ISSN: 0976-3104

SPECIAL ISSUE (SCMSA)
Purushotham and Tripathy

ARTICLE OPEN ACCESS

CLUSTER ANALYSIS USING HYBRID SOFT COMPUTING TECHNIQUES

Swarnalatha Purushotham*, BK. Tripathy
VIT University, Vellore-Tamil Nadu, INDIA

ABSTRACT

In a given set of data values with several attributes, similar data points can be clubbed together using a clustering architecture that uses global prototypes. These subset prototypes are exchanged so that a communication link is established between different clustering units. In this paper, a detailed clustering methodology is developed by combining both rough and fuzzy set techniques. This methodology shall be used to formulate a grouping of randomly generated unsupervised data that considers the integration of collaborative clustering in fuzzy data sets.

Received on: 30th-Nov-2015
Revised on: 11th-March-2016
Accepted on: 26th-March-2016
Published on: 10th–Aug -2016

KEY WORDS

Cluster Centre Matrix, Fuzzy Membership, Lower and Upper Approximation, Objective Function, Satellite Images

***Corresponding author: Email:** pswarnalatha@vit.ac.in **Tel.** +40-9001010010; **Fax:** +40-9001010012

INTRODUCTION

In Cluster is a collection of data objects which are similar to one another within the same cluster but dissimilar to the objects in other clusters. The problem is to group N patterns into c possible clusters with high intra-class similarity and low interclass similarity by optimizing an objective function. In objective function-based clustering algorithms, the goal is to find a partition for a given value of c. The c-means algorithm represents each cluster by its center of gravity [1].

The aim of collaborative clustering is to make different clustering methods collaborate, in order to reach at an agreement on the partitioning of a common dataset. As different clustering methods can produce different partitioning of the same dataset, finding a consensual clustering from these results is often a hard task. The collaboration aims to make the methods agree on the partitioning through a refinement of their results. This process tends to make the results more similar. In this paper, after the introduction of the collaboration process, we present different ways to integrate collaboration into already existing methodologies.

The implementation of fuzzy clustering has to be dealt with imprecise data that takes into consideration soft computing algorithms like c-means clustering. The fuzzy data is specifically used to deal with overlapping of data points. Whereas, the rough c-means incorporates the idea of vagueness and it is used cluster imprecise data.

Rough sets are purposed at defining clusters in terms of upper and lower approximations, which are identified by a pair of parameters while computing cluster prototypes. It is to be noted that RCM assigns objects into two distinct regions, viz., lower and upper approximations, such that objects in lower approximation ensures that the object is absolutely in the cluster while those in the upper approximation indicate possible inclusion in it. Since there is no concept of membership involved, therefore any measure of closeness of patterns to the clusters cannot be determined.

The paper [2] deals with a comparative study using RIFCM [3] with other related algorithms from their suitability in analysis of satellite images with other supporting techniques which deals with proving the superiority of RIFCM with RBP in clustering with other clustering methods and other supporting metrics with and without refined which integrates judiciously RIFCM with RBP. Finally, the superiority of the RIFCM using RBP is demonstrated, along with a comparison with other related algorithms, on satellite images with NASA.org images(Hills, Drought) and national geographic photographic images(Freshwater, Freshwater valley). Several papers have used image segmentation through clustering with various applications in view [4], [5], [6], [7], [28]. A family of clustering algorithms has been established with the use of the kernel function instead of the Euclidean distance [8], [9], [10], [11], [12], [13]. Algorithms have been devised to use mode as the measure of central tendency instead of mean some more clustering algorithms have been devised [14]. Using the possibilistic approach to clustering some algorithms have been proposed [15], [16], [17], [18]. Using covering based rough sets instead of basic rough sets some algorithms have been devised [19]. Some efforts have been done to improve the speed of existing algorithms like in [20]. Clustering of time series data is done in [21]. The initial assignment of input is done arbitrarily in almost all the above algorithms. But using genetic algorithms like the firefly algorithm an algorithm is proposed in [22].

In this paper, we present a novel collaborative clustering through the use of rough–fuzzy sets that is further expanded by means of incorporation of fuzziness powered grouping [23]. The use of rough sets is designed at restricting the effect of uncertainty among patterns that belongs to the upper and lower approximations, during collaboration between the modules. Incorporation of membership, in the RCM framework, is seen to enhance the robustness of clustering as well as collaboration. FRFCM framework is designed such that it is structured at finding data set collaboration.

CLUSTERING ALGORITHMS

Here we are going to describe different clustering algorithms, like c-means, fuzzy c-means, rough c-means and ant colony clustering. We are going to compare and contrast between them.

A. Hard C-means Clustering: Literature Overview

In this algorithm we partition N objects into c clusters. During each iteration centroids of each cluster is calculated. The algorithm goes as follows:

STEP 1: Fix c (2 ≤ c < n) and initialize the $U^{(0)}$ matrix

STEP 2: For r = 0, 1, 2 . . . do

Calculate the c center vectors $v_i^{(r)}$, i =1, 2…with $U^{(r)}$

STEP 3: Calculate the updated characteristic functions (for all i, k) using the formula

$$\chi_{ik}^{(r+1)} = \begin{cases} 1, \; if \; d_{ik}^{(r)} = \min\{d_{jk}^{(r)}\}, \; for \; all \; j = 1, 2, \ldots c \\ 0, \; otherwise \end{cases} \quad (1)$$

STEP 4: If $\left\| U^{(r+1)} - U^{(r)} \right\| < \varepsilon$ (the pre-assigned value then STOPS

Else r = r+1 and go to STEP 2 (The $\| \; \|$ norm here is the Euclidean norm)

The paper must have proposed system, results, discussion to infer the quality of the research paper. All the figures, equations and etc. must be in high resolution and in good quality.

B. Fuzzy C-Means (FCM) [5]

FCM is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by [4] and improved by [5]) is frequently used in pattern recognition. It is based on minimization of the following objective function.

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{c} \mu_{ij}^{\,m} \left\| x_i - c_j \right\|^2, 1 \le m < \infty \quad (2)$$

where μ_{ij} is the degree of membership of the object x_i in the jth cluster, c_i is centre of the i[th] cluster, and ||*|| is any norm expressing the similarity between data and center [6].

The Fuzzy c-means algorithm has the following steps:

STEP 1: Fix c ($2 \le c \le n$) and select a value m'

Initialize the partition matrix $U^{(0)}$

$$v_i^{(r)}, i = 1, 2, \ldots c$$

For r = 0, 1, 2… Do

STEP 2: Calculate the 'c' centres using the formula

$$v_{ij} = \frac{(\sum_{k=1}^{n} \mu_{ik}^{m'} . x_{kj})}{\sum_{k=1}^{n} \mu_{ik}^{m'}} \quad (3)$$

STEP 3: Update the partition matrix for the r[th] step $U^{(r)}$ to $U^{(r+1)} = (\mu_{ik}^{(r+1)})$, where

Taking $I_k = \{i \mid 2 \le c \le n; d_{ik}^{(r)} = 0\}$

$$\mu_{ik}^{(r+1)} = \left[\sum_{j=1}^{c} \left(\frac{d_{ik}^{(r)}}{d_{jk}^{(r)}} \right)^{2/(m'-1)} \right]^{-1}, if\ I_k = \phi, \quad (4)$$

$$= 0, \text{where} \quad i \in I_k^{'} = \{1, 2, \ldots c\} - I_k$$

STEP 4: If $\left\| U^{(r+1)} - U^{(r)} \right\| \le \varepsilon_L$ STOP

Else go to STEP 2,

where ε_L is a termination parameter lying in (0, 1) and k represents the iteration step. This procedure terminates after J_m reaches a local minimum.

C. Rough C-Means (RCM)[3]

The rough set model was introduced by Pawlak in 1982 [24] as another model of imprecision and since then has been found to be useful in many practical situations. [25]. The concept depends upon classification of the universe of discourse, which is equivalent to the notion of equivalence relation on it. For mathematical reasons, Pawlak took equivalence relations to define the model. Here, every subset of the universe is associated with two crisp sets called its lower and upper approximation and the region in between is the region of uncertainty being called as the boundary region associated with the set. The set is said to be rough if the lower and upper approximations are not equal and definable otherwise. Suppose X is a subset of a universe U and R is an equivalence relation defined over U. Then the lower and upper approximations of X with respect to R are denoted by $\underline{R}X$ and $\overline{R}X$ being defined as follows:

$\underline{R}X = \{x \in U \mid [x]_R \subseteq X\}$ and $\overline{R}X = \{x \in X \mid [x]_R \cap X \neq \phi\}$

X is R-rough iff $\underline{R}X \neq \overline{R}X$ and R-definable otherwise.

A schematic diagram for the different notions associated with the definition of rough set is presented in **[Figure- 1]**.

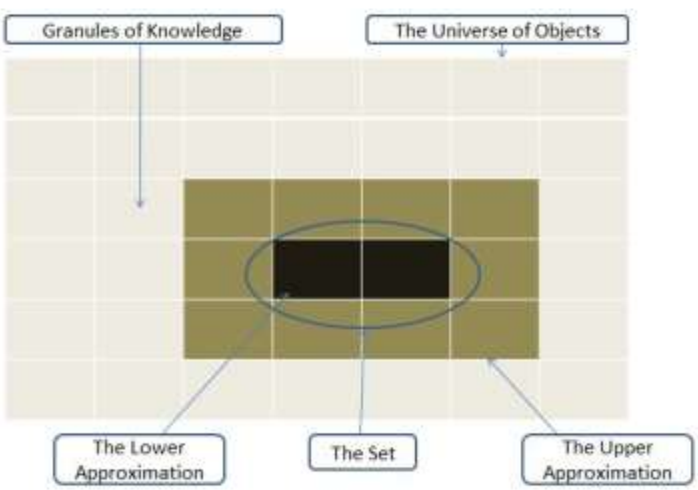

Fig: 1.Rough Set

THE IIOAB JOURNAL
ISSN: 0976-3104

In the rough *c*-means algorithm, the concept of *c*-means is further devised such that the cluster of data sets can be identified as an interval in rough data sets. A rough set *X* is characterized by its lower and upper approximations B*X* and B*X*, respectively, with the following properties.

1. An object x_k can be a part of at most *one* lower approximation.

2. If $x_k \in \underline{B}X$ of cluster X then simultaneously it also belongs to $\overline{B}X$.

3. If x_k is not a part of any lower approximation then it belongs to two or more upper approximations.

i. In both the lower and upper approximations;
ii. Only in lower approximation;
iii. Only in upper approximation of more than one clusters.

When a cluster contains object in both lower and upper approximations then cluster prototype has to be generated using both the weighing factor. When a cluster contains objects only in its lower or in its upper approximation, the cluster prototype is computed in the classical manner without scaling down by w_{low} and w_{up} . This prohibits drifting of prototypes from their desired location. This explains the formulation of the prototype by RCM in the equation. Note that the computation of the new cluster prototype is weighted by w_{low} and w_{up} only when both its approximations are nonempty. The actual algorithm is outlined as follows:

Assign initial means v_i for the *c* clusters.

STEP 1: Assign each data object (pattern) x_i to the lower approximation $\underline{B}U_i$ or upper approximation $\overline{B}U_i$, $\overline{B}U_j$ of cluster pairs U_i and U_j by computing the difference in its distance $d_{ik} - d_{jk}$ from the cluster centroid pairs v_i and v_j .

STEP 2: Let d_{ik} be minimum and d_{jk} be the next to minimum.

If $d_{ik} - d_{jk}$ is less than some threshold, then $x_k \in \overline{B}U_i, x_k \in \overline{B}U_j$, and x_k cannot be a member of any lower approximation

Else $x_k \in \underline{B}U_i$ such that distance d_{ik} is minimum over the *c* clusters.

STEP 3: Compute new centre for each cluster U_i using (5).

STEP 4: Repeat Steps 2 to 4 until convergence, i.e., there are no more new assignments of objects.

It is observed that the performance of the algorithm is dependent on the choice of w_{low} , w_{up} and the threshold. We use $w_{up} = 1 - w_{low}$, $0.5 < w_{loe} < 1$ and 0 < threshold < 0.5.

HYBRID CLUSTERING

In this section we introduce collaborative rough-fuzzy c-means algorithm, this is done by collaboration between different partitions or subpopulations.

A. Rough Fuzzy C-Means (RFCM)[1][8]

This algorithm allows us to incorporate fuzzy membership value u_{ik} of a sample x_k to a cluster mean v_i relative to all other means v_j for all $j \neq i$, instead of distance d_{ik} from the centroids. Fuzzy membership enables efficient handling of overlapping partitions while rough set deals with uncertainty, vagueness and incompleteness in terms of upper and lower approximation [8].

$$\mathbf{v}_i = \begin{cases} w_{\text{low}} \frac{\sum_{\mathbf{x}_k \in \underline{B}U_i} \mathbf{x}_k}{|\underline{B}U_i|} + w_{\text{up}} \frac{\sum_{\mathbf{x}_k \in (\overline{B}U_i - \underline{B}U_i)} \mathbf{x}_k}{|\overline{B}U_i - \underline{B}U_i|}, & \text{if } \underline{B}U_i \neq \emptyset \wedge \overline{B}U_i - \underline{B}U_i \neq \emptyset \\ \frac{\sum_{\mathbf{x}_k \in (\overline{B}U_i - \underline{B}U_i)} \mathbf{x}_k}{|\overline{B}U_i - \underline{B}U_i|}, & \text{if } \underline{B}U_i = \emptyset \wedge \overline{B}U_i - \underline{B}U_i \neq \emptyset \\ \frac{\sum_{\mathbf{x}_k \in \underline{B}U_i} \mathbf{x}_k}{|\underline{B}U_i|}, & \text{otherwise} \end{cases} \quad (5)$$

Incorporation of membership in the RCM framework enhances the robustness of the algorithm. Previously in RCM, one never had the idea of how similar a sample was to the given cluster in the absence of any similarity index. RFCM solves this problem with the help of membership values. Following are the steps of the algorithm.

THE IIOAB JOURNAL
ISSN: 0976-3104

STEP 1: Assign initial means v_i for the c clusters.

STEP 2: Calculate u_{ik} for c clusters and N data objects.

STEP 3: Assign each data object x_k to the lower approximation $\underline{B}U_i$ or upper approximations $\overline{B}U_i$, $\overline{B}U_j$ of cluster pairs U_i and U_j by computing the difference in its distance $u_{ik} - u_{jk}$ from the cluster centroid pairs v_i and v_j.

STEP 4: Let u_{ik} be maximum and u_{jk} be the next to maximum.

If $u_{ik} - u_{jk}$ is less than some threshold, then $x_k \in \overline{B}U_i$ and $x_k \in \overline{B}U_j$ and x_k cannot be a member of any lower approximation, else $x_k \in \underline{B}U_i$ such that membership u_{ik} is maximum over the *c* clusters.

STEP 5: Compute new centre for each cluster U_i using (6).

STEP 6: Repeat Steps 2 to 5 until convergence, i.e., there are no more new assignments of objects.

As indicated earlier we use $w_{up} = 1 - w_{low}$, $0.5 < w_{low} < 1$, m = 2 and 0< threshold < 0.5.

B. Hybrid FCM and RFCM (FRFCM-Fuzzy Rough Fuzzy C-Means)

Let us consider a dataset divided into P subpopulations or modules. Divide and conquer strategy is used to cluster this dataset. Each module or subpopulation is clustered individually to discover its structure. Collaboration is incorporated by exchanging information between the modules regarding local partitions in terms of collection of prototype computed within the modules. This strategy enables efficient handling of large datasets [9]. Hence this algorithm has strong communication levels resulting in presentation of information in small granules of prototypes.

Number of samples in the boundary region of clusters depend on the threshold value, higher the threshold value greater the number. Hence stronger collaboration between different modules is achieved resulting in the movement of clusters towards each other. This implies that the cluster modules are moving independently towards each other due to overlapping regions of corresponding clusters. Since the modules correspond to partitions from same large dataset it stabilizes the data towards efficient determination of globally existent structure.

There exists two phases in the algorithm.

Generation of FCM or RFCM clusters within the modules, without collaboration. Here we employ 0.5<w_{low}<1, thereby giving importance to samples lying within the lower approximation of clusters while

1. computing their prototype locally.

$$\mathbf{v}_i = \begin{cases} w_{low} \dfrac{\sum_{\mathbf{x}_k \in \underline{B}U_i} u_{ik}^m \mathbf{x}_k}{\sum_{\mathbf{x}_k \in \underline{B}U_i} u_{ik}^m} + w_{up} \dfrac{\sum_{\mathbf{x}_k \in (\overline{B}U_i - \underline{B}U_i)} u_{ik}^m \mathbf{x}_k}{\sum_{\mathbf{x}_k \in (\overline{B}U_i - \underline{B}U_i)} u_{ik}^m}, & \text{if } \underline{B}U_i \neq \emptyset \wedge \overline{B}U_i - \underline{B}U_i \neq \emptyset \\ \dfrac{\sum_{\mathbf{x}_k \in (\overline{B}U_i - \underline{B}U_i)} u_{ik}^m \mathbf{x}_k}{\sum_{\mathbf{x}_k \in (\overline{B}U_i - \underline{B}U_i)} u_{ik}^m}, & \text{if } \underline{B}U_i = \emptyset \wedge \overline{B}U_i \neq \emptyset \\ \dfrac{\sum_{\mathbf{x}_k \in \underline{B}U_i} u_{ik}^m \mathbf{x}_k}{\sum_{\mathbf{x}_k \in \underline{B}U_i} u_{ik}^m}, & \text{otherwise} \end{cases} \quad (6)$$

2. Collaborative FCM or RFCM between the clusters, computed locally for each module of the large dataset. Now we use 0<w_{low}<0.5 with a lower value providing higher precedence to samples lying in the boundary region of the overlapping clusters.

   a) In collaborative FCM, a cluster U_i may be calculated with an overlapping cluster U_j.
   b) Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij} and the cluster centers c_j. This procedure converges to a local minimum or saddle point of J_m.
   c) In case of collaborative RFCM U_i can be considered for merging with U_j

   $$\text{if } \sum_{x_k \in \underline{B}U_i} u_{ik} \leq \sum_{x_k \in (\overline{B}U_i - \underline{B}U_i)} u_{ik}$$

   and v_i is closest to v_j in the feature space being the maximum among all overlapping clusters.

COMPUTER SCIENCE

The entire algorithm is summarized below.

STEP 1: Split the large dataset into P modules.

STEP 2: For each module p=1, , , , , ,P do

STEP 3: For each module p do collaboration.

a) Assign each pattern x_k to lower or upper approximation of the C (= c*P) collaborative FCM or RFCM clusters,

with 0<w_{low}< 0.5.

b) Merge overlapping clusters pairs while

i. Compute new prototype for merged clusters U_i and U_j as the mean of v_i and v_j.

ii. Reduce number of clusters C by one.

iii. Reassign each pattern x_k to lower or upper approximation of the C collaborative RCM or RFCM clusters.

COMPARATIVE ANALYSIS

The density of the collaboration between clusters or imprecise set of data values can be statistically evaluated in terms of separate indices and PSNR and RMSE values. These indices are Davies-Bouldin Index, Partition Co-efficient Index, Classification Entropy and Silhouette Statistical Index. In this section, we bring upon these measures to position our collaborative clustering framework. It is also needed to be mentioned that the number of clusters in a module needs to remain fixed but is generalized to be unique both before and after collaboration. Memberships of data objects are computed both before and after collaboration, with respect to the cluster prototypes. Since overlapping clusters can very well be merged by means of collaboration, the final cardinality of indices within different modules are often indication enough towards the degree of the clustering. This lead us to use the maximum membership value max $u_{ik(p)}$ of a data point x_k of module *p*, to one of the clusters U_i, during our computation of separate indices. We consider the four indices *Davies-Bouldin Index, Partition Co-efficient Index, Classification Entropy and Silhouette Statistical Index*
The DB [25] is a function of the ratio of the sum of within cluster distance to between-cluster separation. Another index used for measuring clustering efficiency is the D index [26]. The method with lower value of the index bears the greater potential of clustering. Let $\{x_1,...x_{|ck|}\}$ be a set of patterns lying in a cluster U_k. Then, the Davies Bouldin index is defined as

$$DB = \frac{1}{c}\sum_{j=1}^{c}\max\left\{\frac{d_w(U_i)+d_w(U_j)}{d(U_i,U_j)}\right\} \quad (7$$

for 1< j, i < c and within-cluster distance $d_w(U_i)$ is minimized while the between-cluster separation $d(U_i,U_j)$ gets maximized.

Silhouette Index, *S*, computes for each point a width depending on its membership in any cluster where c_i is the average distance between points *i* and all other points in its own cluster. Here, b_i is the minimum of the average dissimilarities between *i* and points in other clusters. Negative index showcases the stability in the collaboration; lower the magnitude greater is the amount of grouping [27].

$$S_k = \frac{1}{N}\sum_{i=1}^{N}\frac{b_i - a_i}{\max(a_i,b_i)}$$

$$S = \frac{1}{c}\sum_{k=1}^{c}S_k$$

The Partition co-efficient is the measure of overlap between the clusters. This index value directly corresponds to the degree of partition achieved [28]. If u_{ij} is taken as membership of data point j in cluster i and c as the number of clusters, then the index is defined as

$$PC(c) = \frac{1}{N}\sum_{i=1}^{c}\sum_{j=1}^{N}(u_{ij})^2$$

A similar index named Classification Entropy is designed such that the stability of clustering methodology is calculated. The imprecise data is handled and its fuzziness is considered [28]. The negative entropy indicates a stable arrangement among data sets. With all probabilistic cluster partitions c obeying the rule 0 <1-PC(c) <CE(c), the classification entropy is defined as

$$CE(c) = -\frac{1}{N}\sum_{i=1}^{c}\sum_{j=1}^{N}u_{ij}\log(u_{ij})$$

Cluster validity index values **[Table-1]** are calculated on the data set given by [(1,3), (1.5,3.2), (1.3,2.8), (3,1)]

iv.

TABLE 1: Cluster validity index values

Index	FCM	RCM	RFCM	FRFCM
Davies-Bouldin	1.956	2.608	1.872	1.870
Silhouette	-0.370	-0.790	-0.352	-0.350
Partition Coefficient Index	0.984	4.570	4.486	4.605
Classification Entropy	0.017	-0.693	-0.685	-0.681

In **[Figure- 2]**, the different index values are plotted and the aforementioned indices are shown in black, red, green and blue colour respectively. The x-axis values correspond to the algorithms specified.

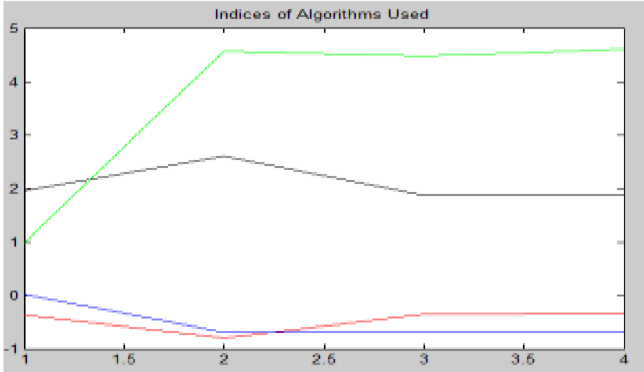

Fig: 2. Comparative Analysis of Clustering Algorithms using different indices

RMSE and PSNR Values

Peak Signal-to-Noise Ratio can be characterized as PSNR, which is the relation with the majority likely power of a signal and the power of corrupting distortions that influence the fidelity of its demonstration.

The PSNR value can be computed through mean squared error (MSE). For an example distortion-free m x n monochrome image 'I' with its noisy approximation 'K'. The RMSE of a model prediction is defined as the square root of the mean squared error: Hence, the PSNR is defined as where MAX_I is the most possible 0's and 1's values of an image. And it will be replaced with 255, as and when the 0's and 1's are given using 8 bits per model. And MSE will become '0'; when the distortion is null indicating that the two input images are same.

Here, MAX_I is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. In the absence of noise, the two images I and K are identical, and thus the MSE is zero. In this case the PSNR is undefined. And the performance of PSNR and RMSE values as per **Table- 2** & **Figure- 3** of scanned cerebral image [29] [30] resulted efficient result of using hrbrid clustering algorithm given in **Figure-8** compared to other clustering algorithms as per

THE IIOAB JOURNAL ISSN: 0976-3104

Original brain image- **Figure-4,** FCM-**Figure- 5** , RCM-**Figure-6,**RFCM-**Figure- 7** and FRFCM-**Figure-8.**

TABLE: 2. PSNR AND RMSE VALUES

Metric/Cluster Techniques	FCM	RCM	RFCM	FRFCM
PSNR	7.8409	9.856	9.9526	8.5964
RMSE	7.0379	6.8821	6.8127	9.421

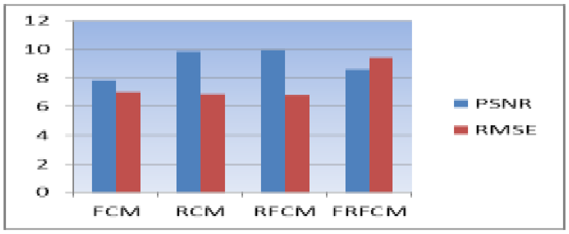

Fig: 3. Performance of PSNR and RMSE Values

RESULTS AND DISCUSSION

The final output of hybrid algorithm is devised to represent membership values in FRFCM framework in scanned cerebral image. The following set of images depicts the level of clustering and finally groups the points with same attributes to bring forward high definition precision with efficient results using metrics of Davies-Bouldin Index, Partition Co-efficient Index, Classification Entropy and Silhouette Statistical Index and PSNR & RMSE values(8.5964 & 9.4210) for FRFCM clustering algorithm in comparison with FCM, RCM, RFCM clustering algorithms.

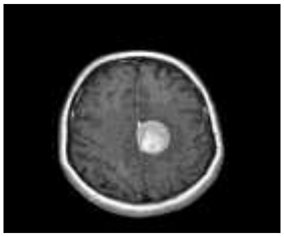

Fig: 4. Original Brain Scan Image

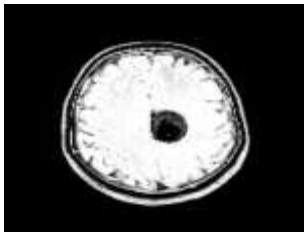

Fig: 5. Clustered Image applying FCM

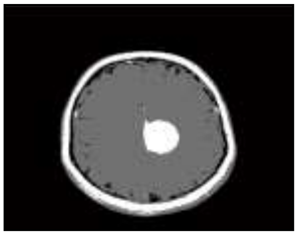

Fig 6: Clustered Image after Applying RCM

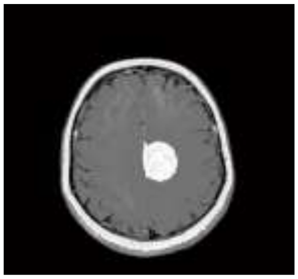

Fig: 7. Clustered Image After applying RFCM

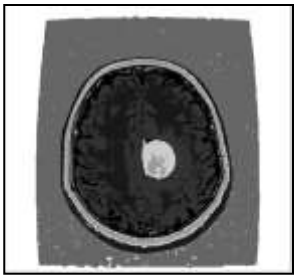

Fig: 8. Clustered Image after applying FRFCM

CONCLUSION

In this paper we have introduced a new clustering algorithm called the Fuzzy Rough Fuzzy C-Means (FRFCM) which is a first of its kind where a hybrid model is formed by combining three models. We know that the hybrid models are more efficient than their individual components. Here we could establish experimentally that as we increase the level of hybridization the efficiency further increases. For this purpose we have taken several measuring indices and also we have taken established images and could show that the segmentation shows improved results than the individual models in the form of individual or another two level hybrid model. This mow opens up a direction of research in which we can increase the hybridization to further higher levels. It would be interesting to find whether this trend will continue. Of course the complexity of these higher levels hybrid algorithms increase.

FINANCIAL DISCLOSURE

No financial support was received to carry out this project.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers and the editor-in-chief of the journal for their valuable guidance which has improved the quality and presentation of the paper.

CONFLICT OF INTERESTS

Authors declare no conflict of interest.

REFERENCES

[1] S Mitra, H. Banka and W. Pedrycz, [2006] Rough–Fuzzy Collaborative Clustering, IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics, 36(4):68–73.
[2] BK Tripathy, P Swarnalatha. [2013] A Comparative Study of RIFCM with Other Related Algorithms from Their Suitability in Analysis of Satellite. Images using Other Supporting Techniques, Kybernetes, 43(1).
[3] R Bhargav, B. K. Tripathy, A Tripathy, E. Verma, Raj Kumar, P Swarnalatha: Rough Intuitionistic Fuzzy C-Means Algorithm and a Comparative Analysis, COMPUTE"13, Aug 22-24, Vellore, Tamil Nadu, India Copyright 2013 ACM 978-1-4503-2545-5/13/08.
[4] P Swarnalatha, B K Tripathy [2016]: ANALYSIS OF DEPTH USING CLUSTERING TECHNIQUES WITH BIT PLANE FILTER FOR SATELLITE IMAGERY, IIOABJ, 7(5): 108-125.
[5] B. K. Tripathy, 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.
[6] BK Tripathy, P Swarnalatha [2015] A Comparative analysis of Depth computation of Leukemia Images using a refined Bit Plane and Uncertainty based clustering Techniques, Cybernetics and Information Technologies,15(1):126- 146.
[7] B. K. Tripathy, P Swarnalatha. [2013] A Novel Fuzzy C-Means Approach with Bit Plane Algorithm for Classification of Medical Images, IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology [ICECCN 2013], pp.360 – 365
[8] Deepthi PH., B.K.Tripathy [2016]: Kernel Based Spatial Fuzzy C-Means for Image Segmentation, the IIOAB journal, 7(5): 150-156
[9] B.K.Tripathy, A Ghosh, G.K. Panda. [2012] Adaptive K-means Clustering to Handle Heterogeneous Data using Basic Rough Set Theory, In the Proceedings of Second Intl. Conf. [CCIT-2012], Bangalore, Advances in Computer Science and Information Technology Network Communications, LNICST series, Springer, 84:193 – 201.
[10] B.K. Tripathy, A Ghosh, G K. 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, Coimbatore, INDIA, pp.1–5
[11] B. K.. Tripathy, R Bhargav [2013]: Kernel Based Rough-Fuzzy C-Means, PReMI, ISI Calcutta, December, LNCS 8251, pp.148–157.
[12] B. K. Tripathy, A Tripathy, K.Govindarajulu and R. Bhargav [2014]: On kernel Based rough Intuitionistic Fuzzy C-means algorithm and a comparative analysis, ICACNI 2014, Advanced computing networking and informatics, vol.1, Smart innovation systems and technologies, 27: 349–359.
[13] B. K. Tripathy, A Tripathy, K.Govindarajulu and R. Bhargav [2014]: On kernel Based rough Intuitionistic Fuzzy C-means algorithm and a comparative analysis, ICACNI 2014, Advanced computing networking and informatics, vol.1, Smart innovation systems and technologies, 27: 349–359.
[14] B. K. Tripathy, A Goyal, PA Sourav [2016] A comparative analysis of rough intuitionistic fuzzy k-mode algorithm for clustering categorical data, Research Journal of Pharmaceutical, Biological and Chemical Sciences Volume 7, Issue 5, pp. 2787-2802
[15] B. K. Tripathy, A Tripathy, K.Govindarajulu [2014] Possibilistic rough fuzzy C-means algorithm in data clustering and image segmentation, proceedings of the IEEE ICCIC2014, pp.981-986.
[16] B. K. Tripathy, A Tripathy, K.Govindarajulu [2014] Possibilistic rough fuzzy C-means algorithm in data clustering and image segmentation, proceedings of the IEEE ICCIC2014, pp.981-986.
[17] B. K. Tripathy, A Tripathy, K.Govindarajulu [2014] Possibilistic rough fuzzy C-means algorithm in data clustering and image segmentation, proceedings of the IEEE ICCIC2014, pp.981-986.

[18] B. K. Tripathy, A Tripathy, K Govindarajulu. [2015]: On PRIFCM Algorithm for Data Clustering, Image Segmentation and Comparative Analysis, Souvenir of the 2015 IEEE International Advance Computing Conference, IACC 2015, Article number-7154725, pp.333–336.
[19] P Lingras, G Peters.[2012] Applying rough sets to clustering. Rough Sets: Selected Methods and Applications in Management and Engineering, Advanced Information and Knowledge Processing, DOI 10.1007/978-1-4471-2760-4_2, © Springer-Verlag London Limited.
[20] B. K. Tripathy, S Sobti, V Shah [2014] A refined rough fuzzy clustering algorithm, proceedings of the IEEE ICCIC2014, pp.776–779
[21] B. K. Tripathy, P Prabhavathy [2015]: An Intelligent Clustering Approach for Sequential Data, presented at ICIMR 2015, Bhubaneswar.
[22] B. K. Tripathy, A Tripathy, K.Govindarajulu [2014] Possibilistic rough fuzzy C-means algorithm in data clustering and image segmentation, proceedings of the IEEE ICCIC2014, pp.981-986.
[23] P Lingras, G Peters.[2012] Applying rough sets to clustering. Rough Sets: Selected Methods and Applications in Management and Engineering, Advanced Information and Knowledge Processing, DOI 10.1007/978-1-4471-2760-4_2, © Springer-Verlag London Limited.
[24] Z Pawlak [1982] Rough Sets, International Journal of Information and Computer Sciences, 11:341–356.
[25] S Petrovic.[2003] A Comparison Between the Silhouette Index and the Davies-Bouldin Index in Labelling IDS Clusters. NIS Lab, Department of Computer Science and Media Technology, pp.1-12.
[26] JC Dunn [1973] A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, pp.32-57.
[27] P Maji, SK Pal [2007] Rough Set Based Generalized Fuzzy C-Means Algorithm and Quantitative Indices. Center for Soft Computing Research Indian Statistical Institute, Fundamenta Informaticae 80:475–496.
[28] W Pedrycz [2002] Collaborative fuzzy clustering, Pattern Recognition. Letters, vol. 23, no. 14,1675–1686, Dec., Article in a conference 7670.
[29] LC Robert, JV Dave, JC Bezdek. [1986] Efficient Implementation of the Fuzzy C-Means Algorithm", Mathematics Department, Utah State University, IEEE Transactions of Pattern Analysis and Machine Intelligence, PAMI-8(2).
[30] Swarnalatha P, B. K. Tripathy BK, Nithin, P L and D Ghosh [2014]:: Cluster Analysis Using Hybrid Soft Computing Techniques, CNC-2014International Conference of Network and Power Engineering ,Proceedings of Fifth CNC-2014,pp. 516–524.

ABOUT AUTHORS

Swarnalatha Purushotham is an Associate Professor, in the School of Computer Science and Engineering, VIT University, at Vellore, India. She pursued her Ph.D degree in Image Processing and Intelligent Systems. She has published more than 60 papers in International Journals/International Conference Proceedings/National Conferences. She is having 15+ years of teaching experiences. She is a member of IACSIT, CSI, ACM, IACSIT, IEEE (WIE), and ACEEE. She is an Editorial board member/reviewer of International/ National Journals and Conferences. Her current research interest includes Image Processing, Remote Sensing, Artificial Intelligence and Software Engineering.

B.K.Tripathy is a senior professor in the school of computer sciences and engineering, VIT University, Vellore, since 2007. He has produced 26 PhDs, 13 M. Phils and 03 M.S students so far. He has published around 400 technical papers in different international journals, conference proceedings. Dr. Tripathy has guest edited some journals of repute. He has edited five research volumes for the IGI publications and Taylor and Francis publications. Dr. Tripathy has written two books on Soft Computing and Computer Graphics. He is in the editorial board or review panel of over 70 international journals including Springer, Science Direct, IEEE and World Scientific publications. He is a life member/ senior member/member of 21 international forums including ACM, IEEE, ACEEE, IRSS, ISRR and CSI. His current interest includes Fuzzy Sets and Systems, Rough sets and Knowledge Engineering, Multiset Theory, List Theory, Data clustering and Database Anonymization, Content Based Learning, Neighbourhood Systems, Soft Set Analysis, Image Processing, Cloud Computing, Social Internet of Things, Big Data Analytics, Multi Criteria Decision Making and Social Network Anonymization.