow P\left(M_{fnc} \times \mathfrak{R}^{sc}\right); \\ M_2(U,V) &= \{(U, J(U))\} \end{aligned} \tag{46}$$

Then

$$\begin{aligned} T_m(U,V) &= M_2 \circ M_1(U,V) \\ &= \left\{\left(U', V'\right) \mid U' \in F(V),\ V' = J(U)\right\} \end{aligned} \tag{47}$$

Definition 3. When any $(U(k), V(k))_{k=1}^{l} \in M_{fnc} \times \mathfrak{R}^{sc}$, there is an initial value of $(U(0), V(0))$, l is the number of iterations, $(U(k), V(k)) = T(U(k-1), V(k-1))$, and then $(U(k), V(k))$ is called the iterative sequence of BFWCOM algorithm.

If the nested implicit relationship among the membership function and the clustering prototype and the parameter m is not considered, take the partial derivative of formula (22) with respect to m:

$$\begin{aligned} &\frac{\partial J_m(U,V)}{\partial m} \\ &= \sum_{i=1}^{c}\sum_{k=1}^{n}\left(\beta_{ik} u_{ik}^m \log\left(u_{ik}\right) D_{ik} + \frac{\alpha}{N_R} u_{ik}^m \log\left(u_{ik}\right) \gamma_i\right) \\ &= \sum_{i=1}^{c}\sum_{k=1}^{n}\left\{\left(\beta_{ik} u_{ik}^m D_{ik} + \frac{\alpha}{N_R} u_{ik}^m \gamma_i\right) \log\left(u_{ik}\right)\right\} < 0 \end{aligned} \tag{48}$$

It can be seen from formula (48) that the increase of $J_m(U,V)$ is monotonously decreasing and $J_m(U,V) \ge 0$. According to the monotone bounded sequence convergence theorem, $(U(k), V(k))$ must converge to a point in the solution set. Meanwhile, there are different optimal fuzzy C partitions for different m values. Therefore, the BFWCOM algorithm has a clustering validity problem for the weighted index m. In other words, it is necessary to determine the value of m, so that the corresponding clustering result is the most effective and reasonable. The parameter m has important influence on FCM algorithm.

According to the above statement, the core theorem of BFWCOM algorithm is as follows.

Procedure BFWCOM (X,c,m,l)
X:Data item matrix
c:The number of clusters
m:Membership degree weighted exponent
l:the number of iterations
n:=the number of data items in X;
1: Initialize all $\beta's$ with 1s;
2: Calculate data iteme characteristics using formula (42);
3: Set up the initial assignment of the random number matrix U, and normalize it using the constraint formula (14);
4: Calculate the clustering prototype V using formula (35);
5: **for** iter:=1 to N **do**
6: Update U and normalize it using formula (30);
7: **for** c: =1 to c **do**
8: **for** d: =1 to p **do**
9: **for** for k: =1 to n **do**
10: $e_{cdk} := |x_{kd} - v_{cd}|$;
11: **end for**
12: sort residual;
13: mark each residual with the number of sequences X the ordered sequence;
14: **for** k: =1 to n **do**
15: calculate the characteristics of β_{ckd} using formulas (40) (41), or uniform weighting;
16: **end for**
17: **end for**
18: **for** k: =1 to n **do**
19: Calculate the characteristics of β_{ck} with the maximum of the multiple values obtained informula (40),(41);
20: **end for**
21: Update the prototype using formula (35);
22: **end for**
23: **end for**
24: Return U, V;
25: End procedure;

Algorithm 1: Pseudocode of BFWCOM clustering algorithm.

Core theorem: let $X = \{x_1, x_2, \cdots, x_n\}$, $x_i \in \Re^p$; c is the number of clusters 1<c<n, for the formula

$$J_m(U,V) = \sum_{i=1}^{c}\sum_{k=1}^{n}\left(\beta_{ik}(u_{ik})^m D_{ik} + \frac{\alpha}{N_R}u_{ik}^m\gamma_i\right) \quad (49)$$

where $D_{ik} = \|y_k - \eta_k - v_i\|^2$; $\gamma_i = \sum_{r\in N_i}\|y_r - \eta_r - v_i\|^2$; $V = \{V_1, V_2, \cdots, V_c\} \subset \Re^p$; m $(1 < m < \infty)$ is weight factor; $U = [u_{ij}]_{c\times n} \in M_{fnc}$. Let $V(0) \in \Re^{sc}$ and $U(0) = F(V(0))$; $(U(0), V(0))$ is the initial value of the mapping T. The iterative sequence of BFWCOM terminates at a point in the solution set of $\min(J_m)$ (local convergence theorem).

4. Algorithm Experiment

The experiment consists of three steps.

First, IRIS data are classified using FCM, FCMS, FWCOM, ARKFCM, FRFCM, and BFWCOM. Then, the advantages and disadvantages of the proposed algorithm in this paper are compared with other algorithms according to the running results.

Second, these algorithms are all applied to the specific composite image processing. By image denoising, the advantages and disadvantages of the new algorithm are compared with other ones.

Third, the advantages and disadvantages of the proposed algorithm are verified by the real image processing.

Experimental environment: software environments: Windows 7 Ultimate and Software MATLAB; hardware environments: (1) processor: Intel(R) Pentium(R) CPU G2020@2.90GHz 2.90GHz; (2) installed memory: 8.00GB; (3) system type: 64-bit operating system.

In all experiments, the fuzzy parameter value is uniformly set: m=2. For FCMS, FCMS1, FCMS2, FLICM, ARKFCM, FRFCM, and BFWCOM, the iteration stops when the Frobenius norm of continuous V matrix difference is less than 10^{-4}. The loss function adopts $\varepsilon = 0.5$, $\alpha = 6.0$, $\beta = 1.0$, and $\delta = 1.0$. The weighted functions (40)-(41) adopt $p_c = 0.5$,

$p_l = 0.2$, and $p_a = 0.2$. For the calculated terminal prototype, we use the Frobenius norm of the difference between the real center matrix and the terminal prototype matrix to measure the clustering performance.

4.1. IRIS Data Used for Comparison of Classification Effectiveness. Iris data set, collected by Fisher in 1936, is a commonly used data set for classification experiment. Iris, also called Iris flower data set, is a data set for multivariate analysis. It contains 150 data sets, divided into 3 classes, 50 data per class, each data containing 4 attributes. Aiming at the commonly used IRIS data, six algorithms, namely, FCM, FCMS, FWCOM, ARKFCM, FRFCM, and BFWCOM, are used for comparison. The comparison process is as follows.

4.1.1. Classification Error Rate (CER). The CER is referred to as the probability that a sample that should belong to one of the class c is incorrectly assigned to other classes. In this paper, $X = \{x_i^{(j)} \mid 1 \le i \le n; 1 \le j \le c\}$; that is to say, the data are divided into c spaces, each space being a class. Regardless of how it is divided, there is always a misclassification phenomenon that an individual of a class is assigned to other classes. Therefore, the good and bad of cluster analysis algorithm can be equivalent to the division of the lowest average misclassification rate.

$$CER = \left(1 - \frac{Number\ of\ correctly\ classified\ pixels}{Total\ number}\right) \times 100\% \quad (50)$$

4.1.2. Deviation Degree (DD) Calculation

Definition 4. Sum of squares for error is also called residual sum of squares (RSS) or sum of squares within groups. After fitting the appropriate model based on n observations, the remaining that is underfitting ($e_i = x_i - \overline{x}$) is called the residual, of which $\overline{x}$ represents the average value of n observations $\overline{x} = \sum_{i=1}^{n} x_i$. The sum of the squares of all the n residuals is called residual sum of squares (RRS) or sum of squares for error (SSE), $SSE = \sum_{i=1}^{n}(x_i - \overline{x})^2$.

Definition 5. In the set, $X = \{x_i^{(j)} \mid 1 \le i \le n; 1 \le j \le c\}$, $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^T \in \Re^p$, j is the concrete class to which x_i belongs to. The mean of sum of the j-th sample is as follows: $\overline{x}_j = (1/n_j)\sum_{l=1}^{n_j} x_l^{(j)}$, of which n_i is the total number of data in class j. The mean of sum of squares of total deviation for all the data samples, also called the mean of the residual sum of squares, is as shown below:

$$S = \frac{1}{p}\sum_{i=1}^{p}\left\{\frac{1}{n}\sum_{j=1}^{c}\sum_{l=1}^{n_i}\left(x_{li}^{(j)} - \overline{x}_{ji}\right)^2\right\} \quad (51)$$

S, the mean of sum of squares of total deviation, can reflect the deviation degree between all measured data and the mean value of the data it classifies. The function above is called Deviation degree (DD).

It is obviously seen that the smaller the deviation degree, the better the clustering. In Table 2, CER is calculated with function (50), DD is calculated with function (51), and the clustering center is calculated with corresponding center formulas of algorithms.

Table 3 shows that BFWCOM, the new algorithm, which performs better than FCM, FCMS, FWCOM, ARKFCM, and FRFCM in classification, for the residual sum of squares is the smallest, and the misclassification rate is the lowest.

4.2. Comparison of Experimental Results Using Composite Images. In this part, synthetic images are used to verify the performance of the proposed algorithm. In addition to BFWCOM, the other six algorithms are also taken into consideration, namely, FCMS, FCMS1, FCMS2, FLICM, ARKFCM, and FRFCM. For all fuzzy algorithms, the size of the neighboring window is 3×3. In order to compare the segmentation performance, segmentation accuracy (SA) is used as a quantitative index [52].

$$SA = \frac{Number\ of\ correctly\ classified\ pixels}{Total\ number} \times 100\% \quad (52)$$

As shown in Figure 2, the composite image contains five regions, namely, five classes. The gray values of these five regions are, respectively, 0, 60, 120, 180, and 255. Figure 2 shows the different clustering results of several methods under the condition of images with Gaussian noise with the variance: $\sigma_g = 20$. It can be observed that BFWFCM produces a better segmentation than the other six fuzzy clustering algorithms. Visually, they can achieve acceptable segmentation results under Gaussian noise. Figure 3 shows the segmentation accuracies (SAs) of different methods for composite images under different noises, of which σ_g is the variance of Gaussian noise. As can be seen from the first line of Table 4, for synthetic images with different noise levels, the segmentation accuracy of our method is always higher than that of other methods.

As shown in Figure 3, "salt & pepper" noise is added to the synthetic images above to carry out the denoising experiment. This figure shows the clustering denoising results of images with 20% "salt & pepper" noise. It can be observed that, under this condition, BFWFCM produces a better segmentation than the other six fuzzy clustering algorithms. Visually, they can achieve acceptable segmentation results under the "salt & pepper" noise. Table 4 shows the segmentation accuracies (SAs) of different methods for synthetic images under different "salt & pepper" noises. From Table 4, it can be seen that, for synthetic images with different levels of "salt & pepper" noise, the segmentation accuracy of our method is always higher than that of other methods.

4.3. Comparison of Experimental Results Using Brain MR and Lena Images. In this section, we experimented with real images, particularly magnetic resonance (MR) images and Lena images, in order to prove the performance of the proposed BFWCOM. Meanwhile, some well-known image segmentation methods, such as FCMS, FCMS1, FCMS2,

Table 3: Classification effect of FCM algorithm, FCMS algorithm, and BFWCOM algorithm on IRIS data.

Algorithm	the number of misclassification	CER	Clustering center	DD
FCM	16	10.7%	$v_1 = (5.0040, 3.4141, 1.4828, 0.2535)$ $v_2 = (5.8887, 2.7610, 4.3636, 1.3972)$ $v_3 = (6.7748, 3.0523, 5.6465, 2.0534)$	0.08
FCMS	10	6.7%	$v_1 = (4.9940, 3.3649, 1.4833, 0.2537)$ $v_2 = (5.7787, 2.7851, 4.2636, 1.3882)$ $v_3 = (6.6648, 3.0573, 5.6485, 2.0564)$	0.06
FWCOM	8	5.3%	$v_1 = (4.9940, 3.3649, 1.4833, 0.2537)$ $v_2 = (5.7787, 2.7851, 4.2636, 1.3882)$ $v_3 = (6.6648, 3.0573, 5.6485, 2.0564)$	0.04
ARKFCM	7	4.7%	$v_1 = (4.9840, 3.3549, 1.4533, 0.2497)$ $v_2 = (5.7817, 2.8051, 4.2539, 1.3082)$ $v_3 = (6.3948, 3.0583, 5.3485, 2.0821)$	0.04
FRFCM	5	3.3%	$v_1 = (4.9941, 3.3650, 1.4863, 0.2577)$ $v_2 = (5.7767, 2.7853, 4.2615, 1.2782)$ $v_3 = (6.4648, 3.0893, 5.3485, 2.1564)$	0.037
BFWCOM	5	3.3%	$v_1 = (4.9969, 3.3649, 1.4623, 0.2158)$ $v_2 = (5.7723, 2.7852, 4.1582, 1.2782)$ $v_3 = (6.4490, 2.9253, 5.4026, 2.0110)$	0.035

Table 4: Comparison of segmentation accuracies (%) on synthetic images degraded by different noise.

Noise	FCMS	FCMS1	FCMS2	FLICM	ARKFCM	FRFCM	BFWCOM
Gaussian (σ_g=10)	99.91	99.93	99.96	99.95	99.97	99.98	99.98
Gaussian (σ_g=20)	96.15	96.40	96.44	97.68	97.85	98.89	99.35
Gaussian (σ_g=25)	71.50	94.74	94.90	98.84	98.97	99.32	99.68
Salt & pepper (5%)	96.64	96.64	96.64	97.34	98.44	99.28	99.87
Salt & pepper (10%)	93.17	93.17	93.17	94.23	95.84	97.84	99.77
Salt & pepper (20%)	86.07	86.17	87.05	93.68	94.56	95.78	98.12

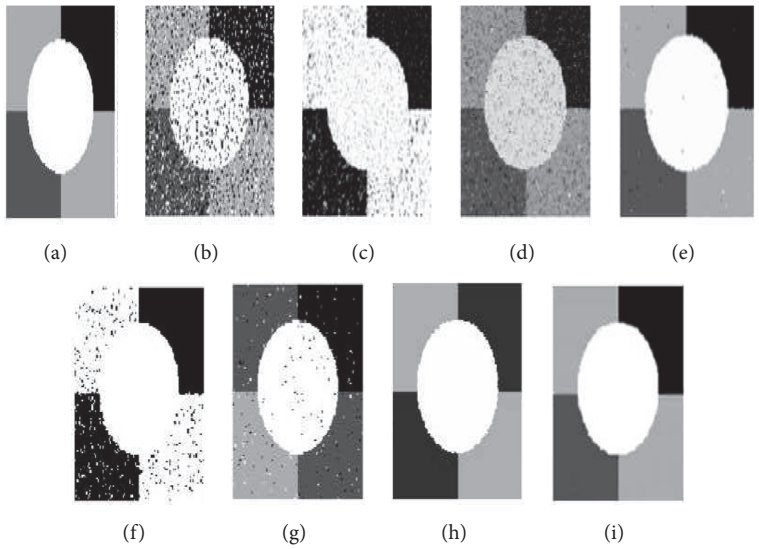

Figure 3: Comparison of segmentation results on the synthetic image from denoising base on salt & pepper. (a) Original image; (b) the image added salt & pepper (zero mean and 20% variance); (c) FCMS result; (d) FCMS1 result; (e) FCMS2 result; (f) FLICM result; (g) ARKFCM result; (h) FRFCM result; (i) BFWCOM result.

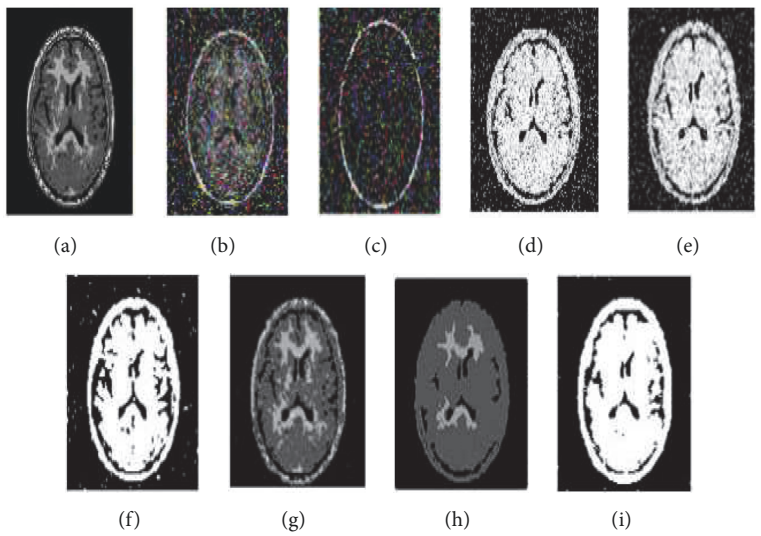

Figure 4: Comparison of segmentation results on the brain CT image from denoising base on salt & pepper. (a) Ground truth; (b) the image added salt & pepper (zero mean and 20% variance); (c) FCMS result; (d) FCMS1 result; (e) FCMS2 result; (f) FLICM result; (g) BFWCOM result.

FLICM, ARKFCM, and FRFCM, were used as comparison methods. In Figures 2 and 3, the denosing results of synthetic images have been compared. In addition to this, in Figures 4 and 5, we first use gauss and "salt & pepper" mixture noise to erode the real images [53] and then use these 7 algorithms to denoise real images with noise. Eventually, the comparison is made on the denoising results of real images. The selection of such real images is mainly carried out from two aspects. One is to select MR image for testing; the other is to select the head image of Lena for testing. Such selection also reflects the comparison between medical image and natural portrait.

4.3.1. Denoising Experimental Results of MR Image. Figures 4(c)–4(i) show the denoising segmentation results of different algorithms for images with 20% "salt & pepper" noise. Figures 5(c)–5(i) show the Gaussian denoising segmentation results of various algorithms with. The segmentation results show that the proposed BFWCOM algorithm can achieve excellent segmentation, while the other fuzzy methods are more or less affected by noise. The FRFCM algorithm is a relatively excellent method; however, it merely considers the influence of traditional fuzzy c-means clustering algorithm and spatial neighborhood, but fails to consider the typicality of data

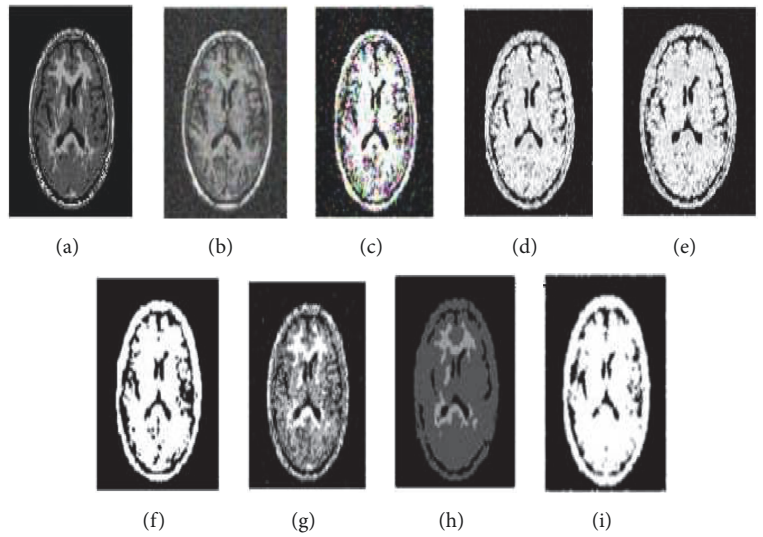

Figure 5: Comparison of segmentation results on the brain CT image from denoising base on Gaussian. (a) Ground truth; (b) the image added Gaussian (zero mean and 20% variance); (c) FCMS result; (d) FCMS1 result; (e) FCMS2 result; (f) FLICM result; (g)FRFCM result; (i) BFWCOM result.

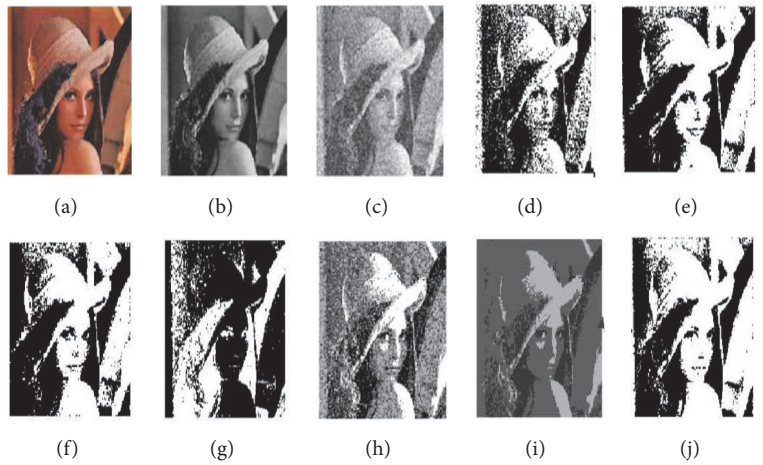

Figure 6: Comparison of segmentation results on the Lena image from denoising base on Gaussian. (a) Original Lena image; (b) gray image; (c) the image added Gaussian (zero mean and 20% variance); (d) FCMS result; (e) FCMS1 result; (f) FCMS2 result; (g) FLICM result; (h) ARKFCM result; (i) FRRCM result; (j) BFWCOM result.

attributes. Thus, it is less delicate than BFWCOM algorithm in image denoising.

4.3.2. Denoising Experimental Results of Lena's Head Images. The parameters in the BFWCOM include the fuzzy control parameter m, size of the neighborhood N, and similarity window S. Roughly speaking, the larger the size of N, the smoother the final results, the larger the similarity window, and the more structural information. However, it might use considerably large neighborhood windows and similarity windows for oversmoothing. In the experiment, we find that the 3×3 windows of N and S are applicable to brain MR image. For the real Lena images (Figure 6), a larger window is needed to measure the similarity. Under this circumstance, 5×5 or 7×7 windows of S are applicable. In this paper, 5×5 windows are adopted.

Figure 6 is a comparison of denoising results on the Lena image with added Gaussian noise with 20% variance. As can be seen from the figure above, the proposed BFWCOM algorithm has more obviously advantage than FCMS, FCMS1, FCMS2, FLICM, and ARKFCM. The FRRCM algorithm achieves a very good performance in the processing effect but is not better than BFWCOM algorithm in detail.

In order to further verify the denoising effect of the BFWCOM algorithm proposed in this paper, the performance of the proposed algorithm is compared and analyzed based on the objective data. Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR) [54, 55], and signal-to-noise ratio

Table 5: Noise-free Lena image denoising effect evaluation index (n=3; $\delta = 0.2$).

Denosing method	MSE	PSNR/db	SNR/db
Image with noise	640.5461	20.0653	13.4480
FCMS	138.6119	26.7128	21.0557
FCMS1	179.6752	26.1314	20.0553
FCMS2	182.4737	25.5188	19.8616
FLICM	107.1739	26.1670	20.5099
ARKFCM	91.4922	28.8986	22.2414
FRRCM	67.5066	29.5425	23.8853
BFWCOM	54.9023	30.7349	24.6808

(SNR) [56] are used for numerical calculation of the image after denoising.

The function for MSE is shown as below:

$$MSE = \frac{1}{M \times N}\left[\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(j,i) - \widehat{f}_{ji}\right)^2\right] \quad (53)$$

The functions for PSNR and SNR are shown as below:

$$PSNR = 10 \times \log_{10}\left(\frac{255^2}{MSE}\right) = 20 \cdot \log_{10}\frac{255}{MSE} \quad (54)$$

$$SNR = 10 \times \log_{10}\left\{\frac{\left[\sum_{i=1}^{M}\sum_{j=1}^{N} f(j,i)^2\right]}{(M \times N)} \cdot MSE\right\} \quad (55)$$

of which $M \times N$is the size of test image; $f(j,i)$ and $\widehat{f}_{ji}$ represent the noise-free image and denoised image, respectively. In order to compare the performance of the proposed algorithm more comprehensively, the MSE function (53), the PSNR function (54) and SNR function (55) are used for comparison of the image denoised by algorithms in Figure 6. The calculation results are shown in Table 5.

The MSE here is referred to as the mean of the squared sum of gray level difference values of pixel points corresponding to the original image and denoised image. The smaller the value is, the better the denoised image quality is. The unit of PSNR is dB. The higher the value of PSNR is, the better the denoising effect is. According to function (53) and function (54) and by limiting$(f(j,i) - \widehat{f}_{ji})^2 \geq 1$, $PSNR \in (0, 48.12)$ after calculation. When the PSNR value here is above 40, it has been close to the original image. Therefore, 30.7349, the PSNR value of the image denoised by the BFWCOM algorithm, is quite good. SNR means signal-to-noise ratio. The larger its value, the better. By the comparison in Table 4, it can be seen that the denoising effect of BFWCOM algorithm in Lena image is better than that of other algorithms.

5. Conclusion

This paper proposes a fast and robust BFWCOM algorithm for image segmentation to improve image segmentation quality and reduce the impact of image noise. The following things are done:

(1) Improve segmentation effect by using the local spatial information of the image and the typicality of the data item.

(2) BFWCOM uses the membership filtering algorithm to exploit the local spatial constraint. Since noise points will have low compatibility in all clusters, the membership functions obtained by this algorithm more approach to the concept of typicality, making their impact on clustering negligible. Therefore, this algorithm is naturally more immune to noise.

(3) This algorithm has additional advantages in calculation. It is a natural mechanism, assigning "fuzzy labels" to data in each iteration. Therefore, it can be used for more complex pattern recognition.

(4) The convergence of BFWCOM is proved by mathematical theory, providing a theoretical preparation for the algorithm.

(5) Experimental results show that the proposed BFWCOM can provide better segmentation results without adjusting parameters for different grayscale.

Data Availability

The FCMS algorithms of this study are available with "Pattern Recognition with Fuzzy Objective Function Algorithms" reference name at https://ieeexplore.ieee.org/document/996338. The FCMS1 and FCMS2 algorithms of this study are available with "Robust image segmentation using FCM with spatial constraints based on new kernel-induced dis12 VOLUME 4, 2016 tance measure" reference name at https://ieeexplore.ieee.org/document/1315771. The FLICM algorithms of this study are available with "Change Detection in Synthetic Aperture Radar Images Based on Deep Neural Network's" reference name at https://ieeexplore.ieee.org/document/7120131?arnumber=7120131. The IRIS data that support the findings of this study are available with "IRIS" reference name at http://archive.ics.uci.edu/ml/index.php.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is funded Natural Science Foundation of Hebei Province, China (Grant no. E2017203351). The work is also funded by Scientific Research Foundation of Yanshan University.

References

[1] S. Guizani, "A k-means clustering-based security framework for mobile data mining," *Wireless Communications and Mobile Computing*, vol. 16, no. 18, pp. 3449–3454, 2016.

[2] L. Wan, T. Zhang, Y. Xiang, and H. You, "A robust fuzzy c-means algorithm based on bayesian nonlocal spatial information for SAR image segmentation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 3, pp. 896–906, 2018.

[3] R. Gharieb, G. Gendy, and H. Selim, "A hard c-means clustering algorithm incorporating membership kl divergence and local data information for noisy image segmentation," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 4, pp. 758–769, 2018.

[4] Y. O. Ouma and M. Hahn, "Pothole detection on asphalt pavements from 2D-colour pothole images using fuzzy c-means clustering and morphological reconstruction," *Automation in Construction*, vol. 83, pp. 196–211, 2017.

[5] S. S. Khan and S. M. K. Quadri, "Structure identification and IO space partitioning in a nonlinear fuzzy system for prediction of patient survival after surgery," *International Journal of Intelligent Computing and Cybernetics*, vol. 10, no. 2, pp. 166–182, 2017.

[6] F. Masulli and S. Rovetta, "Soft transition from probabilistic to possibilistic fuzzy clustering," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 4, pp. 516–527, 2006.

[7] H. Cao, H.-W. Deng, and Y.-P. Wang, "Segmentation of M-FISH images for improved classification of chromosomes with an adaptive fuzzy C-means clustering algorithm," *IEEE Transactions on Fuzzy Systems*, vol. 20, no. 1, pp. 1–8, 2012.

[8] A. Javed, Y.-C. Kim, M. C. K. Khoo, S. L. D. Ward, and K. S. Nayak, "Dynamic 3-D MR visualization and detection of upper airway obstruction during sleep using region-growing segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 2, pp. 431–437, 2016.

[9] V. Grau, A. U. J. Mewes, M. Alcañiz, R. Kikinis, and S. K. Warfield, "Improved watershed transform for medical image segmentation using prior information," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 447–458, 2004.

[10] M. Gong, H. Li, X. Zhang, Q. Zhao, and B. Wang, "Nonparametric statistical active contour based on inclusion degree of fuzzy sets," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 5, pp. 1176–1192, 2016.

[11] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[12] D. Mahapatra, "Semi-supervised learning and graph cuts for consensus based medical image segmentation," *Pattern Recognition*, vol. 5, no. 63, pp. 700–709, 2017.

[13] Z. Li and J. Chen, "Superpixel segmentation using linear spectral clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '15)*, pp. 1356–1363, Boston, MA, USA, June 2015.

[14] 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.

[15] D. Pathak, P. Krahenbuhl, and T. Darrell, "Constrained convolutional neural networks for weakly supervised segmentation," in *Proceedings of the 15th IEEE International Conference on Computer Vision (ICCV '15)*, pp. 1796–1804, December 2015.

[16] B. Wang and Z. Tu, "Affinity learning via self-diffusion for image segmentation and clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 2312–2319, June 2012.

[17] S. Kim, C. D. Yoo, S. Nowozin et al., "Image segmentation using higher-order correlation clustering," *IEEE Transactions On Pattern Analysis & Machine Intelligence*, vol. 36, no. 9, pp. 1761–1774, 2014.

[18] J. Pont-Tuset, P. Arbeląäez, J. T. Barron et al., "Multiscale combinatorial grouping for image segmentation and object proposal generation," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, no. 1, pp. 128–140, 2017.

[19] M. A. Hasnat, O. Alata, and A. Tremeau, "Joint color-spatial-directional clustering and region merging (JCSD-RM) for unsupervised RGB-D image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 11, pp. 2255–2268, 2016.

[20] C. G. Bampis, P. Maragos, and A. C. Bovik, "Graph-driven diffusion and random walk schemes for image segmentation," *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 35–50, 2017.

[21] P. K. Saha, S. Basu, and E. A. Hoffman, "Multiscale opening of conjoined fuzzy objects: Theory and applications," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 5, pp. 1121–1133, 2016.

[22] J. C. Dunn, "Some recent investigations of a new fuzzy partitioning algorithm and its application to pattern classification problems," *Journal of Cybernetics*, vol. 4, no. 2, pp. 1–15, 1974.

[23] J. C. Dunn, "Well-separated clusters and optimal fuzzy partitions," *Journal of Cybernetics*, vol. 4, no. 1, pp. 95–104, 1974.

[24] J. C. Bezdek, "Cluster validity with fuzzy sets," *Journal of Cybernetics*, vol. 3, no. 3, pp. 58–73, 1974.

[25] J. C. Bezdek, "Numerical taxonomy with fuzzy sets," *Journal of Mathematical Biology*, vol. 1, no. 1, pp. 57–71, 1974.

[26] J. C. Bezdek and J. D. Harris, "Convex decompositions of fuzzy partitions," *Journal of Mathematical Analysis and Applications*, vol. 67, no. 2, pp. 490–512, 1979.

[27] J. C. Bezdek, "A convergence theorem for the fuzzy c-means clustering algorithms," *IEEE Transactions PAMI*, vol. 2, no. 1, pp. 1–8, 1980.

[28] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY, USA, 1981.

[29] J. C. Bezdek, "Hybrid modeling in pattern recognition and control," *Knowledge-Based Systems*, vol. 8, no. 6, pp. 359–371, 1995.

[30] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: the fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.

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

[32] Y. A. Tolias and S. M. Panas, "Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained

membership functions," *IEEE Transactions on Systems, Man, and Cybernetics—Part A:Systems and Humans.*, vol. 28, no. 3, pp. 359–369, 1998.

[33] 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.

[34] D. L. Pham, "Spatial models for fuzzy clustering," *Computer Vision and Image Understanding*, vol. 84, no. 2, pp. 285–297, 2001.

[35] A. W. C. Liew, S. Leung H, and H. Lau W, "Fuzzy image clustering incorporating spatial continuity," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 147, no. 2, pp. 185–192, 2000.

[36] L. SzilĺćGyi, Z. Benyĺő, S. SzilĺćGyii M et al., "MR brain image segmentation using an enhanced fuzzy c-means algorithm," in *Proceedings of the International Conference of the IEEE Engineering in Medicine & Biology Society*, pp. 724–726, 2003.

[37] K. Wu and M. Yang, "Alternative c-means clustering algorithms," *Pattern Recognition*, vol. 35, no. 10, pp. 2267–2278, 2002.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] M. Gong, Z. Zhou, and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2141–2151, 2012.

[44] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, and A. K. Nandi, "Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 5, pp. 3027–3041, 2018.

[45] A. M. Saranathan and M. Parente, "Uniformity-based superpixel segmentation of hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1419–1430, 2016.

[46] Z. Zhao, L. Cheng, and G. Cheng, "Neighbourhood weighted fuzzy c-means clustering algorithm for image segmentation," *IET Image Processing*, vol. 8, no. 3, pp. 150–161, 2014.

[47] F.-F. Guo, X.-X. Wang, and J. Shen, "Adaptive fuzzy c-means algorithm based on local noise detecting for image segmentation," *IET Image Processing*, vol. 10, no. 4, pp. 272–279, 2016.

[48] J. M. Leski, "Fuzzy *c*-ordered-means clustering," *Fuzzy Sets and Systems*, vol. 286, pp. 114–133, 2016.

[49] P. J. Huber, *Robust Statistics*, John Wiley & Sons, New York, NY, USA, 1981.

[50] R. R. Yager, "On ordered weighted averaging aggregation operators in multicriteria decision making," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, no. 1, pp. 183–190, 1988.

[51] K. Siminski, "Fuzzy weighted C-ordered means clustering algorithm," *Fuzzy Sets and Systems*, vol. 318, pp. 1–33, 2017.

[52] E. Menaka and K. S. Suresh, "Improving segmentation accuracy for detecting deforestation using texture feature derived from landsat 8 oli multispectral imagery," *European Journal of Remote Sensing*, vol. 48, no. 1, pp. 169–181, 2015.

[53] Mathworks, "Natick M. Image Processing Toolbox," http://www.mathworks.Com.

[54] S. Sanjith and R. Ganesan, "Fusion of DWT – DCT algorithm for satellite image compression," *International Journal of Applied Engineering Research*, vol. 10, no. 59, pp. 130–137, 2015.

[55] S. Sanjith, R. Ganesan, and R. S. Isaac, "Experimental analysis of compacted satellite image quality using different compression methods," *Advanced Science, Engineering and Medicine*, vol. 7, no. 3, pp. 227–233, 2015.

[56] F.-B. Zhou, C.-G. Li, and H.-Q. Zhu, "Research on threshold improved denoising algorithm based on lifting wavelet transform in UV-Vis spectrum," *Spectroscopy and Spectral Analysis*, vol. 38, no. 2, pp. 506–510, 2018.