Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 12, No. 4, 2024, 497-510.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

Fuzzy Clustering of Noisy Images Using a Gaussian Kernel and Spatial Information with Automatic Parameter Tuning and C+ Means Initialization

Mohsen Erfani Haji Pour*

Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.

Article Info

Article History:
Received 11 September 2024
Revised 08 December 2024
Accepted 21 December 2024

DOI:10.22044/jadm.2024.15043.2606

Keywords:
Fuzzy Clustering, Noisy Images, Spatial Information, Gaussian Kernel.

**Corresponding author: m.erfanihajipour97@mail.um.ac.ir (M. Erfani Haji Pour).*

Abstract

The segmentation of noisy images remains one of the primary challenges in image processing. Traditional fuzzy clustering algorithms often exhibit poor performance in the presence of high-density noise due to insufficient consideration of spatial features. In this paper, a novel approach is proposed that leverages both local and non-local spatial information, utilizing a Gaussian kernel to counteract high-density noise. This method enhances the algorithm's sensitivity to spatial relationships between pixels, thereby reducing the impact of noise. Additionally, a C+ means initialization approach is introduced to improve performance and reduce sensitivity to initial conditions, along with an automatic smoothing parameter tuning method. The evaluation results, based on the criteria of fuzzy assignment coefficient, fuzzy segmentation entropy, and segmentation accuracy, demonstrate a significant improvement in the performance of the proposed method.

1. Introduction

In recent decades, significant advancements in image processing and algorithm analysis have drawn the attention of researchers towards the development of methods for segmentation grayscale, color, and medical images [1]. Precise segmentation of various image components is crucial in diverse fields such as medicine, machine vision, and industrial image processing. Historically, several algorithms have been introduced in this domain, with the Fuzzy Clustering Algorithm (FCM) being one of the most notable unsupervised clustering algorithms. Since this algorithm does not require training samples, it has been widely used in data mining, machine learning, and pattern recognition. The core idea of the FCM algorithm is to partition a set of samples into clusters using an objective function, maximizing intra-cluster similarity while minimizing inter-cluster similarity [2],[34]. However, this algorithm performs well only on noise-free images. When applied to noisy images, its performance deteriorates. Therefore, there is considerable room for improving the noise robustness aspect of the algorithm. In many applications, during data collection and transmission, sensitive images are often heavily contaminated with irregular noise, and the FCM algorithm fails to perform effectively in noisy image segmentation due to insufficient consideration of spatial information in the images. Consequently, many modified versions of this algorithm have been proposed to address the issue of inadequate noise robustness in the FCM algorithm. The following are some of the methods proposed in recent years. The authors in [3], [4] and [17] focus on methods based on local spatial information (LSI). These methods use statistical information from each pixel neighborhood to correct pixel membership with cluster center similarity, thereby performing segmentation while eliminating noise. The authors in [5] introduce a method for adaptively adjusting window parameters based on spatial information of the image, measuring flexible window size.

In the works by the authors in [6] and [7], methods based on non-local spatial information (NLSI) are examined, where spatial information is obtained by finding features in images with similar neighborhood configurations. NLSI enhances the algorithm's noise robustness by considering the interrelationships between pixel neighborhoods. The authors in [8] present a method that combines LSI and NLSI, simultaneously utilizing the features of both to achieve high performance. In the works by the authors in [14] and [15], methods based on superpixels are reviewed, which reduce computational complexity. The authors in [16] introduce a method involving non-negative representation learning with an adaptive graph, imposing a non-negative constraint on self-representation learning. This method improves both data representation and structure, with reduced sensitivity to noise. The authors in [18] propose an adaptive fuzzy mean segmentation algorithm using sparse subspace clustering (SSC), a new distance metric, fuzzy weight, and prior entropy, offering better initial setting guidance, superior image region separation, and feature correction. The authors in [19] present a new model based on bias field, which offers higher computational speed, greater segmentation accuracy, and better robustness compared to classical models.

The authors in [20] develop the Rough K-Means method, which incorporates the Levy-Cauchy arithmetic optimization algorithm and utilizes the CIELab color space. The authors in [21] introduce a method that adapts local information weights and uses the Gold Panning algorithm for optimizing initial cluster centers. The authors in [22] propose a framework for kernel fuzzy clustering with Fourier superpixels, providing higher segmentation accuracy. The authors in [23] propose a method that improves semi-supervised probabilistic C-means clustering by incorporating supervised information and feature weighting to enhance robustness to noise and imbalanced data.

The authors in [24] presents a method that uses the classification and regression tree CART algorithm as a substitute for ECS-AMRFC for unsupervised image segmentation. The authors in [25] propose a noise-robust fuzzy image clustering method that uses fuzzy Euclidean distance instead of regular Euclidean distance and suppresses noise using Bregman divergence. In the works by the authors in [26], the proposed method includes shape priority and connectivity as threshold criteria and utilizes fuzzy sets for feature analysis in the boundaries of color images. The authors in [27] present a geometrically adaptive fuzzy cluster set model with gradient-weighting, demonstrating superior performance using the semi-implicit optimization algorithm. The authors in [28] introduce a multi-objective fuzzy clustering algorithm (K2 MORFC) with the help of Kriging. Kullback-Leibler-based fitness functions and edge information are developed for spatial constraints and membership similarity, and the K2EA evolutionary algorithm is used for optimization. The authors in [29] propose a method that segments images through decomposition, aggregation, and merging of fuzzy superpixels, based on a low-dimensional feature space. The authors in [30] introduce a multi-objective fuzzy clustering algorithm using region information, applying complementary fitness functions considering intra-class compactness, inter-class separation, and regional consistency, and employing an incremental evolutionary framework with Kriging for optimization. The authors in [31] propose a method that employs quadratic polynomial level fuzzy C-means clustering for image segmentation in scenarios with uneven brightness, weak edges, and strong noise. This method applies modifications to the C-means clustering algorithm and improves image segmentation issues by utilizing local pixel neighborhood membership. The authors in [32] present a method for RGB-Depth image segmentation using Henry's gas solubility optimization and fuzzy clustering, based on the NYU Depth V2 dataset, with superior performance compared to K-means, fuzzy C-means, chaotic gravitational search, and J-Segmentation methods. The authors in [33] propose the SUFEMO method for radiological image segmentation of COVID-19 patients, utilizing the fuzzy electromagnetic algorithm, type-2 fuzzy logic, and superpixels, which offer reduced computational load and more accurate early disease detection.

In line with the topics discussed, this article also proposes an image segmentation algorithm to improve segmentation algorithms for noisy images. This algorithm aims to reduce noise effects and increase segmentation accuracy by utilizing LSI and NLSI along with Gaussian kernels instead of the Euclidean distances used in previous similar research, and by employing the C+ means initialization approach. One of the challenges of image segmentation algorithms based on NLSI is the tuning of parameters such as the smoothing parameter. The proposed method in this research addresses this issue and proposes an automatic tuning method for the smoothing parameter.

The remainder of this article is organized as follows: Section 2, introduces Gaussian kernel functions, the traditional FCM algorithm, and methods based on LSI and NLSI. Section 3, presents the proposed method and its implementation. Section 4, demonstrates the segmentation results of the algorithm on color, grayscale, and medical images, and discusses various performance analyses of the algorithm. Section 5 offers suggestions for future research, and finally, in Section 6 the discussion and conclusions are provided.

2. Related works

2.1. Gaussian kernel functions

One of the most commonly used functions in clustering non-linear data is kernel functions. In this part, the kernel function will be introduced in detail.

Let H be a feature space and $\Phi: x \to H$ be a mapping. The kernel function between $\Phi(y)$ and $\Phi(x)$ is defined as follows:

$$K(x,y) = \langle \Phi(x), \Phi(y) \rangle \tag{1}$$

In Equation (1), $\langle \Phi(x), \Phi(y) \rangle$ represents the inner product operator of mapping x and y in the feature space. Therefore, we have:

$$\begin{aligned}\|\Phi(x)-\Phi(y)\|^2 &= (\Phi(x)-\Phi(y))^T(\Phi(x)-\Phi(y)) \\ &= K(x,x) - 2K(x,y) + K(y,y)\end{aligned} \tag{2}$$

One of the most commonly used kernel functions is the Gaussian kernel function, which is defined as follows:

$$K(x,y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{3}$$

In Equation (3), x and y represent the data of ith and jth respectively, and σ>0 is the bandwidth of the kernel. Also, in Equation (3), we have $k(x,x)=1$ and $k(y,y)=1$. Considering these equations, we can rewrite Equation (2) as follows [9]:

$$\|\Phi(x)-\Phi(y)\|^2 = 2 - 2K(x,y) \tag{4}$$

2.2. Fuzzy clustering algorithm

Fuzzy clustering is a method for assigning a data point a specific membership value to several clusters. Let us assume we want to divide an image with N pixels into k clusters. First, the standard Fuzzy C-Means (FCM) objective function is introduced as Equation (5):

$$J = \sum_{j=1}^{N}\sum_{i=1}^{k} u_{ij}^m \|x_j - c_i\|^2 \tag{5}$$

In Equation (5), u_{ij} represents the membership value of pixel x_j in the ith cluster with center c_i and m is the fuzzifier parameter. In this equation, $\|.\|$ denotes the Euclidean distance. Using the Lagrange method, the membership values u_{ij} and the centers c_i are updated as follows until reaching the minimum value of the objective function [2]:

$$u_{ij} = \frac{1}{\sum_{r=1}^{k}\left(\frac{\|x_j - c_i\|}{\|x_j - c_r\|}\right)^{\frac{2}{m-1}}} \tag{6}$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m} \tag{7}$$

2.3. Non-local spatial information (NLSI)

one of the strategies for improving the robustness of the FCM method is considering NLSI. In this method, when segmenting an image, besides considering the neighborhood data of the central pixel, information from different blocks of neighboring pixels with similarity to the current location is also utilized, forming NLSI. This method complements the FCM method's robustness to noise to some extent. The formula for its calculation is as follows:

$$\bar{x}_j' = \frac{1}{w_j}\sum_{\substack{x_{q\in R_j} \\ q\neq j}} x_q.\exp\left(\frac{-gauss[M(x_q)-M(x_j)]_\sigma^2}{h^2}\right) \tag{8}$$

$$W_j = \sum_{\substack{x_{q\in R_j} \\ q\neq j}} \exp\left(\frac{-gauss[M(x_q)-M(x_j)]_\sigma^2}{h^2}\right) \tag{9}$$

In Equation (8), R_j represents a search window with dimensions S × S centered at pixel x_j. $M(x_q)$ and $M(x_j)$ are neighboring windows with dimensions $l \times l$ centered at pixels x_q and x_j. In this equation, $gauss[M(x_q)-M(x_j)]_\sigma^2$ denotes the weighted Gaussian distance with standard deviation ($\sigma = 4$) and h as the smoothing parameter. w_j is the normalization coefficient calculated using Equation (9) [6],[7], and [8].

2.4. Local fuzzy factor

A clustering method based totally on local fuzzy factor has been proposed to save you immoderate smoothing of the image or membership matrix and to enhance the robustness of the algorithm. This approach modifies the membership matrix of the FCM algorithm through local operations to measure local similarity and maintain image information. The formula for calculating this method is as follows:

$$G_{ir} = \sum_{r \in N_j} \frac{1}{d_{rj}+1}(1-u_{ir})^m \left\| x_r - c_i \right\|^2 \tag{10}$$

In Equation (10), N_j represents a neighborhood window of size r × r, and d_{rj} denotes the Euclidean distance between pixels x_r and x_j [3], [8], [10], and [11].

2.5. Fuzzy clustering algorithm based on LSI and NLSI

As mentioned, the traditional fuzzy clustering algorithm does not exhibit high robustness. Therefore, the use of both LSI and NLSI has been proposed to address this issue. using local neighborhood information when the image noise is enormously weak can partially recover the correct content of the image. However, with increased noise, more local neighboring pixels become contaminated with noise, making it hard to get the actual content of the image from this data. One suggestion to mitigate this challenge is to increase the radius of the local neighborhood. However, this results in excessive image smoothing, resulting in distorted image segmentation. To address this challenge, the authors in [8] proposed an approach, in which the original image undergoes processing using a method based on LSI and NLSI. As mentioned, in this approach, when segmenting the image, information from other blocks of neighboring pixels similar to the current location is also utilized, enabling the algorithm to leverage more information and reduce the impact of noise. In the next stage, the processed image is smoothed using a local fuzzy factor approach, where the local spatial impact on the central pixel is absolutely considered. The non-local and local spatial operations enhance the method's robustness and preserve more information. To compute the LSI and NLSI of the image, we first use equations (8) and (9) to calculate the NLSI. Then, using this data and Equation (10), we will rewrite the equation for the local fuzzy factor as follows:

$$G_{ir} = \sum_{r \in N_j} \frac{1}{d_{rj}+1}(1-u_{ir})^m \left\| \bar{x}_r^{'} - c_i \right\|^2 \tag{11}$$

Now, we can express the objective function as follows:

$$J = \sum_{i=1}^{k}\sum_{j=1}^{N}[u_{ij}^m \left\| \bar{x}_j^{'} - c_i \right\|^2 + \sum_{r \in N_j} \frac{1}{d_{rj}+1}(1-u_{ir})^m \left\| \bar{x}_r^{'} - c_i \right\|^2] \tag{12}$$

3. Proposed Method

In this paper, inspired by the methods mentioned above, an approach has been attempted to propose a method for image segmentation improvement. The relevant actions are as follows:

1. Algorithm based on NLSI has parameters, and the proper setting of these parameters is very effective in the performance of the algorithm. One of these parameters is the smoothing parameter (h). In this paper, a method to automatically tune this parameter according to Image features is proposed to improve the performance of the algorithm by accurately tuning this parameter.
2. Using a Gaussian kernel instead of Euclidean distances increases the sensitivity of the image segmentation algorithm to spatial relationships between pixels and helps to mitigate noise with higher densities. Therefore, considering that Euclidean distances have been used in the proposed LSI and NLSA based approach in reference [8], a Gaussian kernel can be employed instead of these distances to improve the algorithm's robustness against high-density noise.
3. In all clustering algorithms, initializing the initial values of membership values and cluster centers is required for the clustering process, which can significantly affect the algorithm's performance. In most cases, the initial values of these parameters are randomly initialized. In this paper, the approach of initializing the initial values with the C+ means is proposed to improve the algorithm's performance.

3.1. Automatic tuning of the smoothing parameter

The algorithm based on NLSI for image segmentation has influential parameters such as the smoothing parameter (h). This parameter is primarily determined empirically. If the value of h is tremendous, it leads to losing image details and edge information when obtaining MLSI. Conversely, if its value is too small, the results will be influenced by noise [8]. Therefore, in this article, it is proposed to automatically determine the value of h based on the pixel information of the image. Initially, the feature function is defined as follows:

$$F_j = \sqrt{\frac{1}{N_k}\sum_{i \in z_j^l}(x_j - x_i)^2}, \; j = 1, 2, ..., N \tag{13}$$

In Equation (13),z_j^l represents an l×l neighborhood centered at pixel x_j, and N_k is the cardinality. The values of the feature function indicate the difference between pixels and their neighborhoods. If its value is big, the area is heterogeneous, indicating a difference among the central pixel and its neighborhood. In other words, the area is closely infected with noise. Therefore, h should have a higher value. In this article, several experiments have been conducted considering proportional values to the feature function values, indicating that selecting the maximum obtained values for pixels can lead to better performance. Hence, the parameter h is determined as follows:

$$h = F_{max} \tag{14}$$

3.2. Fuzzy clustering algorithm based on Gaussian kernel with LSI and NLSI

As observed, in the fuzzy clustering algorithm based on LSI and NLSI, Euclidean distances, similar to the FCM method, are still utilized. Although this approach is computationally simple, employing Euclidean distance can lead to unreliable results in image segmentation when affected by high-density noise, outliers, etc. Therefore, in this article, the use of more robust distance metrics, such as the Gaussian kernel in the objective function, is suggested to reduce the impact of outliers and high-density noise on clustering results. Smoothing is a fundamental feature of the Gaussian kernel, which entails estimating mean values. This feature aids in noise reduction in images. When the Gaussian kernel is applied to an image, each pixel is calculated based on the pixel values in its vicinity (especially nearby points) with specific weights. These weights are calculated based on the Gaussian probability distribution. Another essential feature in Gaussian kernel smoothing is that points further away have less weight. In other words, points closer to the target point are more influential in the averaging process, while points farther away have less weight.

To apply the Gaussian kernel instead of Euclidean distances in the fuzzy clustering algorithm based on LSI and NLSI, equations (11) and (12) are rewritten as follows, based on Equation (4):

$$G_{ir} = \sum_{r\in N_j} \frac{1}{d_{rj}+1}(1-u_{ir})^m (1-K(\bar{x}_r^{'}, c_i)) \tag{15}$$

$$J = \sum_{i=1}^{k}\sum_{j=1}^{N}[u_{ij}^m (1-K(\bar{x}_j^{'}, c_i)) + \sum_{r\in N_j} \frac{1}{d_{rj}+1}(1-u_{ir})^m (1-K(\bar{x}_r^{'}, c_i))] \tag{16}$$

In these equations, $K(.,.)$ represents the Gaussian kernel function, calculated based on Equation (3).

3.3. C+ means initializing approach

As mentioned, in all clustering algorithms, initializing membership values and cluster centers is necessary for the clustering process. Most of the time, these initial values are randomly assigned. In this paper, it is proposed to use the initial value approach of C+ means instead of random initialization. In this approach, instead of randomly initializing membership values and cluster centers in the proposed clustering algorithm, first the traditional fuzzy clustering algorithm is executed for the desired image, and then the output of this algorithm, i.e. the obtained membership values and cluster centers, are considered as the initial values for the proposed algorithm. Performing this process reduces the sensitivity of the algorithm to initial conditions.

3.4. Calculating the iterative relationships of the algorithm

The objective function of the proposed method was introduced in Equation (16). To perform the image segmentation process using the proposed method, the membership values and cluster centers must be updated until the objective function reaches its minimum value. In this part, using the Lagrange method, the equations related to the membership values and cluster centers will be calculated. First, the Lagrange function is defined as follows:

$$L = \sum_{i=1}^{k}\sum_{j=1}^{N}[u_{ij}^m (1-K(\bar{x}_j^{'}, c_i)) + \sum_{r\in N_j} \frac{1}{d_{rj}+1}(1-u_{ir})^m (1-K(\bar{x}_r^{'}, c_i))] + \sum_{j=1}^{N}\lambda_j (1-\sum_{j=1}^{N} u_{ij}) \tag{17}$$

By partially differentiating the Lagrange function concerning u_{ij} and c_i , we will have:

$$A_{irj} = \frac{1}{d_{rj}+1}(1-u_{ir})^m (1-K(\bar{x}_r^{'}, c_i)) \tag{18}$$

$$B_{ij} = (1-K(\bar{x}_j^{'}, c_i)) + \sum_{r\in N_j} A_{irj} \tag{19}$$

$$B_{lj}^{'} = (1-K(\bar{x}_j^{'}, c_l)) + \sum_{r\in N_j} \frac{1}{d_{rj}+1}(1-u_{ir})^m (1-K(\bar{x}_r^{'}, c_l))] \tag{20}$$

$$u_{ij} = \frac{1}{\sum_{l=1}^{k}(\frac{B_{ij}}{B_{lj}^{'}})^{\frac{1}{m-1}}} \tag{21}$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ij}^m K(\bar{x}_j^{'}, c_i)\bar{x}_j^{'}}{\sum_{j=1}^{N} u_{ij}^m K(\bar{x}_j^{'}, c_i)} \tag{22}$$

The pseudo-code for the proposed method is provided in Algorithm (1).

4. Simulation and results

In this part, we delve into simulating the proposed method and compare the results obtained with seven other algorithms. First, a brief introduction to these 7 algorithms will be provided. The FCM method stands as the primary and conventional approach to fuzzy clustering. Next, the FLICM [3] method incorporates LSI, while FCM_SNLS [7] integrates NLSI. Meanwhile, CGFFCM [12] introduces weight adjustments based on image features. The FALRCM [13] method leverages KL data for the membership matrix but is confined by various parameters. Moving on, FSC_LNML [6] incorporates information from fuzzy and image subspace. The FWCW_IT2PFCM_SIC [2] algorithm is also a probabilistic type-2 fuzzy clustering algorithm with adaptive spatial constraints and local feature and cluster weighting. Lastly, FCM_LNIS [8] integrates both LSI and NLSI using Euclidean distance.

Algorithm 1: The proposed method.

Start,
Inputs:
- $\boldsymbol{I}$: Noisy image,
- $\boldsymbol{k}$: number of clusters,
- $\boldsymbol{m}$: fuzziness parameter,
- $\boldsymbol{l}$: neighborhood window,
- $\boldsymbol{S}$: search window,
- $\boldsymbol{\sigma}$: kernel bandwidth,
- $\boldsymbol{t_{max}}$: maximum number of iterations,
- $\boldsymbol{\varepsilon}$: minimum error,

Step 1: Normalize $\boldsymbol{I}$ within the range [0,1].
Step 2: Calculate the smoothing parameter ($\boldsymbol{h}$) using Equations (13) and (14).
Step 3: Calculate the NLSI part ($\boldsymbol{\bar{x}'_j}$) using Equations (8) and (9).
Step 4: Execute the FCM algorithm using Equations (5) to (7) and calculate the initial values $\boldsymbol{u_{ij}}^{(\boldsymbol{0})}$ and $\boldsymbol{c_i}^{(\boldsymbol{0})}$.
Step 5: Set iteration counter $\boldsymbol{a}$ = 1.
- ***Step 6:*** Repeat,
- ***Step 6.1***: Calculate cluster centers $\boldsymbol{c_i}^{(\boldsymbol{a})}$ using Equation (22),
- ***Step 6.2***: Calculate membership values $\boldsymbol{u_{ij}}^{(\boldsymbol{a})}$ using Equation (21),
- ***Step 6.3***: Calculate local fuzzy factor $\boldsymbol{G_{ir}}$ using Equation (15),
- ***Step 6.4***: Calculate the objective function $\boldsymbol{J}^{(\boldsymbol{a})}$ using Equation (16),
- ***Step 6.5***: $\boldsymbol{a}$ = $\boldsymbol{a}$ + 1,
- ***Step 6.6:*** Check if $\|\boldsymbol{J}^{(\boldsymbol{a})}-\boldsymbol{J}^{(\boldsymbol{a-1})}\| \leq \boldsymbol{\varepsilon}$ or $\boldsymbol{a} > \boldsymbol{t_{max}}$,
- ***Step 6.7***: If step 6.6 holds true, then The Repeat iteration ends; otherwise return to Step 6.1.

Outputs:
Membership matrix $\boldsymbol{u_{ij}}$, cluster centers $\boldsymbol{c_i}$.
End.

4.1. incorporating noise to images
To evaluate the effectiveness of the proposed algorithm 1, it is imperative to introduce combined noise into the images and subsequently assess the method's robustness. This procedure can be facilitated by utilizing Algorithm (2).

Algorithm 2: The Image Noise Addition Algorithm.

Start,
Input:
O: Original image,
Step 1: Normalize ***O*** within the range [0,1],
Step 2: Adjust the noise density ***σ₁***,
Step 3: Combine Gaussian noise with mean zero and variance σ1 with image ***O***,
Step 4: Combine salt-and-pepper noise with density ***σ₁*** with image ***O***,
Step 5: Combine uniform noise with density ***σ₁*** with image ***O***,
Step 6: Normalize the pixel values of the image to between [0, 255] ,

Output:
Noisy image ***I***.
End.

In the subsequent experiments, all noisy images are described as follows:
Mixture with σ1 × 100% noise.

4.2. The Adjustment of Constant Parameters
For conducting experiments, several parameters in image segmentation algorithms need to be adjusted. In this report, the minimum error ε, maximum iterations t_{max}, and fuzziness coefficient m are set to 10^{-4},100, and 2 respectively. In the proposed algorithm, for achieving the best results, the parameters of the search window s, neighborhood window l, and kernel bandwidth σ are set to 15, 7, and 10, respectively. Additionally, the required parameters in the comparison algorithms are adjusted according to the references for comparing the proposed algorithm.

4.3. Assessment metrics
To demonstrate the effectiveness of the proposed algorithm and compare it with other algorithms, this article will use three assessment metrics: Fuzzy Assignment coefficient (V_{PC}), Fuzzy Segmentation Entropy (V_{PE}), and segmentation accuracy (SA).

1. Fuzzy Assignment (V_{PC}): This metric serves as an indicator of membership proximity. A higher value indicates a more favorable classification outcome and greater uniqueness in pixel features. The formula for this metric is expressed as follows:

$$V_{PC} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} u_{ij}^2 \times 100\% \tag{23}$$

2. Fuzzy Segmentation Entropy (V_{PE}): A smaller value indicates a better classification outcome. The formula for this metric is as follows:

$$V_{PE} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} u_{ij}\ln(u_{ij}) \times 100\% \tag{24}$$

3. Segmentation Accuracy (SA): This metric suggests the ratio of successfully segmented

pixels to the full number of pixels. The formula for this metric is as follows:

$$SA = \frac{\sum_{i=1}^{k} A_i \cap C_i}{\sum_{r=1}^{k} C_r} \times 100\% \quad (25)$$

Where A_i stands for the ith segmented region obtained, and C_i represents the ith region in the reference image segmentation. A higher value of this metric signifies a higher level of accuracy in the segmentation process.

4.4. Experiments

4.4.1. Gray-scale images

In this part, we assess the performance of the algorithms on grayscale images.

Figure 2 showcases the segmentation results of image Sy1 into two categories. Figure 3 displays the segmentation outcomes of image Sy2 into three categories following the introduction of a 5% noise. Furthermore, Tables 1-4 provide a detailed overview of the algorithms' performance under varying levels of noise.

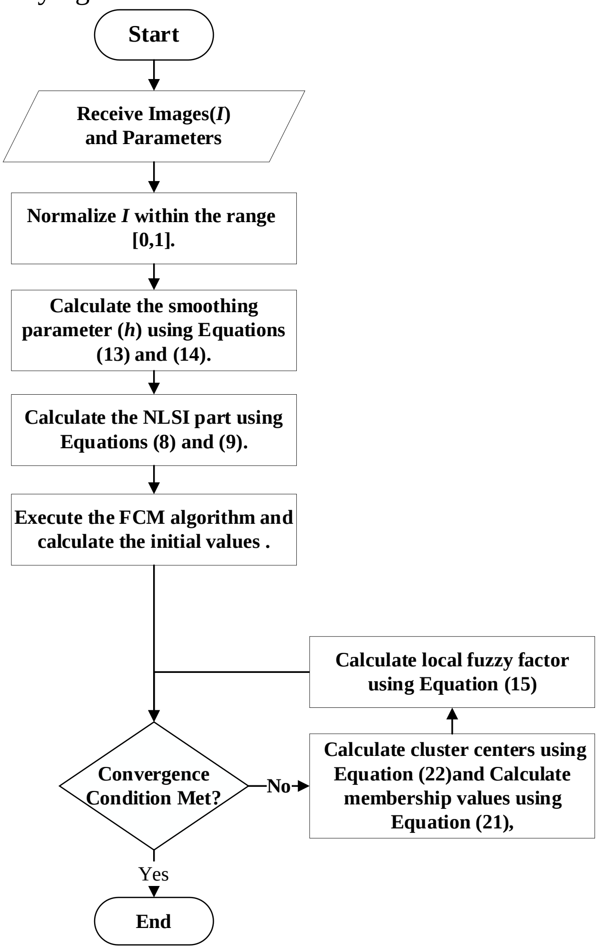

Figure 1. Flowchart of the proposed method

In Tables 1 and 2, the evaluation has been conducted based on the criteria of V_{PC} and V_{PE}. The traditional FCM algorithm has shown relatively acceptable performance but lacks robustness against increasing noise density. The FCM_SNLS algorithm demonstrates robustness to noisy images by increasing the flexibility of non-local spatial information. However, its stability decreases with increasing noise density. The local fuzzy factor in FLICM does not perform well with increasing in noise density. The feature-based clustering method CGFFCM performs poorly against image noise due to its dependency on input feature information. In contrast, FALRCM, FCM_LNIS, and FSC_LNML provide higher values for $V_{PC}$and V_{PE} by utilizing complicated spatial information extraction operations, which helps improve clustering results. As observed, the average values of V_{PE} and V_{PC} for the proposed algorithm are 98.40 and 3.11, respectively. Experimental results for the proposed algorithm indicate superior segmenting performance compared to most of the mentioned algorithms. Furthermore, by examining these two criteria, it can be seen that the proposed algorithm exhibits better and higher robustness to increasing noise density. Various algorithms have also been evaluated using the SA metric. The overall performance of FCM and CGFFCM has been weak, with average SA values of 88.32 and 72.69, respectively. These algorithms are sensitive to noise and have low robustness due to not considering spatial information .CGFFCM based on image feature data for decomposition and analysis, with the input to the algorithm being the image features rather than the image itself. As shown in Figures 2 and 3, most of the image information is lost in the noise, resulting in significant errors in the analysis results. FCM_SNLS and FLICM perform well against low density noise. For example, during the decomposition of Sy1, their SA values could exceed 99%, which is similar to the performance of the proposed algorithm. but, with a non-stop increase in noise density, the overall performance of these algorithms regularly decreases. This is because LSI and NLSI have limited power against high-density noise, leading to reduced accuracy of decomposition and analysis. The values of FSC_LNML and FCM_LNIS are very near to the results of the proposed algorithm. Cconsistent with the experimental data, the average values of these strategies and the proposed method in the SA index can exceed 97%. The last experiment demonstrates that the proposed algorithm in this paper effectively solves the problem of segmenting images containing noise through image information processing using LSI and NLSI, utilizing the Gaussian kernel and initial c+ means assignment approach.

4.4.2. Color Images

In this part, the performance of algorithms for color images has been evaluated. Figure 4 shows the segmentation results of color images from the BSDS500 database with 5% noise applied. These images were randomly selected from this database. Additionally, Table 4 represents the performance of the mentioned algorithms in terms of the SA metric. In this part, color images are used to test the segmentation of the proposed algorithm. Different color images from the BSDS500 database have been used to perform these tests. As shown in Figure 4, the segmentation results of FCM and CGFFCM have low noise and segmentation performance, indicating that these two algorithms have weak effects on noisy images.

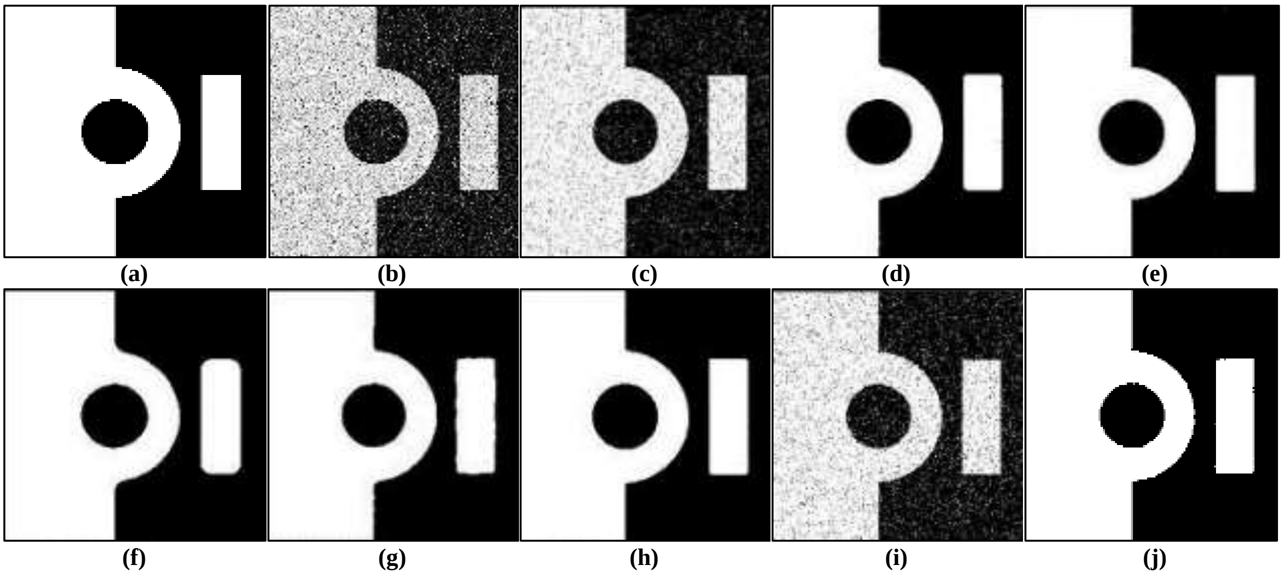

Figure 2. Segmentation results of image Sy1 into two categories. a) original image. b) Image with 5% noise. c) FCM method. d) FLICM method. e) FCM_SNLS method. f) FCM_LNIS method. g) FSC_LNML method. h) FALRCM method. i) CGFFCM method. j) Proposed method.

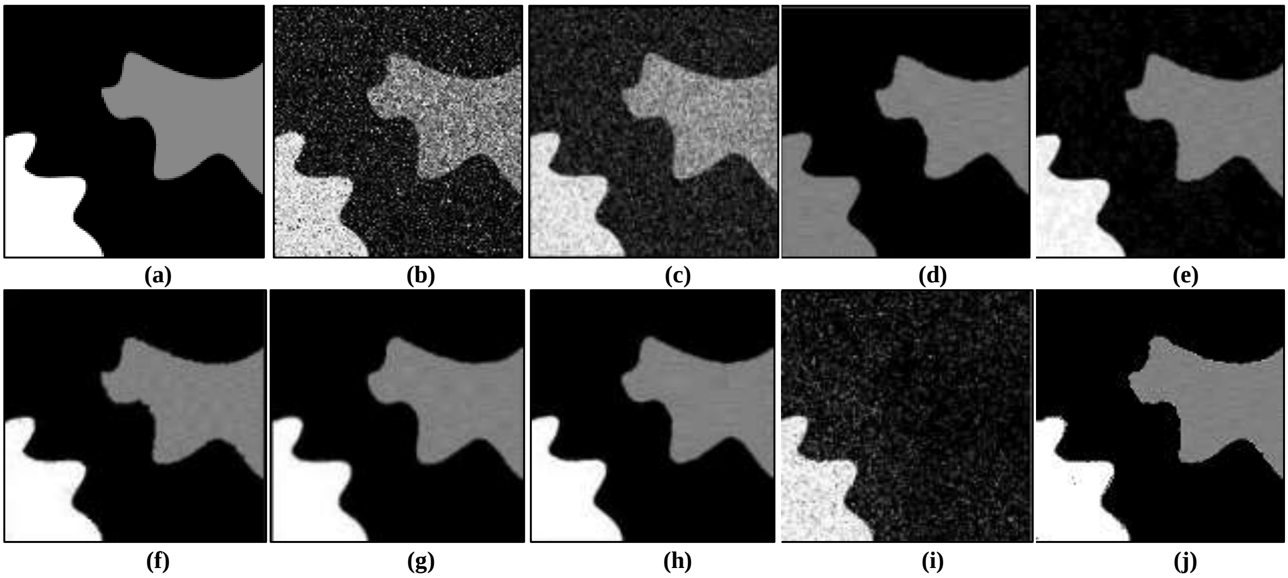

Figure 3. Segmentation results of image Sy2 into two categories. a) original image. b) Image with 5% noise. c) FCM method. d) FLICM method. e) FCM_SNLS method. f) FCM_LNIS method. g) FSC_LNML method. h) FALRCM method. i) CGFFCM method. j) Proposed method.

Table 1. Comparison of the V_{PC} metric for grayscale images.

Images	noise	CGFFCM	FALRCM	FSC_ LNML	FCM_ LNIS	FCM_ SNLS	FILCM	FCM	Proposed method
Sy1	1%	93.23	97.60	98.30	**99.82**	99.64	97.70	98.55	99.21
	5%	75.66	97.94	98.16	**99.81**	97.73	96.86	92.59	99.13
	10%	69.81	98.87	97.91	**99.81**	94.17	95.59	89.07	99.02
	15%	68.08	95.80	98.49	97.78	89.88	94.14	88.60	**98.85**
Sy2	1%	82.31	96.83	98.77	**99.54**	98.01	96.44	92.25	98.96
	5%	76.46	97.64	98.30	98.50	90.95	60.55	86.01	**98.60**
	10%	68.96	95.33	97.34	97.46	82.90	57.03	86.52	**97.79**
	15%	58.05	94.14	96.28	**98.34**	77.78	54.15	87.59	95.66
Average		74.07	96.77	97.94	**98.88**	91.38	81.56	90.15	98.40

Table 2. Comparison of the V_{PE} metric for grayscale images.

Images	noise	CGFFCM	FALRCM	FSC_ LNML	FCM_ LNIS	FCM_ SNLS	FILCM	FCM	Proposed method
Sy1	1%	11.90	2.52	**1.16**	1.60	1.97	3.86	3.63	1.31
	5%	39.13	4.56	1.52	1.82	4.59	5.93	13.59	**1.49**
	10%	46.82	7.56	2.14	2.03	11.09	8.67	18.71	**1.79**
	15%	49.03	9.25	3.12	3.70	17.85	11.56	19.48	**2.23**
Sy2	1%	34.50	4.23	2.95	2.46	4.85	7.35	15.60	**1.91**
	5%	45.13	6.89	3.69	3.59	18.27	61.50	25.77	**2.91**
	10%	57.31	8.59	5.93	4.74	32.07	69.22	24.83	**4.65**
	15%	73.01	10.11	8.22	**5.08**	40.74	75.39	22.92	8.62
Average		44.60	6.71	3.59	3.13	16.43	30.43	18.06	**3.11**

Table 3. Comparison of the SA metric for grayscale images.

Images	noise	CGFFCM	FALRCM	FSC_ LNML	FCM_ LNIS	FCM_ SNLS	FILCM	FCM	Proposed method
Sy1	1%	99.23	85.99	98.36	99.59	99.10	99.04	99.49	**99.96**
	5%	93.41	99.32	99.40	99.55	99.35	99.48	98.10	**99.90**
	10%	83.83	98.57	99.41	99.49	99.60	99.14	86.15	**99.70**
	15%	76.90	97.95	99.43	99.48	98.13	99.05	77.50	**99.49**
Sy2	1%	74.49	98.58	97.52	98.42	**99.48**	98.76	97.74	98.18
	5%	53.03	67.24	96.35	98.43	**98.47**	73.53	77.49	97.76
	10%	49.70	95.48	95.99	97.24	90.11	81.42	67.93	**97.49**
	15%	50.97	94.29	94.65	**97.07**	80.98	83.04	62.22	95.79
Average		72.69	92.18	97.64	**98.65**	95.65	91.68	83.32	98.53

Table 4. Comparison of the SA Metric for Images from the BSDS500 Database.

Images	CGFFCM	FALRCM	FSC_ LNML	FCM_ LNIS	FCM_ SNLS	FILCM	FCM	FWCW_ IT2PFCM_SIC	Proposed method
3063	59.95	89.84	78.56	93.45	**94.41**	79.66	69.51	80.21	75.03
24063	69.94	83.76	89.52	89.69	79.38	89.15	68.63	96.88	**96.90**
42049	32.50	84.01	93.49	94.87	93.04	92.44	68.11	94.76	**95.34**
118035	76.54	**96.48**	82.87	94.32	94.38	71.89	76.37	95.22	95.59

Except for #3063, FLICM has shown better segmentation effects. FCM_SNLS has very good results in segmentation of #118035 and #42049. However, both algorithms perform poorly against noise and produce weak visual effects.

FALRCM, except for the failure in precise object-background segmentation in #3063, has performed well in segmenting other images.

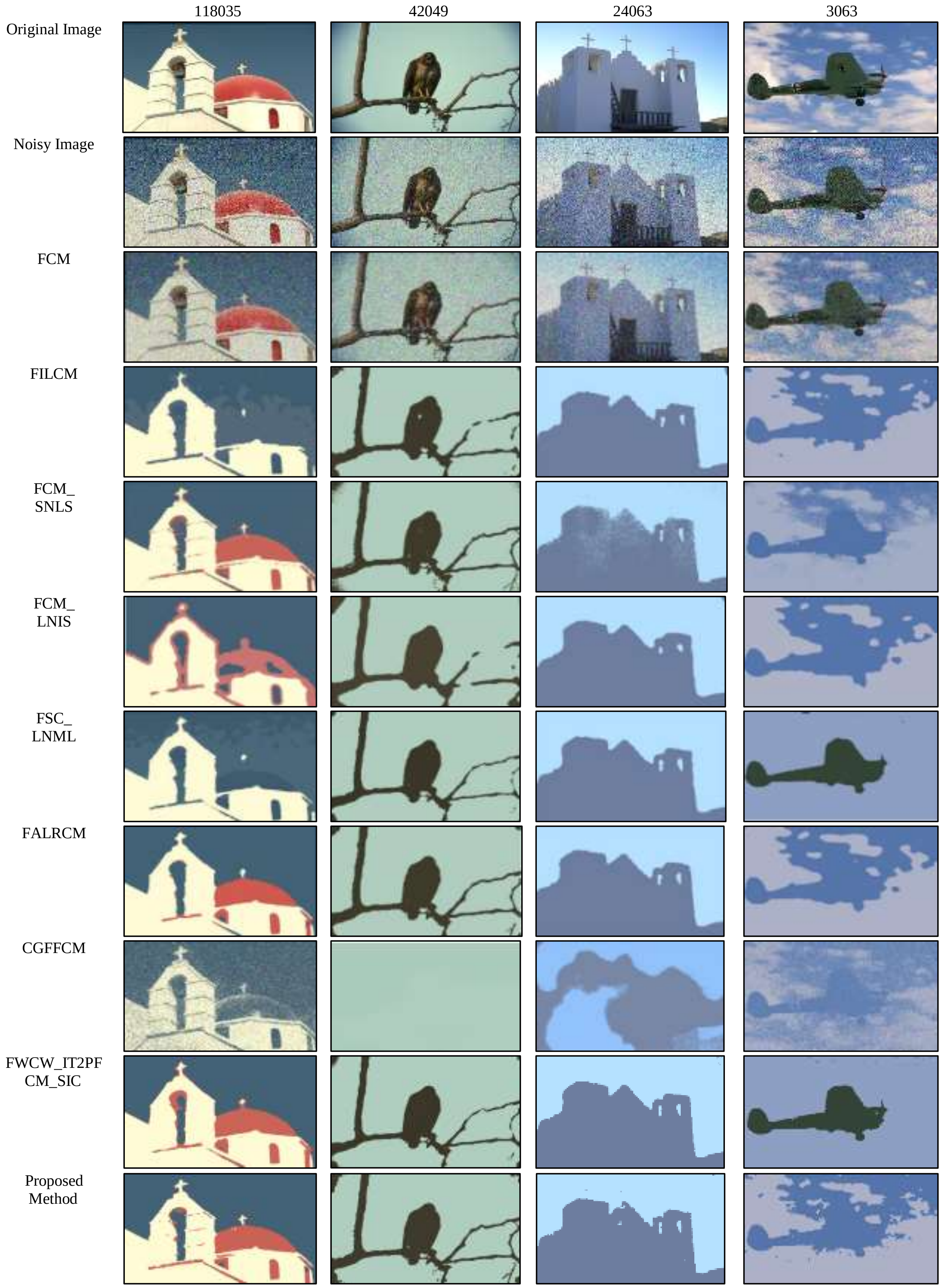

Figure 4. Segmentation results of selected color images from BSDS500 database.

The proposed algorithm, similar to the successful algorithms mentioned, has high SA values and excellent performance, except for #3063. The results indicate that the proposed method has better noise suppression capabilities, and considering comprehensive LSI and NLSI along with Gaussian kernel can improve segmentation accuracy. Experimental data shows that this proposed method not only has a significant impact on noise suppression but also is capable of precise target segmentation.

4.4.3. MR Images

In this part, the overall performance of the proposed method for MR images has been evaluated. Figure 5 shows the segmentation effects of 3 MR images with numerous applied noise levels.

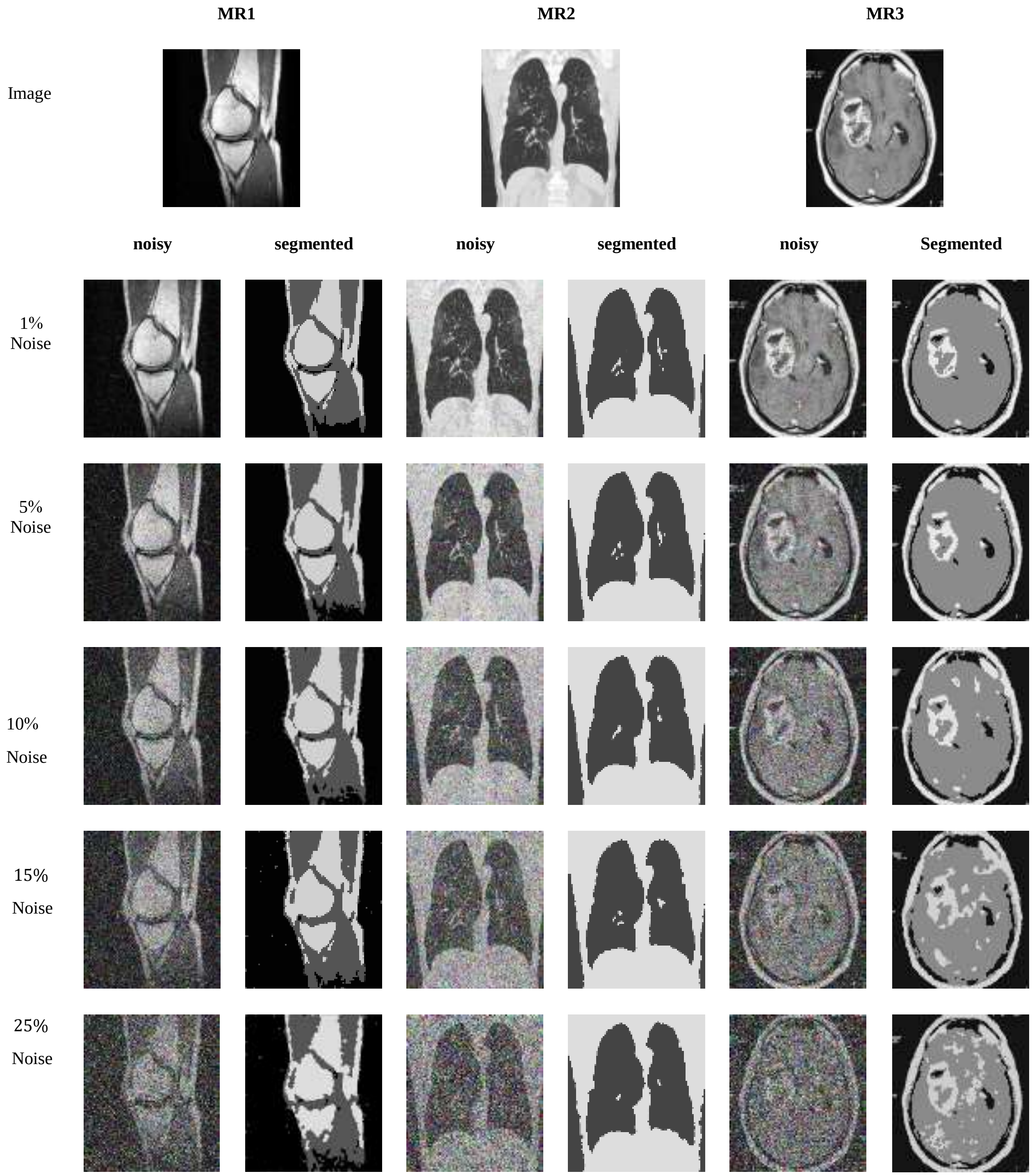

Figure 5. Segmentation Results of MR Images by the Proposed Algorithm.

Additionally, Table 5 represents the performance of the proposed algorithm with increasing noise levels. Due to the fact that MR images contain a lot of details and are important for medical diagnosis, the performance of the algorithm is very important. All the segmentation results obtained in this section are almost noise-free. In image #MR1, the bone region is correctly segmented. However, the bottom area is detected as background. In image #MR2, the lung region is well separated and recognizable. In image #MR3, the algorithm has performed well in separating white matter and gray matter. The boundaries are clearly defined, and only a few more random structures than noise points and some narrower areas have been mistakenly segmented. Considering the results in Figure 5 and Table 5 ($V_{PE}$and V_{PC} metrics), it can be observed that the proposed algorithm exhibits very high robustness against increasing noise density. Even with severe noise, the algorithm performs very well, and the segmentation of target structures is done effectively.

5. Suggestions

In future research, it is recommended to conduct a more extensive examination of various types of noise, including mixed noise, to better understand their impact on image segmentation. Furthermore, integrating novel approaches, such as deep networks, could enhance the performance of the proposed algorithm. Investigating the potential applications of the algorithm in specific domains, such as medical or industrial imaging, could also be valuable and provide direction for future studies.

Table 5. Comparison of V_{PE} and V_{PC} Metrics in MR Images

Images	Noise	V_{PE}	V_{PC}
MR1	1%	5.93	96.92
	5%	8.88	95.72
	10%	11.30	94.70
	15%	12.67	94.17
	25%	14.57	93.67
MR2	1%	2.81	98.32
	5%	3.68	98.00
	10%	4.5	97.70
	15%	5.14	97.43
	25%	6.55	96.87
MR3	1%	11.17	93.94
	5%	14.41	92.65
	10%	17.11	91.46
	15%	20.52	89.79
	25%	24.36	83.23

6. Discussion and Conclusion

In this paper, a method for image segmentation in the presence of noise using non-local and local spatial information along with Gaussian kernel and an initialization approach with the means c+ value is proposed. Simulation results on grayscale and color images from the BSDS500 database, as well as on MR medical images with noise, demonstrate that the proposed method illustrate highly desirable performance. In contrast to previous studies, considering the outcomes results, our proposed algorithm has demonstrated higher robustness to noise and achieved to a higher level of accuracy in segmentation. Specifically, it performs better than other algorithms in lowering the influence of noise for better segmentation in color images with 5% noise. Additionally, in noisy MR images, our algorithm has shown robustness against various types of noise and also achieved high accuracy in segmentation. This work illustrates that in grayscale, color, and even noisy MR medical images, the proposed method exhibits high accuracy in segmentation and can serve as an effective solution for many image-based diagnostic processes.

References

[1] Y. Chen, Z. Wang, and X. Bai, "Fuzzy sparse subspace clustering for infrared image segmentation," IEEE Transactions on Image Processing, vol. 32, pp. 2132-2146, 2023.

[2] T. Wei, X. Wang, J. Wu, and S. Zhu, "Interval type-2 possibilistic fuzzy clustering noisy image segmentation algorithm with adaptive spatial constraints and local feature weighting & clustering weighting," International Journal of Approximate Reasoning, vol. 157, pp. 1-32, 2023.

[3] 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.

[4] C. Feng, W. Li, J. Hu, K. Yu, and D. Zhao, "BCEFCM_S: Bias correction embedded fuzzy c-means with spatial constraint to segment multiple spectral images with intensity inhomogeneities and noises," Signal Processing, vol. 168, p. 107347, 2020.

[5] 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 Transactions on Geoscience and Remote Sensing, vol. 55, no. 9, pp. 5057-5068, 2017.

[6] T. Wei, X. Wang, X. Li, and S. Zhu, "Fuzzy subspace clustering noisy image segmentation algorithm with adaptive local variance & non-local information and mean membership linking,"

Engineering Applications of Artificial Intelligence, vol. 110, p. 104672, 2022.

[7] F. Zhao, "Fuzzy clustering algorithms with self-tuning non-local spatial information for image segmentation," Neurocomputing, vol. 106, pp. 115-125, 2013.

[8] J. Wu, X. Wang, T. Wei, and C. Fang, "Full-parameter adaptive fuzzy clustering for noise image segmentation based on non-local and local spatial information," Computer Vision and Image Understanding, vol. 235, p. 103765, 2023.

[9] S. Zeng, X. Wang, X. Duan, S. Zeng, Z. Xiao, and D. Feng, "Kernelized mahalanobis distance for fuzzy clustering," IEEE Transactions on Fuzzy Systems, vol. 29, no. 10, pp. 3103-3117, 2020.

[10] Lei, Tao, and Asoke K. Nandi. *Image Segmentation: Principles, Techniques, and Applications*. John Wiley & Sons, 2022. (chapter 2)

[11] Caponetti, Laura, and Giovanna Castellano. *Fuzzy logic for image processing: A gentle introduction using Java*. Springer International Publishing, 2017. (chapter 7)

[12] A. G. Oskouei, M. Hashemzadeh, B. Asheghi, and M. A. Balafar, "CGFFCM: Cluster-weight and Group-local Feature-weight learning in Fuzzy C-Means clustering algorithm for color image segmentation," Applied Soft Computing, vol. 113, p. 108005, 2021.

[13] Q. Wang, X. Wang, C. Fang, and J. Jiao, "Fuzzy image clustering incorporating local and region-level information with median memberships," Applied Soft Computing, vol. 105, p. 107245, 2021.

[14] Y. Liu, E. Li, S. Wang, Y. Zhu, and W. Zhu, "Superpixel segmentation of high-resolution remote sensing image based on feature reconstruction method by salient edges," Journal of Applied Remote Sensing, vol. 17, no. 2, pp. 026516-026516, 2023.

[15] C. Wu and J. Zhao, "Robust superpixel-based fuzzy possibilistic clustering method incorporating local information for image segmentation," The Visual Computer, vol. 40, no. 11, pp. 7961-8000, 2024.

[16] Y.-P. Zhao, X. Dai, Z. Wang, and X. Li, "Subspace clustering via adaptive non-negative representation learning and its application to image segmentation," IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 8, pp. 4177-4189, 2023.

[17] M. Ray, N. Mahata, and J. K. Sing, "Uncertainty parameter weighted entropy-based fuzzy c-means algorithm using complemented membership functions for noisy volumetric brain MR image segmentation," Biomedical Signal Processing and Control, vol. 85, p. 104925, 2023.

[18] S. Song, Z. Jia, F. Shi, J. Wang, and D. Ni, "Adaptive fuzzy weighted C-mean image segmentation algorithm combining a new distance metric and prior entropy," Engineering Applications of Artificial Intelligence, vol. 131, p. 107776, 2024.

[19] G. Weng and B. Dong, "A new active contour model driven by pre-fitting bias field estimation and clustering technique for image segmentation," Engineering Applications of Artificial Intelligence, vol. 104, p. 104299, 2021.

[20] A. Das, A. Namtirtha, and A. Dutta, "Lévy–Cauchy arithmetic optimization algorithm combined with rough K-means for image segmentation," Applied Soft Computing, vol. 140, p. 110268, 2023.

[21] D. Wei, Z. Wang, L. Si, C. Tan, and X. Lu, "An image segmentation method based on a modified local-information weighted intuitionistic Fuzzy C-means clustering and Gold-panning Algorithm," Engineering Applications of Artificial Intelligence, vol. 101, p. 104209, 2021.

[22] L. Chen, Y.-P. Zhao, and C. Zhang, "Efficient kernel fuzzy clustering via random Fourier superpixel and graph prior for color image segmentation," Engineering Applications of Artificial Intelligence, vol. 116, p. 105335, 2022.

[23] H. Yu, X. Xu, H. Li, Y. Wu, and B. Lei, "Semi-supervised possibilistic c-means clustering algorithm based on feature weights for imbalanced data," Knowledge-Based Systems, vol. 286, p. 111388, 2024.

[24] F. Zhao, Z. Tang, Z. Xiao, H. Liu, J. Fan, and L. Li, "Ensemble CART surrogate-assisted automatic multi-objective rough fuzzy clustering algorithm for unsupervised image segmentation," Engineering Applications of Artificial Intelligence, vol. 133, p. 108104, 2024.

[25] H. Lohit and D. Kumar, "Modified total Bregman divergence driven picture fuzzy clustering with local information for brain MRI image segmentation," Applied Soft Computing, vol. 144, p. 110460, 2023.

[26] P. Karthick, S. Mohiuddine, K. Tamilvanan, S. Narayanamoorthy, and S. Maheswari, "Investigations of color image segmentation based on connectivity measure, shape priority and normalized fuzzy graph cut," Applied Soft Computing, vol. 139, p. 110239, 2023.

[27] P. Shi, L. Guo, H. Cui, and L. Chen, "Geometric consistent fuzzy cluster ensemble with membership reconstruction for image segmentation," Digital Signal Processing, vol. 134, p. 103901, 2023.

[28] F. Zhao, Z. Xiao, H. Liu, Z. Tang, and J. Fan, "A knee point driven Kriging-assisted multi-objective robust fuzzy clustering algorithm for image segmentation," Knowledge-Based Systems, vol. 271, p. 110522, 2023.

[29] T. C. Ng, S. K. Choy, S. Y. Lam, and K. W. Yu, "Fuzzy superpixel-based image segmentation," Pattern Recognition, vol. 134, p. 109045, 2023.

[30] F. Zhao, Z. Tang, H. Liu, Z. Xiao, and J. Fan, "A reliable region information driven kriging-assisted multiobjective rough fuzzy clustering algorithm for color image segmentation," Expert Systems with Applications, vol. 228, p. 120419, 2023.

[31] C. Wu and Z. Wang, "A modified fuzzy dual-local information c-mean clustering algorithm using quadratic surface as prototype for image segmentation," Expert Systems with Applications, vol. 201, p. 117019, 2022.

[32] N. K. Yadav and M. Saraswat, "A novel fuzzy clustering based method for image segmentation in RGB-D images," Engineering Applications of Artificial Intelligence, vol. 111, p. 104709, 2022.

[33] S. Chakraborty and K. Mali, "SUFEMO: A superpixel based fuzzy image segmentation method for COVID-19 radiological image elucidation," Applied Soft Computing, vol. 129, p. 109625, 2022.

[34] M. Fateh and E. Kabir, "Color Reduction in Hand-drawn Persian Carpet Cartoons before Discretization using image segmentation and finding edgy regions," *Journal of AI and Data Mining*, vol. 6, no. 1, pp. 47-58, 2018, doi: 10.22044/jadm.2017.908.

خوشه‌بندی فازی تصاویر نویزدار با استفاده از کرنل گاوسی و اطلاعات مکانی با تنظیم پارامترهای خودکار و مقداردهی اولیه میانگین+C

محسن عرفانی حاجی پور*

گروه مهندسی برق کنترل، دانشگاه فردوسی مشهد، مشهد، ایران

ارسال ۲۰۲۴/۰۹/۱۱؛ بازنگری ۲۰۲۴/۱۲/۰۸؛ پذیرش ۲۰۲۴/۱۲/۲۱

چکیده:

تقسیم‌بندی تصاویر نویزدار یکی از چالش‌های اصلی در پردازش تصویر است. الگوریتم‌های خوشه‌بندی فازی سنتی اغلب در حضور نویز با چگالی بالا به دلیل در نظر گرفتن ناکافی ویژگی‌های مکانی، عملکرد ضعیفی از خود نشان می‌دهند. در این مقاله، یک رویکرد جدید پیشنهاد شده است که از اطلاعات مکانی محلی و غیر محلی و یک کرنل گاوسی برای مقابله با نویز با چگالی بالا استفاده می‌کند. این روش حساسیت الگوریتم را به روابط مکانی بین پیکسل‌ها افزایش می‌دهد و در نتیجه تاثیر نویز را کاهش می‌دهد. علاوه بر این، یک رویکرد مقداردهی اولیه میانگین +C برای بهبود عملکرد و کاهش حساسیت به شرایط اولیه، همراه با یک روش تنظیم پارامتر هموارسازی خودکار معرفی شده است. نتایج ارزیابی، بر اساس معیارهای ضریب تخصیص فازی، آنتروپی تقسیم بندی فازی، و دقت تقسیم بندی، بهبود قابل توجهی را در عملکرد روش پیشنهادی نشان می‌دهد.

کلمات کلیدی: خوشه‌بندی فازی، تصاویر نویزدار، اطلاعات مکانی، کرنل گوسی.