6th IAPR International Conference on

Computer Vision and Image Processing (CVIP-2021)

Paper ID : 106

Paper Title: Adaptive Rough-Fuzzy Kernelized Clustering Algorithm for Noisy Brain MRI Tissue Segmentation

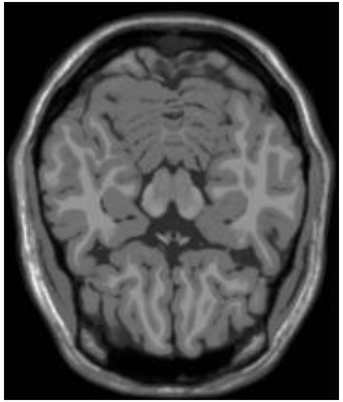
List of Authors:

- Rudrajit Choudhuri
- Amiya Halder

Presenter Information:
Rudrajit Choudhuri
Final Year Undergraduate Student
Department of Computer Science & Engineering
St. Thomas' College of Engineering & Technology
Kolkata, West Bengal, India

INTRODUCTION

- Image segmentation is a method which segregates a digital image into various homogeneous regions called segments thereby reducing the complexity of the image and aiding in better analysis.
- Segmentation is crucial in biomedical image processing for quantification and analysis of various regions, partitioning and anomaly detection.
- Magnetic Resonance Imaging (MRI) provides significant advantage when it comes to studying the human brain and helps in early detection of abnormalities in its tissues and organs.
- Segmentation becomes a challenging problem to tackle mainly due to the complex structure of the human brain.
- Clustering of image data points is a decent approach for medical image segmentation but its performance drops in the presence of noise corruptions, intensity inhomogeneities and other artifacts.



ORIGINAL IMAGE

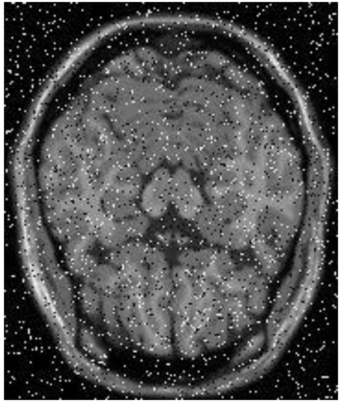


IMAGE CORRUPTED BY 8% SALT & PEPPER NOISE

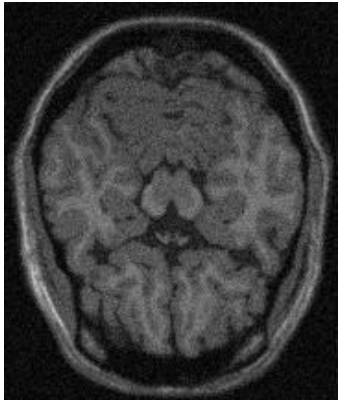


IMAGE CORRUPTED BY 9% GAUSSIAN NOISE & 40% INHOMOGENEITY

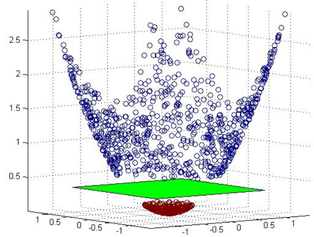
- Noise corruptions in medical images degrade the image to such an extent that segmentation or any further image processing task fails to achieve desired results.
- The reason for these corruptions lie in the faulty acquisition, storage or transmission of an MR Image.
- An adaptive rough-fuzzy set based kernelized C means clustering algorithm is presented in this research which is aimed at robust segmentation of brain magnetic resonance imaging (MRI) data.

GROUNDWORK

FUZZY SET THEORY

K-MEANS &
FUZZY C MEANS

KERNEL METHODS



ROUGH SET THEORY

FUZZY SET THEORY

- Fuzzy set logic was developed to come up with uncertainty handling techniques stemming from ambiguous class definitions.
- Incorporation of fuzzy set logic satisfactorily deals with the uncertainties in imprecise decision making or selection of multiple options in a collection.
- In the domain of image segmentation, assimilating fuzzy sets allow data points to belong in a particular cluster upto a certain degree.
- Partial membership allowance to a data point and accounting for respective importance of each class is a major benefit of amalgamating fuzzy sets in a clustering algorithm.
- A fuzzy set mapping for a set L is an $L \rightarrow [0, 1]$ mapping and depending on the use case, the membership function can be designed.

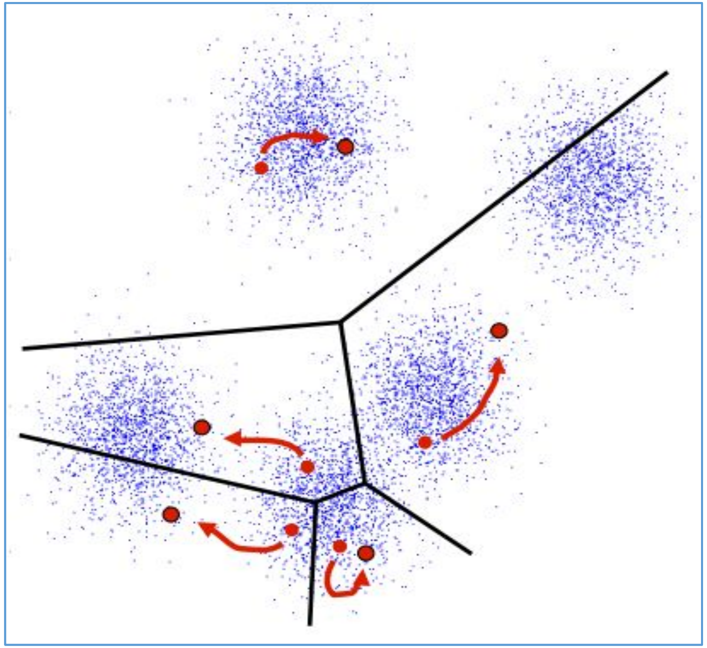
K-MEANS & FUZZY C MEANS

- K-Means is an iterative clustering algorithm.

- It partitions N data points into K disjoint subsets S_i so as to minimize the sum of squares criterion (objective function)

$$J = \sum_{i=1}^{K} \sum_{n \in S_i} |x_n - \mu_i|^2$$

where x_n is a vector representing n^{th} data point & μ_i is the geometric centroid of the data points in S_i.



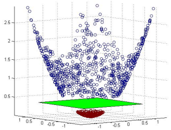
K-MEANS & FUZZY C MEANS

- FCM is an iterative optimization algorithm used for clustering data points to different classes.
- It incorporates fuzzy set logic for membership computation which is necessary for assigning category to a data point.
- Consider Y = (y1, y2. . . yn) to be an image consisting of N pixels which are to be segmented into C clusters, where image pixels in Y denote multispectral features.
- FCM iteratively minimizes the cost function by assigning higher membership values to pixels which are nearer to the cluster centroid.

$$J = \sum_{q=1}^{c} \sum_{p=1}^{N} \mu_{qp}^{m} ||y_p - v_q||^2$$

$$\mu_{qp} = \frac{1}{\sum_{j=1}^{c} \left(\frac{||y_p - v_q||}{||y_p - v_j||} \right)^{\frac{2}{m-1}}}$$

$$v_q = \frac{\sum_{p=1}^{N} \mu_{qp}^{m} y_p}{\sum_{p=1}^{N} \mu_{qp}^{m}}$$



KERNEL METHODS

- Kernel methods are a commonly used trick in the domain of pattern recognition.
- The reason behind using kernel methods is due to the ease in structuring and separating data points by mapping them to a higher dimensional space.
- Using kernel methods assimilated with vector algebra is a smarter alternative to the rigorous computations involved in mapping functions.
- It bridges the gap between linearity and nonlinearity for any algorithm that can be represented in the form of scalar products between vectors.
- A kernel function represents a scalar product in a feature space and is of the form: $K(x, y) = \langle \psi(x), \psi(y) \rangle$

ROUGH SET THEORY

- Rough set theory is based on approximation space substructure, which in effect is a pair <U, R>
- U is a non-empty set & R is an equivalence relation on U.
- Using R the set U is decomposed into disjoint categories such that two elements (x, y) are in same category if xRy holds.
- U/R = {X1, X2, ..., Xn} is the quotient set of U by relation R, Xi denotes equivalence class of R.
- Two elements of U are indistinguishable if they belong to same equivalence class.
- It is rigorous to describe an arbitrary set accurately in <U, R>

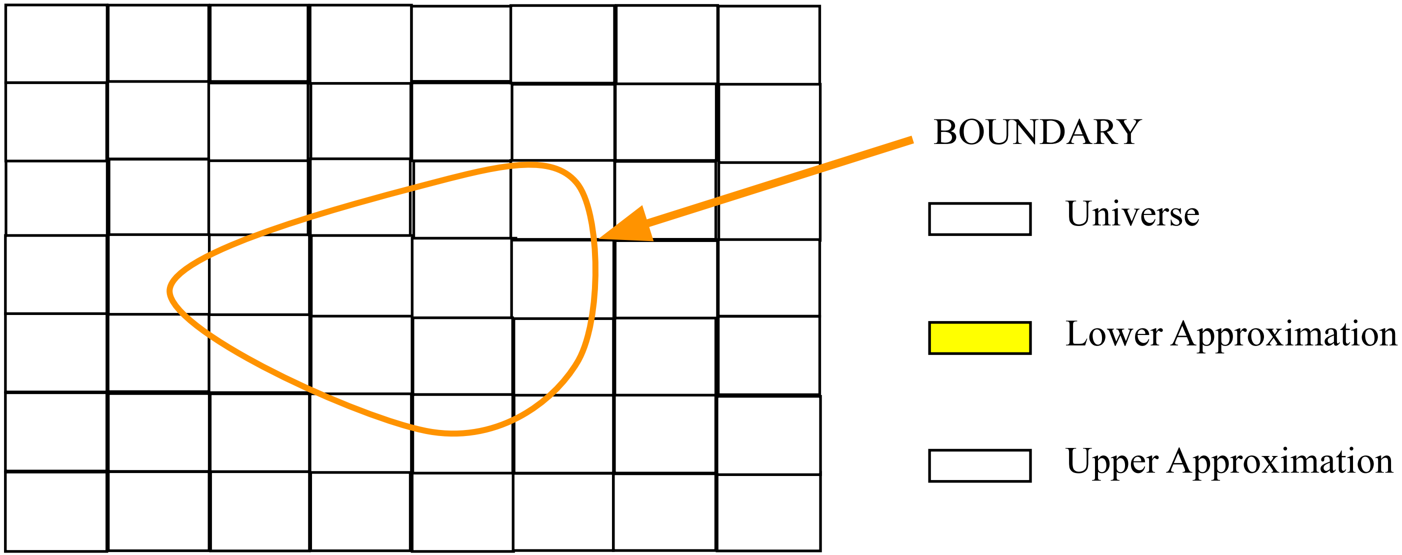
ROUGH SET THEORY

- A set $X \in 2^U$ can be characterized by a pair of lower and upper approximations.

$$\underline{R}(X) = \bigcup_{X_i \subseteq X} X_i; \quad \overline{R}(X) = \bigcup_{X_i \cap X \neq \emptyset} X_i$$

- $\underline{R}$(X) is the lower approximation i.e., the union of every elementary set that are subsets of X
- $\overline{R}$(X) is the upper approximation, i.e. the union of every elementary set having non empty intersections with X.
- Given the lower and upper approximations, [$\underline{R}$(X), $\overline{R}$(X)] is the representation of X in the approximation space < U, R > and is known as the rough set of X.
- The use of rough sets tackles ambiguities, imperceptibility and vagueness through approximations.

ROUGH SET THEORY



EXISTING METHODS

- Kernelized FCM
- Enhanced FCM
- Fast Generated FCM
- Rough Fuzzy C Means
- Contextual Kernelized Rough Fuzzy C Means
- Local Intensity Based Fuzzy Clustering Model

PROPOSED ALGORITHM

- The algorithm amalgamates fuzzy set and rough set theories along with an adaptive substructure based kernelized distance metric to form a reliable clustering approach for seamless segmentation.

- Rough and fuzzy sets help in dealing with uncertainties caused due to vagueness, imperceptibility, ambiguities and overlappingness while clustering data points.

- Incorporation of an adaptive substructure prudently includes the influence of spatial features and neighborhood details while clustering data points thus compensating for noise corruptions and inhomogeneities of intensity values in the image.

- The kernelized distance metric leverages linear separability and allows for easy segregation by mapping data points to a higher dimensional space while preserving fine details and spatial features.

ADAPTIVE KERNELIZED DISTANCE METRIC

- The proposed approach considers spatial information around an image pixel along with its intensity value as a feature.

- To achieve this, a 3×3 window centered at the concerned pixel (yi) is picked and the average pixel intensity value (yi) of the window is computed and is used as a data point for clustering.

- The adaptive substructure accounts for local spatial information during segmentation which compensates for inhomogeneities leading to an efficient segregation.

- The proposed approach uses a circular kernel:

$$K(x,y) = \frac{2}{\pi} cos^{-1}\left(\frac{-||x-y||}{\sigma}\right) - \frac{2}{\pi}\frac{||x-y||}{\sigma}\sqrt{1-\left(\frac{||x-y||}{\sigma}\right)^2}$$

where σ is the tuning parameter for adjusting the kernel.

OBJECTIVE FUNCTION

- Consider $\underline{R}(\xi i)$, $\overline{R}(\xi i)$ to be the lower approximation and upper approximation of the cluster (ξi), and $H(\xi i) = \{\overline{R}(\xi i) - \underline{R}(\xi i)\}$ to be the boundary region corresponding to the cluster (ξi).

- The objective function that the proposed method aims to optimize is:

$$J_{ARKFCM} = \begin{cases} \omega \times J_L + (1-\omega) \times J_B & \text{if } \underline{R}(\xi_i) \neq \phi, H(\xi_i) \neq \phi \\ J_L, & \text{if } \underline{R}(\xi_i) \neq \phi, H(\xi_i) = \phi \\ J_B, & \text{if } \underline{R}(\xi_i) = \phi, H(\xi_i) \neq \phi \end{cases}$$

$$J_L = \sum_{k=1}^{c} \sum_{\overline{y_i} \in \underline{R}(\xi_i)} \mu_{ki}^{q} (1 - K(\overline{y_i}, v_k)) + \sum_{k=1}^{c} \lambda_i \sum_{\overline{y_i} \in \underline{R}(\xi_i)} (1 - \mu_{ki})$$

$$J_B = \sum_{k=1}^{c} \sum_{\overline{y_i} \in H(\xi_i)} \mu_{ki}^{q} (1 - K(\overline{y_i}, v_k)) + \sum_{k=1}^{c} \lambda_i \sum_{\overline{y_i} \in H(\xi_i)} (1 - \mu_{ki})$$

- In the presented technique, each cluster has a crisp lower approximation space along with a defined fuzzified boundary region. Lower approximation is responsible for the final partition fuzziness.
- According to the rough set theory, if a data point is in the lower approximation space, it definitely belongs to the concerned cluster.
- Therefore, the weights of these data points should be independent of the influence of other clusters and should not be coupled with other clusters based on similarity measures.
- On the contrary, if a data point lies in the boundary region, it has potential chance of belonging to a different cluster.
- Therefore, the data points lying in the boundary region should not directly influence the centroid of the cluster.
- Effectively, the presented technique bifurcates data points into two different classes – lower approximation space and boundary region.
- The technique only fuzzifies data points which are in the boundary region.

MEMBERSHIP FUNCTION & CLUSTER CENTROID

- The cost function on optimization with respect to the membership function and cluster centroid leads to necessary and sufficient conditions for the membership function and the cluster centroid to be at the minimal saddle point.

$$\mu_{ki} = \frac{(1 - K(\overline{y_i}, v_k))^{-\frac{1}{q-1}}}{\sum_{j=1}^{c}(1 - K(\overline{y_i}, v_k))^{-\frac{1}{q-1}}}$$

$$v_k^{/} = \begin{cases} \omega \times \frac{\sum_{\overline{y_i} \in \underline{R}(\xi_i)} \mu_{ki}^{q} \overline{y_i} K(\overline{y_i}, v_k)}{\sum_{\overline{y_i} \in \underline{R}(\xi_i)} \mu_{ki}^{q} K(\overline{y_i}, v_k)} + (1-\omega) \times \frac{\sum_{\overline{y_i} \in H(\xi_i)} \mu_{ki}^{q} \overline{y_i} K(\overline{y_i}, v_k)}{\sum_{\overline{y_i} \in H(\xi_i)} \mu_{ki}^{q} K(\overline{y_i}, v_k)} & \text{if } \underline{R}(\xi_i) \neq \phi, H(\xi_i) \neq \phi \\ \frac{\sum_{\overline{y_i} \in \underline{R}(\xi_i)} \mu_{ki}^{q} \overline{y_i} K(\overline{y_i}, v_k)}{\sum_{\overline{y_i} \in \underline{R}(\xi_i)} \mu_{ki}^{q} K(\overline{y_i}, v_k)}, & \text{if } \underline{R}(\xi_i) \neq \phi, H(\xi_i) = \phi \\ \frac{\sum_{\overline{y_i} \in H(\xi_i)} \mu_{ki}^{q} \overline{y_i} K(\overline{y_i}, v_k)}{\sum_{\overline{y_i} \in H(\xi_i)} \mu_{ki}^{q} K(\overline{y_i}, v_k)}, & \text{if } \underline{R}(\xi_i) = \phi, H(\xi_i) \neq \phi \end{cases}$$

ω tunes the influence of the lower approximation space.

Algorithm 1: PROPOSED ALGORITHM

Input: Brain MRI Image
Output: Segmented MRI Image

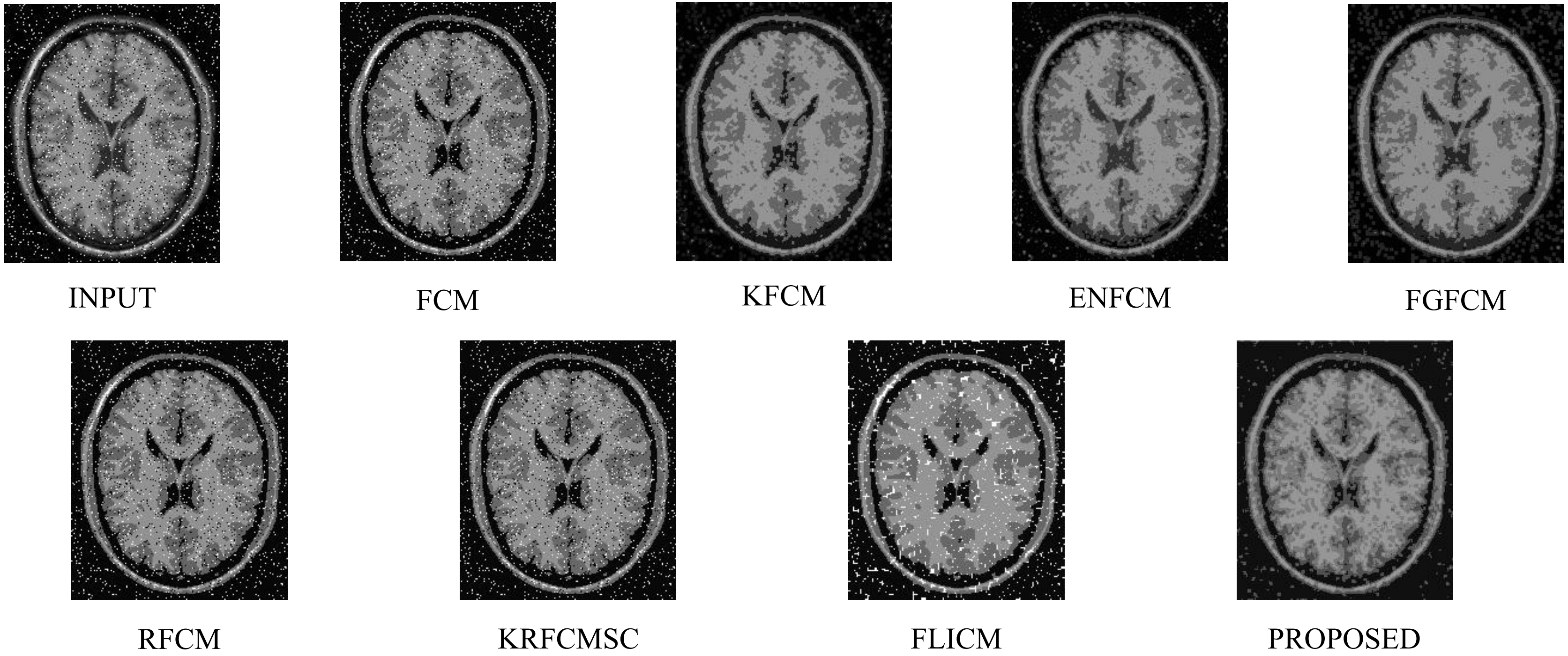
Initialize number of clusters(C), maximum iterations (M), fuzziness control parameter (q) $[q > 1]$, and threshold $T > 0$

Set initial membership values μ^0 is 0 and select C distinct data points as cluster centroids, set iteration counter (IC) as 1.

while $IC < M$ **do**
- Compute membership values μ_{ki} using Eqn. 12 for all pixels and all clusters
- **for** $i = 1 \ldots N$ **do**
  - Maximum Membership value is assigned to each pixel y_i as follows:
  - $\mu_{gi} = max(\mu_{ji})$
  - $g = j$ corresponding to max(μ_{ji})
  - **for** $j = 1 \ldots C$ **do**
    - **if** $|\mu_{ji} - \mu_{gi}| \leq T$ **then**
      - assign y_i to upper approximation spaces $\overline{R}(\xi_g)$ and $\overline{R}(\xi_j)$
    - **if** $y_i \notin$ *upper approximation spaces* **then**
      - assign y_i to lower approximation space $\underline{R}(\xi_j)$
- Compute cluster centroids based on Eqn. 14.
- **if** $|v_k^{/} - v_k| < T$ **then**
  - Stop
- $IC = IC + 1$

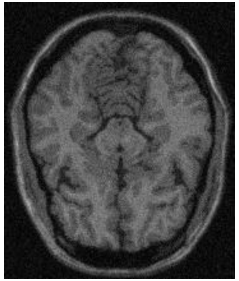
RESULT ANALYSIS

- Three different T1-weighted data volumes have been collected containing 51 images each.
- The first, second and third image volumes contain 7% gaussian noise, 8% salt and pepper noise, 9% gaussian noise with 40% inhomogeneity respectively.
- The performance of the proposed method is compared with the methods in literature.
- Quantitative metrics including partition coefficient, partition entropy and tissue segmentation accuracy are computed to analyze the performance of different methods.
- The results demonstrate the resiliency of the proposed method.
- The generated output along with the metric values defend the better performance of the presented method.

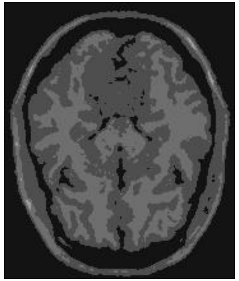


8% SALT & PEPPER NOISE CORRUPTED BRAIN MRI(F90 SLICE) SEGMENTED VIA DIFFERENT METHODS

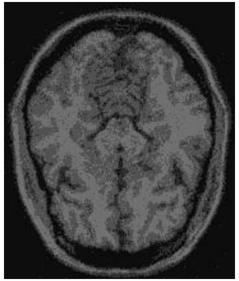
CVIP 2021
IIT ROPAR



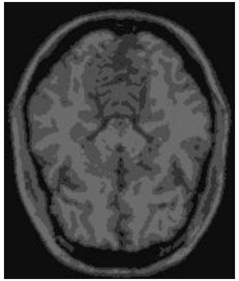
INPUT



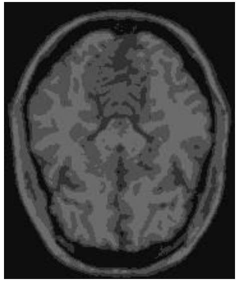
FCM



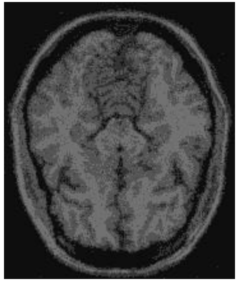
KFCM



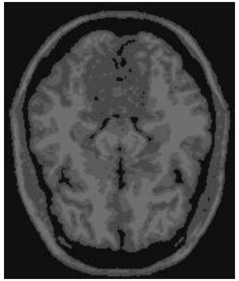
ENFCM



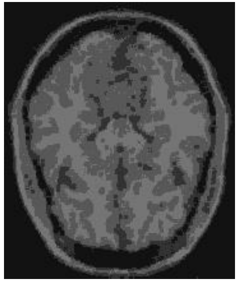
FGFCM



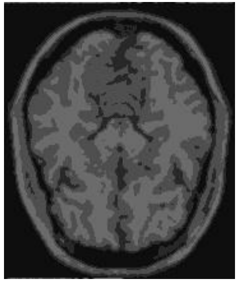
RFCM



KRFCMSC



FLICM



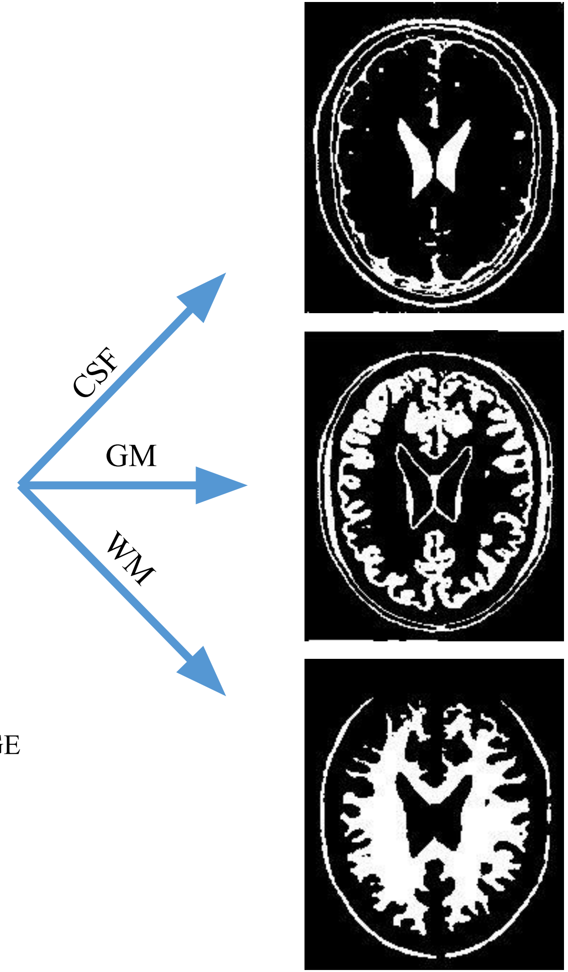
PROPOSED

9% GAUSSIAN NOISE CORRUPTED BRAIN MRI(F60 SLICE) WITH 40% INHOMOGENEITY SEGMENTED VIA DIFFERENT METHODS

BRAIN MRI - F95 SLICE
WITH 9% GAUSSIAN
NOISE & 40%
INHOMOGENEITY

AFTER
SEGMENTATION

SEGMENTED BRAIN IMAGE



CVIP 2021
IIT ROPAR

METHOD	50	55	60	65	70	75	80	85	90	95	100
FCM	0.8559	0.8558	0.8318	0.8476	0.8438	0.8505	0.8463	0.8557	0.865	0.8666	0.8731
KFCM	0.8036	0.8461	0.8375	0.8381	0.8341	0.8398	0.8289	0.8474	0.8566	0.8615	0.865
ENFCM	0.775	0.7877	0.7989	0.8008	0.7923	0.7949	0.7957	0.7993	0.8012	0.8174	0.8098
FGFCM	0.7707	0.7809	0.7921	0.7924	0.7905	0.7978	0.7955	0.8003	0.8083	0.8102	0.8155
RFCM	0.8101	0.8303	0.845	0.8351	0.8351	0.8393	0.8478	0.8579	0.8668	0.8515	0.8521
KRFCMSC	0.8093	0.8118	0.8283	0.8376	0.8294	0.8386	0.8392	0.8242	0.8607	0.861	0.8596
FLICM	0.8187	0.8374	0.8478	0.8488	0.8404	0.8478	0.8501	0.855	0.8612	0.8625	0.8705
MRKFCM	0.866	0.8695	0.8814	0.8825	0.8767	0.8748	0.8726	0.8708	0.8817	0.8788	0.8845

Table 1: Comparative Vpc values for different algorithms over 8% salt-and-pepper noise corrupted Brain MRI images

METHOD	50	55	60	65	70	75	80	85	90	95	100
FCM	0.2823	0.2834	0.3251	0.289	0.2963	0.2842	0.2921	0.2749	0.2576	0.2552	0.2423
KFCM	0.3774	0.3058	0.3146	0.3131	0.3211	0.3093	0.3414	0.297	0.2796	0.2695	0.2636
ENFCM	0.4234	0.4024	0.3821	0.3781	0.3908	0.384	0.384	0.377	0.3721	0.348	0.3556
FGFCM	0.4364	0.4182	0.3994	0.3979	0.4011	0.3878	0.3917	0.3836	0.3699	0.3649	0.3557
RFCM	0.3568	0.3216	0.2946	0.3185	0.3181	0.3094	0.291	0.2726	0.2558	0.2893	0.2833
KRFCMSC	0.3721	0.3595	0.3345	0.3221	0.3432	0.3164	0.3172	0.3488	0.2769	0.278	0.2761
FLICM	0.3452	0.3112	0.2913	0.289	0.3042	0.2913	0.2879	0.2787	0.2667	0.2647	0.2496
MRKFCM	0.2632	0.2582	0.2435	0.2431	0.2509	0.2517	0.2526	0.2532	0.2433	0.2499	0.2414

Table 2: Comparative Vpe values for different algorithms over 8% salt-and-pepper noise corrupted Brain MRI images

TISSUE SEGMENTATION ACCURACY

METHOD	CSF	WM	GM
FCM	75.01	75.76	65.69
KFCM	73.82	75.3	66.74
ENFCM	86.93	86.02	80.24
FGFCM	86.96	86.11	74.87
RFCM	83.61	82.3	75.87
KRFCMSC	85.64	85.91	75.98
FLICM	86.44	86.21	80.67
PROPOSED	87.42	88.34	81.95

Table 3: 7% Gaussian noise corrupted F-100 Brain MRI images

METHOD	CSF	WM	GM
FCM	86.58	86.57	79.86
KFCM	86.34	85.30	79.28
ENFCM	86.74	86.32	80.00
FGFCM	86.89	86.31	80.02
RFCM	86.34	85.30	79.28
KRFCMSC	86.68	86.48	79.89
FLICM	86.87	86.44	80.31
PROPOSED	87.08	86.21	79.38

Table 4: 8% Salt & Pepper noise corrupted F-90 Brain MRI images

METHOD	CSF	WM	GM
FCM	80.785	76.981	69.028
KFCM	86.903	83.003	80.543
ENFCM	77.702	83.614	78.171
FGFCM	77.623	84.006	77.979
RFCM	84.39	82.735	77.555
KRFCMSC	83.442	83.822	79.447
FLICM	80.677	81.908	78.69
PROPOSED	87.171	85.877	82.822

Table 5: 9% Gaussian noise corrupted F-95 Brain MRI images with 40% IIH

CONCLUSION

- Experimental results demonstrate the efficiency of the proposed method in the domain of segmentation.

- The algorithm is straightforward, and its wide scale application and resiliency makes it a rightful candidate to be integrated with hardware to form an embedded system for real time computer aided diagnosis.

- In future, the algorithm can be scaled with different pre and post processing techniques for better performance and for applications in domains of star region clustering, remote sensing and others.

THANK YOU