ISSN: 0976-3104
SPECIAL ISSUE (SCMDSA)
Hudedagaddi and Tripathy

THE IIOAB JOURNAL

ARTICLE OPEN ACCESS

APPLICATION OF SPATIAL IFCM IN LEUKAEMIA CELLS

Deepthi P Hudedagaddi* and Balakrushna Tripathy

SCOPE, VIT University, Vellore, Tamil Nadu-632014, INDIA

ABSTRACT

Managing nature of images induced by noise with the fusion of fuzzy algorithms has been a challenge. Traditional Fuzzy C Means (FCM) calculation is amazingly noise sensitive and fails in giving great results. Subsequently, as an answer for this issue Tripathy et. al presented a change of fuzzy c means which includes calculation that consolidated spatial data and intuition. The spatial capacity is the summation of the considerable number of estimations of the participation elements of the pixel's neighborhood which is under consideration. This methodology ended up being a right answer for the noise affectability issues when contrasted with FCM. This calculation incorporates an intuitionistic component in the enrollment capacity of the current spatial FCM (sFCM). Intuitionism manages the wavering segment that emerges because of less data and insufficient learning. This is superior to anything existing calculations. This spatial IFCM(sIFCM) has been connected on leukaemia images.

Received on: 09th-Sep-2015
Revised on: 27th-Sep-2015
Accepted on: 25th– March-2016
Published on: 16th–May-2016

KEY WORDS

Fuzzy C Means; Intuitionistic Fuzzy set; Clustering; Spatial Fuzzy C Means, Intuitionistic Fuzzy C Means

***Corresponding author: Email:** deepthiph@gmail.com **Tel:** +91-9986387435

INTRODUCTION

Analysis of images is a noteworthy part for diagnosis and treatment of diseases, research studies and more. At present, the calculation is assuming a vital part because of the expanding size and number of medicinal images. The choice of strategies fundamentally relies on imaging modalities, its particular application and different variables. For example, the brain tissue has distinctive necessities from other organs. Restorative image segmentation computerizes the particular radiological capacity. The objective of segmentation is to investigate the representation of an image into important and simpler divisions. It alludes to post a computerized image into various sections which are fundamentally developed with sets of pixels. Every pixel in the locale of interest comprises of some essential qualities and processed property, known as power, surface and shading. Yet, no single division strategy yields satisfactory results for each therapeutic image.

In image understanding and vision based machine intelligence, image segmentation plays an important role by dividing an image into multiple homogeneous segments that are more suitable to further analysis. Among different image segmentation methods, the most popular and intensively explored ones are the approaches using data clustering algorithms. The targets of maximizing the similarities of pixels in each segment and minimizing the similarities of pixels from different segments in image segmentation problems also constitute exactly the requirement of traditional data clustering if we regard the pixels as the data points to be clustered. However, the visual similarities between pixels in an image and their calculation need to be well designed. Partitioning pixels into spatial homogeneous regions should be considered in the clustering algorithms.

Spatial data mining (SDM), or learning revelation in spatial database, alludes to the extraction of understood information, spatial relations, or different examples not expressly put away in spatial databases. SDM comprises of removing information, spatial connections and some other properties which are not unequivocally put away in the database. SDM is utilized to discover verifiable regularities, relations between spatial information and/or non-spatial information. We can consequently see the immense significance of spatial connections in the examination process. Worldly perspectives for spatial information are likewise a main issue, however they are infrequently considered. Information mining strategies are not suited to spatial information since they do not bolster area information nor the certain connections between items. Thus, it is important to grow new strategies including

THE IIOAB JOURNAL
ISSN: 0976-3104

spatial connections and spatial information taking care of. Computing these spatial connections is tedious, and an immense volume of information is created by encoding geometric area. Spatial information mining incorporates different errands and, for every undertaking, various distinctive techniques are frequently accessible, whether computational, factual, visual, or some blend of them. Conventional FCM algorithm fails to provide appropriate results on images in the presence of noise. Spatial FCM (sFCM) and spatial IFCM(sIFCM) [1,2], a bi-step process, incorporates spatial information of the pixel in consideration. Though it fails in completely removing the distortion of noise, the algorithm proves to be less sensitive to noise as compared to traditional FCM. However,it does not eliminate the distortion caused by noise completely. In addition to the uncertain based clustering methods, hybrid clustering algorithms have been developed [3-7].

SPATIAL CLUSTERING

Cluster investigation is generally utilized for information examination, which arranges an arrangement of information things into gatherings (or groups) so that things in the same gathering are like each other and not quite the same as those in other groups. Many diverse grouping techniques have been produced in different exploration fields, for example, insights, design acknowledgment, information mining, machine learning, and spatial analysis [5].

Clustering is a noteworthy unsupervised learning procedure. Fuzzy C-Means clustering is an understood delicate division technique and it suitable for therapeutic image segmentation. In any case, this customary calculation is computed by iteratively minimizing the separation between the pixels and to the bunch focuses. Spatial relationship of neighboring pixel is a guide for segmentation of images. These neighboring pixels are exceedingly related the same element information. In spatial space the enrollment of the neighbor focused are indicated to get the group dissemination insights. In view of this insights to figure the weighting work and connected into the enrollment function [6].

Spatial clustering techniques can be apportioning or hierarchical, density-based, or framework based. Regionalization is a remarkable kind of spatial clustering that tries to total spatial things into spatially contiguous gatherings (i.e., areas) while redesigning an objective limit. Various geographic applications, for instance, climate zoning, scene examination, remote recognizing picture division, frequently require that groups are geographically circumscribing. Existing regionalization frameworks that rely on a clustering thought can be planned into three events: (1) multivariate (non-spatial) gathering took after by spatial planning to amend clusters into regions (2) gathering with a spatially weighted uniqueness measure, which considers spatial properties as a variable in confining bunches and (3) contiguity constrained grouping that approves spatial contiguity in the midst of the gathering strategy[8].

EXISTING METHODS

Fuzzy models are incorporated in analyzing spatial data.

Fuzzy Clustering

Fuzzy c-mean in view of fuzzy sets [9] is an algorithm proposed by James C. Bezdek [10,11]. In fuzzy clustering (additionally alluded to as soft clustering), information components can fit in with more than one cluster, and connected with every component is an arrangement of enrollment levels. These demonstrate the quality of the relationship between that information component and a specific cluster. Fuzzy clustering is a procedure of relegating these enrollment levels, and after that utilizing them to allocate information components to one or more clusters [12,13].

1. Assign initial cluster centers or means for c clusters.
2. Calculate the distance d_{ik} between data objects x_k and centroids v_i using Euclidean formula

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} \tag{1}$$

3. Generate the fuzzy partition matrix or membership matrix U:

If $d_{ij} > 0$ then

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}} \tag{2}$$

Else

THE IIOAB JOURNAL ISSN: 0976-3104

$\mu_{ik} = 1$

4. The cluster centroids are calculated using the formula

$$V_i = \frac{\sum_{j=1}^{N} (\mu_{ij})^m x_j}{\sum_{j=1}^{N} (\mu_{ij})^m} \qquad (3)$$

5. Calculate new partition matrix by using step 2 and 3
6. If $\|U^{(r)} - U^{(r+1)}\| < \varepsilon$then stop else repeat from step 4.

Usually, for all experimental purpose, m is considered to be 2 and ε to be 0.02.

Spatial Fuzzy C Means

Chuang. et al. [1] clarified that a traditional FCM algorithm does not completely use the spatial data in the image. They altered the present fuzzy c-means (FCM) calculation and created FCM that joins spatial data into the participation capacity for clustering. The spatial function is given by the summation of the membership values in the neighboring of every pixel under consideration. The benefits of the new strategy are: (1) it yields more homogeneous clusters than those of other strategies, (2) it diminishes the spurious blobs, (3) it uproots boisterous spots, and (4) it is less sensitive to noise than other systems. This procedure is an intense strategy for image segmentation and works for both single and several information on spatial data. On comparable lines, spatial IFCM was additionally developed by Tripathy et.al by presenting the intuitionsitic feature.

It implies when two close pixels are considered, relation between them is generally high. Since the neighboring pixels offer comparative force, the likelihood of them gathering into a same group is comparatively high. The spatial FCM calculation exploits this criteria. A spatial function is characterized as

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik} \qquad (4)$$

where $NB(x_j)$ refers to the neighborhood pixels of x_j. A 5x5 equally weighted mask centered on pixel x_j has been used for this purpose. The spatial function h_{ij} represents the degree of likeliness that x_j is in the i^{th} cluster. The value of the spatial function for a pixel is high for a particular cluster if most of the neighborhood pixels belong to the same cluster. It is included in the membership function as:

$$u'_{ij} = \frac{u_{ij}^{p} h_{ij}^{q}}{\sum^{c} u^{p} h^{q}}]) \qquad (5)$$

here p and q indicate the relative weightage of the initial membership and the spatial function respectively. In case of a noisy image, the spatial function reduces the number of misclassified pixels by taking the neighboring pixels into account.

The sFCM clustering algorithm has two-steps. For every iteration, the conventional FCM algorithm is followed in the first step. Here the distance formula being used is radial-based kernel distance. Later, the centroid and the membership functions are updated. In the second step, the spatial function h_{ij} is calculated and then the new membership function (5) is computed.

Intuitionsitic FCM

The Intuitionistic fuzzy c-means proposed by T. Chaira [14,15] brings in to account a new parameter that helps in increasing the accuracy of clustering. This parameter is known as the hesitation value.

1. Assign initial cluster centres or means for c clusters.
2. Calculate the distance d_{ik} between data objects x_k and centroids v_i using Euclidean formula(1).
3. Generate the fuzzy partition matrix or membership matrix U:

If $d_{ij}>0$ then compute μ_{ik}

Else

$\mu_{ik}=1$

1. Compute the hesitation matrix π

THE IIOAB JOURNAL ISSN: 0976-3104

2. Compute the modified membership matrix U' using
   $\mu'_{ik} = \mu_{ik} + \pi_{ik}$

3. The cluster centroids are calculated using the formula
   $$V_i = \frac{\sum_{j=1}^{N} (\mu'_{ij})^m x_j}{\sum_{j=1}^{N} (\mu'_{ij})^m}$$

4. Calculate new partition matrix by using step 2 to 5

If $\|U^{(r)} - U^{(r+1)}\| < \varepsilon$ then stop else repeat from step 4. ε is considered to be 0.02.

.

Spatial IFCM

The spatial Intuitionistic Fuzzy C means (sIFCM) algorithm developed by Tripathy et.al[2] is an extended version of IFCM where the membership function incorporates the spatial function.

The algorithm of spatial IFCM(sIFCM) is as follows.

1. Provide the initial values for the centroids v_i where i =1,…,c

2. Compute the membership function as follows:
   $$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{\frac{2}{m-1}}}$$
   for all i = 1,…,c and j = 1,…,N

3. Compute the hesitation value as:
   $\pi_{ij}(x) = 1 - u_{ij}(x) - (1 - u_{ij}(x))/(1 + \lambda u_{ij}(x))$
   for all i = 1,…c , λ>0 and j = 1,…,N

4. Compute the membership function as:
   $\mu'_{ik} = \mu_{ik} + \pi_{ik}$
   for all i = 1,…c and j = 1,…,N

5. Calculate the spatial function as
   $$h_{ij} = \sum_{k \in NB(x_j)} u_{ik}$$
   for all i = 1,…c and j = 1,…,N

6. Compute the new membership function which incorporates the spatial function as:
   $$u'_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^{c} u_{kj}^p h_{kj}^q}$$

7. Set $u_{ij} = u_{ij}''$ for all j = 1,…,N and i = 1,…c

8. Calculate the new centroids as follows:
   $$v_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m}$$
   for i=1,..,c

9. If $|u_{ij}(new) - u_{ij}(old)| < \epsilon$ then stop, otherwise go to step 2.

THE IIOAB JOURNAL
ISSN: 0976-3104

Usually m=2 and $\in$ =0.02.

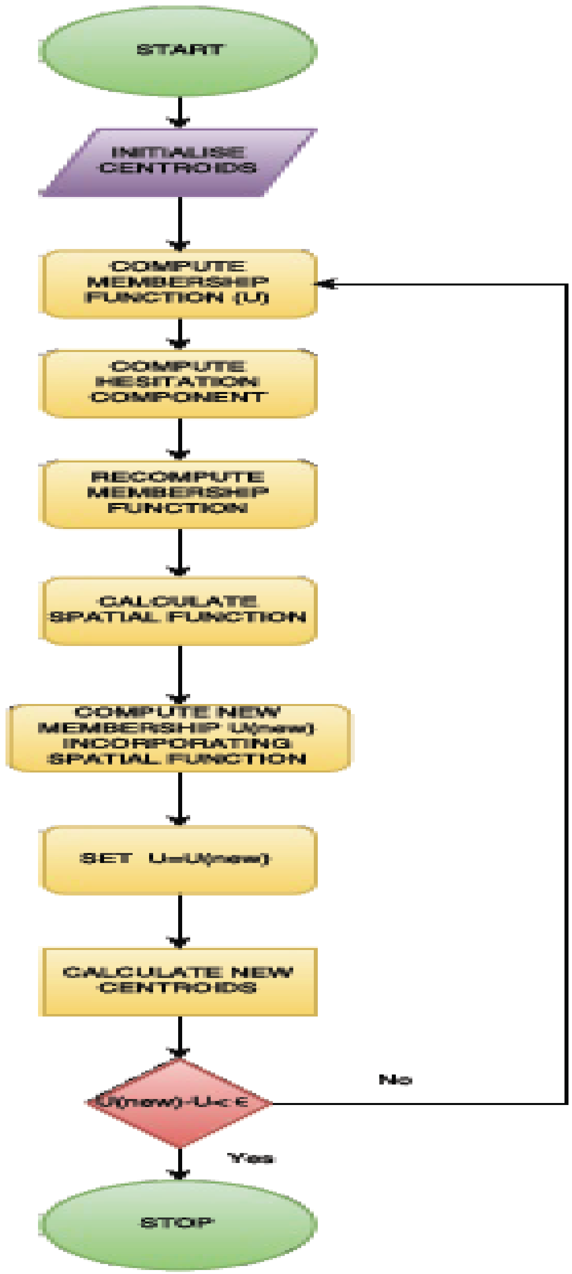

Fig: 1. Spatial IFCM

sIFCM ON LEUKAEMIA IMAGE

It is a well established fact that early detection of cancer cells and estimating their rate of growth plays a crucial role in detection and treatment of cancer/leukaemia. In this paper, we have considered leukaemia cells for the study. Two types of cluster validity functions, fuzzy partition and feature structure, are often used to evaluate the performance of clustering in different clustering methods. The representative functions for the fuzzy partition are partition coefficient V_{pc}[2, 9] and partition entropy V_{pe}[10].

THE IIOAB JOURNAL
ISSN: 0976-3104

The idea of these validity functions is that the partition with less fuzziness means better performance. Hence, the best clustering is achieved with maximum value of V_{pc} and minimum value of V_{pe}. Disadvantages of V_{pc} and V_{pe} are that they measure only the fuzzy partition and lack a direct connection to the featuring property. Other validity functions based on the feature structure are available.

$$V_{pc} = \frac{\sum_j^N \sum_i^c u_{ij}^2}{N} \qquad (6)$$

and

$$V_{pe} = -\frac{\sum_j^N \sum_i^c [u_{ij} \log u_{ij}]}{N} \qquad (7)$$

Clustering is said to be efficient if sample in a cluster are extremely. A good clustering result generates samples that are compacted within one cluster and samples that are separated between different clusters. Minimizing V_{xb} is expected to lead to a good clustering.

$$V_{xb} = \frac{-\sum_j^N \sum_i^c u_{ij} \vee x_j - v_i \vee^2}{N * \left(min_{i \neq k} \left\{ \| v_k - v_i \|^2 \right\} \right)} \qquad (8)$$

DB and D indices indicate the proximity of clusters within and in between. Hence, larger D value and lower DB value indicates a good clustering[16].

As a follow on with the development of spatial fuzzy c means algorithm, Deepthi et. al[17] applied the developed sFCM on leukaemia images to demonstrate it's working. This clustering of leukaemia cells helps in providing information regarding the growth of cancerous cells. This further helps in diagnosis.

Table: 1. Performance indices of sIFCM on leukaemia image

Index	FCM	IFCM	$sIFCM_{1,1}$	$sIFCM_{1,2}$	$sIFCM_{2,1}$
V_{pc}	0.2118	0.2074	0.4762	0.2241	0.2281
V_{pe}	0.0178	0.02585	3.56E-005	1.68E-005	5.90E-006
V_{xb}	0.0168	0.02096	0.05929	0.0069	0.0063
DB	0.467	0.4906	0.4126	0.4208	0.4188
D	2.2254	2.0529	3.5866	3.0203	2.9655

The results in **Table 1** show that the sIFCM succeeds in providing better results than conventional FCM. It is distinctly visible through higher D value and lower DB values, different combinations of p and q values in sIFCM overpower the conventional FCM and IFCM. The partition entropies are also proved to be lower. However, the partition coefficient values are higher in sIFCM. As per the cluster validity measures, V_{xb} should be minimum to prove it the algorithm is providing better clustering. In this measure also, the application of sIFCM is better and hence proves the clusters provided are efficient.

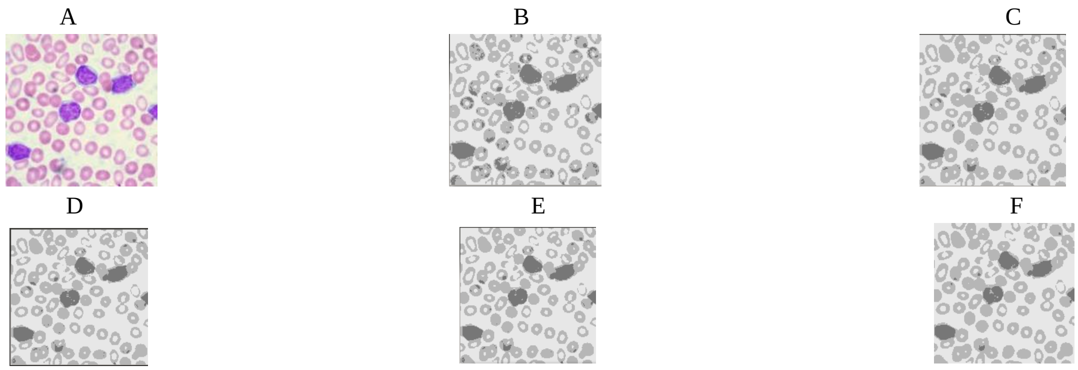

Fig: 2. (a) Original image, Segmented images of leukaemia using (b)FCM; (c)IFCM (d)sIFCM$_{2,1}$; (e)sIFCM$_{1,1}$; (f)sIFCM$_{1,2}$

The segmented images with the application of sIFCM provide better clarity and understanding of presence of leukemia cells than that of conventional FCM and IFCM. As an extension, the sIFCM as shown above, is applied to leukaemia image. The leukaemia cells are clustered based on sIFCM and the images are shown. On trial and error method, different values were provided for p and q values in spatial membership function. The results of three combinations are given. However, the appropriate result depends on the image and the application.

CONCLUSION AND FUTURE WORKS

The drawbacks in one clustering algorithm paves a way for developing novel algorithms or modifications for the same algorithm. However, it can be seen that if developing algorithm for one application is one arena, finding the applications of same algorithm in various other domains is another arena. This paper has made an attempt to apply the spatial IFCM algorithm which was developed incorporating the drawbacks of IFCM to a cancer cell images. In this manner, algorithms developed should also be reused for several other applications in different domains. The results obtained are also better than conventional algorithms. This application would help in diagnosis or estimating the growth and effect of these cancerous cells. Also, spatial data clustering has provided a wide platform for researchers. The developed algorithms are to be exploited completely and various applications in different domains are to be found.

CONFLICT OF INTEREST

Authors declare no conflict of interest.

ACKNOWLEDGEMENT

None

FINANCIAL DISCLOSURE

No financial support was received to carry out this project.

REFERENCES

[1] KS Chuang, HLTzeng, S Cheren, Jay Wu,Tzong-Jer Chen.[2006] Fuzzy c-means clustering with spatial information for image segmentation *Computerized Medical Imaging and Graphics*, vol. 30,

[2] Tripathy BK, AvikBasu,and SahilGovel.[2014] Image segmentation using spatial intuitionistic fuzzy C means clustering. Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on. *IEEE*

[3] B.K.Tripathy and A.Ghosh.[2013] Data Clustering Algorithmsusing Rough sets, Handbook of research on Computational Intelligence for engineering, Science and Business, *IGI Global Publications*, pp.297-327.

[4] BK.Tripathy, A Ghosh, GK Panda.[2012] Kernel Based K-Means Clustering Using Rough Set, Proceedings of 2012 International Conference on Computer Communication and Informatics (ICCCI -2012), Jan. 10 – 12, (2012), Coimbatore, INDIA, pp.1 -5.

[5] BK Tripathy, A Tripathy, K Govindarajulu, R Bhargav.[2014] On Kernel Based Rough Intuitionistic Fuzzy C-means Algorithm and a Comparative Analysis. In Advanced Computing, Networking and Informatics, *Springer International Publishing*. 1: 349-359.

[6] BK Tripathy and P Swarnalatha.[2014] A Comparative Study of RIFCM with Other Related Algorithms from Their Suitability in Analysis of Satellite Images using Other Supporting Techniques, Kybernetes, .43(1): 53-81.

[7] B.K.Tripathy and R. Bhargav: Kernel Based Rough-Fuzzy C-Means, PReMI, ISI Calcutta, December, LNCS 8251, (2013), pp.148-157.

[8] DianshengGuo, Jeremy Mennis.[2009] Spatial data mining and geographic knowledge discovery—An introduction", Computers, Environment and Urban Systemspp.403-408

[9] LA Zadeh,[196] *Fuzzy Sets, Information and Control*, .8(11): 338 – 353.

[10] Bezdek JC. Cluster validity with fuzzy sets. *J Cybern* 1974;3:58–73.

[11] Bezdek JC.[1975] Mathematical models for systematic and taxonomy. In:proceedings of eigth international conference on numerical taxonomy,San Francisco;, p. 143–66.

[12] P Swarnalatha, BK Tripathy, PL Nithin and D Ghosh.[2014] Cluster Analysis Using Hybrid Soft Computing Techniques, CNC- 2014 International Conference of Network and Power Engineering ,Proceedings of Fifth CNC-2014,pp. 516-524.

[13] R Bhargav, BK Tripathy, A Tripathy, R Dhull, E Verma, and P Swarnalatha. [2013]Rough Intuitionistic Fuzzy C-Means Algorithm and a Comparative Analysis, *ACM conference, Compute* 2013, ACM 978-1-4503-2545-5/13/08.

[14] Chaira, T. (2011). A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images. Applied Soft Computing, 11(2), 1711-1717.

[15] KT Atanassov. [1986] Intuitionistic Fuzzy Sets, Fuzzy sets and Systems, 20(1):87-96.

[16] JC Dunn. [1973]A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, pp. 32-57.

[17] Deepthi P Hudedagaddi,B.K. Tripathy, Application of Spatial FCM in Cancer Detection Cells, Proceedings of NCMCS-2015,Upendranath College of Engineering,India.

ABOUT AUTHORS

Deepthi P Hudedagaddi is pursuing her Masters at Vellore Institue of Technology. She is working fuzzy clustering techniques on spatial data.

Dr. Tripathy, a triple gold medalist, is a senior professor in VIT University. He has supervised 19 PhD s, 13 M. Phil s and 02 M.S degrees. He is a senior member of IEEE, ACM, ACEEE and CSI. and is associated with over 60 international journals,published 320 articles, two research volumes and two books.He is working on Rough sets, Fuzzy sets, Social networks, Data mining, Soft Computing, E-Learning, Granular computing, Multi criteria decision making, Neighbourhood systems, SIoT and Soft Sets.

COMPUTER SCIENCE