Fast Fuzzy C-Means Clustering based on SLIC Superpixel for Color Image Segmentation

Souvik Sarkar[1] Sourav Paul[2]

Research Scholar[1] Assistant Professor[2]

[1,2] Department of Computer Science and Engineernig

Sikkim Manipal Institute of Technology, Sikkim, India

Abstract-Fuzzy c-means (FCM) clustering works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster center and the data point .However, Most of FCM algorithms are time consuming and unable to provide desired segmentation results for color images because the incorporation of local spatial information often leads to a high computational complexity due to the repeated distance computation between pixels within local spatial neighbors and clustering centers. To solve this issue, concept of superpixels using Simple Linear Iterative Clustering (SLIC) which reduces complexity of FCM algorithm.Firstly, Simple Linear Iterative Clustering (SLIC) operation is defined to obtain a superpixel image. Superpixel image provides better local spatial neighborhoods that are helpful for improving color image segmentation. Secondly, based on the obtained superpixel image, the original color image is simplified efficiently and its histogram is computed easily by counting the number of pixels in each region of the superpixel image. Finally, we implement FCM with histogram parameter on the superpixel image to obtain the final segmentation result. Experiments performed on color images prove that proposed algorithm provides good results with less execution time.

Index Term– fuzzy c-means clustering (FCM), color image segmentation, SLIC superpixels.

1. INTRODUCTION:

Image segmentation is the process of partitioning a digital image into multiple segments. Image segmentation is a key step of object recognition and classification in computer vision. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse [1] . Although large number of algorithms used for image segmentation have been proposed, image segmentation remains one of the most challenging research topics because none of them is able to provide a unified framework for achieving fast and effective image segmentation. There are many applications of image segmentation such as medical imaging (MRI), object detection, recognition task, traffic control system etc. It is difficult to propose a general segmentation framework to achieve complex image segmentation tasks due to two reasons. The first one is that image segmentation is a multiple solution problem, i.e., there are multiple best segmentation results for one image. The second is that an image is always complex because of noise, background, low signal to-noise ratio, and intensity nonuniformity.

Machine learning algorithms for Image segmentation can be classified into two types – supervised and unsupervised learning. Supervised approaches can achieve image segmentation by learning the feature to train the models, but they require a lot of training data and label images, such approaches are convolutional neural network (CNN) [2] and fully convolution networks (FCN) [3]. Where as in unsupervised approaches, there are no training data and label images, such approaches are clustering [4], [5], GraphCut [6], active contour model [7], watershed transform (WT) [8], etc. are popular due to their simplicity. In this paper , we mainly use unsupervised technique for image segmentation.

In unsupervised learning, since clustering is useful for both low- and high-dimensional data. Here, we will use clustering for image segmentation by minimizing an objective function [9]. Fuzzy c-means clustering [FCM] also known as soft clustering and k-means clustering also known as soft clustering are same since, they both works by minimizing an objective function. The only difference being the introduction of a vector which expresses the percentage of belonging of a given point to each of the clusters. In FCM, each data point has a probability of belonging to each cluster where as in k-means each data point can only belong to exactly one cluster only. Though FCM is slower than K-means, it improves shortcomings of k-means by increasing the iterations count.

Fuzzy c-means (FCM) clustering was developed by J.C. Dunn in 1973 [10] and improved by J.C. Bezdek in 1981 [11]. Fuzzy c-means has been a very important tool for image segmentation for its effectiveness and simplicity.However, one disadvantage of standard FCM, it is very sensitive to noise and other imaging artifacts, since it does not consider any spatial information in image context. Pham [12] modified the FCM objective function by including a spatial penalty on the membership functions.

The penalty term leads to an iterative algorithm, which is very similar to the original FCM and allows the estimation of spatially smooth membership functions.Ahmed et al. [13] modified the objective function of FCM to compensate for the gray inhomogeneity and to allow the labeling of a pixel to be influenced by the labels in its immediate neighborhood, and they call the algorithm as FCM_S.This algorithm is very effective, but very time consuming, since it needs to compute the neighbourhood in each iteration.In order to reduce the computational loads of FCM_S, Chen and Zhang [14] proposed two variants, FCM_S1 and FCM_S2, which simplified the neighborhood term of the objective function of FCM_S. To replace the neighborhood term of FCM_S, these two algorithms computes the extra mean-filtered image and median-filtered image respectively in advance.

Cai, Chen, and Zhang [15] proposed Fast and robust fuzzy c-means (FGFCM) clustering algorithm that incorporates local information for image segmentation. . But in FGFCM, the clustering is performed on pixels but not the gray level histogram because it is difficult and complex to obtain the histogram of a color image.

Superpixels can capture redundancy in the image, and greatly reduce the complexity of subsequent image processing tasks. The superpixel is usually constructed by grouping similar pixels, and the methods for superpixel extraction can be broadly classified into two groups: graph based[16,17] and gradient-based solutions[18,19].

In order to solve the problem of high computational complexity, we propose a modified FCM algorithm based on superpixels obtained by SLIC [19] andbased on a superpixel image obtained by SLIC , we propose a simple color histogram computational method that can be used to achieve a fast FCM algorithm [20] for color image segmentation.Experiments show that the algorithm can achieve satisfying results quickly.

2. BACKGROUND:

Superpixel is nothing but finding groups of pixels having similar pixel values instead of working with every pixel in an image.For example, instead of looking at the amount of Red, Green, and Blue in each tiny dot, we can look at eyeballs, ears, wheels, and various little repeating things that have no names, but are generically labelled. Various superpixel methods such as mean-shift [21], simple linear iterative clustering (SLIC) [19], and WT [22], are usually considered as pre-segmentation algorithms for improving segmentation results generated by clustering algorithms. Here, we will use the SLIC methods to generate superpixels.

2.1 SLIC superpixels: The SLIC superpixel [19] method groups pixels into region based on the similar pixel values.

/* Initialization */

Initialize cluster centers $C_k=[l_k, a_k, b_k, x_k, y_k]^T$ at the interval S

set label L(i)=-1 for each pixel

set distance d(i)=∞ for each pixel

/* Assignment */

repeat

for each cluster C_k do

for each pixel i in a 2Sx2S region around C_k do

Compute the distance D between C_k and i

If D<d(i)

d(i)=D

L(i)=K

end if

end for

end for

/* Update*/

Compute new cluster centers.

Compute residual error E.

until $E \leq$ threshold.

Where,

N→Number of pixels in image

K→Amount of Superpixels

N/K→Average area of Superpixels

$S=\sqrt{N/k}$ Distance between cluster centers

$$D_{lab} = \sqrt{(lk-li)^2+(ak-ai)^2+(bk-bi)^2}$$

$$D_{xy} = \sqrt{(xk-xi)^2+(yk-yi)^2} \quad (1)$$

$D_S = D_{lab}+D_{xy}$

Here, pixel's color is represented in the CIELAB color space$[l,a,b]^T$ whose range of possible values is known. The pixel's position position $[x,y]^T$.$D_{lab}$is the Euclidean distance of color space between pixel and cluster. Similarly, $D_{xy}$is the Euclidean distance between position of pixel and cluster center and Ds is the 5D Euclidean distance between l_{abxy}color space.

2.2 Fuzzy c-means clustering: Fuzzy c-means (FCM) is a clustering method that allows each data point to belong to multiple clusters with varying degrees of membership is based on the minimization of the following objective function

$$J_m = \sum_{i=1}^{D}\sum_{j=1}^{N} \mu_{ij}^{m}\|x_i-c_j\|^2 \quad (2)$$

 ISSN: 1847-9790 || p-ISSN: 2395-0126

Where

- D is the number of image pixels.
- N is the number of clusters.
- m is fuzzy partition matrix exponent for controlling the degree of fuzzy overlap, with $m > 1$.
- x_i is the ith image pixel.
- c_i is the center of the jth cluster.
- μ_{ij} is the degree of membership of x_i in the jth cluster.For a given data point, x_i, the sum of the membership values for all clusters is one.

FCM performs the following steps during clustering:

1. Randomly initialize the cluster membership values, μ_{ij}.
2. Calculate the cluster centers:

$$c_j = \frac{\sum_{i=1}^{D} \mu_{ij}^m x_i}{\sum_{i=1}^{D} \mu_{ij}^m} \quad (3)$$

3. Update μ_{ij} according to the following:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{N} \left(\frac{\|x_i - c_j\|}{x_i - c_k}\right)^{\frac{2}{m-1}}} \quad (4)$$

4. Calculate the objective function, J_m.
5. Repeat steps 2–4 until J_m improves by less than a specified minimum threshold or until after a specified maximum number of iterations.

3.METHODOLOGY:

Superpixel has been increasingly used in the image processing field, since it can greatly reduce the complexity of post-processing tasks. However, superpixel needs to have several properties before it can be applied in our FCM algorithm. First, superpixel representation should not distort the image details, which means that the superpixel representation should have strong spatial adherence to the object boundaries. Second, superpixels should be compact and regular in shape, because our FCM algorithm will utilize neighborhood relationships among superpixels, and treat every superpixel as an atomic piece of some object with homogeneous intensity.

Normal FCM algorithms perform clustering operation on pixels which reduces its time complexity, while in our proposed algorithm we modified the objective function which help FCM algorithm to perform clustering operation on superpixels.

$$J_m = \sum_{i=1}^{Q} \sum_{j=1}^{N} \mu_{ij}^m \left\| s_i - c_j \right\|^2 \quad (5)$$

Where Q is the number of superpixels and s_i is the ith superpixel in image.

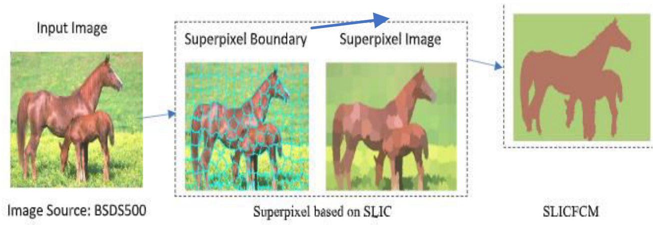
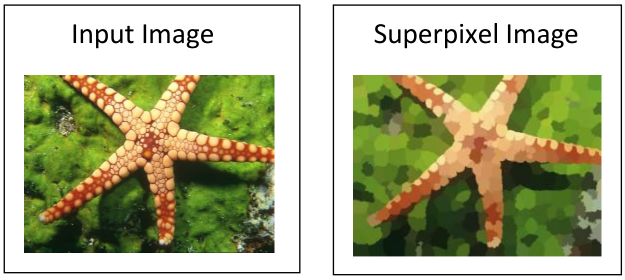


Fig.2. Framework of the proposed algorithm



(a) (b)

Fig.1. Superpixel generated by SLIC method

Here in fig.1 (b) is the superpixel image of the input image from BSDS500 generated by SLIC method.

SLIC-based Fast FCM: The proposed algorithm can be summarized as follows:

1. Initialize the values K ,N, η where K is the desired number of superpixels, and N is the Number of clusters and η is the convergence condition used for SLICFCM.
2. Apply SLIC method to obtain superpixel image
3. Initialize randomly the membership matrix U^0 according to the superpixel image.
4. Set the loop counter b=0.
5. Update the cluster center c_j.
6. Update the membership partition matrix U^b
7. If $\max(U^b - U^{(b+1)}) < \eta$ then stop, otherwise, set b=b+1 and go to Step 5

The compactness parameter of the SLIC algorithm controls the shape of superpixels. A higher value makes superpixels more regularly shaped, that is, a square. A lower value makes superpixels adhere to boundaries better, making them irregularly shaped. The allowed range is (0 Inf). Typical values for compactness are in the range [1,20].

The SLIC approach generates superpixels by clustering pixels based on their color similarity and proximity in the image plane.Itis fast (O(N) whereN is the number ofpixels in the image), allows control over the desired number of superpixels, and has good boundary adherence [19]. So, we select SLIC method to generate superpixels for our FCM algorithm.

We can see SLICFCM result from the fig.2. Firstly, SLIC method is applied to obtain the superpixel image which we can easily understand by looking at superpixel boundary in after that FCM is applied to obtain the segmentation result.

International Research Journal of Engineering and Management Studies (IRJEMS)
Volume: 03 Issue: 04 | April -2019 ISSN: 1847-9790 || p-ISSN: 2395-0126

4.RESULTS & EXPERIMENTS:

We perform the experiments on the color images from the Berkeley Segmentation Dataset and Benchmark (BSDS) [23]. The experiments are conducted on Hp Laptop with Intel (R) core (TM) i5-5200U CPU with 4GB ram.

A. *Comparative Algorithm:*

We used seven comparative algorithms based on clustering for color image segmentation to evaluate the efficiency and effectiveness of the proposed algorithm, these are FCM [9],FLICM [24], FCM_S1 [14], FCM_S2 [14], FGFCM [15], FGFCM_S1[25], FGFCM_S2 [25].

To evaluate the performance of the different clustering algorithms for color image segmentation, we used Jaccard Index [26] also known as intersection over union. The Jaccard coefficient measures similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets. If A is the segmented image and B is the ground truth image, then Jaccard index can be defined as

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

B. *Result on Color Images:*

We tested these comparative algorithms and the proposed SLICFCM ontwocolorimages from BSDS500 toshowtheirrobustness.

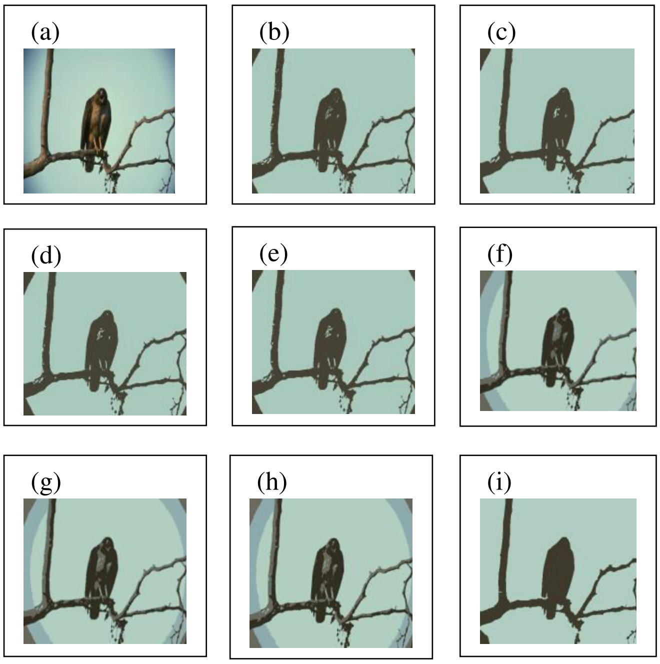


Fig.3. Comparison of segmentation results on color images from BSDS500 using different models (a)Original Image ("42049") (b)FCM result (c) FLICM result (d) FCM_S1 result (e) FCM_S2 result (f) FGFCM result (g) FGFCM_S1 result (h) FGFCM_S2 result (i) **SLICFCM** result.

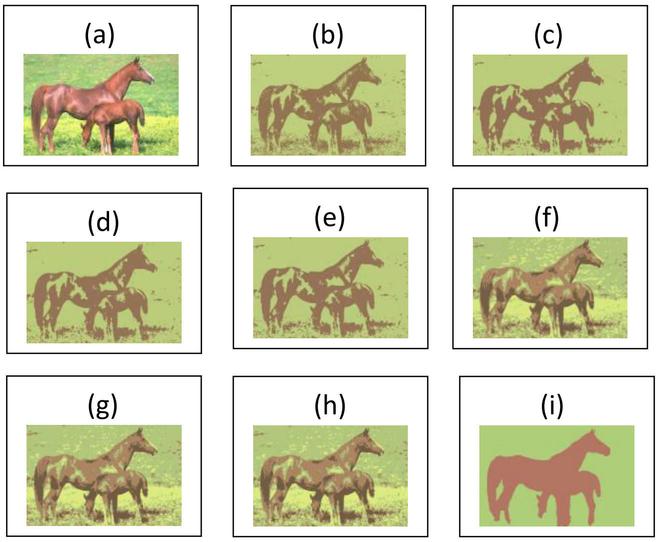


Fig.4. Comparison of segmentation results on color images from BSDS500 using different models (a)Original Image ("113044") (b)FCM result (c) FLICM result (d) FCM_S1 result (e) FCM_S2 result (f) FGFCM result (g) FGFCM_S1 result (h) FGFCM_S2 result (i) **SLICFCM** result.

In practical applications, since it is difficult to propose an algorithm to achieve the best segmentation result for every image in a dataset, researchers usually use the average result onallimagesinthedataset,e.g., BSDS andMSRC,toestimate the algorithm performance. So, we conducted experiments on theBSDS to demonstrate that the proposed SLICFCM is useful for real image segmentation. The BSDS is a popular benchmark that has been widely used by researchers for the task of image segmentation [23].

From Fig.4 and Fig.5 show segmentation of color images using different methods, among them SLICFCM gives the better segmentation result. Since it isdifficult to obtain local spatial information of color images, most of the improved FCM algorithms are only efficient for gray image segmentation.However, FCM is able to segment color image with a shorter time, as local spatial information is neglected in FCM.It is easy to extend FCM S1 and FCM S2 to color image segmentation because image filtering is performed on each channel of color images, respectively. Euclidean distance of pixels (3D vector) is employed in FLICM and FGFCM for color image segmentation, where the local spatial information is computed in each iteration of FLICM. Thus, FLICM has a very high computational complexity for color image segmentation. For FGFCM, the clustering is performed on pixels but not the gray level histogram because it is difficult and complex to obtain the histogram of a color image.

International Research Journal of Engineering and Management Studies (IRJEMS)
Volume: 03 Issue: 04 | April -2019 ISSN: 1847-9790 || p-ISSN: 2395-0126

C. *Execution Time:*

We computed the execution time of these comparative algorithms and proposed SLICFCM to show their efficiency. The execution times of these comparative algorithms and proposed SLICFCM are shown in Table. I where the best values are shown in bold.

TABLE I: Comparison of execution times (in seconds) of different algorithms on tested images.

Image	FCM	FCM_S1	FCM_S2	FGFCM	FGFCM_S1	FGFCM_S2	FLICM	SLICFCM
Fig. 3	2.16	66.15	58.81	2.87	2.41	2.66	585.35	**2.01**
Fig. 4	2.57	81.99	85.99	2.70	2.60	2.59	446.06	**1.98**
Fig. 6	1.65	65.77	64.77	2.60	2.54	2.55	541.06	**1.62**

TABLE II: Comparison of Segmentation Accuracy (SA%) (Jaccard index) of different algorithms on tested images.

Image	FCM	FCM_S1	FCM_S2	FGFCM	FGFCM_S1	FGFCM_S2	FLICM	SLICFCM
Fig. 3	95.38	95.36	95.30	95.38	95.13	95.31	95.36	**95.49**
Fig. 4	94.67	95.00	94.96	95.10	95.09	95.06	95.06	**96.32**
Fig. 6	96.85	96.85	96.81	96.58	96.57	96.57	97.06	**97.25**

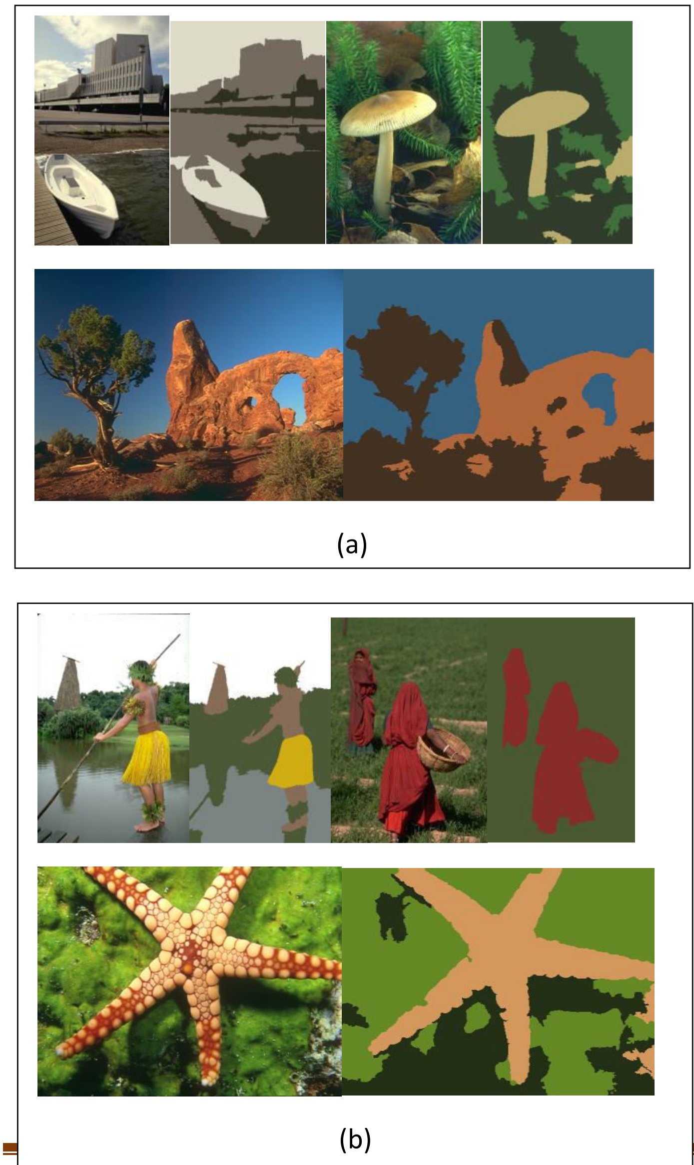


Fig.5. Segmentation results on color images from BSDS500 using SLICFCM methods.

D. Segmentation Accuracy:

We computed the Segmentation Accuracy based on Jaccard index of these comparative algorithms and proposed SLICFCM to show their efficiency. The SA values of these comparative algorithms and proposed SLICFCM are shown in Table. II where the best values are shown in bold.

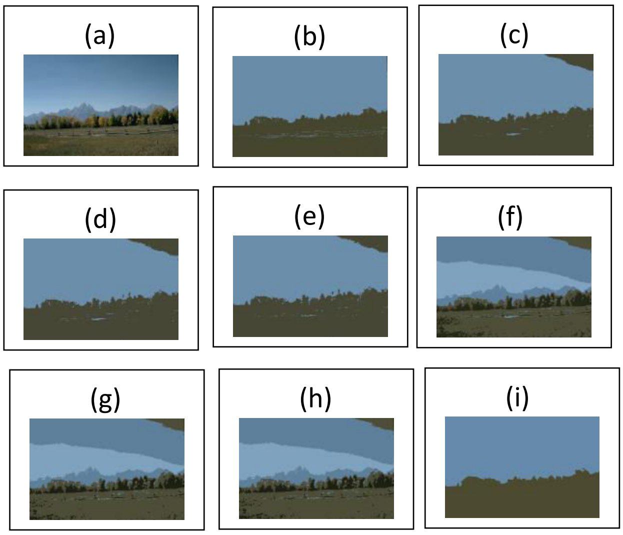


Fig.6. Comparison of segmentation results on color images from BSDS500 using different models (a)Original Image ("113044") (b)FCM result (c) FLICM result (d) FCM_S1 result (e) FCM_S2 result (f) FGFCM result (g) FGFCM_S1 result (h) FGFCM_S2 result (i) **SLICFCM** result.

5. CONCLUSION: In this paper, a modified FCM algorithm SLICFCM which utilizes superpixels as clustering objects has been proposed. The superpixelsimage obtained by SLIC reduces the computational complexity of our method drastically. The proposed SFFCM is tested on color images from BSDS500 dataset. The experimental results demonstrate that the proposed SLICFCM is superior to state-of-the-art clustering algorithms because it provides the best segmentation results and requires the shortest running time.

REFERENCES:

1. Linda G. Shapiro, George C. Stockman,"Computer Vision", pp. 279-325, New Jersey, Prentice-Hall, 2001.
2. A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Proc. Adv. Neural Inf. Proc. Syst. (NIPS), New York, NY, 2012, pp. 1097-1105
3. E. Shelhamer, J. Long and T. Darrell, "Fully convolutional networks for semantic segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 4, pp. 640-651, Apr. 2017.
4. B. Abu-Jamous, R. Fa, A. K. Nandi, Integrative Cluster Analysis in Bioinformatics. Hoboken, NJ, USA: Wiley, 2015.
5. S. Zeng, X. Wang, H. Cui, C. Zheng and D. Feng, "A unified collaborative multi-kernel fuzzy clustering for multiview data," IEEE Trans. Fuzzy Syst., vol. 26, no. 3, pp. 1671-1687, Jun. 2018.
6. J. Ma, S. Li, H. Qin and A. Hao, "Unsupervised multi-class cosegmentation via joint-cut over L1-manifold hyper-graph of discriminative image regions," IEEE Trans. Image Process., vol. 26, no. 3, pp. 1216-1230, Mar. 2017.
7. M. Gong, H. Li, X. Zhang, Q. Zhao and B. Wang, "Nonparametric statistical active contour based on inclusion degree of fuzzy sets," IEEE Trans. Fuzzy Syst., vol. 24, no. 5, pp. 1176-1192, Oct. 2016.
8. M. Bai and R. Urtasun, "Deep watershed transform for instance segmentation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, 2017, pp. 2858-2866.
9. N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," IEEE Trans. Fuzzy Syst., vol. 3, no. 3, pp. 370-379, Aug. 1995.
10. Dunn, J. C. (1973-01-01). "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters". Journal of Cybernetics. **3** (3): 32–57.
11. Bezdek, James C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. ISBN 0-306-40671-3.
12. D.L. Pham, Fuzzy clustering with spatial constraints, in IEEE Proc.Int. Conf. Image Processing, New York, (2002) II-65–II-68.
13. M.N. Ahmed, S.M. Yamany, N. Mohamed, A.A. Farag, T. Moriarty, A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data, IEEE Trans. Med. Imaging., 21(2002) 193–199.
14. S.C. Chen and D.Q. Zhang, Robust Image Segmentation Using FCM With Spatial Constraints Based on New Kernel-Induced Distance Measure, IEEE Trans. Syst.,Man, Cybern. B., 34(4) (2004) 1907-1916.
15. 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, Mar. 2007.
16. A. Ayvaci, and S. Soatto, "Motion segmentation with occlusionson the superpixelgraph", In ICCV, pp.727-734,2009.
17. O. Veksler, Y. Boykov, and P. Mehrani, "Superpixels and supervoxels in an energy optimization framework", In ECCV, pp. 211-224,2010.
18. D. Comaniciu, and P. Meer, "Mean shift: A robust approach toward feature space analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5), pp.603-619,2002
19. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods", IEEE Transactions on Pattern Analysis and MachineIntelligence,34(11),pp.2274-2281,2012.
20. 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," in *IEEE Transactions on Fuzzy Systems*. 10.1109/TFUZZ.2018.2889018
21. D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603-619, May 2002.
22. Z. Hu, Q. Zou and Q. Li, "Watershed superpixel," in Proc. IEEE Int. Conf. Image Process. (ICIP), Quebec City, QC, 2015, pp. 349-353.
23. 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.
24. S. Krinidis, and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm", IEEE Transactions on Image Processing,19(5), pp.1328-1337,2010.
25. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation Weiling Cai, Songcan Chen∗, Daoqiang ZhangPattern Recognition vol 40 (2007) 825 – 838.
26. Jaccard, P. (1901) Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. Bulletin de la Société Vaudoise des Sciences Naturelles 37, 241-272.