Image Segmentation Based on Fuzzy Low-Rank Structural Clustering

Sensen Song, Zhenhong Jia, Jie Yang, *Senior Member, IEEE*, and Nikola Kasabov, *Life Fellow, IEEE*

***Abstract*—Fuzzy clustering is an essential algorithm in image segmentation, and most of them are based on fuzzy c-mean algorithms. However, it is sensitive to noise, center point selection, cluster number, and distance metric. To address this problem, we propose a new fuzzy clustering method based on low-rank representation (LRR) for image segmentation, which integrates low-rank structure with fuzzy theory. First, we improve the morphological reconstruction superpixel method based on edge detection by introducing anisotropy to enhance the image edge. Thus, on the one hand, the improved morphological reconstruction superpixel method can improve its noise-resistance performance; on the other hand, the complexity of the subsequent low-rank computation can be reduced by enhancing the superpixels constructed by the edges. Second, inspired by the fact that rank can represent correlation, we propose the concept of fuzzy low-rank structure, which is not dealing with data directly but with the relationship between data. Specifically, we perform rank minimization on the constructed membership matrix to obtain the optimal matrix. To obtain better clustering results, we added the Frobenius norm of the fuzzy matrix as a fuzzy regularization term in the LRR model to achieve global convergence and obtain a membership matrix with a strong element correlation. Finally, we obtain the final clustering results by clustering the processed membership matrix using a subspace clustering with a low-rank structure constraint. Experiments performed on artificial and real-world images show that the proposed method is more effective and efficient than the current state-of-the-art methods.**

***Index Terms*—Fuzzy clustering, fuzzy low-rank structure, image segmentation, low-rank representation (LRR), superpixel.**

I. Introduction

IMAGE segmentation is a well-known problem in computer vision research, which refers to dividing an image into several disjoint regions based on features, such as grayscale, color, spatial texture, and geometry [1]. These features show consistency or similarity within the same region while displaying significant differences between regions. According to this consensus, many kinds of algorithms for image segmentation have been proposed, clustering algorithms [2], [3], [4], active contour models [5], graph cuts [6], random walkers [7], region merging [8], neural networks [9], etc. Among these algorithms, clustering algorithms are employed to cluster the pixels in an image into several regions and can achieve unsupervised image segmentation without labels. Moreover, they have fewer parameters compared to other classes of algorithms. Furthermore, clustering algorithms are robust and effective in dimensionality reduction of high-dimensional data, so they have obvious advantages in multichannel image segmentation. Therefore, clustering algorithms have attracted much attention.

Among the clustering methods, the fuzzy c-mean (FCM) algorithm is one of the most widely used fuzzy clustering algorithms in image segmentation [10], [11]. Although it has dramatically improved compared to previous fuzzy clustering algorithms, it still has some problems, such as unsatisfactory clustering of images affected by noise, outlier points, and other artifacts. The most severe issue is that the results of FCM-based image segmentation are highly dependent on the center point, the number of clusters chosen, and the distance metric [4], [10], [12]. As a result, some improvements have been developed.

Several further studies have been conducted to solve the problems mentioned previously. The most straightforward strategy is to merge the local spatial information into the objective function to improve the segmentation effect. Inspired by this, FCM algorithms with spatial constraints, such as FCM_S [13], FCM_S1/S2 [14], and deviation-sparse fuzzy c-means with neighbor information constraint (DSFCM) [15], was proposed. However, they have high computational complexity and are not robust to Gaussian noise. Krinidis and Chatzis [11] proposed a fuzzy local information c-mean (FLICM) method to improve the denoising ability. However, it has some shortcomings in identifying class boundary pixels and preserving image details. To overcome this drawback, Gong et al. [16] introduced a kernel metric and a weighted fuzzy factor to enhance the robustness of FLICM (KWFLICM) for identifying class boundary pixels and preserving image details. Moreover, they also provided two other schemes: a variant of the FLICM algorithm (RFLICM) [17] and a fuzzy local similarity measure based on the pixel space attraction model (ADFLICM) [18]. The former uses local variable coefficients instead of fixed spatial distances to extract local texture

Manuscript received 24 March 2022; revised 2 August 2022; accepted 4 November 2022. Date of publication 9 November 2022; date of current version 30 June 2023. This work was supported by the National Science Foundation of China under Grant U1803261 and Grant 62261053, and in part by the International Science and Technology Cooperation Project of the Ministry of Education of the People's Republic of China under Grant 2016–2196. *(Corresponding author: Zhenhong Jia.)*

Sensen Song and Zhenhong Jia are with the College of Information Science and Engineering, Xinjiang University, Urumqi 830046, China, and also with the Key Laboratory of Signal Detection and Processing, Xinjiang Uygur Autonomous Region, Xinjiang University, Urumqi 830046, China (e-mail: song_sen_sen@163.com; jzhh@xju.edu.cn).

Jie Yang is with the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai 200400, China (e-mail: jieyang@sjtu.edu.cn).

Nikola Kasabov is with the Knowledge Engineering and Discovery Research Institute, Auckland University of Technology, Auckland 1020, New Zealand, and also with the George Moore Chair in Data Analytics, Ulster University, Londonderry BT48 7JL, U.K. (e-mail: nkasabov@aut.ac.nz).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TFUZZ.2022.3220925.

Digital Object Identifier 10.1109/TFUZZ.2022.3220925

information. At the same time, the latter adaptively determines the weight factors of adjacent pixel effects to preserve the edges and details of image regions. Following the concept of region-level information, Wang et al. [19] proposed an improved FCM, which combines adaptive local and region-level information with median membership of Kullback–Leibler (KL) information for noisy image segmentation (FALRCM). Also, they presented an improved FCM with adaptive space and intensity constraints and membership linking (FCM_SICM) [20] for noisy image segmentation.

However, the added constraint terms and complex structure easily increase the computational complexity while improving denoising performance. To reduce the computational complexity, Szilagyi et al. [21] enhanced FCM (EnFCM) by linearly weighting sum each pixel's local neighborhood mean grayness with the original image and then clustering them based on the grayness histogram of the summed image. Therefore, the runtime of EnFCM is very short. Cai et al. [22] introduced a local similarity measure that combines spatial and grayscale information to form a nonlinear weighted summed image. Namely, the clustering is performed based on summing the image grayscale histograms. So it is a fast generalized FCM (FGFCM) algorithm and its runtime, similar to EnFCM, is also very short. Lei et al. [10] solved the computational complexity problem by morphological reconstruction (MR) [35] and membership filtering, which is a fast and robust FCM algorithm (FRFCM) due to its removal of repeated distance calculations between pixels and clustering centers within the neighborhood window. Inspired by the superpixel technique [23], [24], [25], [26], [27], and EnFCM, Lei et al. [28] proposed a fast FCM color image segmentation algorithm (SFFCM) based on superpixels. It has two advantages. One is that the proposed watershed transform-based multiscale morphological gradient reconstruction (MMGR-WT) algorithm can provide better superpixel results, which helps to improve the final clustering results. The other is that the color histogram is incorporated into the objective function of FCM, which speeds up the implementation of the algorithm. Although SFFCM is very good for color image segmentation, it requires a manual setting of the number of clusters. From the affiliation perspective, a triangular inequality-based membership scale FCM (MSFCM) is proposed [29], which effectively improves the convergence speed and maintains the accuracy of data clustering.

To address the FCM problems of center points selection, the number of clusters, and the distance metric, researchers have explored new fuzzy frameworks. For example, in KWFLICM, they argued that one of the main reasons for the nonrobustness of FCM is the use of nonrobust Euclidean distance. Therefore, they proposed a non-Euclidean distance measure based on a kernel method to accomplish clustering. Lei Tao et al. [4] proposed a new automatic fuzzy clustering framework (AFCF) to further solve two problems. One employs density peak clustering to adaptively determine the number of clusters. The other uses covariance analysis and Markov random fields to measure the similarity between different superpixel regions instead of Euclidean distance. Fang et al. [30] presented a new active contour image segmentation model based on global and local fuzzy image fitting (FRAGL). They designed two fitted images: a global fuzzy fitted image and a local fuzzy fitted image. The blurring theory is introduced into the active contour model to make it more robust to image segmentation with noise, boundary-blurring, and intensity inhomogeneity. Moreover, they provided a hybrid energy-driven active contour segmentation method (HLFRA) based on the edges of blurred regions for image segmentation with high noise and intensity inhomogeneity [31]. These papers mentioned previously are new fuzzy theory-based image segmentation methods.

To solve the abovementioned problems, we propose an image segmentation algorithm based on fuzzy low-rank structure clustering (FLRSC). Its structure consists of two main parts; the first part is the preprocessing part. We perform superpixel preprocessing on the image before clustering to transform the pixel-level features into region-level features. Its role is to improve the speed of subsequent processing and reduce the computational complexity and incorporate anisotropy to enhance its noise immunity and edge preservation. The second part is a new fuzzy clustering framework that combines fuzzy theory with the low-rank structure to obtain different clusters by minimizing the rank structure for image segmentation. Through experiments, the effectiveness of the proposed algorithm in denoising and image segmentation has been demonstrated. Its main contributions are as follows.

1) We improve the edge detection-based superpixel algorithm, which preserves rich edge details and good noise immunity performance.
2) The image segmentation algorithm based on fuzzy low-rank clustering is a clustering method that does not need to set the center points, the number of clusters, and the distance metric, similar to the evolving clustering method [32], [33]. It has fewer parameters to optimize.
3) We propose a new framework for image segmentation that combines low-rank structure and fuzzy theory to obtain optimal solutions by an iterative weighting algorithm and demonstrate its convergence with better results in image segmentation applications.

The rest of this article is organized as follows. In Section II, we illustrate the motivations of this work. In Section III, we propose our methodology and analyze its superiority. The experimental results on synthetic and natural images are described in Section IV. Finally, Section V concludes this article.

II. Related Work

In this section, we focus on the related algorithms by which our algorithm is motivated, mainly including three aspects of algorithms: first, the superpixel algorithm, second, the fuzzy clustering algorithm, and third, the low-rank matrix representation.

A. Motivation for Using Superpixel

Superpixels refer to pixel blocks with certain visual significance composed of adjacent pixels with similar texture, color, brightness, and other features. They employ the similarity of features between pixels to group pixels and replaces a large number of pixels with a small number of superpixels to express image

features, which greatly reduces the complexity of subsequent image processing, so it is usually used as a preprocessing step in segmentation algorithms. This step has been widely used in computer vision applications, such as image segmentation, pose estimation, object tracking, and object recognition. Several common superpixel segmentation methods include: mean-shift [34], SLIC [23], DBSCAN [24], LSC [25], GMMSP [26], HS [27], and MMGR-WT [28].

These algorithms have their advantages and disadvantages. For example, the mean-shift algorithm has good noise immunity and edge fit. However, it has three parameters to which it is sensitive and remains susceptible to their effects. Although the SLIC algorithm has fast processing speed, uniform superpixel size, and good tightness, the results of subsequent region merging are not satisfactory. The DBSCAN algorithm can be adaptively clustered and can effectively handle noisy points, but it is sensitive to the uniformity of data density. LSC combines local features with a globally optimized objective function to produce more reasonable superpixels. GMMSP generates superpixels from the Gaussian distribution of pixels. MMGR-WT is based on multiscale morphological gradient reconstruction images, but different results are obtained from different edge images.

We found that the MR algorithm [35] is an excellent choice for superpixel generation. The adaptive morphological reconstruction (AMR) algorithm in [36] gives us great inspiration. It is defined as follows:

$$\psi(g,s,m) = \vee_{s\le i\le m}\left\{R_g^{\phi}(f)b_i\right\} \tag{1}$$

where $R_g^{\phi}(f)$ is the compositional morphological closing reconstruction, and ϕ denotes closing reconstruction. b_i is the nested structural element, where i is the scale parameter of the structural element, and $1 \le s \le i \le m, s, m \in N^+$, $i = s, s+1, \ldots, m$. g is a gradient image, $f = \varepsilon_{b_i}(g)$ denotes its morphological expansion reconstruction from f to g, and $f < g$.

AMR can provide different segmentation results by changing the parameter s. We can determine that when the value of s is small, the number of clusters is large, the segmentation area is small, and vice versa. However, although increasing the value of s can reduce the number of regions, the accuracy of segmenting regions will decrease. Although setting the parameter s achieves a good segmentation result in the noise-free image, it is sensitive to noise, and the effect is less satisfactory.

B. FCM Clustering

Given an image $X = x_1, x_2, ...x_i, ..., x_n, x_i \in R_n$, can be viewed as a dataset in an n-dimensional vector space. n is the number of feature vectors (the number of pixels in the image) and c is the number of clusters $(2 \le c < N)$. Then the objective function of FCM is defined as follows:

$$J_m(U,V) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \|x_j - v_i\|^2 \tag{2}$$

where $U = \{u_{ij}^m\}$ is the fuzzy membership of x_j with respect to the clustering center $V = \{v_i\}$, $0 \le u_{ij} \le 1$, and $\sum_{i=1}^{c} u_{ij} = 1$. m is the fuzzification index of matrix U. $\|\cdot\|$ denotes the Euclidean norm. The FCM uses an alternative optimization (AO) scheme [37] to obtain the optimal U and V

$$U^t = \arg\min_U J_m(U, V^t) \tag{3}$$

$$V^t = \arg\min_V J_m(U^t, V) \tag{4}$$

where t is the number of iterations, and FCM usually initializes U^0 or V^0 randomly and then updates U and V until convergence. Finally, a cluster image is formed centered on V to complete image segmentation.

C. Low-Rank Representation (LRR)

In recent years, the subspace clustering algorithm based on LRR [38], [39], [40], [41], [42] has become a prevalent method for data clustering. It takes advantage of the low-rank property of data self-representation to establish the affinity matrix, which can ensure that data belonging to the same subspace can be represented linearly with each other. Moreover, from the global structure of the dataset, LRR can establish the objective function by matrix rank minimization. However, matrix rank minimization is challenging to solve. Therefore, researchers usually adopt norm minimization to approximate rank minimization.

In [43], the formulation of the LRR problem is

$$\min_Z \operatorname{rank}(Z) \quad \text{s.t.} \quad X = AZ. \tag{5}$$

The optimal solution Z of the abovementioned problem is called the LRR of the data X with respect to the dictionary A. The abovementioned optimization problem is difficult to solve, and its solution may not be unique. Therefore, many norm functions are employed to replace low-rank functions, which are transformed into a convex optimization problem [44], [45], [46], [47].

III. Methodology

In this section, we begin with a description of how we construct a fuzzy low-rank structure that overcomes the inherent drawbacks of the FCM algorithm and approximation of the rank minimization problem. First, we improve the superpixel method, which improves the noise immunity performance and reduces the complexity of the rank minimization calculation. Second, we combine the low-rank structure with fuzzy theory. The clustering does not depend on the selection of center points and the number of clusters. Moreover, we provide a new method to solve the rank minimization problem.

A. Superpixel Based on Improved AMR

In computer vision, superpixels are increasingly used in the preprocessing stage of image processing. The main reason is that the application of superpixels can effectively reduce the redundancy of local information in an image, making image processing much less complex. Moreover, individual image pixel does not have any practical significance, and only when combined into different regions can they convey the information they carry to people.

In [28], the effectiveness of the MR superpixel algorithm for subsequent image processing has been demonstrated. It improves the speed of image processing and incorporates the

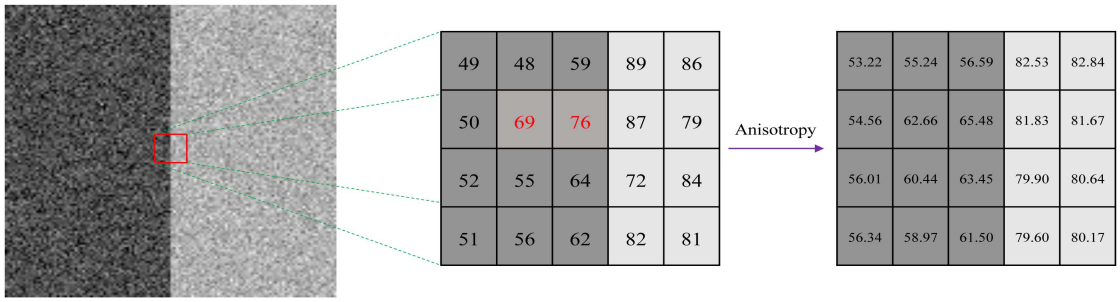

Fig. 1. Performance of edge noise immunity with and without anisotropic algorithms under noisy conditions.

Fig. 2. Comparison of AMR algorithm and improved AMR algorithm. (a) Original image "118035." (b) Superpixels generated by the AMR algorithm (the number of superpixel is 92). (c) Superpixels generated by the improved AMR algorithm (the number of superpixel is 71).

local information of image pixels since superpixels are small regions composed of pixel points whose positions are adjacent and have similar features, such as color, luminance, and texture. However, MR-based superpixels depend on their gradient images because the accuracy of their edges determines the subsequent segmentation results. Moreover, the detection of edges is sensitive to noise. Therefore, to enhance the noise immunity, we improve AMR by adding anisotropy [48] enhance its edge-holding capability, and then (1) becomes

$$\psi(g,s,m) = \vee_{s\le i\le m}\left\{R_g^{\phi}(F)b_i\right\} \tag{6}$$

where $F = \mathrm{div}\{d(\|\nabla f\|\cdot\nabla f)\}$, and ∇ is gradient operator. $d(\|\nabla f\|\cdot\nabla f)$ is the diffusion function that controls the degree of diffusion, and its expression is

$$d\left(\|\nabla f\|\right) = \frac{1}{1+\left(\|\nabla f\|/i\right)^2} \tag{7}$$

where i is the scale parameter.

To demonstrate the effectiveness of the improved AMR algorithm, Fig. 1 shows that our algorithm can effectively filter out the noise and keep the edges smooth under noisy conditions after adding anisotropy. Moreover, in Fig. 2, taking the image "118035" in the BSDS500 database as an example, the original image corresponds to Fig. 2(a), the result of AMR corresponds to Fig. 2(b), and the result of the improved AMR corresponds to Fig. 2(c). This effectively preserves edges, removes smaller blocks of superpixels, and reduces the number of superpixels. Since our purpose is to present an image segmentation algorithm with denoising capability and speed, the improved AMR algorithm is more suitable for our task requirements than other superpixel algorithms.

B. *Fuzzy Low-Rank Structural Clustering*

The FCM-based image segmentation algorithm results are limited by the selection of its center points and the number of clusters, as well as the distance metric. To solve the abovementioned problems, we are required to get rid of the influence of the original fuzzy structure and search for a new fuzzy framework. Moreover, the low-rank subspace clustering algorithm can establish the objective function by minimizing the rank of the matrix and does not require the abovementioned settings. Inspired by this, we present a new fuzzy low-rank structure that combines fuzzy theory and low-rank structure to solve the inherent problem of the FCM-based algorithm and the LRR problem. Concretely, we obtain the membership relationship of each data by extending the number of center points in the FCM algorithm to consider each data as a center point. Then rank minimization is employed to obtain the low-rank structure of the membership matrix. In other words, we are not dealing with the data itself but the relationship between the data. Moreover, to solve the LRR problem, we add a fuzzy regularization term to increase the association between data, making the approximation rank minimization more desirable.

The proposed fuzzy low-rank structure clustering algorithm includes two parts: one is the fuzzy data processing, which clusters the associated data in a subspace. That is, the data with similar membership are grouped. The other is image

segmentation, which processes the optimized membership data through the algorithm of [42] to discover the low-rank structure information hidden in the initial data and ultimately completes the clustering of the initial data, i.e., completes the image segmentation.

1) Fuzzy Low-Rank Processing: The objective function of the FCM algorithm is shown in (2), and the number of its clusters c is fixed and predetermined, but it is not clear to us what is the optimum number of clusters to set for processing images. Although the optimal clustering result can be obtained by an AO method, a predetermined number of clusters may not be appropriate. In our method, we assume that each pixel is the center of a cluster, and then we optimize the clusters to a smaller number. Using (5), a clustering of data from a global structure perspective is achieved by removing redundancy through rank minimization. Specifically, considering the extended membership matrix U as the spatial relationship matrix and the new clustering center V as the conditional constraint, the objective function is as follows:

$$\min_{U} \operatorname{rank}(U) \qquad \text{s.t.} \quad AU = V \tag{8}$$

where A is a "dictionary" that linearly spans the data space. Since the rank function is not convex and it is difficult to find the optimal value, the optimization problem of (8) can be relaxed to the following convex optimization problem:

$$\min_{U} \|U\|_*^2 \qquad \text{s.t.} \quad AU = V \tag{9}$$

where $\|U\|_*^2$ denotes the nuclear norm of U. However, this convex relaxation formulation may not be the best approximation to the original nonconvex problem (8), so we would like to use other relaxations that can better approximate the original problem of (8). By definition, the Schatten $-\,p$ norm is a better relaxation of the rank objective function than the nuclear norm when $p < 1$ [49]. Then the form of (8) can become

$$\min_{U} \|U\|_{S_p}^p + \beta \cdot \|AU - V\|_F^2. \tag{10}$$

Although the constraint of the $\|AU - V\|_F^2$ term can ensure the convergence of the abovementioned equation, the constraint is not very strong, so it may not be able to effectively approximate the rank minimization. To strengthen the constraint of (10), we add a fuzzy regularization term $\|U\|_F^2$, which is essentially a convex quadratic term, so that the model can be more robust and more closely approximates the rank minimization to reach the optimal solution. Also, to prevent the singularity of the gradient, the smoothing regularization term μI [47] is added, and our objective function is

$$J(U) = \min_{U} \|U; \mu I\|_{S_p}^p + \alpha \cdot \|U\|_F^2 + \beta \cdot \|AU - V\|_F^2 \tag{11}$$

where α, β, and $\mu > 0$, and $I \in \mathbb{R}^{n \times n}$ is the identity matrix.

To solve for the abovementioned solution, in the first step, taking the derivative of $J(U)$ with respect to U, and $\|U; \mu I\|_{S_p}^p = Tr((U^T U + \mu^2 I)^{\frac{p}{2}})$ we have

Algorithm 1: Algorithm to Solve Problem (15).

Data: membership matrix U, constraint matrix A, V, parameters $\alpha > 0$, $\beta > 0$, $\rho > 1$, and $0 < p < 1$.

```
Initialize: U_{t=0} is the affinity matrix of A and V, μ_0 > 0;
while not converged do
    Calculate Q_t = p(U_t^T U_t + μ_t^2 I)^{p/2 - 1};
    Calculate Q_t + 2βA^T A + αI;
    Calculate 2βA^T V − αU_t;
    Calculate μ_{t+1} = μ_t / ρ;
    Update U_{t+1}, μ_{t+1};
    Check the convergence conditions ||U_{t+1} − U_t||_∞;
    if ||U_{t+1} − U_t||_∞ ≤ ε then
        break
    end
    t ← t + 1;
end
```

$$\frac{\partial J(U)}{\partial U} = \frac{\partial \left(Tr \left(U^T U + \mu^2 I \right)^{\frac{p}{2}} \right)}{\partial U} + \frac{\partial \left(\alpha \cdot \|U\|_F^2 + \beta \cdot \|AU - V\|_F^2 \right)}{\partial U} \tag{12}$$

and set it to zero

$$p(U^T U + \mu^2 I)^{\frac{p}{2}-1} U + (2\beta A^T A + 2\alpha I)U - 2\beta A^T V = 0. \tag{13}$$

The detailed procedure for the derivation from (12) to (13) is in Appendix A. The abovementioned nonlinear equation is a fixed point equation that can be written as follows:

$$\left(p(U^T U + \mu^2 I)^{\frac{p}{2}-1} + 2\beta A^T A + \alpha I \right) U = 2\beta A^T V - \alpha U. \tag{14}$$

To obtain the optimal solution, we define an iterative algorithm as follows:

$$\left(Q_t + 2\beta A^T A + \alpha I \right) U_{t+1} = 2\beta A^T V - \alpha U_t \tag{15}$$

where $Q_t = p(U_t^T U_t + \mu_t^2 I)^{\frac{p}{2}-1}$, $\mu_{t+1} = \frac{\mu_t}{\rho}$, and $\rho > 1$, ρ is the parameter that controls the descent of μ. We obtain the optimal solution by updating the iterative (15) until convergence. According to the abovementioned discussion, the algorithm to solve problem (15) is summarized in Algorithm 1.

To verify that (11) is convergent, we propose the following Theorem 3.1.

Theorem 3.1: Given $0 < p < 1$, U_t is a sequence generated by Algorithm 1, and it satisfies the following properties.

1) $J(U_{t+1}) \le J(U_t)$, where the equality holds if and only if $U_{t+1} = U_t$.
2) $\lim_{t \to \infty} (U_{t+1} - U_t) = 0$.
3) If there exists a subspace matrix U_t^k,such that $U_t^k \to U$, U satisfies the first-order differential optimality condition $\frac{\partial J(U)}{\partial U} = 0$.

Our proposed algorithm satisfies Theorem 3.1. Its detailed proof procedure is provided in Appendix B. Therefore, our algorithm has good convergence, showing that (11) can better

(a) (b) (c)

Fig. 3. Example of the image segmentation process. (a) Original image "3063." (b) Superpixels generated by the AMR algorithm. (c) Segmentation result after fuzzy low-rank structure clustering.

approximate the rank minimization. Here, we obtain the membership matrix U, which results from global optimization.

2) Image Segmentation: Previously, the optimized membership matrix U was obtained by the fuzzy processing algorithm. Now, we face a new problem to exploit the membership relationship in the matrix U for image segmentation. In [43], we found that the authors introduced a new rank constraint, which consists of two steps to build the relationship matrix and spectral clustering. This constraint allows their model to learn a subspace indicator that captures the different clusters directly from the data without postprocessing. Because it can discover low-rank structural information hidden in the data without postprocessing, it is well appropriate to handle our membership matrix U with the following equation [43]:

$$\min_{G_i|_{i=1}^k} \sum_{i=1}^{k} \left(\sum_j \min\left(s\sigma_j(UG_i), 1\right) \right)^2$$
$$\text{s.t.} \quad G_i|_{i=1}^k \subseteq \{0,1\}^{n\times n}, \sum_{i=1}^{k} G_i = I \tag{16}$$

where $\sigma_j(UG_i)$ represents the jth singular value of the matrix UG_i. The k diagonal matrices $G_1, G_2, \ldots, G_k$ are defined to represent the subspace matrix U_i, and $s > 0$.

To solve problem (16), the authors employed an iteration-based reweighting method instead of the KKT conditions since the constraints are constant [43]. Therefore, problem (16) can be transformed into the following problem:

$$\min_{G_i|_{i=1}^k} Tr(A_i G_i)$$
$$\text{s.t.} \quad G_i|_{i=1}^k \subseteq \{0,1\}^{n\times n}, \sum_{i=1}^{k} G_i = I \tag{17}$$

where $A_i = U^T D_i U$, and $D_i = ((n-1) + \sum_{i=1}^{k} s\sigma_j)\hat{U}\Lambda\hat{U}^T$. Based on the singular value decomposition, we can obtain $UG_i = \hat{U}\sum_i V^T$. Here, n is the number of singular values σ_j $(j = 1, \ldots, n)$ in matrix UG_i, and the Λ_i is a diagonal matrix and its jth diagonal element is $s\sigma_j$, if $s\sigma_j < 1$ or 0 otherwise.

Since $G_i|_{i=1}^k$ are all n by n diagonal matrices, then (17) can be rewritten as

$$\min_{g_{ci}\subseteq\{0,1\}, \sum_{i=1}^{k} g_i = 1} \sum_{i=1}^{k}\sum_{c=1}^{n} a_{ci} g_{ci} \tag{18}$$

where g_{ci} is the cth diagonal element of matrix G_i and a_{ci} is the cth diagonal element of matrix A_i. We can obtain the optimal solution to (17) [43]

$$g_{ci} = \begin{cases} 1, & \arg\min_l a_{cl} \quad 1 \le l \le k \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

Finally, the position of $g_{ci} = 1$ with the row and column of the membership matrix U is restored to the superpixel image to complete the image segmentation. An example of the image segmentation process is shown in Fig. 3.

IV. Experiments

To demonstrate the effectiveness of our proposed FLRSC, we evaluate its results on synthetic noisy images and color images. Qualitative and quantitative comparisons with some state-of-the-art methods are also provided, and experimental discussions and analyses are performed. The following two main effects of FLRSC are verified in the experiments: 1) the ability to denoise images with synthetic noise; 2) feasibility and robustness are tested on images from the Berkeley segmentation dataset (BSDS500) [50].

There are 16 algorithms for comparison: FCM [51], FCM_S1 [14], FCM_S2 [14], EnFCM [21], FGFCM [22], FLICM [11], KWFLICM [16], FRFCM [10], DSFCM [15], FALRCM [19], FCM_SICM [20], SFFCM [28], RSSFCA [52], FRAGL [30], HLFRA [31], and AFCF [4]. These are all algorithms based on fuzzy theory for solving image segmentation problems.

A. Parameters Setting

The parameter settings in these comparison algorithms follow the corresponding original papers except for some general parameter settings, which we set uniformly for a fair comparison. For example, the FCM-based comparison algorithm requires three indispensable parameters: the weighting exponent, the

minimum error threshold, and the maximum number of iterations. Moreover, in our experiments, the values of these parameters are 2, 10^{-4}, and 100, respectively. In addition, the number of clusters is set to 3 and the neighborhood window size is set to 3×3. For FCM _ S1, FCM _ S2, and EnFCM, α is the control parameter associated with local denoising and is generally set to $\alpha = 3.8$. The spatial scale factor and the gray-level scale factor are $\lambda_s = 3$ and $\lambda_g = 5$ for FGFCM. Moreover, there is no other setting for FLICM, KWFLICM [24], or DSFCM except for the three essential parameters and the number of clusters. In FRFCM, the structural element size and the filter window are both set to 3×3. RSSFCA requires a regularization parameter $\gamma = 0.2$. For SFFCM and AFCF, their preprocessing algorithm MMGR-WT has two parameters, which are $r_1 = 2$ and $\eta = 10^{-4}$, respectively. For FALRCM, the control factor $k = 50$ and the KL divergence constraint $\gamma_{\mathrm{KL}} = 0.15$. In FCM _ SICM, the parameters related to the bilateral filter are set to $\sigma_d = 1$ and $\sigma_r = 7$. While FRAGL and HLFRA, the parameter settings are the same as the original article. The structure of FLRSC is different from FCM, so it does not have the three required parameter settings. Its superpixel parameters are the same except that the minimum structural scale s is set differently from AMR. The parameter s can control the size of the superpixel region, which we set to 1 to retain more details. Moreover, the smoothing term parameter μ, the fuzzy regularization term parameter α, and the constraint term parameter β must be greater than zero to ensure convergence. From the iterative (15), it can be noticed that the values of α should not be too large, while β should not be too small in general. Otherwise, it will cause the problem of nonconvergence. Based on experience, we set $\alpha = 0.2$ and $\beta = 1$. The parameters μ and ρ are set as in [47], $\mu_0 = 0.1\|U\|_2$ and $\rho = 1.1$. Furthermore, the value of p should not be too small because too small may not converge to the optimal solution in the experiment. Thus, we set $p = 0.8$. All experiments are performed on a PC workstation with a 3.6 GHz CPU and 8 GB RAM using MATLAB 2019a.

B. Results on Synthetic Images

To demonstrate the robust denoising ability of our proposed FLRSC, three synthetic images of size 256*256 are employed in the experiment. They are multichannel color images with different shapes, corrupted by Gaussian noise (G), salt and pepper noise (SP), and mixed noise, as shown in Figs. 4, 5, and 6, respectively. We evaluate the performance of FLRSC and the comparison method using the segmentation accuracy (SA), defined as the sum of correctly classified pixels divided by the total number of pixels, and the quantitative index score (S), representing the degree of equality between the pixel set A_k and ground truth C_k. They are calculated as follows:

$$S = \sum_{k=1}^{m} \frac{A_k \cap C_k}{A_k \cup C_k} \tag{20}$$

$$SA = \sum_{k=1}^{m} \frac{A_k \cap C_k}{\sum_{j=1}^{m} C_j} \tag{21}$$

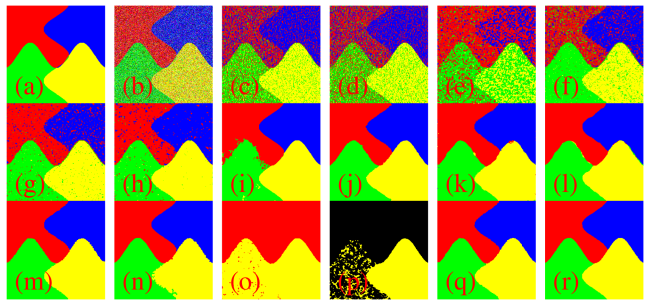

Fig. 4. Comparison of segmentation results on the first symmetric image. (a) Original image. (b) First noisy synthetic image (Gaussian noise, the noise level is 15%). (c) FCM_S1. (d) FCM_S2. (e) EnFCM. (f) FGFCM. (g) FLICM. (h) KWFLICM. (i) FRFCM. (j) DSFCM. (k) FALRCM. (l) FCM_SICM. (m) SFFCM. (n) RSSFCA. (o) FRAGL. (p) HLFRA. (q) AFCF. (r) FLRSC.

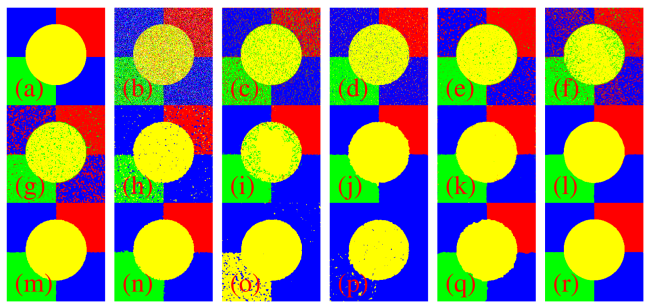

Fig. 5. Comparison of segmentation results on the second symmetric image. (a) Original image (b) Second noisy synthetic image (salt and pepper noise, the noise level is 30%). (c) FCM_S1. (d) FCM_S2. (e) EnFCM. (f) FGFCM. (g) FLICM. (h) KWFLICM. (i) FRFCM. (j) DSFCM. (k) FALRCM. (l) FCM_SICM. (m) SFFCM. (n) RSSFCA. (o) FRAGL. (p) HLFRA. (q) AFCF. (r) FLRSC.

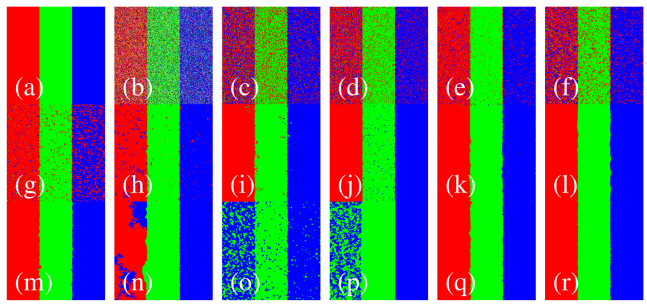

Fig. 6. Comparison of segmentation results on the third symmetric image. (a) Original image (b) Third noisy synthetic image (the mixed noise, the noise level is G 10% and SP 20%). (c) FCM_S1. (d) FCM_S2. (e) EnFCM. (f) FGFCM. (g) FLICM. (h) KWFLICM. (i) FRFCM. (j) DSFCM. (k) FALRCM. (l) FCM_SICM. (m) SFFCM. (n) RSSFCA. (o) FRAGL. (p) HLFRA. (q) AFCF. (r) FLRSC.

where m is the number of classes, A_k represents the set of pixels falling into the kth class in the result, and C_k denotes the set of pixels belonging to the class in the ground truth. Moreover, Tables I and II present the average SA and S values of the proposed algorithm and the comparison algorithms, which are the average results of 100 replicate experiments.

In Figs. 4, 5, and 6, the segmentation results of FCM_S1, FCM_S2, EnFCM, FGFCM, and FLICM exhibit poor performance and are sensitive to Gaussian noise, SP, and their

TABLE I
SEGMENTATION ACCURACIES (SA%) OF 16 ALGORITHMS ON THE FIRST SYNTHETIC IMAGE CORRUPTED BY NOISE, WHERE SP REPRESENTS SALT & PEPPER NOISE AND G REPRESENTS GAUSSIAN NOISE

Methods	G 5%	G 10%	G 20%	SP 10%	SP 20%	SP 30%	SP 10% + G 5%	SP 20% + G 10%	SP 30%+ G 20%
FCM_S1	90.69	85.31	76.38	88.77	80.68	69.37	87.18	69.30	58.91
FCM_S2	91.45	84.72	77.58	98.57	93.61	85.92	88.65	81.32	72.52
EnFCM	89.65	87.68	70.47	86.35	88.59	81.26	88.24	72.20	54.30
FGFCM	95.04	86.54	68.63	95.36	87.22	81.45	91.05	88.41	76.01
FLICM	97.34	90.46	83.87	92.85	85.59	83.84	86.35	85.69	65.12
KWFLICM	98.13	93.16	80.79	99.60	98.44	97.91	97.80	91.26	77.51
FRFCM	99.22	92.53	79.46	**99.84**	99.01	97.54	99.18	97.43	80.23
DSFCM	99.12	96.78	88.67	99.82	99.36	96.35	99.66	92.79	68.92
FALRCM	99.41	95.70	92.15	87.47	88.21	80.55	99.56	98.80	96.55
FCM_SICM	99.65	**98.52**	96.39	99.65	98.48	98.12	**99.84**	99.35	97.73
SFFCM	98.23	97.00	81.65	99.79	98.61	97.83	99.55	96.17	94.00
RSSFCA	97.96	96.21	89.44	99.56	96.42	95.28	99.25	96.73	77.77
FRAGL	55.60	53.14	50.60	52.88	51.63	48.76	78.09	73.59	62.18
HLFRA	67.77	60.93	55.38	78.84	65.26	59.68	90.99	78.85	44.69
AFCF	99.45	96.34	88.98	99.06	97.75	89.56	99.58	97.15	80.39
FLRSC	**99.70**	98.06	**97.19**	99.81	**99.80**	**99.32**	99.77	**99.62**	**98.16**

The best values are highlighted.

TABLE II
SCORES (S%) OF 16 ALGORITHMS ON THE SECOND SYNTHETIC IMAGE CORRUPTED BY NOISE, WHERE SP REPRESENTS SALT & PEPPER NOISE AND G REPRESENTS GAUSSIAN NOISE

Methods	G 5%	G 10%	G 20%	SP 10%	SP 20%	SP 30%	SP 10% + G 5%	SP 20% + G 10%	SP 30%+ G 20%
FCM_S1	79.40	56.43	45.24	81.66	69.79	49.73	69.23	47.85	39.28
FCM_S2	73.98	51.56	44.62	92.30	82.97	73.06	64.24	62.86	52.25
EnFCM	71.28	68.38	57.19	68.16	48.18	42.08	67.78	59.79	37.64
FGFCM	89.62	72.01	57.74	84.48	69.74	52.23	69.06	60.72	39.24
FLICM	88.82	66.17	59.69	82.54	55.72	46.50	73.56	52.06	47.73
KWFLICM	93.43	86.54	67.93	99.26	99.04	98.44	89.22	71.69	46.96
FRFCM	99.58	96.55	91.03	99.25	99.10	91.57	97.52	85.68	80.42
DSFCM	99.64	95.14	92.15	99.39	99.04	98.44	99.11	92.19	77.78
FALRCM	99.65	99.35	96.17	98.10	99.12	98.02	99.64	98.38	94.62
FCM_SICM	99.80	99.08	98.63	**99.67**	99.15	99.01	99.59	99.07	**98.61**
SFFCM	99.08	87.33	81.67	98.93	97.90	94.63	98.96	98.03	84.50
RSSFCA	88.71	69.34	57.71	99.62	98.25	92.52	99.21	76.28	70.22
FRAGL	47.75	44.33	42.98	47.77	47.80	47.68	47.33	46.36	34.54
HLFRA	47.87	52.16	45.39	47.85	47.36	39.82	47.82	47.42	29.27
AFCF	99.10	98.76	80.50	97.32	90.56	81.13	97.60	96.11	79.40
FLRSC	**99.92**	**99.36**	**98.96**	99.31	**99.30**	**99.14**	**99.79**	**99.07**	96.37

The best values are highlighted.

mixtures. As shown in Table I, FCM_S1 uses a mean filter to process local spatial information. Although it will affect the low-level noise, the denoising ability will worsen as the noise intensity increases. FCM_S2 is more effective for SP noise than Gaussian noise due to its use of median filters. FLICM introduces parameter-free local information to improve noise immunity. Then, KWFLICM adds a kernel metric to overcome some shortcomings of FLICM in identifying class boundary pixels and preserving image details. Therefore, its SA and S values are better than those of the FLICM, as seen in Table II. Compared to previous algorithms, EnFCM only uses histograms to reduce computational complexity without considering the denoising problem, resulting in poor noisy image segmentation performance. Furthermore, the FGFCM algorithm outperforms EnFCM because it introduces a new factor as a local (spatial and grayscale) similarity metric for noise immunity and detail preservation.

Furthermore, Figs. 4, 5, 6, Tables I, and II show that FRFCM, DSFCM, FALRCM, FCM SICM, SFFCM, RSSFCA, and AFCF can achieve better segmentation performance in the presence of noise interference. FRFCM is based on MR and member filtering, which solves the computational complexity problem and improves noise immunity performance. In addition, DSFCM imposes a sparse constraint on the deviation between measured and theoretical values, making it easier to identify noise or outliers. As shown in Tables I and II, the values of SA and S of FALRCM do not linearly decrease with increasing noise intensity. However, their denoising effect performs better in a specific noise interval, which we analyzed as being related to the KL information of its local median affiliation. Some of the FCM_SICM metrics are best at one noise level. We can deduce that this is due to its algorithm design for noise image segmentation, which employs a fast bilateral filtering method to obtain local spatial and intensity information. SFFCM can effectively remove Gaussian and SP noise by using adaptive and irregular local spatial neighborhood information. RSSFCA can remove noise by incorporating the Gaussian metric under regularization into the objective function. Moreover, AFCF shows that the superpixel technique and prior entropy can significantly improve SA in noisy images. While FRAGL and HLFRA are active contour-based image segmentation algorithms driven by global and local blur information and blurred region energy, respectively, their effects on noisy synthetic images are unsatisfactory, and the edge information is easily lost. Compared to the abovementioned algorithms, the proposed FLRSC algorithm can effectively filter the noise at the edges and preserve the edges well, demonstrating the effectiveness of FLRSC for noisy image segmentation.

C. Results on Color Images

To demonstrate the effectiveness and superiority of the proposed FLRSC algorithm for image segmentation, we tested and verified it on the BSDS500 database. Moreover, the parameters

TABLE III
PERFORMANCE COMPARISON OF DIFFERENT ALGORITHMS ON THE BSDS500 DATASET

Methods	PRI ↑	VoI ↓	GCE ↓	BDE ↓
FCM	0.69	2.93	0.38	14.47
FCM_S1	0.70	2.87	0.37	14.30
FCM_S2	0.69	2.86	0.36	14.42
EnFCM	0.71	2.90	0.39	14.65
FGFCM	0.69	2.85	0.36	14.24
FLICM	0.70	2.82	0.36	14.12
KWFLICM	0.72	2.80	0.35	14.27
FRFCM	0.73	2.60	0.30	13.98
DSFCM	0.70	2.82	0.36	14.63
FALRCM	0.71	2.62	0.32	14.40
FCM_SICM	0.67	2.87	0.34	14.78
SFFCM	0.72	2.31	0.26	14.35
RSSFCA	0.73	2.28	0.28	14.46
FRAGL	0.66	2.57	0.29	14.23
HLFRA	0.65	2.55	0.27	14.21
AFCF	0.74	2.22	0.22	13.93
FLRSC	**0.76**	**2.14**	**0.21**	**13.88**

The best values are highlighted.

of the comparison algorithms follow the original paper, and the number of clusters is set to 3 in the FCM-based algorithm for a fair comparison.

In this article, we adopt four popular evaluation metric functions, probabilistic rand index (PRI), variation of information (VoI), global consistency error (GCE), and boundary displacement error (BDE), to evaluate the segmentation results of the proposed FLRSC and the comparison algorithms. They evaluate the performance of segmentation results from different aspects, making the evaluation more accurate. Among them, PRI is the ratio of the number of pixels that overlap the result of the image segmentation algorithm and ground truth to the whole number of pixels. VoI defines the distance between the segmentation result of an image segmentation algorithm and the ground truth as the average conditional entropy. Moreover, it can measure the degree of randomness in the segmentation results that the ground truth cannot cover. GCE measures the degree to which the result of the image segmentation algorithm is consistent with the ground truth. BDE calculates the average displacement error between the result of the image segmentation algorithm and the boundary pixels in the ground truth. The larger the PRI value and the smaller the VoI, GCE, and BDE values in the quantitative results, the closer the segmentation results are to the ground truth, indicating the superior performance of the proposed algorithm.

In Figs. 7 and 8, some examples of image segmentation, containing simple backgrounds, complex backgrounds, scenes with similar objects and backgrounds, etc., are shown to demonstrate the effectiveness of the proposed FLRSC algorithm for image segmentation. Moreover, the superiority of FLRSC is intuitively reflected by the values of the four evaluation metrics in Table III. It can be seen that FCM, FCM_S1, and FCM_S2, have similar PRI, VI, GCE, and BDE values in Table III. Moreover, the main difference between them is that FCM_S1 and FCM_S2 add mean and median filters, respectively, which can reduce some

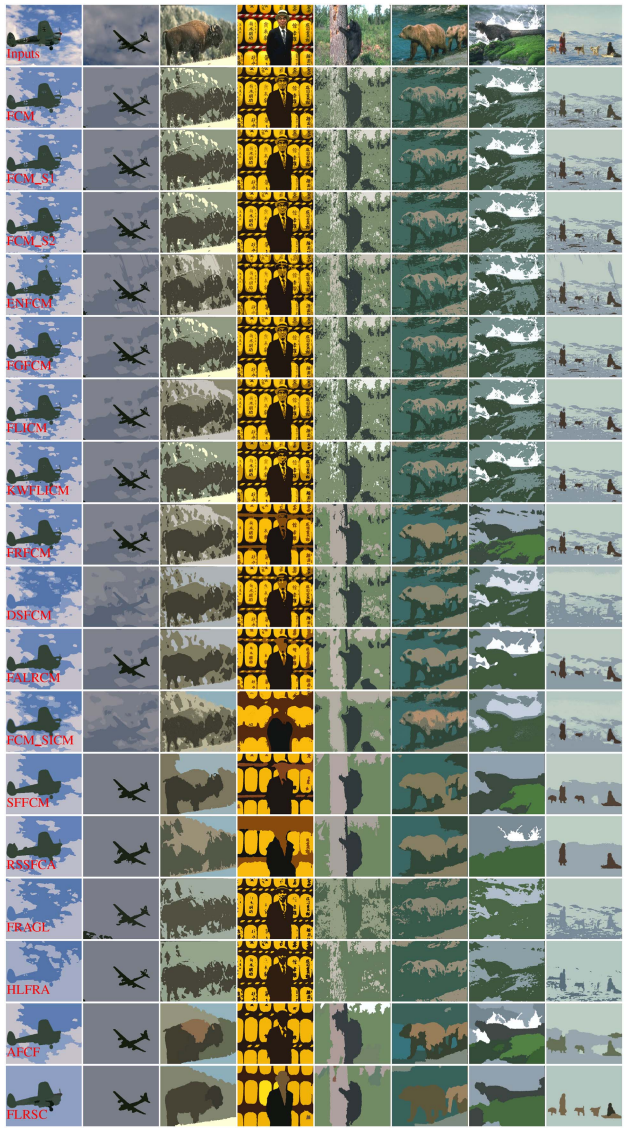

Fig. 7. Visual comparison of segmentation results on relatively simple scene images of BSDS500 using 17 algorithms.

noise. However, there is no significant improvement in image segmentation, as shown in Figs. 7 and 8. Although EnFCM can speed up the image processing, the cost is that its GCE and BDE values perform worse than FCM. Then FGFCM introduces a local similarity measure based on EnFCM, which improves the value of each evaluation metric. FLICM and KWFLICM introduce local information, making the image clustering effect better than before. Furthermore, as shown in Table III, not only are their four evaluation metrics improved, but their denoising effect is superior to previous work. The PRI and BDE values of FRFCM are inferior to those of AFCF and FLRSC, which also shows that preprocessing the image is efficient for fuzzy clustering. DSFCM, RSSFCA, FCM_SICM, FALRCM, and HLFRA add regularization or other information constraints to FCM. However, the four metric values of FCM_SICM perform poorly, while its denoising performance is just second to the FLRSC, indicating its suitability for noisy image processing. In contrast, DSFCM and RSSFCA obtain better segmentation results by

Fig. 8. Visual comparison of segmentation results on relatively complex scene images of BSDS500 using 17 algorithms.

using sparse regularization and self-sparse Gaussian regularization, respectively. FRAGL and HLFRA are active contour models with fuzzy information that obtain binary segmentation images, as shown in Figs. 7 and 8. Therefore, their PRI values are significantly lower than those of other algorithms. The metrics of SFFCM and AFCF perform relatively well, and their denoising effects are also superior. It demonstrates that the superpixel preprocessing of the image can reduce the subsequent computations and noise interference, which results in a promising segmentation result. The proposed FLRSC outperforms the comparison algorithms in all four evaluation metrics, demonstrating that FLRSC can effectively obtain better image segmentation results for color images.

D. Complexity Analysis

The computational complexity of an algorithm is also a valuable metric for assessing its performance. Table IV shows the computational complexity of the proposed FLRSC and the competing methods. N is the number of image pixels, c denotes the number of clusters, and t indicates the number of iterations. ω represents the filter window's size, and q is the image's grayscale value. N' is the number of superpixels significantly less than N, T' is the number of iterations less than t, and $O(M(c))$ is the computational complexity of Newton's method.

TABLE IV
COMPUTATIONAL COMPLEXITY OF DIFFERENT ALGORITHMS

Methods	Computational complexity
FCM	$O(N \times c \times t)$
FCM_S1	$O(N \times \omega^2 + N \times c \times t)$
FCM_S2	$O(N \times \omega^2 + N \times c \times t)$
EnFCM	$O(N \times \omega^2 + q \times c \times t)$
FGFCM	$O(N \times \omega^2 + q \times c \times t)$
FLICM	$O(N \times \omega^2 \times c \times t)$
KWFLICM	$O(N \times (\omega+1)^2 + N \times \omega^2 \times c \times t)$
FRFCM	$O(N \times \omega^2 + q \times c \times t)$
DSFCM	$O(N \times \omega^2 \times c \times t)$
FALRCM	$O(N \times (\omega^2+3) + N \times \omega^2 \times c \times t)$
FCM_SICM	$O(N \times q \times \log(N \times q) + 2N \times \omega^2 + N \times c \times t)$
SFFCM	$O(N \times T' + N' \times c \times t)$
RSSFCA	$O(N \times (M(C)+c) \times t + N \times c \times t)$
FRAGL	$O(N^2 \times \omega^2 \times t + N^2 \times t)$
HLFRA	$O(2N^2 \times \omega \times t + N^2 \times t)$
AFCF	$O(N \times T' + N' \times c \times t \times 2)$
FLRSC	$O(N \times T' + N'^2 \times t + N'^3 \times t)$

Compared to other algorithms, the FILICM, KWFLICM, and DSFCM have relatively high computational complexity, which means that they have to compute the information for each neighborhood. Similarly, the FALRCM algorithm increases the computational complexity due to the addition of KL information and region-level information. To reduce the computational complexity, FCM_S1 and FCM_S2 use filters instead of neighborhood spatial information, and EnFCM and FGFCM employ the grayscale histogram technique to replace the computation for each pixel, thus increasing the computational speed. Moreover, FRFCM is very fast because it removes the repeated distance calculation between pixels and clustering centers in the neighborhood window. Although the FCM_SICM expression appears to be complicated, its computation speed is not correspondingly slow because it has no local information calculation. However, the computational workload of RSSFCA goes up because the Newton algorithm is used in the iterative process. FRAGL and HLFRA are active contour segmentation methods incorporating local and other information. Therefore, there is no doubt that they are time-consuming methods. In contrast, the SFFCM and AFCF algorithms are faster because they employ superpixel preprocessing, which reduces the quantity of data. As seen from Table IV, the computational speed of the proposed FLRSC depends on the number of superpixels N: the fewer the superpixels, the more efficient the computation—however, the fewer the superpixels, the lower the SA.

V. Conclusion

In this article, we explore a new fuzzy clustering algorithm that does not require center point selection, the number of clusters, or the distance metric. Moreover, we propose a new framework for image segmentation based on fuzzy low-rank structure clustering, which combines low-rank structure and fuzzy theory to achieve image segmentation. The proposed FLRSC has two main contributions. One is that the objective function can be established by minimizing the rank of the membership matrix, thus replacing the classical form of the FCM function. Furthermore, it is concerned with membership data relative to the data rather than the data itself. The other contribution is building the objective function with a fuzzy regularization term to guarantee robustness and convergence, and then, a smoothing regularization term is added to prevent the gradient from becoming singular. We also demonstrate its convergence and solve its optimal solution with a weighted iterative algorithm. In addition, the improved superpixel generation algorithm has good noise immunity for superpixel image preprocessing. The experiments on synthetic and color images demonstrate the excellence and robustness of the proposed FLRSC.

The running speed and SA of our algorithm depend on the superpixel processing of the image, which is a limitation for fuzzy low-rank structure processing. In future, we will consider deep learning algorithms to extract image superpixel features and explore new ways of combining fuzzy theory and LRR.

Appendix A
Derivation Process of (12)–(13)

The right term of (12) can be written as

$$\frac{\partial\left(Tr\left(\left(U^TU+\mu^2I\right)^{\frac{p}{2}}\right)\right.}{\partial U}+\frac{\partial\left(\alpha\cdot\|U\|_F^2\right)}{\partial U}+\frac{\partial\left(\beta\cdot\|AU-V\|_F^2\right)}{\partial U}=0. \tag{22}$$

According to the derivative rule of the matrix, its first term becomes

$$\frac{\partial(Tr\left(\left(U^TU+\mu^2I\right)^{\frac{p}{2}}\right)}{\partial U}=\frac{p}{2}\Big(\left(U^TU+\mu^2I\right)^{\frac{p}{2}-1}\cdot\frac{\partial\left(Tr\left(U^TU\right)\right.}{\partial U}. \tag{23}$$

Then, in the next step, we need to solve for $\frac{\partial(Tr(U^TU)}{\partial U}$. By the differential rule of matrices and the properties of traces [53], we have

$$\begin{aligned}d\left(Tr\left(U^TU\right)\right)&=Tr\left(d\left(U^TU\right)\right)\\&=Tr\left(d(U^T)\,U\right)+Tr\left(U^TdU\right)\\&=Tr\left(\left(dU\right)^TU\right)+Tr\left(U^TdU\right)\\&=Tr\left(U(dU)^T\right)+Tr\left(U^TdU\right)\\&=Tr\left(U^TdU\right)+Tr\left(U^TdU\right)\\&=2Tr\left(U^TdU\right).\end{aligned} \tag{24}$$

Since $df(U)=Tr(\frac{\partial(f(U)}{\partial U_T}dU)$, we derive

$$\frac{\partial(Tr(U^TU)}{\partial U}=2\,U. \tag{25}$$

Therefore

$$\frac{\partial\Big(Tr\left(\left(U^TU+\mu^2I\right)^{\frac{p}{2}}\right)}{\partial U}=p(\left(U^TU+\mu^2I\right)^{\frac{p}{2}-1}U. \tag{26}$$

Both the second and third terms are derivations with respect to the F-norm, and

$$\|U\|_F^2=Tr\left(U^TU\right) \tag{27}$$

$$\|AU-V\|_F^2=Tr\left((AU-V)^T(AU-V)\right) \tag{28}$$

we have

$$\frac{\partial(\alpha\cdot\|U\|_F^2)}{\partial U}=\frac{\partial(\alpha\cdot Tr(U^TU)}{\partial U}=2\alpha U \tag{29}$$

and

$$\begin{aligned}&\frac{\partial\left(\beta\cdot\|AU-V\|_F^2\right)}{\partial U}\\&\quad=\frac{\partial\big(\beta\cdot Tr\left(\left(AU-V\right)^T\left(AU-V\right)\right)}{\partial U}\\&\quad=2\beta\left(A^TAU-A^TV\right)\end{aligned} \tag{30}$$

so the (12) is converted to (13)

$$p\left(U^TU+\mu^2I\right)^{\frac{p}{2}-1}U+\left(2\beta A^TA+2\alpha I\right)U-2\beta A^TV=0. \tag{31}$$

Appendix B
Proof of Theorem 3.1

Proof: (1) First, we define an auxiliary function

$$\begin{aligned}f(x,y)=&\left(\mu^2+x^2\right)^{\frac{p}{2}}-\left(\mu^2+y^2\right)^{\frac{p}{2}}\\&-py\left(\mu^2+x^2\right)^{\frac{p}{2}-1}(x-y).\end{aligned} \tag{32}$$

Then, we have

$$\begin{aligned}f(x,y)=&\Big[\left(px^2-2pxy+py^2\right)+\left(2\mu^2+(2-p)x^2\right.\\&\left.+py^2-2\left(\mu^2+x^2\right)^{1-\frac{p}{2}}2\left(\mu^2+y^2\right)^{\frac{p}{2}}\right)\Big]\\&\cdot\left(\mu^2+x^2\right)^{\frac{p}{2}-1}\\=&\,p(x-y)^2\cdot\left(\mu^2+x^2\right)^{\frac{p}{2}-1}\\&+\Big[(2-p)\left(\mu^2+x^2\right)+p\left(\mu^2+y^2\right)\\&-2\left(\mu^2+x^2\right)^{1-\frac{p}{2}}2\left(\mu^2+y^2\right)^{\frac{p}{2}})\Big]\\&\cdot\left(\mu^2+x^2\right)^{\frac{p}{2}-1}.\end{aligned} \tag{33}$$

Since for any nonnegative constants a, b, and $t \in (0,1)$, we have $a^{(1-t)}b^t \leq (1-t)a + tb$, where the equality holds if and only if $a = b$. Therefore, the second term in the abovementioned equation is nonnegative and the first term, $f(x,y) = 0$ when $x = y$, is a nonnegative complete square term. Moreover, $f(x,y) \geq p(x-y)^2 \cdot (\mu^2 + x^2)^{\frac{p}{2}-1}$, which is used to demonstrate the second term.

Let

$$\begin{aligned}
&J(U_t) - J(U_{t+1}) \\
&= Tr\left(U_t^T U_t + \mu^2 I\right)^{\frac{p}{2}} - Tr\left(U_{t+1}^T U_{t+1} + \mu^2 I\right)^{\frac{p}{2}} \\
&\quad + \beta \cdot \left(\|AU_t - V\|_F^2 - \|AU_{t+1} - V\|_F^2\right) \\
&\quad + \alpha \cdot \left(\|U_t\|_F^2 - \|U_{t+1}\|_F^2\right) \\
&= Tr\left(U_t^T U_t + \mu^2 I\right)^{\frac{p}{2}} - Tr\left(U_{t+1}^T U_{t+1} + \mu^2 I\right)^{\frac{p}{2}} \\
&\quad + \beta \cdot Tr\big((AU_t - V)^T (AU_t - V) \\
&\quad - (AU_{t+1} - V)^T (AU_{t+1} - V) \big) \\
&\quad + \alpha \cdot Tr\left(U_t^T U_t - U_{t+1}^T U_{t+1}\right) \\
&= Tr\left(U_t^T U_t + \mu^2 I\right)^{\frac{p}{2}} - Tr\left(U_{t+1}^T U_{t+1} + \mu^2 I\right)^{\frac{p}{2}} \\
&\quad + \beta \cdot Tr\big((AU_t)^T (AU_t) - (AU_t)^T V - V^T (AU_t) \\
&\quad - (AU_{t+1})^T (AU_{t+1}) + (AU_{t+1})^T V + V^T (AU_{t+1}) \big) \\
&\quad + \alpha \cdot Tr\left(U_t^T U_t - U_{t+1}^T U_{t+1}\right) \\
&= Tr\left(U_t^T U_t + \mu^2 I\right)^{\frac{p}{2}} - Tr\left(U_{t+1}^T U_{t+1} + \mu^2 I\right)^{\frac{p}{2}} \\
&\quad + \beta \cdot \big[\|AU_t - AU_{t+1}\|_F^2 + Tr\big((AU_t)^T (AU_{t+1}) \\
&\quad + (AU_{t+1})^T (AU_t) - 2(AU_{t+1})^T (AU_{t+1}) \\
&\quad - (AU_t)^T V - V^T (AU_t) - (AU_{t+1})^T (AU_{t+1}) \\
&\quad + (AU_{t+1})^T V + V^T (AU_{t+1}) \big)\big] \\
&\quad + \alpha \cdot Tr\left(U_t^T U_t - U_{t+1}^T U_{t+1}\right). \qquad (34)
\end{aligned}$$

Considering the properties of the matrix trace, the abovementioned expression can be written as

$$\begin{aligned}
&J(U_t) - J(U_{t+1}) \\
&= Tr\left(U_t^T U_t + \mu^2 I\right)^{\frac{p}{2}} - Tr\left(U_{t+1}^T U_{t+1} + \mu^2 I\right)^{\frac{p}{2}} \\
&\quad - Tr\big(\left(2\beta A^T V - 2\beta A^T A U_{t+1} - \alpha U_t - \alpha U_{t+1}\right) \\
&\quad \cdot (U_t - U_{t+1}) \big) + \beta \cdot \|AU_t - AU_{t+1}\|_F^2. \qquad (35)
\end{aligned}$$

By the (15) and (32), we obtain

$$\begin{aligned}
&J(U_t) - J(U_{t+1}) \\
&= Tr\left(U_t^T U_t + \mu^2 I\right)^{\frac{p}{2}} - Tr\left(U_{t+1}^T U_{t+1} + \mu^2 I\right)^{\frac{p}{2}} \\
&\quad - Tr\left(Q_t (U_t - U_{t+1})\right) + \beta \cdot \|AU_t - AU_{t+1}\|_F^2 \\
&= Tr\big(\left(U_t^T U_t + \mu^2 I\right)^{\frac{p}{2}} - \left(U_{t+1}^T U_{t+1} + \mu^2 I\right)^{\frac{p}{2}} \\
&\quad - Q_t U_{t+1} (U_t - U_{t+1}) \big) + \beta \cdot \|AU_t - AU_{t+1}\|_F^2 \\
&= Tr\left(f(U_t, U_{t+1})\right) + \beta \cdot \|AU_t - AU_{t+1}\|_F^2. \qquad (36)
\end{aligned}$$

Now, it is known that $f(U_t, U_{t+1}) = 0$ and $\|AU_t - AU_{t+1}\|_F^2$ only when $U_t = U_{t+1}$, then (30) implies that $J(U_{t+1}) \leq J(U_t)$.

(2) By $Q_t = p(U_t^T U_t + \mu^2 I)^{\frac{p}{2}-1}$ and $J(U_{t+1}) \leq J(U_t)$, we have

$$\begin{aligned}
Q_t^{\frac{p}{p-2}} &= \left(\frac{p}{2\beta}\right)^{\frac{p}{p-2}} (U_t^T U_t + \mu^2 I)^{\frac{p}{2}} \\
&\leq \left(\frac{p}{2\beta}\right)^{\frac{p}{p-2}} J(U_t) \leq \left(\frac{p}{2\beta}\right)^{\frac{p}{p-2}} J(U_0). \qquad (37)
\end{aligned}$$

We assume that $(\frac{p}{2\beta})^{\frac{p}{p-2}} J(U_0) = W_0$, then $Q_t \geq W_0^{\frac{p-2}{p}} = W_1$. Therefore,

$$\begin{aligned}
f(U_t, U_{t+1}) &\geq p(U_t - U_{t+1})^2 \cdot \left(U_t^2 + \mu^2 I\right)^{\frac{p}{2}-1} \\
&\geq W_1 p (U_t - U_{t+1})^2 \qquad (38)
\end{aligned}$$

and

$$\begin{aligned}
J(U_t) - J(U_{t+1}) &= Tr\left(f(U_t, U_{t+1})\right) \\
&\quad + \beta \cdot \|AU_t - AU_{t+1}\|_F^2 \\
&\geq Tr\left(W_1 p (U_t - U_{t+1})^2\right) \\
&\quad + \beta \cdot \|AU_t - AU_{t+1}\|_F^2 \\
&\geq Tr\left(W_1 p (U_t - U_{t+1})^2\right). \qquad (39)
\end{aligned}$$

By summing these inequalities, we obtain

$$\begin{aligned}
\sum_{i=0}^{n} Tr\left(W_1 p (U_t - U_{t+1})^2\right) &\leq J(U_0) - J(U_{t+1}) \\
&\leq J(U_0). \qquad (40)
\end{aligned}$$

Thus, we draw the conclusion that

$$\lim_{t \to \infty} (U_t - U_{t+1}) = 0. \qquad (41)$$

(3) By the abovementioned properties (1) and (2), we have

$$0 \leq J(U_{t+1}) \leq J(U_t). \qquad (42)$$

Then it can be seen that $J(U)$ is monotonically decreasing, and the lower boundary is zero and the upper boundary is $J(U_0)$. If there exists $U_t^k \to U$, then by the iterative algorithm, U_{t+1}^k will also converge to a matrix $U^\star$, and $U = U^\star$.

Moreover

$$\begin{aligned}
J(U) - J(U^\star) &\geq Tr(W_1 p (U - U^\star)^2) \\
&\quad + \beta \cdot \|AU - AU^\star\|_F^2. \qquad (43)
\end{aligned}$$

If we assume its limit is J, then $J(U) = J(U^\star) = J$. From (36), it follows:

$$\|AU - AU^\star\|_F^2 \leq J(U) - J(U^\star) = 0. \qquad (44)$$

Thus, we can obtain

$$A(U - U^\star) = 0. \qquad (45)$$

Since

$$Tr\left(U^TU+\mu^2I\right)^{\frac{p}{2}}+\alpha\cdot\|U\|_F^2+\|AU-V\|_F^2 = Tr\left(\left(U^\star\right)^TU^\star+\mu^2I\right)^{\frac{p}{2}}+\alpha\cdot\|U^\star\|_F^2+\|AU^\star-V\|_F^2 \tag{46}$$

we also further derive

$$Tr\left(U^TU+\mu^2I\right)^{\frac{p}{2}}-Tr\left(\left(U^\star\right)^TU^\star+\mu^2I\right)^{\frac{p}{2}} = \alpha\cdot\|U^\star\|_F^2-\alpha\cdot\|U\|_F^2. \tag{47}$$

Considering the iterative relationship

$$\left(Q_t+2\beta A^TA+\alpha I\right)U_{t+1}=2\beta A^TV-\alpha U_t. \tag{48}$$

When $t\to\infty$, the abovementioned equation becomes

$$\left(\left(Q_t+2\beta A^TA+\alpha I\right)U^\star\right)=\left(2\beta A^TV-\alpha U\right). \tag{49}$$

By multiplying $(U-U^\star)^T$ with both sides of (49), it becomes

$$\left(U-U^\star\right)^T\left(Q_t+2\beta A^TA+\alpha I\right)U^\star = \left(U-U^\star\right)^T\left(2\beta A^TV-\alpha U\right). \tag{50}$$

Thus, we have

$$Tr\left[\left(U-U^\star\right)^T\left(Q_t+2\beta A^TA+\alpha I\right)U^\star\right] = Tr\left[\left(U-U^\star\right)^T\left(2\beta A^TV-\alpha U\right)\right]. \tag{51}$$

Obviously,

$$Tr\big[(U-U^\star)^TQ_t+\left(U-U^\star\right)^T\left(2\beta A^TA-2\beta A^TV\right) +\left(U-U^\star\right)^T\left(\alpha U^\star+U\right)\big]=0. \tag{52}$$

By the equality $A(U-U^\star)=0$ and (36), we obtain

$$Tr(f(U,U^\star))=0 \tag{53}$$

which concludes $U=U^\star$. Therefore, when $k\to\infty$ and $U_t^k=U$, U satisfies the first-order differential optimality condition $\frac{\partial J(U)}{\partial U}=0$.

REFERENCES

[1] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Comput. Vis., Graph., Image Process.*, vol. 29, no. 1, pp. 100–132, 1985.

[2] G. B. Coleman and H. C. Andrews, "Image segmentation by clustering," *Proc. IEEE*, vol. 67, no. 5, pp. 773–785, May 1979.

[3] Z. Yang, F.-L. Chung, and W. Shitong, "Robust fuzzy clustering-based image segmentation," *Appl. soft Comput.*, vol. 9, no. 1, pp. 80–84, 2009.

[4] T. Lei, P. Liu, X. Jia, X. Zhang, H. Meng, and A. K. Nandi, "Automatic fuzzy clustering framework for image segmentation," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 9, pp. 2078–2092, Sep. 2020.

[5] K. Zhang, L. Zhang, K.-M. Lam, and D. Zhang, "A level set approach to image segmentation with intensity inhomogeneity," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 546–557, Feb. 2016.

[6] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[7] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006.

[8] R. Nock and F. Nielsen, "Statistical region merging," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 11, pp. 1452–1458, Nov. 2004.

[9] G. Dong and M. Xie, "Color clustering and learning for image segmentation based on neural networks," *IEEE Trans. Neural Netw.*, vol. 16, no. 4, pp. 925–936, Jul. 2005.

[10] 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 Trans. Fuzzy Syst.*, vol. 26, no. 5, pp. 3027–3041, Oct. 2018.

[11] S. Krinidis and V. Chatzis, "A robust fuzzy local information c-means clustering algorithm," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1328–1337, May 2010.

[12] J. Arora, K. Khatter, and M. Tushir, "Fuzzy c-means clustering strategies: A review of distance measures," *Softw. Eng.*, vol. 731, pp. 153–162, 2019.

[13] 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 Trans. Med. Imag.*, vol. 21, no. 3, pp. 193–199, Mar. 2002.

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

[15] Y. Zhang, X. Bai, R. Fan, and Z. Wang, "Deviation-sparse fuzzy c-means with neighbor information constraint," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 1, pp. 185–199, Jan. 2019.

[16] 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 Trans. image Process.*, vol. 22, no. 2, pp. 573–584, Feb. 2013.

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

[18] H. Zhang, Q. Wang, W. Shi, and M. Hao, "A novel adaptive fuzzy local information C-means clustering algorithm for remotely sensed imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5057–5068, Sep. 2017.

[19] Q. Wang, X. Wang, C. Fang, and J. Jiao, "Fuzzy image clustering incorporating local and region-level information with median memberships," *Appl. Soft Comput.*, vol. 105, 2021, Art. no. 107245.

[20] Q. Wang, X. Wang, C. Fang, and W. Yang, "Robust fuzzy C-means clustering algorithm with adaptive spatial & intensity constraint and membership linking for noise image segmentation," *Appl. Soft Comput.*, vol. 92, 2020, Art. no. 106318.

[21] L. Szilagyi, Z. Benyo, S. M. Szilágyi, and H. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 1, 2003, pp. 724–726.

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

[23] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[24] J. Shen, X. Hao, Z. Liang, Y. Liu, W. Wang, and L. Shao, "Real-time superpixel segmentation by DBSCAN clustering algorithm," *IEEE Trans. image Process.*, vol. 25, no. 12, pp. 5933–5942, Dec. 2016.

[25] Z. Li and J. Chen, "Superpixel segmentation using linear spectral clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1356–1363.

[26] Z. Ban, J. Liu, and L. Cao, "Superpixel segmentation using Gaussian mixture model," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4105–4117, Aug. 2018.

[27] X. Wei, Q. Yang, Y. Gong, N. Ahuja, and M.-H. Yang, "Superpixel hierarchy," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 4838–4849, Oct. 2018.

[28] T. Lei, X. Jia, Y. Zhang, S. Liu, H. Meng, and A. K. Nandi, "Superpixel-based fast fuzzy C-means clustering for color image segmentation," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 9, pp. 1753–1766, Sep. 2019.

[29] S. Zhou, D. Li, Z. Zhang, and R. Ping, "A new membership scaling fuzzy C-means clustering algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 9, pp. 2810–2818, Sep. 2021.

[30] J. Fang, H. Liu, J. Liu, H. Zhou, L. Zhang, and H. Liu, "Fuzzy region-based active contour driven by global and local fitting energy for image segmentation," *Appl. Soft Comput.*, vol. 100, 2021, Art. no. 106982.

[31] J. Fang, H. Liu, L. Zhang, J. Liu, and H. Liu, "Region-edge-based active contours driven by hybrid and local fuzzy region-based energy for image segmentation," *Inf. Sci.*, vol. 546, pp. 397–419, 2021.

[32] N. Kasabov and Q. Song, "DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction," *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 2, pp. 144–154, Apr. 2002.

[33] N. Kasabov, *Evolving Connectionist Systems: The Knowledge Engineering Approach*. Berlin, Germany: Springer, 2007.

[34] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 8, pp. 790–799, Aug. 1995.
[35] É. O. Rodrigues, L. Torok, P. Liatsis, J. Viterbo, and A. Conci, "k-MS: A novel clustering algorithm based on morphological reconstruction," *Pattern Recognit.*, vol. 66, pp. 392–403, 2017.
[36] T. Lei, X. Jia, T. Liu, S. Liu, H. Meng, and A. K. Nandi, "Adaptive morphological reconstruction for seeded image segmentation," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5510–5523, Nov. 2019.
[37] D. Corne et al., *New Ideas in Optimization*. New York, NY, USA: McGraw-Hill, 1999.
[38] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.
[39] G. Liu and S. Yan, "Latent low-rank representation for subspace segmentation and feature extraction," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 1615–1622.
[40] J. Chen and J. Yang, "Robust subspace segmentation via low-rank representation," *IEEE Trans. Cybern.*, vol. 44, no. 8, pp. 1432–1445, Aug. 2014.
[41] Z. Liu, W. Ou, W. Lu, and L. Wang, "Discriminative feature extraction based on sparse and low-rank representation," *Neurocomputing*, vol. 362, pp. 129–138, 2019.
[42] X. Zhu, S. Zhang, Y. Li, J. Zhang, L. Yang, and Y. Fang, "Low-rank sparse subspace for spectral clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 8, pp. 1532–1543, Aug. 2019.
[43] F. Nie, W. Chang, Z. Hu, and X. Li, "Robust subspace clustering with low-rank structure constraint," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 3, pp. 1404–1415, Mar. 2022.
[44] S. Zhang, L. Jiao, F. Liu, and S. Wang, "Global low-rank image restoration with Gaussian mixture model," *IEEE Trans. Cybern.*, vol. 48, no. 6, pp. 1827–1838, Jun. 2018.
[45] Q. Shen, Y. Chen, Y. Liang, S. Yi, and W. Liu, "Weighted Schatten p-norm minimization with logarithmic constraint for subspace clustering," *Signal Process.*, vol. 198, 2022, Art. no. 108568.
[46] J. Miao and K. I. Kou, "Color image recovery using low-rank quaternion matrix completion algorithm," *IEEE Trans. Image Process.*, vol. 31, pp. 190–201, 2022.
[47] C. Lu, Z. Lin, and S. Yan, "Smoothed low rank and sparse matrix recovery by iteratively reweighted least squares minimization," *IEEE Trans. Image Process.*, vol. 24, no. 2, pp. 646–654, Feb. 2015.
[48] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
[49] F. Nie, H. Huang, and C. Ding, "Low-rank matrix recovery via efficient Schatten p-norm minimization," in *Proc. 26th AAAI Conf. Artif. Intell.*, 2012, pp. 655–661.
[50] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.
[51] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy C-means clustering algorithm," *Comput. Geosci.*, vol. 10, no. 2/3, pp. 191–203, 1984.
[52] X. Jia, T. Lei, X. Du, S. Liu, H. Meng, and A. K. Nandi, "Robust self-sparse fuzzy clustering for image segmentation," *IEEE Access*, vol. 8, pp. 146182–146195, 2020.
[53] X. Zhang, *Matrix Analysis and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

Sensen Song received the B.S. degree in electronic information science and technology from Xinjiang Normal University, Urumqi, China, in 2014, and the M.S. degree in information and communication engineering from the Department of Information Science and Engineering, Xinjiang University, Urumqi. He is currently working toward the Ph.D. degree in computer science and technology with the Department of Information Science and Engineering, Xinjiang University. His research interests are in the area of image processing.

Zhenhong Jia received the B.S. degree in physics from Beijing Normal University, Beijing, China, in 1985, and the M.S. and Ph.D. degrees in optics from Shanghai Jiao Tong University, Shanghai, China, in 1987 and 1995, respectively.

He is currently a Professor with the Autonomous University Key Laboratory of Signal and Information Processing Laboratory, Xinjiang University, China. His research interests include digital image processing, optical information detection, and machine learning.

Jie Yang (Senior Member, IEEE) received the B.S. degree in automatic control and the M.S. degree in pattern recognition and artificial intelligence from Shanghai Jiao Tong University, Shanghai, China, in 1985 and 1988, respectively, and the Ph.D. degree in computer science from the Department of Computer Science, Hamburg University, Hamburg, Germany, in 1994.

He is currently a Professor with the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University. His major research interests include object detection and recognition, data fusion and data mining, and medical image processing.

Nikola Kasabov (Life Fellow, IEEE) received the M.Sc. degree in computing and electrical engineering, and the Ph.D. degree in mathematical sciences from the Technical University of Sofia, Bulgaria, in 1971 and 1975, respectively. He is Founding Director of KEDRI and Professor at Auckland University of Technology, Ulster University, and Bulgarian Academy of Sciences; Honorary Professor at the University of Auckland, Peking University and Teesside University; and Doctor Honoris Causa of Obuda University. His major research interests include computational intelligence, neural networks, bioinformatics, neuroinformatics, with more than 700 publications in these fields.