Volume 9, No. 1, Jan. - 2018 (Special Issue)
International Journal of Mathematical Archive-9(1), 2018, 169-175
Available online through www.ijma.info ISSN 2229 – 5046

IMAGE SEGMENTATION USING PARTICLE SWARM OPTIMIZATION BASED INTUITIONISTIC FUZZY CLUSTERING

S. PARVATHAVARTHINI*[1]
[1]Kongu Engineering College, Perundurai, Tamilnadu, 638060, India.

N. KARTHIKEYANI VISALAKSHI[2]
[2]NKR Government Arts College for Women, Namakkal, Tamil Nadu, 637001, India.

S. SHANTHI[3]
[3]Kongu Engineering College, Perundurai, Tamil Nadu, 638060, India.

J. MADHANMOHAN[4]
[4]Erode Sengunthar Engineering College, Perundurai, Tamil Nadu, 638057, India.

E-mail: [1]varthinis@gmail.com, [2]karthichitru@yahoo.co.in, [3]shanthis@kongu.ac.in and [4]madhan79in@gmail.com

ABSTRACT

Image segmentation leads to the identification of several patterns from an image. The main aim of segmentation is to remove the noise and fuzziness from the image and identify certain regions of interest by grouping the pixels into several clusters based on their features. This study presents a novel method for segmenting satellite images using Particle Swarm optimization and Intuitionistic fuzzy clustering algorithm by extracting several statistical features from the image. The performance of the proposed methodology is evaluated over images from QuickBird satellite and the regions are clustered based on the spectral contrast. The optimal number of clusters is evaluated using silhouette coefficient.

Keywords *– Intuitionistic Fuzzy C Means, Particle Swarm Optimization, Image Segmentation, Satellite images, Entropy, Silhouette Index.*

1. INTRODUCTION

In recent days, the necessity for data mining has grown to a great extent. Data mining leads to the invention of patterns and relationships among data. Clustering plays a prominent role in grouping data into clusters based on some distance measures. The usefulness of clustering is observed in several applications like image segmentation, categorization and retrieval.

Image segmentation has become the need of the hour due to its wide variety of applications in the field of medical diagnosis, object detection, pattern recognition, bioinformatics, satellite image analysis, business, geology, etc. Nowadays, all these fields contain myriad of data that is difficult to analyze using man power. So there is a need for segregating data into meaningful non-overlapping parts and extract the specific region of interest from the vast set of images.

The first and most common type of clustering algorithm is K-Means which is a hard clustering algorithm. A hard clustering algorithm allocates an object to exactly one cluster. On the other hand, a soft clustering algorithm like Fuzzy C Means [4] allots an object to various clusters based on their membership values. Atanassov [2] introduced another notion of fuzzy set called Intuitionistic Fuzzy Set (IFS) that takes both membership and non-membership values into account. It also introduces a third parameter called hesitancy factor denoting the inability of user in determining the belongingness or the presence of noise in the data.

CONFERENCE PAPER
International Conference dated 08-10 Jan. 2018, on Intuitionistic Fuzzy Sets and Systems (ICIFSS - 2018), Organized by Vellalar College for Women (Autonomous), Erode, Tamil Nadu, India.

The medical images may belong to different modalities and there is a high possibility of noise or misguidance due to the poor illumination. So Intuitionistic Fuzzy (IF) clustering has the utmost ability to deal with this uncertainty and vagueness. It considers both the membership and non-membership degree for computation and the third parameter called hesitancy degree paves way to represent the user's ignorance or the uncertainty in deciding whether or not an object belongs to a cluster. It represents the state of 'may be' or the unknown state.

Particle Swarm Optimization (PSO) [15] is a well-known optimization technique which simulates the stochastic behavior of bird flocking. All the clustering algorithms select the initial set of seed points in a random manner. But this leads to local convergence and the results get stuck in local optimal values. To avoid this, an optimization algorithm can be employed to select the optimal initial centers that can be passed as input to the clustering algorithm. This work introduces a novel algorithm by combining Tchaira method of Intuitionistic Fuzzy C Means (IFCM) algorithm along with PSO to achieve better results than the existing methods and thrives for global optimum.

In order to enhance the quality of image segmentation, our contributions include

- Developing a highly scalable algorithm by hybridizing IFCM with PSO
- Considering five statistical measures along with the pixel values as the input for IFCM clustering algorithm.
- Ability to deal with satellite images that have a high possibility of noise and produce efficient segments.

2. RELATED WORKS

There are several works carried out in the area of image processing by utilizing the Intuitionistic fuzzy clustering. An Intuitionistic fuzzy approach for detecting tumor/hemorrhage in medical images is proposed by Chaira et al [8]. The IF representation of image is created and it is clustered using IFCM. The resulting image is histogram thresholded to remove unwanted pixels and finally edges are detected for finding the tumor region.

An image is represented as several fuzzy sets with the membership functions for symbolizing the foreground and background and then converted to IFS by Ananthi *et al.* [1]. The gray scale images are segmented using IFS. The entropy is calculated to find the threshold. The value that minimizes the entropy is taken as the threshold for segmenting the image. Chaira [7] made use of type II fuzzy set for enhancing medical images. This is achieved by constructing an IFS using Hamacher T conorm resulting in a clearer and brighter image. Ching-Wen-Huang [11] improved Chaira's method by adding the neighborhood attraction. This method used the membership of a pixel and its neighbors to modify the membership value and finally applied GA to improve the performance of image segmentation algorithm.

T Chaira modified the objective function of IFCM to include the IF entropy and utilized this to minimize the entropy of the histogram of an image. The IFS is created using Yager generating function. Tchaira [6] proposed an IFS approach for color region extraction from remote sensing images. The image is represented as IFS using the Sugeno generating function. The hesitation degree is considered as the uncertainty in the image.

A new clustering algorithm which considered the car data set is built by Xu & Wu [26] for clustering both IFS and Interval-valued IFS. This proved to be more efficient with numerical datasets. A robust IFCM and kernel version of IFCM is presented by kaur *et al.* [15] with a new distance metric incorporating the distance variation of data-points within each cluster. Bhargava *et al.* [5] hybridized rough set with IFS in order to describe a cluster by its centroid and its lower and upper approximations. The method introduces modified Rough FCM with the membership of IFCM. Shanthi, & Bhaskaran. [24] processed a set of mammogram images to detect and classify breast cancer by finding the region of interest and separating the affected part.

An artificial bee colony algorithm is designed by Naser *et al.*, [17] for effectively grouping the social networks by collecting people with common interests. Balasubramaniam *et al.* [3] segmented nutrition deficiency in incomplete crop images using IFCM. The missing pixels in the incomplete images are imputed using IFCM algorithm. The resulting membership matrix efficiently portrayed the deficiency region of the crop. Izakian *et al.* [14] utilized Fuzzy C Means and Particle Swarm Optimization to cluster moving objects or trajectories. Data is represented using discrete cosine transform. Salmeron *et al.* [22] adapted Hebbian-based FCM learning and hybridized with PSO to calculate the severity of arthritis.

Saxena *et al.* [23] reviewed the different methods for clustering, along with the measures for finding similarity the evaluation criteria and discussed the applications of clustering in various domains. Hein *et al.* [10] developed a fuzzy particle swarm reinforcement learning to create self-organized fuzzy controllers and applied it to benchmark datasets.

Nobile *et al.* [18] proposed Fuzzy self tuning PSO to control the parameters of PSO using fuzzy logic. The efficiency of the algorithm is proved by testing it against twelve benchmark functions.

Oliveira *et al.* [25] presented a homogeneous cluster ensemble based on particle swarm clustering algorithm. Initially, many base partitions are taken from the data and they are given as input to the consensus function and genetic selection operators are used to decide the final partition. Izakian *et al.* [13] combined fuzzy PSO with FCM to minimize the objective function leading to a global solution.

Parvathavarthini *et al.* [19] combined IFCM with cuckoo search and applied it to several real time datasets for validating the cluster structure using various cluster indices. Parvathavarthini *et al.* [20] hybridized IFCM with crow search optimization producing very low error rates for real time datasets. In general, PSO is found to be efficient for both real time and image datasets.

3. BACKGROUND

3.1 Fuzzy Set And Intuitionistic Fuzzy Set

Fuzzy sets are designed to manipulate data and information possessing non-statistical uncertainties. A fuzzy set is represented by Zadeh [25] as follows

$$FS = \{< x, \mu_{FS}(x) > | \, x \in X\}$$

where $\mu_{FS} : X \rightarrow [0, 1]$ and $\nu_{FS} : X \rightarrow [0, 1]$ and $\nu_{FS}(x)=1 - \mu_{FS}(x)$. Here μ_{FS} is the membership value and ν_{FS} is the non-membership value.

An Intuitionistic Fuzzy Set proposed by Atanassov [2] can be symbolized as below

$$IFS = \{< x, \mu_{IF}(x), \nu_{IF}(x) > | \, x \in X\}$$

where $\mu_{IF} : X \rightarrow [0, 1]$ and $\nu_{IF} : X \rightarrow [0, 1]$ define the degree of membership and non-membership, respectively and

$$\pi_{IF}(x) = 1 - \mu_{IF}(x) - \nu_{IF}(x) \quad \text{such that } 0< \mu_{IF}(x) + \nu_{IF}(x)<1$$

where π_{IF} is the hesitancy value used to represent the uncertainty.

3.2 Particle Swarm Optimization

Particle swarm optimization (PSO) [16] is a population-based stochastic optimization technique inspired by bird flocking and fish schooling which is based on iterations/generations. Each particle has an initial position and it moves towards a better position with a velocity. The positions represent the solutions for the problem. Initially, the position and velocity matrices are assigned random values.

Consider the population or swarm size as m and the particle dimension as n. Let Velocity be represented as $Velo_i =\{v_1, v_2, \ldots, v_n\}$ and position be represented as $Xpos_i =\{x_1, x_2, \ldots, x_n\}$ where i= 1 to n. For every iteration, these two vectors are updated.

The fitness is evaluated by calculating the objective function for each particle in the swarm. The individual best performance is termed as p_{best} and it is updated by comparing fitness values of each iteration with that of the previous iteration. The overall best position attained by any particle with the overall minimum fitness (in case of minimization problems like clustering) is chosen as the g_{best}. The inspiring feature of PSO is that it exempts the possibility of the solution getting stuck in the local optima and tries to reach the global optima by converging in less number of iterations.

3.3 Intuitionistic Fuzzy Clustering

The first task for IFCM algorithm [9] is to convert crisp data into fuzzy data which in turn would be converted to Intuitionistic fuzzy data.

The intuitionistic fuzzification converts the intermediate fuzzy dataset to intuitionistic fuzzy dataset. The hesitancy factor can be found by adding the membership and non-membership degrees and subtracting the sum from one.

The IFCM clustering method proposed by Tchaira [9] contains an objective function with two terms. Initially, the Intuitionistic fuzzy Euclidean distance is calculated and the membership values are found using the following formula

$$U_{ij} = \frac{1}{\sum_{r=1}^{C}\left(\frac{dis(d_j^{'}, v_i)}{dis(d_j^{'}, v_r)}\right)^{\frac{2}{m-1}}}, 1 \le i \le C, 1 \le j \le n, m = 2 \tag{1}$$

CONFERENCE PAPER
International Conference dated 08-10 Jan. 2018, on Intuitionistic Fuzzy Sets and Systems (ICIFSS - 2018), Organized by Vellalar College for Women (Autonomous), Erode, Tamil Nadu, India.

The centroids are updated using the following formula

$$V_i = \frac{\sum_{k=1}^{n} u_{ik}^{m} x_k}{\sum_{k=1}^{n} u_{ik}^{m}} \text{ where } 1 \le i \le C \quad (2)$$

Now, the hesitation degree is calculated and new membership values are calculated using

$$U_{ik}^{*} = U_{ik} + \pi_{ik}. \quad (3)$$

Then the cluster centers are updated using this new membership values and thus the first term of objective function is given by

$$J_m(x, y) = \sum_{i=1}^{c}\sum_{j=1}^{n} U_{ij}^{m} \| X_j - V_i \|, 1 \le m \le \infty \quad (4)$$

The second term depends on the entropy value and is given by $J_2 = \sum_{i=1}^{C} \pi_i^{*} e^{1-\pi_i^{*}}$ (5)

where $\pi_i^{*} = \frac{1}{N}\sum_{k=1}^{n} \pi_{ik}, k \in [1, N]$ and π_{ik} is the hesitation degree of the k[th] element in cluster 'i'.

Thus the new objective function is formulated by combining Eq. 4 and Eq. 5

$$J = \sum_{i=1}^{c}\sum_{j=1}^{n} U_{ik}^{*m} \| X_j - V_i \| + \sum_{i=1}^{C} \pi_i^{*} e^{1-\pi_i^{*}} \quad (6)$$

4. PROPOSED METHODOLOGY

In order to cluster various regions in an image, each pixel value along with five statistical features are considered. A 3 X 3 window is moved over the image pixels to obtain these statistical features. The statistical features utilized here include mean, median, mode, standard deviation and kurtosis. The image is converted into intuitionistic fuzzy representation and thus every pixel becomes three values contributing to membership, non-membership and hesitancy values. So, a total of eight attributes are passed on to the clustering algorithm.

The lambda value has to be fixed for each image. The value of lambda is chosen as the one which maximizes the entropy value. Entropy is the amount of fuzziness present in any given dataset and it is calculated as

$$IFE = \frac{1}{N \times M}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} \frac{2\mu_i(d_j)\nu_i(d_j) + \pi_i^2(d_j)}{\pi_i^2(d_j) + \mu_i^2(d_j) + \nu_i^2(d_j)} \quad (7)$$

where N and M are the rows and columns of the dataset.

The crisp data is converted into fuzzy data using the following equation

$$\mu_i(d_j) = \frac{d_{ij} - \min(d_j)}{\max(d_j) - \min(d_j)} \quad (8)$$

where d_{ij} is the current cell of the matrix under consideration and $\min(d_{ij})$ indicates the minimum value in the dataset matrix and $\max(d_{ij})$ indicates the maximum value in the dataset matrix.

Then the fuzzy data is converted to Intuitionistic fuzzy data as follows:

$$\mu_i(d_j; \lambda) = 1 - (1 - \mu_i(d_j))^{\lambda} \quad (9)$$

$$\nu_i(d_j; \lambda) = 1 - (1 - \mu_i(d_j))^{\lambda(\lambda+1)} \; where\, \lambda \in [0,1] \quad (10)$$

For utilizing PSO to select the optimal initial centroids, the parameters like number of particles, c1, c2, inertia weight, the maximum number of iterations, the number of clusters c, the problem dimension D and the fuzziness parameter m are initialized. The position and velocity matrices are randomly initialized.

For each particle, the distance measure is computed and the membership values of each object to various clusters are computed using Eq. 1. After finding hesitancy values, the fitness of each particle is evaluated using Eq. 6. The personal best value p_{best} for each particle and the overall best performance g_{best} for the entire swarm is found.

Subsequently, the velocity and position matrices are updated using the following equations

$$Velo(k+1)=wt.Velo(k) +(c1.rand1). (p_{best}(k) - Xpos(k)) + (c2.rand2).(g_{best}(k) - Xpos(k)) \quad (11)$$

$$Xpos(k+1) = Xpos(k) + Velo(k+1) \quad (12)$$

where c1 and c2 are user-defined constants, wt denotes the inertia weight, rand1 and rand2 are the random values from 0 to 1. These steps are repeated until g_{best} attains stability or until the maximum number of iterations is met. The particle that has the global best value with minimum cost is taken as the initial set of centroids for the execution of the IFCM algorithm.

Again, the new membership values are calculated using distance matrix Eq. 3. Then the centroids are updated using Eq. 2 by substituting the new membership values. For each iteration, the value for the second term of criterion function is also calculated using Eq. 5 and the entropy value is found. The IFCM algorithm is repeated till it converges. Finally, the entire combination of PSO and IFCM is run to reach the maximum number of iterations. The resulting cluster indices are found and the pixels which fall into the same cluster are grouped together to get the segmented image.

5. RESULTS AND DISCUSSION

The images taken from Quick Bird [12] satellite are to be segmented based on the color regions. Vegetation images exhibit the crop growth in various phases from planting through to harvest, helps in monitoring the changes as the season progresses and assessing abnormalities such as weed patches, soil compaction, watering problems, etc. The presence of green vegetation can be identified through the spectral contrast and some characteristics like vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) can be calculated. From the image, the number of clusters cannot be predicted well in advance. But the clustering algorithms need to specify the clusters in prior. Thus the experiment is carried out with different ‘k’ values and the optimum number of clusters has been found.

Silhouette [21] refers to a method of interpretation and validation of clusters of data. The technique provides a succinct graphical representation of how well each object lies within its cluster. One with the highest silhouette value identifies the optimum number of clusters. Table 1 presents the silhouette values for the IFCM and PSO-IFCM and it is clearly observed that the PSO-IFCM has the higher values producing optimal results.

To select an optimal clustering scheme, the criteria to be considered include compactness and separation. Compactness is minimizing the distance of objects within a cluster and separation is the measure of how each cluster is well-separated from other clusters.

The formula for silhouette coefficient is given by $SC = \frac{1}{N}\sum_{i=1}^{N} s(x)$ (13)

s(x) can be calculated from the formula $s(x) = \frac{b(x) - a(x)}{\max\{a(x), b(x)\}}$ (14)

where a(x) denotes compactness and b(x) indicates separation

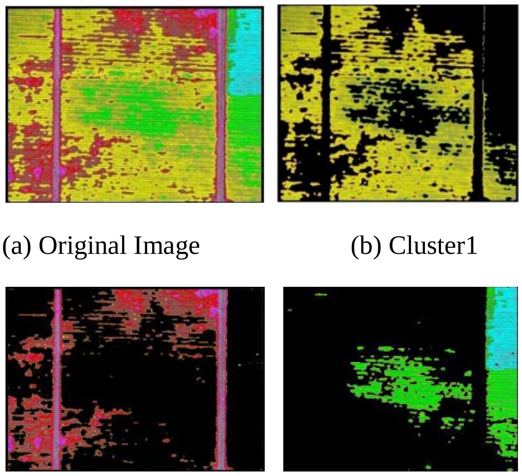
Table-1: Comparison of IFCM and PSO-IFCM using Silhouette Measure

		Results of IFCM			Results of PSO-IFCM		
S. No	**Image**	No of Clusters	Silhouette Index	Optimum No of clusters	No of Clusters	Silhouette Index	Optimum No of clusters
1	**Irrigation pattern**	3	0.7362	3	3	0.7334	4
		4	0.7032		4	0.7431	
		5	0.6757		5	0.6733	
2	**Crop health variations**	3	0.6080	4	3	0.5971	5
		4	0.6205		4	0.5994	
		5	0.6133		5	0.6299	

The results of the segmentation for Irrigation pattern image in Fig. 1 shows the segmentation using IFCM algorithm and it is visible that there are 3 clusters. Fig. 2 shows that there are 4 clusters found when using the PSO-IFCM. Thus the efficiency of the proposed method can be observed. Similarly Fig. 4 shows amore clear segmentation of the crop health variations image using PSO-IFCM when compared to IFCM in Fig. 3.

6. CONCLUSION

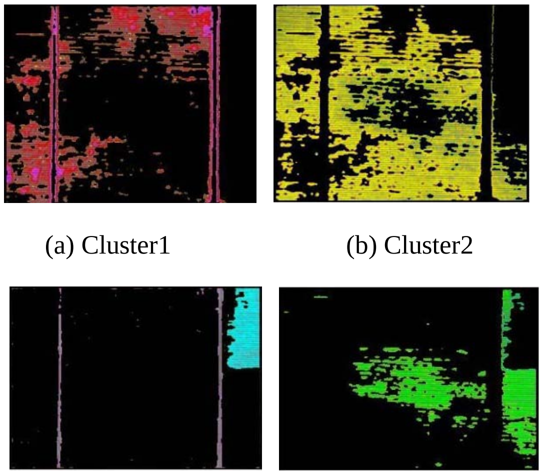
The proposed PSO-IFCM is a novel attempt to incorporate the ability to deal with noise and reach global optimal results. The results are demonstrated with satellite images and the quality of clustering is evaluated using an internal index named silhouette index. The segments produced differentiated every color region from others clearly. In future, the algorithm may be applied for identifying urbanization, ground water detection images, etc.



(a) Original Image (b) Cluster1

(c) Cluster2 (d) Cluster3

Figure-1: Image segmented using IFCM



(a) Cluster1 (b) Cluster2

(c) Cluster3 (d) Cluster4

Figure-2: Image Segmented using PSO-IFCM

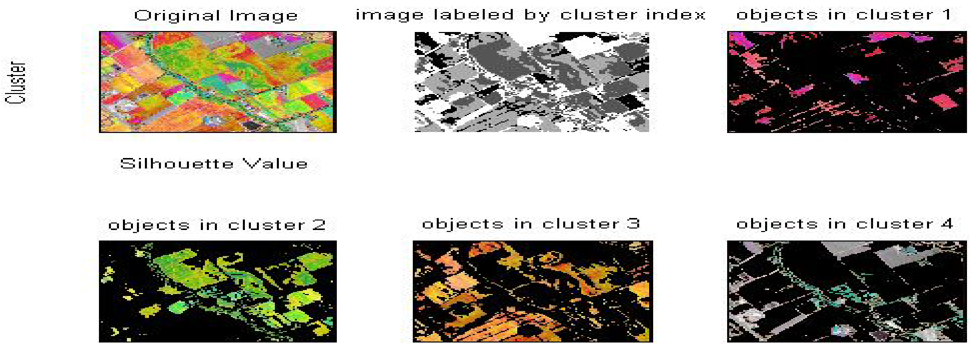


Figure-3: Crop Health Variations clustered using IFCM

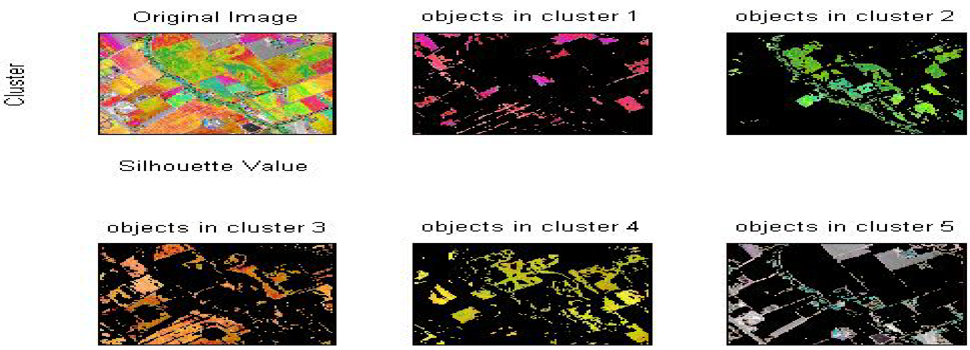


Figure-4: Crop Health Variations clustered using PSO-IFCM

REFERENCES

1. Ananthi, V. P., Balasubramaniam, P., & Lim, C. P., Segmentation of gray scale image based on intuitionistic fuzzy sets constructed from several membership functions. Pattern Recognition, Vol. 47, No. 12, 3870-3880, 2014.
2. Atanassov, K. T., Intuitionistic fuzzy sets: past, present and future. In EUSFLAT Conf., pp. 12-19, 2003.
3. Balasubramaniam P., and V. P. Ananthi, Segmentation of nutrient deficiency in incomplete crop images using intuitionistic fuzzy C-means clustering algorithm , *Nonlinear Dynamics*, Vol. *83*, No. 1-2, pp.849-866, 2016.
4. Bezdek J. C., Ehrlich R., Full W., FCM: The fuzzy c-means clustering algorithm, Computers & Geosciences, vol. 10, no.(2-3), pp. 191-203, 1984.
5. Bhargava R., Tripathy B. K., Tripathy A, Dhull R, Verma E., and Swarnalatha P, Rough intuitionistic fuzzy C-means algorithm and a comparative analysis , In *Proceedings of the 6th ACM India Computing Convention*, p. 23. ACM, 2013.
6. Chaira T, Intuitionistic fuzzy set approach for color region extraction, 2010
7. Chaira, T., An improved medical image enhancement scheme using Type II fuzzy set, *Applied soft computing* 25, pp 293-308, 2014.
8. Chaira, T., & Anand, S., A novel intuitionistic fuzzy approach for tumour/hemorrhage detection in medical images, 2011
9. Chaira T., A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images , *Applied Soft Computing*, Vol. 11, No. 2, pp.1711-1717, 2011.
10. Hein D., Hentschel A., Runkler T. and Udluft S., Particle swarm optimization for generating interpretable fuzzy reinforcement learning policies, Engineering Applications of Artificial Intelligence, Vol. 65, pp.87-98, 2017.
11. Huang, C. W, *et al.* Intuitionistic fuzzy c-means clustering algorithm with neighborhood attraction in segmenting medical image, *Soft Computing* Vol. 19, No. 2, pp 459-470, 2015.
12. http://www.satimagingcorp.com
13. Izakian A. and Abraham H., Fuzzy C-means and fuzzy swarm for fuzzy clustering problem , *Expert Systems with Applications*, Vol. *38, No.* 3, pp.1835-1838, 2011.
14. Izakian Z., Mesgari M.S., and Abraham A., Automated clustering of trajectory data using a particle swarm optimization , Computers, Environment and Urban Systems, Vol. 55, pp.55-65, 2016.
15. Kaur P,. Soni A. K, and Gosain A., Robust Intuitionistic Fuzzy C-means clustering for linearly and nonlinearly separable data , Proceedings of *Int. Conf. on Image Information Processing (ICIIP),* IEEE, pp. 1-6, 2011.
16. Kennedy J. and Eberhart R., Particle swarm optimization, *IEEE International Conference Proceedings on Neural Networks*, vol. 4, pp. 1942-1948, 1995.
17. Naser A.M.A and Alshattnawi S., An Artificial Bee Colony (ABC) Algorithm for Efficient Partitioning of Social Networks. *International Journal of Intelligent Information Technologies (IJIIT)*, Vol. *10, No.*4, pp.24-39, 2014.
18. Nobile M. S., Cazzaniga P., Besozzi D., Colombo R., Mauri G. and Pasi G., Fuzzy Self-Tuning PSO: A Settings-Free Algorithm for Global Optimization, Swarm and Evolutionary Computation, 2017.
19. Parvathavarthini S., Karthikeyani N., Shanthi S., and Mohan J M, Cuckoo-search based Intuitionistic Fuzzy Clustering Algorithm, Asian Journal of Research in Social Sciences and Humanities, Vol. 7, No. 2, pp.289-299, 2017.
20. Parvathavarthini S., Karthikeyani N., Shanthi S., and Lakshmi K., Crow-Search-Based Intuitionistic Fuzzy C-Means Clustering Algorithm , Developments and Trends in Intelligent Technologies and Smart Systems, p.129, 2017.
21. Rousseeuw P. J., Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, Vol. *20*, pp.53-65, 1987.
22. Salmeron J.L., Rahimi S.A., Navali A. M. and Sadeghpour A., Medical diagnosis of Rheumatoid Arthritis using data driven PSO–FCM with scarce datasets, Neurocomputing, Vol. 232, pp.104-112, 2017.
23. Saxena A., Prasad M., Gupta A., Bharill N., Patel O.P, Tiwari A.M.J., Ding, W. and Lin C.T., A review of clustering techniques and developments , Neurocomputing, Vol. 267, pp.664-681, 2017.
24. Shanthi S, and Bhaskaran V.M., Intuitionistic fuzzy C-means and decision tree approach for breast cancer detection and classification , *European Journal of Scientific Research*, Vol. 66, No. 3, pp.345-351, 2011.
25. Valente de Oliveira J., Szabo A. and de Castro L. N., Particle Swarm Clustering in clustering ensembles, Applied Soft Computing, Vol. 55, pp.141-153, 2017.
26. Xu, Z and Wu J., Intuitionistic fuzzy C-means clustering algorithms, *Journal of Systems Engineering and Electronics*, Vol. 21, No. 4, pp.580-590, 2010.

Source of support: Proceedings of UGC Funded International Conference on Intuitionistic Fuzzy Sets and Systems (ICIFSS-2018), Organized by: Vellalar College for Women (Autonomous), Erode, Tamil Nadu, India.

