Tech Science Press

Unstructured Oncological Image Cluster Identification Using Improved Unsupervised Clustering Techniques

S. Sreedhar Kumar[1], Syed Thouheed Ahmed[2,*], Qin Xin[3], S. Sandeep[4], M. Madheswaran[5] and Syed Muzamil Basha[2]

[1]Dr. T. Thimmaiah Institute of Technology, VTU, KGF, Karnataka, India
[2]School of Computing & Information Technology, REVA University, Bengaluru, India
[3]Faculty of Science and Technology, University of the Faroe Islands, Faroe Islands, Denmark
[4]K S School of Engineering, Bengaluru, India
[5]Muthayammal Engineering College, Rasipuram, Tamil Nadu, India
*Corresponding Author: Syed Thouheed Ahmed. Email: syed.edu.in@gmail.com

Abstract: This paper presents, a new approach of Medical Image Pixels Clustering (MIPC), aims to trace the dissimilar patterns over the Magnetic Resonance (MR) image through the process of automatically identify the appropriate number of distinct clusters based on different improved unsupervised clustering schemes for enrichment, pattern predication and deeper investigation. The proposed MIPC consists of two stages: clustering and validation. In the clustering stage, the MIPC automatically identifies the distinct number of dissimilar clusters over the gray scale MR image based on three different improved unsupervised clustering schemes likely improved Limited Agglomerative Clustering (iLIAC), Dynamic Automatic Agglomerative Clustering (DAAC) and Optimum N-Means (ONM). In the second stage, the performance of MIPC approach is estimated by measuring Intra intimacy and Intra contrast of each individual cluster in the result of MR image based on proposed validation method namely Shreekum Intra Cluster Measure (SICM). Experimental results show that the MIPC approach is better suited for automatic identification of highly relative dissimilar clusters over the MR cancer images with higher Intra closeness and lower Intra contrast based on improved unsupervised clustering schemes.

Keywords: Magnetic resonance image; unsupervised clustering scheme; intra intimacy; intra contrast; iLIAC; shreekum intra cluster measure; medical image clustering

1 Introduction

Cluster based image segmentation is a significant and mathematical process in the MR image analysis system for deeper investigation, enhancement, tumor predication and pattern identification. Generally, it is defined as a process of dividing MR image pixels into different numbers of dissimilar sub regions based on pixel intensity similarity [1]. The goal of cluster based image separation is to

simplify or change the representation of an image into a version that is more meaningful and easier to investigate and identify. Recently, many of the researchers have been reported in [2], the cluster based segmentation process is applied in many medicine related application likely medical image segmentation, tumor or cancer predication, medical image enhancement, medical image compression, pattern identification, medical image classification and medical image retrieval. The result of the cluster based medical image separation is a finite number of dissimilar groups that jointly concealments the complete medical image and the quality of the clustering result depend on the superiority of the medical image quality. The major problem in the existing clustering schemes such as semi-supervised and unsupervised methods [3] is that to predetermine the appropriate number of clusters in the unstructured MR image pixel set and respectively the clustering quality is based on predetermined number of clusters. To overcome these issues, in this paper a new clustering technique called Medical Image Pixels Clustering, it intentions to automatically separate finite number of dissimilar patterns in the MR image based on different improved unsupervised clustering schemes without predetermined knowledge for deeper investigation, enhancement, pattern predication and analysis.

2 Literature Reviews

Several methods are available for cluster based MR image segmentation process including k-means, fuzzy C-means, neural network, fuzzy clustering and hierarchical clustering methods reported in [4–7]. The k-means technique is a semi-supervised partitioned clustering technique and is an iterative procedure that directly decomposes the MR image pixel set into many dissimilar clusters or regions by minimizing the criterion function (e.g., sum-of-square-error) [8]. Many of the authors suggested problem in the K-Means technique is that the entire segmentation result quality of MR image is based on predetermined k number of centroid pixel values. In [9], the authors Jianwei et al. have reported an improved K-Means technique MR brain image segmentation. The improved K-Means scheme is used to identify K distinct clusters over the disordered MR brain image with higher accuracy compared to existing scheme.

Another popular method called fuzzy c-means clustering (FCM) technique was reported in [10,11]. This method is suited to partition the noise-free image into a finest number of groups. Many researchers suggested that the drawback with this method is that it failed to segment images corrupted by noise or inaccurate edges. In [12] the authors Yogita et al. have reported a detail survey of fuzzy C-means (FCM) with intensity inhomogeneity correction and noise robustness. They are discussed how the FCM schemes is better suitable to identify distinct tissues such as cerebrospinal fluid, gray matter and white matter over the MR brain image. The authors Senthilkumar et al. [13] have presented a modified fuzzy C-means clustering scheme to identify the normal and abnormal tissues likely white matter, gray matter, cerebrospinal and tumor part respectively over the MRI brain image. The clustering scheme consists of pre-processing and segmentation stages. In the pre-processing stage, the authors are applied wrapping based curvelet transform over the MR brain image and removed the noise. Similarly, they are applied improved fuzzy C-Means technique [14,15] and segmented the normal and abnormal tumor cells over the MR brain image based on spatial information. In [16], the authors Jinn et al. have reported a hierarchical genetic algorithm with fuzzy learning vector quantization network to partition a multi-spectral MR brain image. The evaluation of this approach was based on a real case of a MR brain image of an individual suffering from meningioma.

The author's Chong et al. [17] have presented hybrid clustering scheme combined with morphological operations to improve the performance of MR image segmentation and reduced the non-brain tissue in the brain image. Firstly, the authors applied wiener filter and morphological operations over

the MR image due to remove the non-brain tissue. Next, they are used combination of K-Means++ and kernel-based fuzzy C-Means algorithm to identify distinct tumor regions in the MR image without noise. In [18,19], the authors Kalyanapu et al. have presented a clustering scheme namely unified iterative partitioned fuzzy clustering (U-IPFC). The U-IPFC scheme uses to identify distinct tissues over the MR brain image with good accuracy. The authors in [18,19] have claimed that the U-IPFC has produced higher accuracy result compared to FCM and K-means schemes. The authors Arul et al. in [20] presented a hierarchical clustering based segmentation (HCS) scheme to identify the distinct groups in hierarchy manner over the dynamic contrast enhanced magnetic resonance (DCSMR) image pixel set. The authors claimed that the HCS scheme is acted a semi-quantitative analytical tool to discover the DCEMR images. Next, the same authors Arul et al. in [21] have extended the detailed research of MR image segmentation based on hierarchical clustering scheme. The authors have experimented HCS scheme over the Multi-parametric Magnetic Resonance Imaging (MPMRI) and identified finite number of dissimilar tissue patterns by sequence of merging process. Another author Filipovych et al. in [22] reported hierarchical clustering scheme based image segmentation and it uses to identify predetermined number of dissimilar clusters in the tree manner over the gray scale image.

3 Proposed Image Pixel Clustering Approach

This section describes detailed study of the MIPC approach of image pixels classification. The MIPC scheme consists of two stages clustering and validation. The first stage automatically identifies the distinct number of highly relative clusters over the gray scale image dataset based on three different improved unsupervised clustering schemes iLIAC, DAAC and ONM in distinct manner. The second stage, it estimates the intra cluster intimacy and intra cluster contrast over the result of clustering stage based on the proposed SICM scheme. The stages involved in the MIPC approach are illustrated in the Fig. 1 and the different stages are described in below subsections.
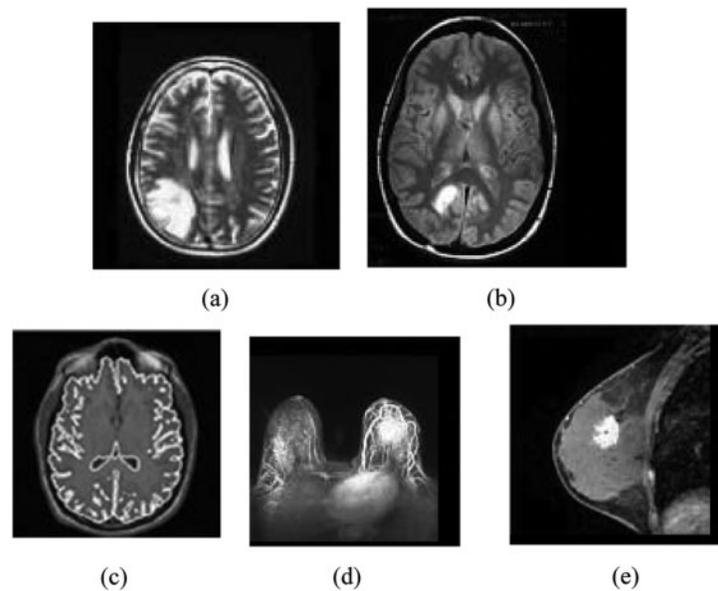
Figure 1: Original MR images: (a) Brain_1, (b) Brain_2, (c) Brain_3 (d) Breast_1 (e) Breast_2

3.1 Clustering Stage

This stage automatically identifies the distinct number of dissimilar clusters on the gray scale image based on three different improved clustering schemes iLIAC [23,24], DAAC [25] and ONM [26] in separate manner. Initially, the digital gray-scale image divides into $(2 * 2)$ sizes of non-overlapping blocks and the image contains n objects plus is defined as $X = x_i$, $x_i = x_{ij}$ for $i = 1, 2, \ldots, n$ and $j = 0, 1, 2, \ldots, d$, where X represents the dataset of MR image with n objects or blocks, x_i represents the ith object or block in dataset X, n denotes the size of MRI image dataset X, x_{ij} is the jth pixel value in i^{th} object in dataset X and d denotes the number of pixels belongs into the each individual block in dataset X. The MIPC approach identifies distinct clusters over the image dataset X using three different improved clustering schemes iLIAC, DAAC and ONM. The clustering schemes are described below subsections.

3.2 MIPC Using iLIAC Scheme

The MIPC approach identifies distinct number of dissimilar clusters over the MRI image dataset $X = x_i$ for $i = 1, 2, \ldots, n$ based on improved agglomerative clustering iLIAC scheme [23] and it consists of three stages feature extraction, control merge cost, clustering. In the feature extraction stage, the iLIAC scheme is extracted single feature value over each individual vector or block in the MR image vector set $X = x_i$ for $i = 1, 2, \ldots, n$ with d pixels $x_i = x_{ij}$ for $j = 0, 1, 2, \ldots, d$ based on statistical mean operation and is defined in the Eq. (1) as

$$\bar{x}_i = \left\{ \frac{\sum_{i=1}^{n} \sum_{j=1}^{d} x_{ij}}{d} \mid \forall x_{ij} \in x_i, \forall x_i \in X \right\} \tag{1}$$

where x_{ij} represents the jth pixel value in i^{th} object that belongs in to the vector set X and d denotes the number of pixel values in i^{th} object in X for $j = 1, 2, \ldots, d$. Next, it computes the control merge costs (ϕ) over the MRI image feature dataset $\bar{X} = \bar{x}_i$, for $i = 0, 1, .., n$ based on standard statistical function and is defined in the Eq. (2) as

$$\phi = (s_d(\bar{X}))^{1/2} \tag{2}$$

where, $s_d(\bar{X})$ denotes the standard deviation of MR image feature dataset $\bar{X} = \bar{x}_i$ and is defined in the Eq. (3) as:

$$s_d(\bar{X}) = \left\{ \left(\frac{1}{n} \left(\sum_{i=0}^{n} |\bar{x}_i - \mu|^2 \right) \right)^{1/2} \mid \forall \bar{x}_i \in \bar{X}, n > 1 \right\} \tag{3}$$

Here $\bar{x}_i$ represents the ith feature or representative value of ith object or block in MRI image dataset X, μ denotes the mean of dataset $\bar{X}$ as computed by

$$\mu = \left\{ \frac{1}{n} \sum_{i=0}^{n} \bar{x}_i \right\} \tag{4}$$

where $\bar{x}_i$ represents the ith object that belongs to the MRI image feature dataset $\bar{X}$ and n denotes the size of the input dataset $\bar{X}$ for $i = 0, 1, 2, \ldots, n$. In the clustering stage, the iLIAC scheme starts with each individual object in $\bar{X} = \bar{x}_i$ for $i = 0, 1, \ldots, n$ as an individual cluster. Firstly, it constructs the upper triangular distance matrix Ud_{ij} over the dataset $\bar{X}$ for $i = 0, 1, 2, \ldots, n, j = i + 1, \ldots, n$ and is

defined in Eq. (5) as

$$Ud_{ij} = d(\bar{x}_i, \bar{x}_j) \; |\forall \bar{x}_i, \bar{x}_j \in \bar{X} \atop {i=0,\ldots,n-1 \atop j=i+1,\ldots,n-1}} \tag{5}$$

where, $(d(\bar{x}_i, \bar{x}_j))$ is the Euclidean distance between i^{th} and j^{th} clusters that belong to the input cluster set $\bar{X}$ is defined as in Eq. (6), where $\bar{x}_i$ and $\bar{x}_j$ indicate i^{th} and j^{th} clusters in the cluster set $\bar{X}$. Subsequently, it identifies the closest cluster pair $(\bar{x}_i, \bar{x}_j)$ with a minimum merge cost Δd over the matrix Ud_{ij} which is defined a

$$d(\bar{x}_i, \bar{x}_j) = \left\{ ((\bar{x}_i - \bar{x}_i)^2)^{1/2} |\forall \bar{x}_i, \bar{x}_j, \in \bar{X} \right\} \tag{6}$$

$$\Delta d = \min_{\substack{i = 0, 1, \ldots, n-1, \\ j = i+1, \ldots, n-1}} \{Ud_{ij}\} \tag{7}$$

Next, the identified closest clusters pair $(\bar{x}_i, \bar{x}_j)$ with minimum merge cost Δd is compared with optimum merge cost. If the minimum merge cost Δd of cluster pair $(\bar{x}_i, \bar{x}_j)$ is lesser than control merges cost (ϕ) then it is merge the cluster pair $(\bar{x}_i, \bar{x}_j)$ into a single cluster $\bar{x}_{ij}$. Later it updates the merged cluster $\bar{x}_{ij}$ into $\bar{x}_i$ by standard statistical average method and is defined in Eq. (8) as

$$\bar{x}_i = \left\{ \frac{\bar{x}_i + \bar{x}_j}{2} \; |\bar{x}_i, \bar{x}_j \in \bar{X} \right\} \tag{8}$$

Then, updates the merged cluster $\bar{x}_i$ status by c_{ij} into c_i, where c_i denotes the status of the ith cluster and subsequently it modifies the size of merged cluster $\bar{x}_i$ by

$$N_i = N_i + N_j \tag{9}$$

where, N_i and N_j represent the number of related objects in i^{th} and j^{th} clusters respectively. After, deletes the jth cluster in the input cluster set X including its status c_j and size N_j respectively. Then, it reduces the input cluster set size to $\{n = n - 1\}$. The above process is repeated until the minimum merge cost of the cluster pair Δd exceeds the control merge cost (ϕ). Finally, the iLIAC produces appropriate number of distinct clusters in the cluster set C over the MR image vector set X and is defined as $C = c_l$, for $l = 0, 1, 2, \ldots, K$, where c_l denotes the lth cluster with N similar objects or blocks that belongs to the resulting cluster C and K represents the number of distinct clusters in the cluster set C for $l = 1, 2, \ldots, K$.

Algorithm 1: iLIAC

Input: MR Image dataset X with n Objects or Blocks
Output: Classification Result C with K clusters $\{c_1, c_2, \ldots, c_K\}$
     Begin
1. Represent each object in MR image dataset $X = x_i$ into single value $\bar{X} = \bar{x}_i$ using Eq. (1).
2. Calculate the optimal merge cost (ϕ) on X using Eqs. (2)–(4)
3. Consider each object in $\bar{X}$, $i = 1, 2.., n$ as a cluster in the image dataset
4.   Construct the upper triangular distance matrix Ud_{ij} on $\bar{X}$ using Eq. (5)
5.   Find the closest clusters pair $(\bar{x}_i, \bar{x}_j)$ with minimum merge cost ∇d on Ud_{ij} using Eq. (7)
6.   If $\Delta d < \varphi$ then the closest cluster pair $(\bar{x}_i, \bar{x}_j)$ is merged as a single cluster $\bar{x}_{ij}$
7.   Update the merged cluster $\bar{x}_{ij}$ into $\bar{x}_i$ using Eq. (8)

(Continued)

Algorithm 1: Continued

---

8.   Obtain the status of merged cluster $\bar{x}_i$ into its respective resulting cluster C_i
9.   Update the size of merged cluster $\bar{x}_i$ by $N_i = N_i + N_j$
10.   Delete the jth cluster in the input cluster set $\bar{X}$
11.   Reduce input cluster set size $\bar{X}$ by one $(n-1)$
12.   Repeat the steps from 4 to 9 until minimum merge cost ∇d not exceed control merge cost
13.   Obtain the distinct clusters result in cluster set C
 End

---

4 MIPC Using DAAC Scheme

Similarly, the MIPC approach is tested the same MRI image dataset $X = x_i$ for $i = 1, 2, \ldots, n$ using DAAC scheme [24]. It consists of two stages Distinct Representative Object Count (DROC) and Clustering. The DROC traces the count of distinct representative objects over the MRI image dataset $X = x_i$ based on occurrence of each individual object in dataset. It consists of three steps, in the first step, it represents the each object in the dataset $X = x_i$ for $i = 1, 2, \ldots, n$ with d features $f = 0, 1, \ldots, d$ into single value $\bar{X} = \bar{x}_i$ based on a statistical mean operation, where $\bar{x}_i$ is the representative value of ith object in MRI image dataset X and is defined in Eq. (10) as

$$\bar{x}_i = \left\{ \sum_{i=1}^{n} \frac{1}{d} \sum_{f=1}^{d} x_{if} \, |\forall x_{if} \in x_i, \; \forall x_i \in X \right\} \tag{10}$$

where x_{if} represents the fth feature in i^{th} object that belongs to the MR image dataset X. Next, the DROC scheme measures the tally of each object occurrence $COO(\bar{x}_i)$ in dataset $\bar{X} = \bar{x}_i$, for $i = 0, \ldots, n$ and is defined in Eq. (11) as:

$$COO(\bar{x}_i) = \sum_{j=i+1}^{n} |\bar{x}_i - \bar{x}_j| \, \left| \forall \bar{x}_i, \bar{x}_j \in \bar{X}, \; where \begin{cases} 1 & |\bar{x}_i - \bar{x}_j| < T \\ 0 & |\bar{x}_i - \bar{x}_j| > T \end{cases} \right. \tag{11}$$

where, $\bar{x}_j$ denotes the representative value of ith object that belongs to the MRI image dataset X, n denotes the size of $\bar{X}$ and T is the threshold value that limits the similarity between i^{th} and j^{th} representative values. If the difference of ith and j^{th} values is lesser than T, it means the jth value is similar to i^{th} value that belongs to the representative dataset $\bar{X}$. Finally, it estimates the sum of K distinct representative objects over the representative dataset $\bar{X}$ of MR image vector set X and is defined in Eq. (12) as

$$K = \left\{ \sum_{i=1}^{n} COV_i | \forall COV_i \in COV, \begin{cases} 1 & COV_i >= MO \\ 0 & COV_i < MO \end{cases} \right\} \tag{12}$$

Here, COV_i denotes the sum of occurrence of ith vector in X and MO represents the maximum occurrence threshold and it uses to limit the count of K distinct representative objects with maximum existence in the MRI image dataset X. In the clustering stage, first, it calculates the upper triangular distance matrix Ud_{ij} for input cluster set $X = x_i$ for $i = 1, 2, \ldots, n$ through Euclidean distance metric and it estimated by

$$Ud_{ij} = \left\{ \begin{array}{c} d(x_i, x_j) \\ i = 0, 1, \ldots, n \\ j = i+1, \ldots, n \end{array} \quad |\forall x_i, x_j \in x \right\} \tag{13}$$

where, n denotes the number of clusters in the input cluster set X and $d(x_i, x_j)$ is the Euclidean distance between i^{th} and j^{th} clusters in the cluster set X and is computed as

$$d(x_i, x_j) = \sqrt{\sum_{f=0}^{d} |x_{if} - x_{if}|^2} \tag{14}$$

In this, x_{il} denotes the fth feature in the ith cluster that belongs to the cluster set X and d represents the number of features in cluster $x_i = x_{il}$ for $f = 1, 2, \ldots, d$. Next, the DAAC scheme traces the adjoining clusters pair (x_i, x_j) with lowest merging cost ϖ on the distance matrix Ud_{ij} and is expressed in the Eq. (15) as:

$$\varpi = \begin{array}{c} Min \\ i = 0, 1, 2, \ldots, n, \\ j = i+1, \ldots, n \end{array} \{d(x_i, x_j) \ |\forall \ d(x_i, x_j) \in Ud_{ij}, \ \forall x_i, x_j \in X\} \tag{15}$$

where, $d(x_i, x_j)$ denotes the Euclidean distance between i^{th} and j^{th} MR image vectors in the MR image dataset or vector set (X). The Eq. (15) finds the adjoining clusters pair (x_i, x_j) with lowest merge cost ϖ and then compare the number of clusters does not exceed the sum of representative value K. If the number of clusters i is not exceed the K, then the adjoining cluster pair (x_i, x_j) is combined into a same cluster x_{ij} which subsequently computes the centroid over the new cluster x_i using Eq. (16) and is defined as:

$$x_i = \left\{ \sum_{f=1}^{d} \frac{1}{2}(x_{if} + x_{jf}) \ |\forall x_{if} \in x_i, \ x_{jf} \in x_j \right\} \tag{16}$$

Next, updates the combined cluster x_i status into respective c_i through $c_i \cup c_j \rightarrow c_i$, where c_i denotes the status of the ith cluster and subsequently it modifies the size of combined cluster x_i by $m_i \cup m_j \rightarrow m_i$, where, m_i and m_j represent number of related objects in i^{th} and j^{th} clusters respectively. After, it removes the jth cluster in the input cluster set X including its status C_j and size N_j respectively and reduces the input cluster set size by one. The above process is repeated until the number of dissimilar clusters in the cluster set is equal to K and afterward the results with K district clusters are defined as $\{c_1, c_2, \ldots, c_K\}$.

5 MIPC Using ONM Scheme

Similarly, in this subsection, the MIPC approach is partitioned the MRI image dataset into distinct number of different clusters based on improved partitioned clustering ONM scheme [25,26]. It consists of two stages likely dissimilar spatial centroid vector (DSCV) and partitioning respectively. In the DSCV stage, the ONM approach identifies the distinct number of centroid vectors over input MRI image vector set $X = x_i$ based on occurrence of objects in the dataset X. First, it computes rate of repetition of each spatial vector $OV(X_i)$ over the dataset $X = x_i$, for $i = 0, \ldots, n$ and is defined in Eq. (17) as:

$$OV(X_i) = \sum_{j=i+1}^{n} |x_i - x_j| \ \bigg| \ \forall x_i, x_j \in X, \ where \ \left\{ \begin{array}{ll} 1 & |x_i - x_j| < T \\ 0 & |x_i - x_j| > T \end{array} \right\} \tag{17}$$

Algorithm 2: DAAC

Input: MR Image Vector set X Containing n Vectors $x_0, x_1, \ldots, x_n$ with d pixels and Threshold (MO)
Output: Generate K Distinct Clusters $C = \{c_1, c_2, \ldots, c_K\}$
  Begin
  1. Represent each object in MRI image dataset X into single dimensional $\bar{X}$ using Eq. (10)
  2. Measure the count of occurrence of each individual vector $COV(\bar{X}_i)$ in $\bar{X} = \bar{x}_i$ for
     $i = 0, 1, 2, \ldots, n$ as described in Eq. (11)
  3. Identify representative objects in X based on count of object occurrences $COV(\bar{X}_i)$ and threshold
     MO as described in Eq. (12)
  4. Count (sum) the number of representative vectors in X using Eq. (12) and obtain the count in
     N
  5. Consider each vector as an individual cluster in the input dataset $X = x_i$ for $i = 0, 1, \ldots, n$
  6. Compute the upper triangular matrix Ud_{ij} as given in Eq. (13).
  7. Find the closest clusters pairs (x_i, x_j) with minimum merge cost ϖ over Ud_{ij} as given in Eq. (15).
  8. Merge the closest cluster pairs (x_i, x_j) into single cluster x_{ij} as described in Eq. (16)
  9. Update the newly merged cluster x_{ij} into x_i as described in Eq. (16)
  10. Update the status of newly merged cluster x_i in c_i by $c_i \cup c_j \rightarrow c_i$
  11. Update the size of newly merged cluster by $m_i = m_i + m_j$
  12. Delete j^{th} cluster (x_i), cluster status (c_j) and its size (m_j) respectively.
  13. Reduce X size by one.
  14. Repeat steps 6 to 13 until the size of the cluster set n is equal to K
  15. Obtain the final clustering result in C
  End

---

where, x_i and x_j represent i^{th} and j^{th} vectors that belongs in to the MR image vector set X, n denotes the size of X and T is the threshold that limit the similarity distance between i^{th} and j^{th} vectors. If the difference of ith and j^{th} objects is lesser than T, it means that the jth object is similar to i^{th} object or vector that belongs to the MR image dataset X. In the second step, it finds the distinct number of different Centroid Vector (CV) in dataset X based on object occurrence $OV(x_i)$ and is computed by

$$CV = \left\{ \left\{ \begin{matrix} OV(x_i) \\ i = 1, \ldots, n \end{matrix} \right\} \forall OV_i \in OV, \quad where \; \{x_i \quad OV_i \geq CC\} \right\} \tag{18}$$

In this, OV_i denotes the rate of occurrence of ith vector in X and CC represents the Control Centroid that intends to dynamically identify the appropriate number of spatial centroid vector in MRI image dataset X and is determined in form of $CV = CV_l$, for $l = 1, \ldots, N, f = 1, 2, \ldots, d$ and $l = 1, \ldots, N$, where, CV_l is in the partitioning stage, the ONM approach divides the MR image vector set into optimum number of N discrete clusters based on distinct centroid vectors. The clustering stage consists of three steps. In the first step, it measures the distance of each individual vector in vector set X over the N centroid vectors in $CV = CV_l$ for $l = 1, 2, \ldots, N$ and $f = 0, 1, \ldots, d$ based on Euclidean distance and is defined in Eq. (19) as

$$D(X, CV) = \{d(x_i, CV) | \forall x_i \in X, \forall CV_l \in CV\} \tag{19}$$

where, $d(x_i, CV_l)$ represents the Euclidean distance between i^{th} vector in X and l^{th} centroid in CV and is computed by

$$d(x_i, CO_l) = \left\{ \left(\sum_{f=1}^{D} (x_{if} - CO_{lf})^2 \right)^{1/2} \right\} \tag{20}$$

Here, x_{if} denotes the fth feature of ith vector in X and CV_{lf} represents the fth feature of lth centroid vector. Second step, it finds the closest centroid vector of each individual object in dataset $X = x_i$ with minimum Euclidean distance which computed at step 1 and respectively it assign the ith object in X into its closest l^{th} cluster in cluster set $C = c_l$ for $l = 0, 1, \ldots, N$ and is defined in Eq. (21) as

$$C_l = \left\{ \min \left\{ \frac{D(x_i, CV_l)}{l = 0, 1, \ldots, N} \right\} \mid \forall x_i \in X \right\} \tag{21}$$

In the last step, it modifies the centroid of each individual cluster in cluster set $C = c_l$, for $l = 0, \ldots, N$ and $c_l = c_{lj}$ for $j = 0, 1, \ldots, R$ and is defined in Eq. (22) as:

$$CV_l = \left\{ \frac{1}{R_l} \sum_{j=0}^{R_l} c_{lj} \mid \forall c_{li} \in c_l, \forall c_l \in C \right\} \tag{22}$$

In this, c_{ij} denotes the jth object in l^{th} cluster in cluster set C and R_l is the size of lth cluster in cluster set C.

Algorithm 3: ONM

Input: MR Image vector set X Containing n vectors $x_0, x_1, \ldots, x_n$ with d features and Threshold (CC)
Output: Cluster set C Containing N Clusters $\{c_1, c_2, \ldots, c_N\}$

    Begin
1. Measures the occurrence of each vector $OV(x_i)$ in X as described in Eq. (18)
2. Find distinct number of centroid vector $CV = CV_l$ for $l = 0, \ldots, N$ on input dataset X based on object occurrence $OV(x_i)$ and Control Centroid (CC) as expressed in Eqs. (17) and (18)
3. Measure the distance of X over the N distinct centroids $CV = CV_l$ for $l = 0, \ldots, N$ based on Euclidean distance as described in Eqs. (19) and (20)
4. Divide the input dataset X into distinct number of clusters $C = c_l$ for $l = 0, \ldots, K$ based on distinct number of centroid objects by using Eq. (21).
5. Update the centroids in $CV = CV_l$ by using Eq. (22).
6. Repeat the steps from 4 to 5 until current iteration result is similar to previous iteration result.
7. Obtain the clustering result in C.

End

---

6 Cluster Validation Stage

This stage presents, the MIPC scheme estimates the closeness and separation among the data objects in each individual cluster in the cluster set of MR image vector set based on proposed cluster validation scheme (SICM). The proposed (SICM) is an improved version of existing validation techniques as reported in [27–29] and it aims to validate the quality of each individual cluster in the cluster set of MR image that identified by MIPC scheme based on probability concept. The SICM consists of two measures Intra Intimacy (II) and Intra Contrast (IC). The II measure uses to estimate the closeness of each individual vector with other vectors in the same cluster $OC(c_{li})$, where, c_{li} represents the ith object in the lth cluster in cluster set C with K clusters and the vector closeness

$VC(c_{li})$ measure is defined in the Eq. (23) as

$$
VC(c_i) = \begin{cases} \left\{ \frac{1}{|c_l|} \sum_{i=1}^{|c_l|} \left\{ \left\{ \frac{1}{|c_l|} \sum_{j=1}^{|C_l|} \sum_{f=1}^{4} |c_{lif} - c_{ljf}| \right\} \times 100 \right\} \right\}, \forall c_{lif} \in c_{li}, \forall c_{lj} \in c_l, \forall c_l \in C \\ where \begin{cases} 1 & |c_{lif} - c_{ljf}| <= \theta \\ 0 & |c_{lif} - c_{ljf}| > \theta \end{cases}, \begin{cases} 1 & \sum_{f=1}^{4} |c_{lif} - c_{ljf}| >= 2 \\ 0 & \sum_{f=1}^{4} |c_{lif} - c_{ljf}| < 2 \end{cases} \end{cases} \tag{23}
$$

where, c_{lif} is the fth pixel value in j^{th} vector in the lth cluster that belongs into the cluster set C for $l = 0, 1, 2, \ldots, K$, $|c_l|$ is the size of the lth cluster for $j = 0, 1, 2, \ldots, N$, θ denotes the predetermined threshold or constant that uses to limit the difference between two objects. Next, the IC calculates the overall intra cluster intimacy ICI among the cluster set C based on individual cluster closeness $VC(c_{li})$ within the same cluster set and is defined in the Eq. (24) as

$$
ICI(C) = \left\{ \frac{1}{K} \sum_{l=0}^{K} VC(c_l) \right\} \tag{24}
$$

Similarly, the intra contrast measure aims to estimate the intra disparity among the vectors within the same cluster in the cluster set. First, it measures the intra disparity $VD(c_l)$ of each individual vector $c_l = c_{lj}$ for $j = 0, 1, 2, \ldots, N$ with other vectors within the same cluster in the cluster set $C = c_l$ for $l = 0, 1, 2, \ldots, K$ and it defined in the below given Eq. (25) as:

$$
VD(c_i) = \begin{cases} \left\{ \frac{1}{|c_l|} \sum_{i=1}^{|c_l|} \left\{ \left\{ \frac{1}{|c_l|} \sum_{j=1}^{|C_l|} \sum_{f=1}^{4} |c_{lif} - c_{ljf}| \right\} \times 100 \right\} \right\}, \forall c_{lif} \in c_{li}, \forall c_{lj} \in c_l, \forall c_l \in C \\ where \begin{cases} 1 & |c_{lif} - c_{ljf}| <= \theta \\ 0 & |c_{lif} - c_{ljf}| > \theta \end{cases}, \begin{cases} 1 & \sum_{f=1}^{4} |c_{lif} - c_{ljf}| >= 2 \\ 0 & \sum_{f=1}^{4} |c_{lif} - c_{ljf}| < 2 \end{cases} \end{cases} \tag{25}
$$

Subsequently, the IC measure estimates the overall intra cluster contrast $ICC(C)$ over the cluster set C with K distinct clusters based on intra vector disparity $VD(c_l)$ of each individual cluster in the cluster set $C = c_l$ for $l = 0, 1, 2, \ldots, K$ and is computed by,

$$
ICC(C) = \left\{ \frac{1}{K} \sum_{l=0}^{K} VD(c_l) \right\} \tag{26}
$$

Algorithm 4: SICM

Input: Resulting Cluster C Containing K Distinct Cluster $C = \{c_1, c_2, \ldots, c_K\}$
Output: Overall Intra Cluster Intimacy $ICI(C)$ and Intra Cluster Contrast $ICC(C)$
   Begin
    1. Compute the closeness of each individual vector $VC(c_{li})$ with other vectors $c_l = c_{lj}, j = 0, 1, 2, \ldots, N$ in the same cluster c_l as expressed in Eq. (24)
    2. Evaluate the overall intra cluster closeness of resulting cluster C based on $VC(c_l)$ for $l = 1, 2, \ldots, K$ using Eq. (25) and the result is obtained in $ICI(C)$.

(Continued)

3. Compute the intra disparity $OD(c_{li})$ of each individual vector c_{ii} with other objects in the same cluster c_l in the cluster set $C = c_l$ for $l = 1, \ldots, K$ based on Eq. (26)

4. Calculate the overall inter cluster contrast $ICC(C)$ of resulting cluster $C = c_l$ with K distinct cluster based on intra disparity of each individual cluster $OD(c_{li})$ in the cluster set C using Eq. (26).

End

7 Complexity Analysis

This section discovers the computational complexity of MIPC approach has tested over MR image dataset by three different improved unsupervised clustering schemes namely iLIAC, DAAC and ONM. The MIPC system consumes time $O(nd)$ to split the digital MR image X into n non overlapping blocks or vectors with d pixels, where n is the number of vectors or blocks or vectors in the input digital MR image vector set X and is describes as $X = x_i$ for $i = 0, 1, 2, \ldots, n$, $x_i = x_{if}$ for $f = 0, 1, 2, \ldots, d$. Ahmed et al. [30] have presented automatic segmentation and detection of brain tumor is a notoriously complicated issue in magnetic methods are limited for detection of tumor in multimodal brain MRI. This work analyses the segmentation performance of existing state of art method improved Fuzzy C-Means clustering (FCMC) method and marker-controlled watershed method to carry out accurate brain tumor detection and enhance the segmentation results. Next, the complexity analysis of MIPC system is performing in the clustering stage including different clustering schemes iLIAC, DAAC and ONM respectively as described in the below.

7.1 MIPC (iLIAC)

First, it requires time $O(nd)$ to extract the single feature over each individual vector or block in the MR image vector set $X = x_i$ with n vectors based on Eq. (1) and the extracted features are obtained in dataset $\bar{X} = \bar{x}_i$ for $i = 0, 1, 2, \ldots, n$. Next, it consumes $O(n)$ time to compute the control merge cost (ϕ) over the MR image feature dataset ($\bar{X}$) with n data elements. Afterward, in the every iteration the iLIAC clustering scheme needs time $O((n(n-1)/2) + 1 + 1)$ to construct upper triangular distance matrix $Ud(\bar{X})$ over the cluster set ($\bar{X}$) with n clusters, identifies closest cluster pair (x_i, x_j) and update the cluster set ($\bar{X}$) respectively. The MIPC system needs time $O((n(n-1)/2) + 1 + 1)$ for $(n - K)$ iterations to identify the appropriate number of dissimilar clusters based on iLIAC scheme without user input. Overall the MIPC (iLIAC) system consumes time $O((((n(n-1))/2) + 1 + 1)(n - K) + (nd))$ to process and identifies applicable number (K) of dissimilar clusters over the MR image vector set ($\bar{X}$).

7.2 MIPC (DAAC)

In the first stage, the (DAAC) clustering scheme needs time $O(nK)$ to identify number of distinct representative objects over the MR image feature set $\bar{X} = \bar{x}_i$ with n objects based on DROC method, where, K is the number of representative objects in image feature set ($\bar{X}$). Next stage, it consumptions time $O((n(n-1)/2) + 1 + 1)$ to build upper triangular matrix over the MR image vector set X, identifies closest vector pair (x_i, x_j) with higher similarity and update the vector set X. Overall the MIPC (DAAC) scheme is required time $O(((n(n-1)/2) + 1 + 1)(n - K) + (nd) + (nK))$ to identify finest number (K) of dissimilar clusters that belongs into the MR image vector set X without predetermined knowledge, where, $(n - K)$ is the number of iterations.

7.3 MIPC (ONM)

Initially, the (ONM) clustering scheme consumptions $O(ndK)$ time to identify appropriate number of dissimilar centroid vectors over the MR image vector set $X = x_i$, $x_i = x_{ij}$, for $i = 0, 1, 2, \ldots, n$ and $j = 0, 1, \ldots, d$ based on DCV method, where, K is the number of centroid vectors that belongs in to the vector set X. In the partitioning stage, the ONM scheme takes time $O(ndKr)$ to iteratively split the MR image vector set X into finest number of K distinct highly relative clusters, where, r is the number of iterations. As a whole, the MIPC (DAAC) system has required time $O(ndKr + ndK)$ to identify finest number (K) of dissimilar clusters that belongs into the MR image vector set X.

8 Results & Discussions

This section presents the MIPC approach, experimented on MR gray scale medical images based on three different improved unsupervised clustering schemes iLIAC, DAAC and ONM respectively. For the experimental purpose, we have taken 100 natural 100 2-D gray scale MR medical images with different sizes such as $(120 * 120)$, $(124 * 124)$ and $(130 * 130)$ respectively and the grey values in the range 0–255.

A subset of this dataset containing ten sample standard MR brain and breast images via, Brain_1, Brain_2, Brain_3, Breast_1 and Breast_2 are reported as representative in this subsection. The sample MRI images are used in many research experiments as reported in (Lai & Huang 2011; Qi et al. 2015; Yong & Shuying 2007). Fig. 1 shows the five standard MRI gray scale images Brain_1, Brain_2, Brain_3, Breast_1 and Breast_2 as illustrated in Figs. 2a–2e respectively. In this experiment, each block of size $(2 * 2)$ is considered as a vector and hence each sample image contains 3844, 4225, 3600, 3844 and 4225 vectors respectively.
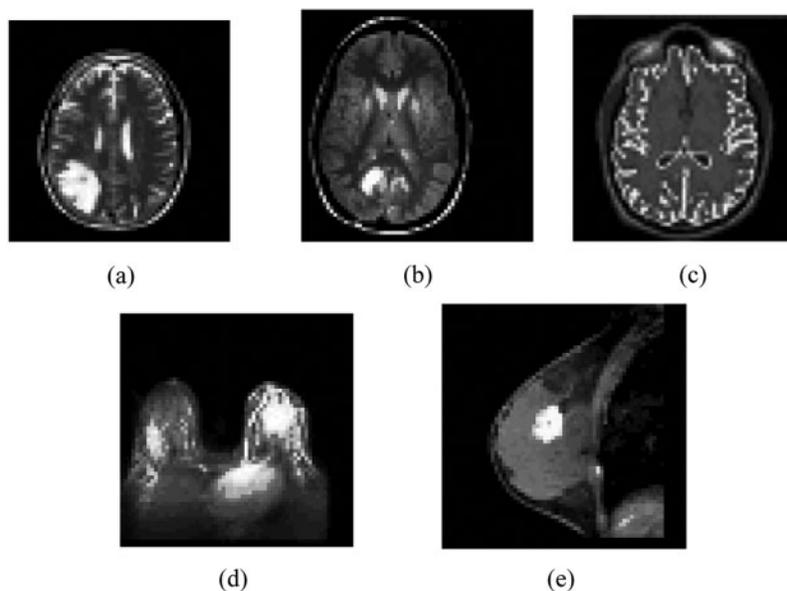
Figure 2: Result of the MIPC scheme tested on the ten gray scale images using iLIAC approach indicated in Fig. 1: (a) Result of brain_1 (b) Result of brain_2 (c) Result of brain_3 (d) Result of breast_1 (e) Result of breast_2

Firstly, the MIPC approach identifies distinct number of dissimilar clusters over the seven gray scale medical image datasets based on iLIAC scheme. Initially, it computes the control merge cost over seven gray scale MR images and the results are obtained in Tab. 1 as 7.87, 7.51, 7.71, 7.85, 7.44 respectively. Then it followed by computation of upper triangular distance matrix and in the case of sample gray scale MRI image datasets are presented in Fig. 2. The clustering scheme could identify 24, 25, 24, 25 and 25 distinct clusters over the MRI images in the Fig. 2. The results are incorporated in the Tab. 1. Fig. 3 demonstrates the clustering result of the iLIAC scheme has tested the MRI images likely Brain_1, Brain_2, Brain_3, Breast_1 and Breast_2 as obtained in Figs. 2a–2e respectively.

Table 1: Result of MIPC scheme tested on seven gray scale MRl images using iLIAC clustering algorithm

Sample MR images	Result of MIPC (iLIAC) scheme	
	Control merge cost	Number of distinct clusters identified
Brain_1	7.87	24
Brain_2	7.51	25
Brain_3	7.71	24
Breast_1	7.85	25
Breast_2	7.44	25
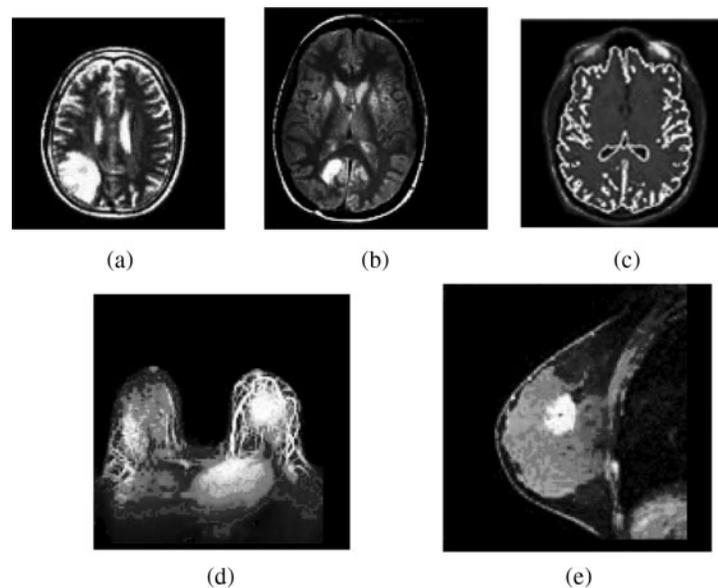
Figure 3: Result of the MIPC scheme tested on the ten gray scale MR images using DAAC approach indicated in Fig. 2: (a) Result of brain_1 (b) Result of brain_2 (c) Result of brain_3 (d) Result of breast_1 (e) Result of breast_2

Similarly, the MIPC approach detects distinct number of unrelated clusters on same five MR image datasets based on DAAC scheme. Primarily, it automatically traces the distinct representative objects over the five MR images as illustrated in Fig. 2 based on frequency of maximum occurrence (MO = 15) and the count of distinct representative objects are obtained in Tab. 2 as 33, 27, 33, 39, 27 respectively. The Maximum Occurrence is a predetermined threshold which used to dynamically find the appropriate number of distinct representative objects in dataset. Then it followed by sequence of merging process and divides the each individual image dataset into distinct number of dissimilar clusters based on count of representative objects as presented in Tab. 3. In the case of sample gray scale image datasets presented in Fig. 3, the clustering scheme could identify 33, 27, 33, 39 and 27 distinct clusters. The resulting clusters of the clustering scheme are incorporated in the Tab. 2. Fig. 3 demonstrates the clustering result of the MIPC (DAAC) on five gray scale MR images Brain_1, Brain_2, Brain_3, Breast_1 and Breast_2 as obtained in Figs. 3a–3e, 3 respectively.

Table 2: Result of MIPC (DAAC) scheme tested on five gray scale MR images

Sample MR images	Result of MIPC (DAAC) scheme with (MO = 15)	
	Count of representative objects (N)	Number of distinct clusters identified (N)
Brain_1	33	33
Brain_2	27	27
Brain_3	33	33
Breast_1	39	39
Breast_2	27	27

Table 3: Result of MIPC (ONM) scheme tested on five MR images

Sample MR images	Result of MIPC (ONM) scheme with (CC = 15)	
	Number of spatial centroid objects	Number of distinct clusters identified
Brain_1	33	33
Brain_2	27	27
Brain_3	33	33
Breast_1	39	39
Breast_2	27	27

In the same way, the MIPC approach divides the MR image dataset into distinct number of discrete clusters based on ONM scheme. In the beginning, it robotically traces the distinct number spatial centroid objects on each individual gray scale MR image dataset based on control centroid (CC = 15) and the results are incorporated in Tab. 3. The Control Centroid (CC) is a user defined threshold that is used to generate the spatial centroid objects in dataset dynamically. Then it followed by iterative process and divides the each individual image dataset into distinct number of dissimilar clusters based

on spatial centroid objects as presented in Tab. 3. The resulting clusters of the five gray scale MR images are incorporated in the Tab. 3. Fig. 4 demonstrates the clustering result of the MIPC (ONM) on five gray scale medical images Brain_1, Brain_2, Brain_3, Breast_1 and Breast_2 as obtained in Figs. 4a–4e respectively.
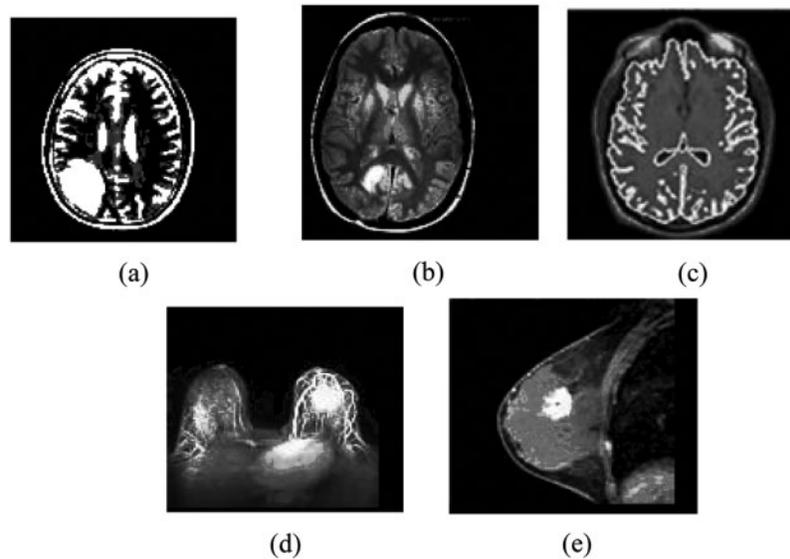
Figure 4: Result of the MIPC scheme tested on the ten MR images using DAAC approach indicated in Fig. 2: (a) Result of brain_1 (b) Result of brain_2 (c) Result of brain_3 (d) Result of breast_1 (e) Result of breast_2

The performance of the MIPC approach with three improved clustering schemes has been validated based on improved SICM schemes. It calculates the intra intimacy and intra cluster contrast over the each individual cluster in cluster set of MR images which tested by MIPC approach and the clustering results as shown in Tabs. 1–3 respectively. Initially, it measures the size of each individual cluster over the results of the five gray scale medical images Brain_1, Brain_2, Brain_3, Breast_1 and Breast_2 respectively. Next, it estimates the intra closeness (OC) and intra disparity (OD) in % among the individual cluster of these sample medical image datasets results based on the centroid of the each individual cluster.

Then, it followed to calculate the overall intra intimacy $ICI(C)$ in % over the results of the MIPC approach with three different clustering schemes iLIAC, DAAC and ONM respectively. Subsequently, it produced 60.06, 56.43, 53.37, 73.39, 77.92; 77.28, 88.27, 77.27, 82.51, 85.39; 72.14, 73.58, 70.215, 79.17, 83.39 for the sample gray scale image datasets Brain_1, Brain_2, Brain_3, Breast_1 and Breast_2 respectively. The estimated results of sample medical image datasets as obtained in Tab. 4. Similarly, the overall intra cluster contras $ICC(C)$ is calculated over the clustering results of MR images which obtained by MIPC scheme based on intra disparity measures.

The validation results of MR images which tested by iLIAC, DAAC and ONM clustering schemes are obtained in Tab. 5 as 39.93, 43.56, 46.62, 26.60, 22.075; 22.71, 11.72, 22.72, 17.48, 14.60 and 27.85, 26.41, 29.78, 20.82, 16.60 respectively. It is clearly shown in the performance measurement results as illustrated in Figs. 4, 5, and 6 that the proposed SICM has flawlessly estimated intra cluster intimacy and intra cluster contrast over the result of MR cancer image. Accordingly to the

performance measurement results, that the DAAC clustering schemes has identified appropriate number of dissimilar groups (Normal & Abnormal regions) over the MR cancer images with good accuracy compared to ONM and iLIAC schemes without predetermined input. Similarly, the ONM scheme has produced better clustering results with higher intra closeness and lower intra contrast compared to iLIAC scheme.

Table 4: Comparison of intra closeness measures among results of MR images with tested by iLIAC, DAAC and ONM clustering schemes

MR image datasets	Overall intra cluster intimacy measure $ICI(C)$ in %		
	MIPC (iLIAC)	MIPC (DAAC)	MIPC (ONM)
Brain_1	60.06	77.28	72.145
Brain_2	56.43	88.27	73.58
Brain_3	53.37	77.27	70.215
Breast_1	73.39	82.51	79.17
Breast_2	77.92	85.39	83.89

Table 5: Comparison of intra separation measures among results of MR images with tested by with iLIAC, DAAC and ONM clustering schemes

MR image datasets	Overall intra contrast measure $ICC(C)$ in %		
	MIPC (iLIAC)	MIPC (DAAC)	MIPC (ONM)
Brain_1	39.93	22.71	27.85
Brain_2	43.56	11.72	26.41
Brain_3	46.62	22.72	29.78
Breast_1	26.60	17.48	20.82
Breast_2	22.075	14.60	16.60
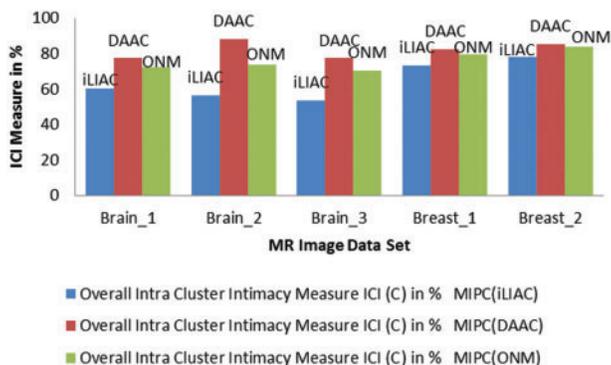
Figure 5: Comparisons of (ICI) performance measure over clustering results of MR images tested by improved unsupervised clustering schemes iLIAC, DAAC and ONM
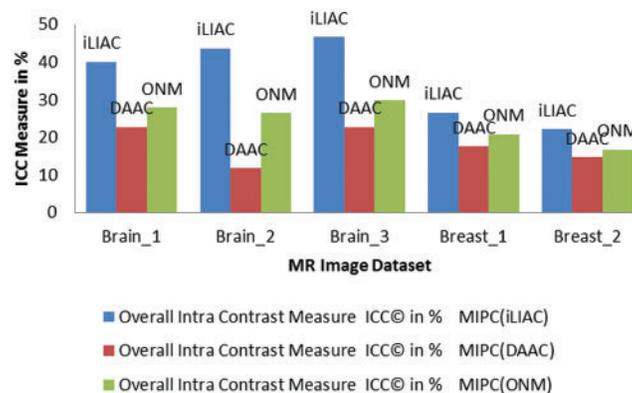
Figure 6: Evaluations of (ICC) performance measure over clustering results of MR images tested by improved unsupervised clustering schemes iLIAC, DAAC and ONM

9 Conclusion

This article presents Inherent Image Pixels Classification using three different improved unsupervised clustering schemes iLIAC, DAAC and ONM. The MIPC approach is aimed to trace the dissimilar pattern over the gray scale medical image through automatic identification of the distinct number of highly relative clusters in the medical image dataset based on improved unsupervised cluster schemes for deeper investigation and analysis. First, the MIPC approach automatically identifies the distinct number of dissimilar clusters over the medical image dataset based on three different clustering schemes iLIAC, DAAC and ONM in the separate manner. Next, the results of the MR images are validated based on proposed SICM scheme. We tested the MIPC approach with three improved unsupervised clustering schemes on five gray scale cancer MR images likely Brain_1, Brain_2, Brain_3, Breast_1 and Breast_2. According to the experimental results, the MIPC approach is more efficient and effective for automatic identification of the maximum number of highly relative clusters including normal and abnormal regions over the gray-scale MR cancer image with higher intra intimacy and lower intra contrast. After conducting various experiments, we concluded that the MIPC approach is better suitable to identify appropriate number of dissimilar regions (normal & abnormal), improving clusters quality and validate the clustering result for plateful to investigate (normal & abnormal regions) the dissimilar patterns in the MR cancer images.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1]   https://en.wikipedia.org/wiki/Medical_image_computing#Clustering, 2021.
[2]   https://en.wikipedia.org/wiki/Image_segmentation, 2021.

[3]   S. Sreedhar Kumar and M. Madheswaran, "A brief survey of unsupervised agglomerative hierarchical clustering schemes," *International Journal of Engineering & Technology (UAE)*, vol. 8, no. 1, pp. 29–37, 2019.

[4]   S. Sreedhar Kumar, M. Deepak and P. Karthik, "Reconstruction of MR image using sparse signal sequences in frequency domain," *International Journal of Innovating Technology and Exploring Engineering (IJITEE)*, vol. 9, no. 3, pp. 895–902, 2020.

[5]   K. Dhawan, *Medical Image Analysis*. Wiley Inter-Science Publications, 2003.

[6]   J. Alfredo, F. Costa, C. Jackson and G. De Souza, "Image segmentation through clustering based on natural computing techniques," in *Image Segmentation, Dr. Pei-Gee Ho (Edition)*, In Tech, 2011.

[7]   A. Jain, N. Murty and J. Flynn, "Data clustering: A review," *ACM Computer Surveys*, vol. 31, no. 3, pp. 264–323, 1999.

[8]   L. Davies and W. Bouldin, "Cluster separation measure," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 1, no. 2, pp. 95–105, 1979.

[9]   L. Jianwei and L. Guo, "An improved k-means algorithm for brain image segmentation," in *3rd Int. Conf. on Mechatroincs, Robotics and Automation (ICMRA 2015)*, Atlantis Press, Netherland, pp. 1087–1090, 2015.

[10] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press, 1981.

[11] J. Bezdek, R. Ehrlich and F. William, "FCM: The fuzzy c-means clustering algorithm," *Computers and Geosciences*, vol. 10, pp. 191–203, 1984.

[12] K. Yogita, M. Milind and M. Mushrif, "FCM clustering algorithm for segmentation of brain MR images," *Advances in Fuzzy Systems*, vol. 12, no. 3, pp. 1–14, 2016.

[13] C. Senthilkumar and R. Gnanamurthy, "A fuzzy clustering based MRI brain image segmentation using back propagation neural networks," *Cluster Computing*, vol. 22, no. 5, pp. 12305–12312, 2019.

[14] J. Liu, M. Li, J. Wang, F. Wu and T. Liu, "A survey of MRI-based brain tumor segmentation methods," *Tsinghua Science Technology*, vol. 19, no. 6, pp. 578–595, 2014.

[15] N. Kwak and H. Choi, "Input feature selection for classification problems," *IEEE Transaction on Neural Networking*, vol. 13, no. 1, pp. 143–159, 2014.

[16] Y. Jinn-Yi and C. Fu, "A hierarchical genetic algorithm for segmentation of multi-spectral human-brain MRI," *Expert System with Application*, Wiley, vol. 34, pp. 1285–1295, 2008.

[17] Z. Chong, S. Xuanjing, H. Cheng and Q. Qingji, "Brain tumor segmentation based on hybrid clustering and morphological operations," *International Journal of Biomedical Imaging*, vol. 2019, pp. 1–12, 2019.

[18] S. Kalyanapu and K. Bhaskar, "Segmentation of MR brain images using unified iterative partitioned fuzzy clustering," *International Journal of Recent Technology and Engineering*, vol. 8, no. 1, pp. 2755–2758, 2019.

[19] N. Arul, S. Pettitt and C. Wright, "Hierarchical clustering based segmentation (HCS) aided interpretation of the DCE MR images of the prostate," *Conference on Medical Image Understanding Analysis*, vol. 17, pp. 1–6, 2015.

[20] N. Arul, M. Laura, S. Lynne, N. Sarah and L. Wright, "Hierarchical cluster analysis to aid diagnostic image data visualization of MR and other medical imaging modalities," in *Imaging Mass Sepctrometry*, Humana Press, US, pp. 95–123, 2017.

[21] G. Jorge, A. Hector and B. Carlos, "Dynamic image segmentation method using hierarchical clustering," *Progress in Pattern Recognition, Image Analysis, Computer Vision and Application*, US, pp. 177–184, 2009.

[22] S. Filipovych, M. Resnick and C. Davatzikos, "Semi-supervised cluster analysis of imaging data," *NeuroImage*, vol. 54, no. 3, pp. 2185–2197, 2011.

[23] M. Gunashree, S. T. Ahmed, M. Sindhuja, P. Bhumika and B. Anusha, "A new approach of multilevel unsupervised clustering for detecting replication level in large image set," *Procedia Computer Science*, vol. 171, pp. 1624–1633, 2020.

[24] S. T. Ahmed, M. Sandhya and S. Sharmila, "A dynamic MooM dataset processing under TelMED protocol design for QoS improvisation of telemedicine environment," *Journal of Medical Systems*, vol. 43, no. 8, pp. 1–12, 2019.

[25] G. Reddy, M. Reddy, K. Lakshmanna and D. S. Rajput, "Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis," *Evolutionary Intelligence*, vol. 13, no. 2, pp. 185–196, 2020.

[26] V. Kolisetty and D. S. Rajput, "A review on the significance of machine learning for data analysis in big data," *Jordanian Journal of Computers and Information Technology*, vol. 6, no. 1, pp. 56–578, 2020.

[27] S. M. Basha and D. S. Rajput, "Aspects of deep learning: Hyper-parameter tuning, regularization, and normalization," in *Intelligent Systems*, Apple Academic Press, Singapore, pp. 171–186, 2019.

[28] S. M. Basha and D. S. Rajput, "A roadmap towards implementing parallel aspect level sentiment analysis," *Multimedia Tools and Applications*, vol. 78, no. 20, pp. 29463–29492, 2019.

[29] S. M. Basha and D. S. Rajput, "Survey on evaluating the performance of machine learning algorithms: Past contributions and future roadmap," in *Deep Learning and Parallel Computing Environment for Bioengineering systems*, Academic Press, India, pp. 153–164, 2019.

[30] S. T. Ahmed, S. Sreedhar Kumar, B. Anusha, P. Bhumika and M. Gunashree, "A generalized study on data mining and clustering algorithms," in *Int. Conf. on Computational Vision and Bio Inspired Computing*, Cham, Springer, pp. 1121–1129, 2018.