



New Clustering Algorithm for Classification of Brain MRI Regions by using m -Universal Metric Technique

Nezhad, A. D. ^{*1}, Beizavi, S. ², and Hekmatimoghaddam, S. H. ³

^{1,2}*School of Mathematics, Iran University of Science and Technology, Iran*

³*Department of Laboratory Sciences, School of Paramedicine, Shahid Sadoughi University of Medical Sciences, Iran.*

E-mail: dehghannezhad@iust.ac.ir

** Corresponding author*

Received: 2 May 2017

Accepted: 31 December 2017

ABSTRACT

In this paper we introduce an m -dimensional ($m \geq 2$) distance metric over a given space to define a new universal metric space. An algorithm for segmentation is purposed in this paper that uses universal metric space as generalized K -means clustering technique to segment brain and to separate the brain tumors in MR brain images. This article analyzes the segmentation method and separate brain tumor region based on the proposed algorithm. On the basis of this algorithm, target parameter m that is representing the similarity measure in K -means algorithm, is changed from case $m = 2$ to cases $m = 3$ to $m = 5$. Brain tumor is separated for different values of m . Finally, to evaluate the proposed algorithm. The results of this procedure for 50 patients, (the selected from MRI center of Yazd city) are compared with segmentation that was performed by the radiologist. Based on these criteria, the proposed algorithm showed an accuracy of about 98.91 in case $m = 5$. So the

Nezhad, A. D., Beizavi, S. and Hekmatimoghaddam, S. H.

use of this method for processing magnetic resonance images can give acceptable results for separating tumor from normal brain tissue.

Keywords: U_m -metric, generalized K -means clustering technique, segment brain.

1. Introduction

The theory of metric spaces plays a major role in different fields of mathematics and applied sciences. The conventional definition of a metric formalizes the concept of distance. In practical situations, one wants to measure how separated the members of a set of more than two elements. In this regard, there have been many attempts to give generalizations of the ordinary metric spaces. Mustafa and Sims introduced a new class of generalized metric spaces called G -metric spaces (Mustafa 2006) as follows.

Definition 1.1. Let X be a nonempty set and $G : X \times X \times X \rightarrow \mathbb{R}^+$, be a function satisfying:

$$(G1) \ G(x, y, z) = 0 \text{ if } x = y = z,$$

$$(G2) \ 0 < G(x, x, y), \text{ for all } x, y, z \in X \text{ with } x \neq y,$$

$$(G3) \ G(x, x, y) \leq G(x, y, z), \text{ for all } x, y, z \in X \text{ with } z \neq y,$$

$$(G4) \ G(x, y, z) = G(x, z, y) = \dots = G(y, z, x) \text{ (symmetry in all three variables), and}$$

$$(G5) \ G(x, y, z) \leq G(x, a, a) + G(a, y, z) \text{ for all } x, y, z, a \in X \text{ (rectangle inequality). Then the function } G \text{ is called a generalized metric, or, more specifically, a } G\text{-metric on } X, \text{ and the pair } (X, G) \text{ is called a } G\text{-metric space.}$$

Since the properties of G -metric spaces (see Dhage 1994) has been studied and developed by many authors (see e.g, Dehghan 2010 and Dehghan 2011). Also in 2014, Dehghan Nezhad introduces the notion of a U_n -metric space, (see Dehghan 2014 and Dehghan 2017).

Medical imaging is one of the important processes of medical science and it is an anatomical visualization of the human body. Image segmentation is one of the important processes in medical image analysis. In fact, image segmentation plays an important role in medical applications.

Image segmentation is an essential subject in computer vision. In image segmentation, an image is partitioned into separate subsets, that any subset specifies significant section of the image. This technique has many applications in the fields of scene analysis, object recognition and tracking moving objects. But one of the most important applications of this technique is in medical image analysis, that it used to separate the brain tumor region in MR images.

Nezhad, A. D., Beizavi, S. and Hekmatimoghaddam, S. H.

There are several methods in image segmentation, one of these methods is k -means clustering algorithm. This method is an unsupervised algorithm, that is a technique in the pixel based algorithm. In this paper, the following topics will be discussed: In section 2 our proposed method is given. In section 3, the results obtained are compared with similar methods. We evaluate the proposed method in section 4. Finally, in section 5, concluding remarks are given.

2. Methods

For $m \geq 2$, let Y^m denotes the cartesian product $\underbrace{Y \times \cdots \times Y}_{m\text{-times}}$. We begin with the following definition.

Definition 2.1. Let Y be a non-empty set. Let $U : Y^m \rightarrow \mathbb{R}^+$ be a function that satisfies the following conditions:

- (1) $U_m(y_1, \dots, y_m) = 0$ if $y_1 = \dots = y_m$.
- (2) $U_m(y_1, \dots, y_m) > 0$ for all y_1, \dots, y_m with $y_i \neq y_j$, for some $i, j \in \{1, \dots, m\}$.
- (3) $U_m(y_1, \dots, y_m) = U_m(y_{\pi_1}, \dots, y_{\pi_m})$, for every permutation (π_1, \dots, π_m) of $(1, 2, \dots, m)$.
- (4) $U_m(y_1, y_2, \dots, y_{m-1}, y_{m-1}) \leq U(y_1, y_2, \dots, y_{m-1}, y_m)$ for all $y_1, \dots, y_m \in Y$.
- (5) $U_m(y_1, y_2, \dots, y_m) \leq U_m(y_1, a, \dots, a) + U(a, y_2, \dots, y_m)$, for all $y_1, \dots, y_m, a \in Y$.

The function U_m is called a universal metric of dimension m , or more specifically a U_m -metric on Y , and the pair (Y, U_m) is called a U_m -metric space.

The following useful properties of a U_m -metric are easily derived from the axioms.

Example 2.1. Let (Y, d) be Euclidean metric space, then (Y, U_m) and (Y, U'_m) are U_m -metric spaces, where

$$U_m(y_1, \dots, y_m) = \frac{2}{m(m-1)} \sum_{1 \leq i < j \leq m} d(y_i, y_j)$$

$$U'_m(y_1, \dots, y_m) = \max \{d(y_i, y_j) : 1 \leq i < j \leq m\}$$

New Clustering Algorithm for Classification of Brain MRI Regions by using m -Universal Metric Technique

Example 2.2. Let Y be a real inner product space of dimension m . Equip Y with the standard U_m -metric. $U(y_1, \dots, y_m) = \det((\langle y_i, y_j \rangle)_{m \times m})^{1/2}$, where $\langle \cdot, \cdot \rangle$ denotes the inner product on Y .

Definition 2.2. Let (Y, U_m) be a U_m -metric space. Then for $y_0 \in Y$, $r > 0$, the U_m -ball with center y_0 and radius r is

$$B_{U_m}(y_0, r) = \{v \in Y : U_m(y_0, v, \dots, v) < r\}. \tag{1}$$

Proposition 2.1. Let (Y, U_m) be a U_m -metric space. Then for $y_0 \in Y$, $r > 0$, (i) If $U_m(y_0, y_2, \dots, y_m) < r$, then $y_2, \dots, y_m \in B_{U_m}(y_0, r)$; (ii) If $v \in B_{U_m}(y_0, r)$, then there exists, $\delta > 0$ such that $B_{U_m}(v, \delta) \subseteq B_{U_m}(y_0, r)$.

Proof. proof part (ii). Let $B_{U_m}(y_0, r)$ be a U_m ball with center y_0 and radius r , let $v \in B_{U_m}(y_0, r)$, we have $U_m(y_0, v, \dots, v) < r$. Suppose $\delta = \frac{r - U_m(y_0, v, \dots, v)}{2}$, then we show that $B_{U_m}(v, \delta) \subseteq B_{U_m}(y_0, r)$. According to $U_m(v, y, \dots, y) < \delta$ we have

$$\begin{aligned} U_m(y_0, y, \dots, y) &\leq U_m(y_0, y, \dots, y) + U_m(v, y, \dots, y) \\ &\leq U_m(y_0, v, \dots, v) + \frac{r - U_m(y_0, v, \dots, v)}{2} \\ &= \frac{r + U_m(y_0, v, \dots, v)}{2} < r \end{aligned}$$

□

Proposition 2.2. Let (Y, U_m) be a U_m -metric space. Then for a sequence $\{y_k\} \subseteq Y$ and a point $y \in Y$ the following are equivalent: (1) $\{y_k\}$ is U_m -convergent to y . (2) $U_m(y_k, \dots, y_k, y) \rightarrow 0$. (3) $U_m(y_k, y, \dots, y) \rightarrow 0$.

Remark 2.1. We remark that a G -metric space is evidently a U_3 -metric space. Another topological notions as U_m -Hausdorff, U_m -compact, U_m -normal and ... define similarly as usual.

2.1 Generalized K -means clustering algorithm

Clustering is one of the branches unsupervised learning, that the data are divided into groups whose members are similar to each other. K -means clustering algorithm is one of the best clustering algorithms which is based on minimizing of an objective function. In our proposed method, the measure of similarity between the data, is distance between N data (for $N \geq 2$) using the universal metric.

Nezhad, A. D., Beizavi, S. and Hekmatimoghaddam, S. H.

In this algorithm, the number of clusters k is assumed to be fixed. There is an error function in this algorithm. The conventional K -means algorithm is briefly described in the paper of Jian (see Jian 1988). Let D be a data set with N instances y_1, \dots, y_N and let C_1, C_2, \dots, C_k be the k disjoint clusters of D . Then the error function is defined as

$$E = \sum_{j=1}^k \sum_{y \in C_j} d(y, \nu(C_j)),$$

where $\nu(C_j)$ is the centroid of the cluster (calculated by averaging the observations of each cluster) and $d(y, \nu(C_j))$ denotes the ordinary distance between the point x and $\nu(C_j)$. For a given initial k clusters, the algorithm randomly picks one data point from the dataset D each time and calculates the distance between cluster centers. This data-point is assigned to the cluster that is having least distance from its center and repeatedly changes the membership of the clusters according to the error function until the error function does not change significantly or the membership of the clusters no longer changes.

Here we propose a K -means algorithm that picks $N - 1$ points at a time and calculates the U_m distance between this points and center of clusters. Then, these $N - 1$ points are assigned to that cluster having least distance between the center and $N - 1$ data points. The proposed algorithm is given below:

- (1) Choose integer k , the number of clusters.
- (2) Assume k number of initial seed points.
- (3) Randomly assign the data into k initial cluster C_1, \dots, C_k and determine $\nu(C_1), \dots, \nu(C_k)$.
- (4) Consider a subset $\{y_{i1}, \dots, y_{i(N-1)}\}$ from the data set $\{y_1, \dots, y_N\}$ and calculate the $d_{ij} = U_m(y_{i1}, \dots, y_{i(N-1)}, \nu(C_j))$.
- (5) Let $L_j = \operatorname{argmin}_{1 \leq j \leq k} d_{ij}$ and assign the $N - 1$ points to the cluster L_j .
- (6) Compute the new centroid after assigning all data points to k clusters.
- (7) Repeat steps (4) to (6) until the difference between the previous and current centroid is less than the specified threshold value.
- (8) Repeat steps (2) to (7) with different initial seed points until the algorithm reaches the minimum objective function.

New Clustering Algorithm for Classification of Brain MRI Regions by using m -Universal Metric Technique

Notice that, since the proposed algorithm picks $N - 1$ data points at a time, it reduces the number of iterations and thus the overall complexity of the clustering process. Now we compare the results (clustering of eight cities) of the case $m = 2$ (conventional K -means clustering) and case $m = 3$ (generalized K -means clustering). Following table shows the results of applying the conventional K -means clustering and generalized K -means clustering to our example data of eight points. U_m distance for $m = 3$ is much better (faster) than the Euclidian distance (case $m = 2$). As you seen, there is the minimal error in $m = 3$ and its value is equal 1938. So, in this case, we have the best clustering.

2.2 Experimental result

In the present study, the target parameter in this algorithm is m and according to similarity criterion in clustering for $m = 2$ is the Euclidean distance which results in conventional K -means clustering. Using various contents, results show that for $m > 2$ the algorithm has operated better in tumor detection. Generalized K -means clustering algorithm in image segmentation is used first to detect the tumor area. Regarding the input gray-levels, the number of algorithm clusters is $k = 4$. Figure 1:(a), shows a sample of the brain MRI image and the clustering area for $m = 2, 3, 4$ and 5. To detect the tumor area in the segmentation image, the area which is the brightest one in the input image will be taken into consideration. Figure 1:(b), illustrates the tumor candidate areas for $m = 2, 3, 4$ and 5. The crucial point is that all the existent areas in the image are not a tumor and they need to be omitted. This procedure can be done through Morphological operations, figure 1:(c).

Finally passing this step, tumor area will be separated the other parts. According to figures 1:(d) and 1:(e), it can be seen that the tumor area in $m = 5$ in comparison with $m = 2, 3, 4$ is well separated from the other parts.

Nezhad, A. D., Beizavi, S. and Hekmatimoghaddam, S. H.

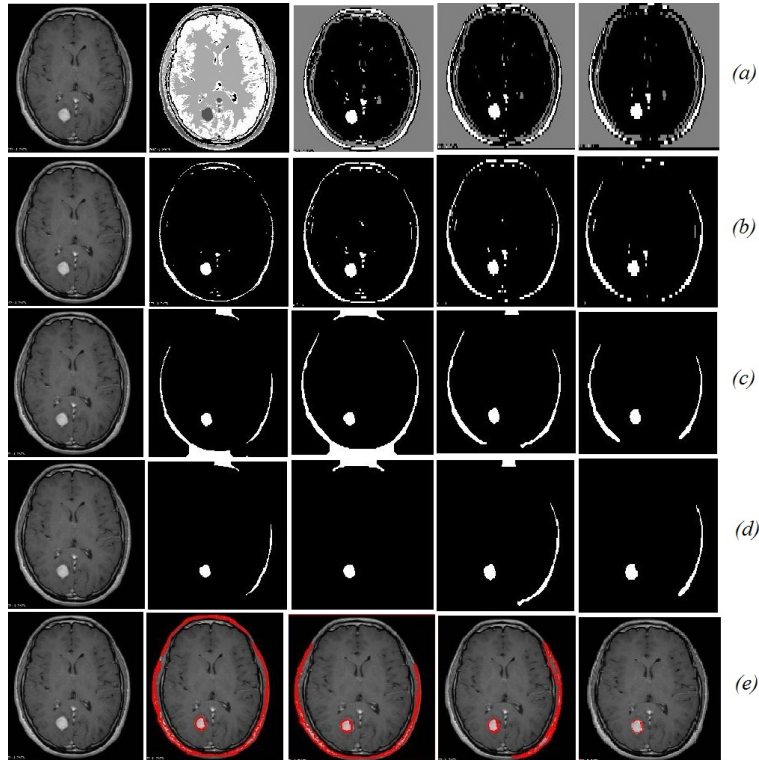


Figure 1: (a) Left: original MRI image, Right: image clustered for $m = 2, 3, 4$ and 5 , (b) Left: original MRI image, Right: tumor candidate for $m = 2, 3, 4$ and 5 , (c) Left: original MRI image, Right: tumor candidate after morphology operation for $m = 2, 3, 4$ and 5 , (d) Left: original MRI image, Right: tumor region for $m = 2, 3, 4$ and 5 , (e) Left: original MRI image, Right: tumor region after using region growing for $m = 2, 3, 4$ and 5 .

3. Compiled results

For our experimental study, we use the well-known simplicity dataset of references H. A Aslam (2013), R. Bhattacharjee (2012), and S. Yadav (2013).

Pillar’s segmentation includes a mechanism for grouping the elements of high-resolution images. This algorithm distributes all initial centroid according to the maximum cumulative distance metric. Figure 2:(a), shows a sample of segmentation using pillar’s method and figure 2:(b), illustrates our proposed method in case $n = 5$.

Figure 3:(a), illustrates practical results for proposed method reference (see Yadav 2013). This paper focuses on a K -means algorithm to diagnose accu-

New Clustering Algorithm for Classification of Brain MRI Regions by using m -Universal Metric Technique

rately the region of the tumor. In this paper, an image is divided into a number of various groups or clusters. Figure 3:(b), shows Experiments demonstrate our proposed method in case $n = 5$.

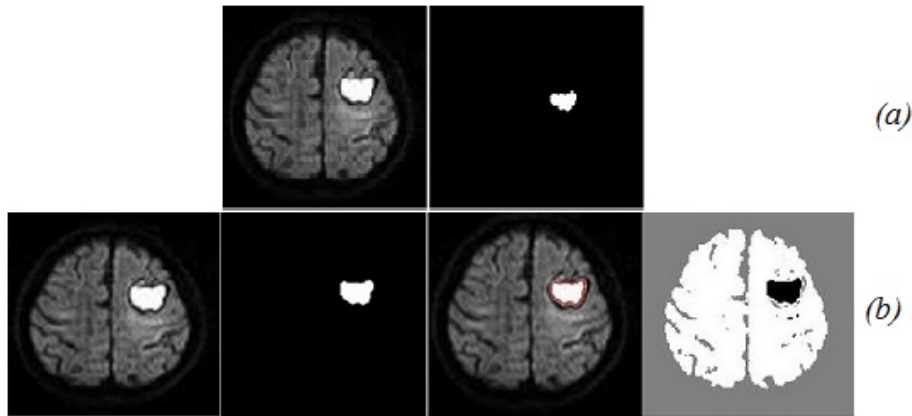


Figure 2: (a) Left: original image, Right: pillar's method, (b) Left: original image, Right: tumor area in our proposed method.

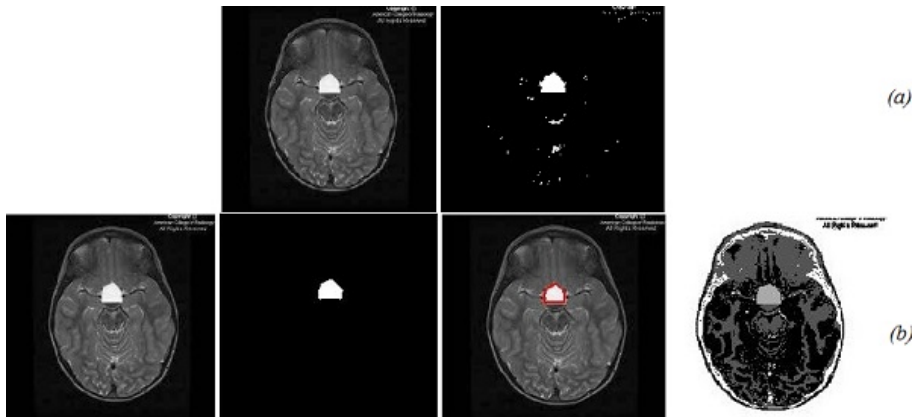


Figure 3: (a) Left: original image, Right: proposed method reference (see Yadav 2013), (b) Left: original image, Right: tumor area in our proposed method.

In reference Bhattacharjee (2012), a novel algorithm is developed to feature out tumor from diseased brain Magnetic Resonance images. This paper, based on a study of quality parameter is selected for de-noising the images. Figures 4:(a) and 5:(a), shows two samples of the brain *MRI* images and tumor segmentation using this algorithm, too figures 4:(b) and 5:(b), illustrates tumor region for case $m = 5$ in our proposed method.

Nezhad, A. D., Beizavi, S. and Hekmatimoghaddam, S. H.

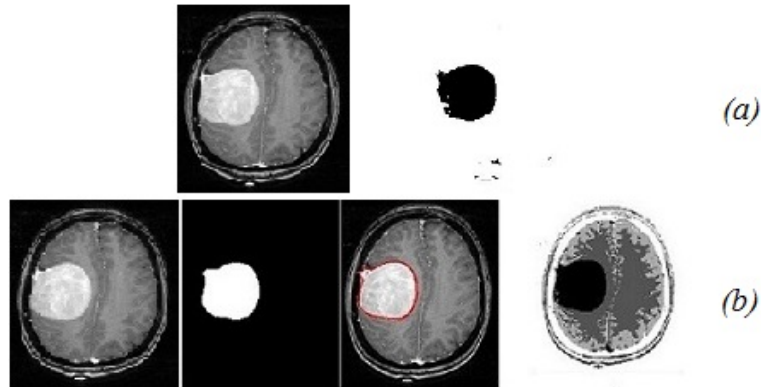


Figure 4: (a) Left: original image, Right: proposed method reference Bhattacharjee (2012), (b) Left: original image, Right: tumor area in our proposed method.

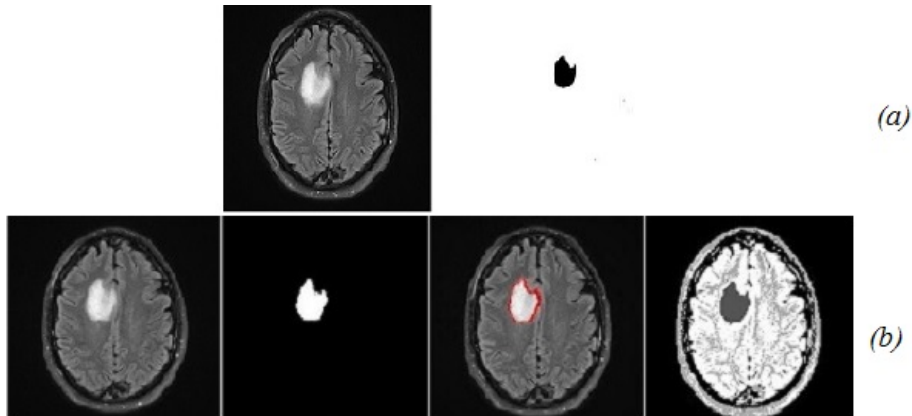


Figure 5: (a) Left: original image, Right: proposed method reference Bhattacharjee 2012, (b) Left: original image, Right: tumor area in our proposed method.

4. Performance analysis and evaluation

To evaluate the proposed algorithm, a dataset contains 50 brain tumor images is selected from the database of MRI center of Yazd city. And ground truth is created by an expert radiologist. Below are some of the images as well as the actual tumor that has been determined by an expert, see figure 6.

New Clustering Algorithm for Classification of Brain MRI Regions by using m -Universal Metric Technique

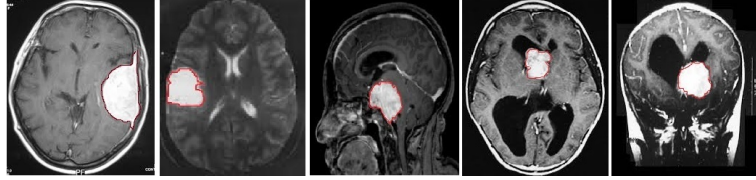


Figure 6: image segmentation by an expert radiologist

For quantitative analysis, number of false negative and false positive are calculated based on the number of pixels of the interested region (ROI). Four parameters true positive (TP), false positive (FP), true negative (TN), false negative (FN) are calculated by the logical AND between ground truth and segmented image. Sensitivity and specificity in terms of brain tumor region could be defined as, sensitivity is the percentage of patients correctly detected with a tumor, whereas specificity is the percentage of patients could not correctly be identified with the tumor.

Formulation as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{4}$$

Now, using the method proposed in this paper, brain tumor in MR images is estimated with different, similarity measure m and area estimates for $m = 2, 3, 4, 5$ is compared with ground truth. After comparing the images in the database, evaluation parameters are calculated to determine the effectiveness of the proposed method for different m .

In the figure 7, you see a number of results brain tumor segmentation with proposed method in this study.

In table 1, amounts of evaluation parameters for different similarity measure (m) are shown.

According to table 1, if m increases then the specificity, sensitivity and accuracy increase. This results show that this algorithm is powerful in the diagnosis of tumor brain.

Nezhad, A. D., Beizavi, S. and Hekmatimoghaddam, S. H.

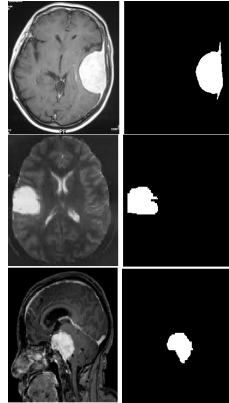


Figure 7: Image segmentation with proposed method

Table 1: Results

m	Specificity	Sensitivity	Accuracy
2	76.08	94.52	93.68
3	76.29	97.37	97.70
4	76.43	98.52	98.21
5	76.58	99.07	98.91

5. Discussion

The principal conclusion from the research in this paper that generalized G -metric spaces into U_m -metric spaces. This conclusion is justified for the following reasons. Firstly, we have shown that U_m -metric spaces as an extension of classical G -metric spaces have been more considered in the recent decade. Secondly, we have shown that such work can be used in other section of mathematics. Finally, the primary motivation for this work has been to develop metrics based tools for applications in program verification in theoretical computer science. A new technique for brain tumor separate from diseased MR images is developed in this work. Disregarding all other structural features and components of an MR image. This developed technique focuses only on tumor detection. A preliminary experiment conducted on the MRI brain image demonstrates encouraging results.

New Clustering Algorithm for Classification of Brain MRI Regions by using m -Universal Metric Technique

References

- Abras, G. N. B. and Virginia, L. (2005). A weighted K-means algorithm applied to brain tissue classification, *JCS*, 5(3):
- Aslam, H. A., Ramashri, T., Imtiaz, M. and Ahsan, A. (2013). A new approach to image segmentation for brain tumor detection using Pillar k-means algorithm, *IEEE Computer Society*, 2(3):
- Bezdek, J. C., Hall, L. O. and Clarke, L. P.(1993). Review of MR image segmentation techniques using pattern recognition, *Medical Physics*, 20(4): 10-33.
- Bhattacharjee, R. and Chakraborty, M.(2012).Brain tumor detection from MR images: image processing, slicing and PCA based reconstruction, *IEEE Computer Society*, 2: 97-101.
- Chowdhury, M.H., Little, W.D. (1995). Image thresholding techniques, *IEEE Pacific Rim Conference on Communications, Computers, and Signal Processing*, Proceedings, pages 585-589.
- Clark, M. C. (1994). *Segmenting MRI volumes of the brain with knowledge-based clustering*, MS Thesis, Department of Computer Science and Engineering, University of South Florida.
- Dehghan, N. A. and Khajuee, N. and Mustafa, Z.(2017).Some New Results on Universal Metric Spaces, *Thai Journal of Mathematics*, 429-449.
- Dehghan Nezhad A. and Mazaheri H., (2010). New results in G-best approximation in G-metric spaces, *Ukrainian Math. J.*, 62(4):648-654.
- Dehghan Nezhad A. and Aral Z., (2011). *The topology of GB-metric spaces*, ISRN. Mathematical Analysis, Hindawi.
- Dehghan Nezhad A. and Khajuee N., (2014). Some new results on complete U_n^* metric space, *Journal of Nonlinear Science and Applications*, pages 216-226.
- Dhage, B. C. (1994). On generalized metric spaces and topological structure II, *Pure. Appl. Math. Sci.*, 37-41.
- Jain, A. and Dubes, R.(1988). *Algorithms for clustering data*, Englewood Cliffs, NJ: Prentice-Hall.

Nezhad, A. D., Beizavi, S. and Hekmatimoghaddam, S. H.

Mustafa, Z. and Sims, B. (2006). A new approach to generalized metric spaces, *Journal of Nonlinear and Convex Analysis*, 7(2): 289-297.

Yadav, S. and Meshram, S.(2013). Brain tumor detection using clustering method, *International Journal of Computational Engineering Research*, 3: 1-14.