ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 8(29), DOI: 10.17485/ijst/2015/v8i29/83379, November 2015* ISSN (Online) : 0974-5645

Identification of Calcification in MRI Brain Images by k-Means Algorithm

A. Naveen* and T. Velmurugan

Department of Computer Science, D. G. Vaishnav College, Arumbakkam, Chennai – 600106, Tamil Nadu, India; naveenking@yahoo.co.in, velmurugan_dgvc@yahoo.co.in

Abstract

Background/Objective: The role of clustering is significant to analyze different kind of applications of its techniques. Similar data are grouped into one and they formed as a cluster. Dissimilar data are grouped into another form in other cluster. Data clustering is an important and active research applied in many fields including multivariate analysis in statistics and some other areas like pattern recognition and machine learning etc. **Methods/Statistical Analysis:** Boundary detection and outlier analysis is an important concept for pre-processing the data. The boundary considers only pixels lying on and near edges and use of gradient or Laplacian to preliminary processing of images. To find the outlier in a group of patterns is a well-known problem in Data Mining (DM). An outlier is a pattern which is different with respect to the rest of the patterns in the data. The k-Means is one of the familiar clustering methods used by different researchers to find the well-formed clusters. Magnetic Resonance Imaging (MRI) uses a magnetic field and radio waves to create detailed images of the organs and tissues within human body. The k-Means algorithm is used to find the tumor by applying the boundary detection and outlier techniques in this research work in MRI brain images. **Findings:** The main goal of this research work is to extract the tumor (Calcification) in an MRI brain image by means of clustering pixels to fortify the quality of clustering algorithm. The results of the MRI brain images are analyzed and identified by the proposed algorithm. The result produced by simple k-Means algorithm is very useful to find the tumor in MRI images perfectly. **Application/Improvements:** The MRI brain images are analyzed and implemented by other methods like classification and some other techniques in future.

Keywords: Image Clustering, Image Preprocessing, k-Means Algorithm, MRI Imagery, Method

1. Introduction

Data mining (DM) is the use of automated data analysis techniques to uncover previously undetected relationships among data items. DM has most suitable algorithms to analyze data stored in a data warehouse in various domains. The DM processes are used to discover the features, patterns and knowledge for decision support system. The major data mining techniques are association, regression, classification, clustering, prediction, sequential analysis and pattern generations are exercised to identify the knowledge which could apply for decision making system[1]. These techniques are applied into the scientific, business, intelligent analysis, educational, statistics, mathematical and medical applications. In this research work, the MRI images are clustered by applying the k-Means clustering algorithm.

The clustering method classifies the data objects into some meaningful groups based on similar feature, attribute and characteristic of the data items. Clustering is classified into two methods namely supervised and unsupervised. In supervised type clustering, cluster criteria are specified by the user. The k-means algorithm is a simple interactive clustering method to find partition for a given dataset with a user specified number of clusters[2]. The k-Means clustering is the most appropriate method for biomedical images to calculate the number of clusters using images of particular areas of the human structure. This research work cluster the MRI brain images based on the colors. Medical image retrieval in the development

**Author for correspondence*

of computer aided system is related with the use of new intelligent capabilities such as multimedia support and data mining to determine the pertinent knowledge for diagnosis in medical fields[3]. MRI is normally used in bio-medical field to perceive and visualize finer details in the internal structure of the human body. This technique is basically used to detect the variances in the tissues which have remote better technique and methods as compared to computed tomography. So, this notion makes this technique a very special one for the brain tumor detection[4].

MRI was announced into clinical medicine and it is extremely used to detect the problems and calcifications in brain images[5]. The information provided in MRI images about the soft tissue composition has intensely improved the quality of brain pathology analysis and treatment. Examples of normal and abnormal MRI Brain Images are shown in Figure 1 and Figure 2.

The description of related work has done in several researches in algorithms, techniques and dataset for MRI brain images. The main points and features of approaches are mentioned below with respective limitations and benefits that make our work effective. Clustering of Brain MRI Image using Data Mining Algorithm is explored by Siji T. Mathew et al.[6] This research article is presented to attempt for clustering brain MRI images using k-Means algorithm. They carry out a comparative study on clustering with k-Means algorithm and Fuzzy C-Means (FCM) algorithm in MRI image dataset. This clustering helps to identify the region of interest in the image. They state that from the clustered image the k-means algorithm gives a clear clustering than the FCM algorithm. They conclude that the clustering time is less for k-Means algorithm than FCM.

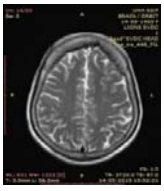

Figure 1. Normal MRI Image.

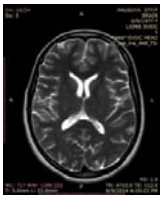

Figure 2. Abnormal MRI Image.

Sunila G, et al.[7] analysis about the various clustering algorithms in their research work. Their work reviews four types of clustering techniques namely k-Means Clustering, Farther first clustering, Density Based Clustering, Filtered cluster. These clustering techniques are implemented and analyzed in WEKA[8] software. They conclude that the k-Means algorithm is best amount the other algorithms and also produced quality results when using large data. Brain tumor extraction of MRI images by using clustering and morphological operations methods are applied in a research work by Ali et al.[9] The preprocessing of MRI images which is in the form T2 weighted modality have been done by using bilateral filter method in order to reduce the noise. In preprocessing, edges amount the different tissues are also maintained. For the extraction of tumor images, the morphological operations are exactly carried out using four techniques. The results show that the morphological method produces the best results than the other methods.

A performance analysis of some clustering algorithms to detect the brain tumor in MRI images is explored by Tamije Selvy et al.[10] The clustering methods used are k-Means, Self-Organizing Map (SOM), Hierarchical Cluster and Fuzzy C-Means (FCM). The k-Means and Hierarchical clustering achieved about 95% result. SOM and FCM achieved a result of 80%. Ashwini Gulhane et al.[11] give an idea of are view of image data clustering techniques in their research. The k-Means algorithm and its limitations are discussed in this paper. M-step clustering is a new approach for clustering is used in this research which may overcomes some limitations of k-Means. To find the optimal cluster by k-Means algorithm and the difficulties of the same algorithm are discussed. The other algorithms like M-step algorithm, N-cut algorithm and a mean shift algorithm which are used for disabling the inadequacy of k-means algorithm is also discussed in their research work for validation.

This research paper is structured and presented as follows. Section II discusses about preprocessing techniques in outlier analysis, boundary detection algorithms and clustering algorithm. The results of image pixel identification in MRI data and performances of clustering algorithm are demonstrated in section III. Finally the work is concluded and findings are given in section IV.

2. Materials and Methods

Clustering is methods of grouping the similar pixels that are create clusters. The clusters are formed based on some meaningful clusters are the goal, and then the resulting clusters should capture the normal structure of the data. The clustering techniques are applied and utilized in cluster analysis that has long been used in a wide variety of grounds like psychology and other social sciences, biology, statistics, pattern recognition, information retrieval, machine learning, and data mining[12–14]. Here the k-Means clustering are discussed in detail. The chief objective of this research work is to cluster MRI brain images. The brief explanation about the dataset and the algorithms are given in this section. The performance of k-means algorithm is measured by extracting the different types of MRI brain images. The processed work this paper is depicted in the block diagram as shown in Figure 3.

2.1 Dataset

This research work uses MRI brain images taken from normal and abnormal patient images at Swamy Vivekananda Diagnostic Centre (SVDC) Hospital in Chennai at D.G. Vaishnav College campus. An abnormality of the MRI brain images was marked by SVDC head. The MRI brain images in DICOM format is taken for analysis. The DICOM is abbreviated as digital imaging and communications in medicine. It is an International Standard for medical images and related information of DICOM. The DICOM file format support the encapsulation of any information object definition. This type of images can hold images with patient information like age, sex, modality, study description, date of image taken, image size and type etc. Normal and abnormal data are processed in this methodology. The MRI brain images were divided into 8 and 16 clusters by the clustering algorithm.

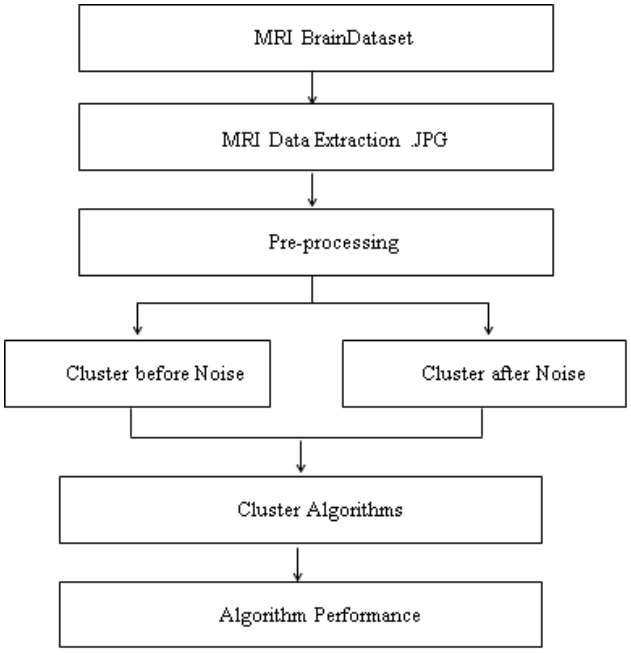

Figure 3. Block Diagram for Methodology.

2.2 Methodology

The primary objective of this work is to discover and predict the performance of clustering algorithm by means of extracting pixels in MRI brain images. The clustering model was built in k-Means algorithms and each of them is used data mining techniques after the removal of noise images and before the removal of noises. In this work, data are used for preprocessing with k-means algorithm to predict the image pixel values and performance of the algorithm. Each of these is discussed in detail in the following topics. Preprocessing the image and noise removal are done by using the Outlier Analysis and Boundary Detection. After the preprocessing, the data set is clustered by k-Means algorithm.

2.3 Preprocessing

Data preprocessing is used to remove the unwanted data and information available in the data set. It is a process of DM which deals with the preparation and transformation of the original dataset. The phrase "Garbage In, Garbage Out" is particularly pertinent to data mining and machine learning approaches to remove the irrelevant information from the data set. The first step of this work is to preprocess the MRI brain images for the efficient use of the algorithms. A number of methods used in this work for the same purpose. The irrelevant information is removed by outlier analysis and boundary detection algorithms. Basically pre-processing is applied to remove unwanted information and distorting as well as ringing effect in order to get the enhanced and much clear image for the efficient use of the data set[15]. From the input image, a small neighborhood of a pixel is converted into a new brightness value in the output image is the usual procedure, which is also applied in this work.

2.4 Outlier Analysis

Outlier analysis has been a very important concept in the actual world data analysis[16]. Outlier analysis is an important branch in data pre-processing and data mining, as

this stage is required in elaboration and mining of data coming from many application fields. Outlier detection techniques are used to minimize the encouragement of outliers in the final model to develop, or as a preliminary pre-processing stage before the information carried by a signal is elaborated. Outliers are patterns in data which behave differently in a particular data set. Outlier analysis in data mining is normally based on distance measures, clustering and spatial methods. An important aspect of an outlier detection technique is the nature for desired outlier. There are three types of outliers are utilized in literature of DM namely Point Outliers, Contextual Outliers and Collective Outliers. The point outlier method is used here for the removal of outliers from the MRI imagery.

2.5 Boundary Detection

Containing cues in an image is called boundaries that are very significant to high level pictorial tasks such as object recognition and scene understanding. Detecting boundaries has been vital problem since the commencement of computer vision. Some critical roles played in the development of boundary detection methods based on the values of datasets along with their evaluation methods. The behavior of data items is responsible for the progress in the problem of boundary detection. They provide not only an objective quantity to judge the value of each newly created algorithm, but also because of the images used in approach. The evaluation ethics they set forth have heavily inclined the researchers during the development of a boundary detection algorithm[17,18].

2.6 The k-Means Algorithm

Since the simplest form of k-Means algorithm, it is used by most of the researchers in the field of data mining. The process of k-Means follows eminent and simple way to classify a given data set by means of a certain number of clusters. The k-means clustering is a method used to divide n patterns $\{x_1, \ldots, x_n\}$ in d dimensional space into k clusters (assume k clusters). The result is a set of k clusters based on k centers, each of which is located at the centroid of the separated dataset. This algorithm can be shortened in the following steps:

Step 1: Give the number of cluster value as k.
Step 2: Randomly choose the k cluster centers
Step 3: Calculate mean or center of the cluster
Step 4: Calculate the distance between each pixel to each cluster center
Step 5: If the distance is near to the center then move to that cluster.
Step 6: Otherwise move to next cluster.
Step 7: Re-estimate the center.
Step 8: Repeat the process until the center doesn't move

In this method, the k-means clustering algorithm fixes the k value as 8 and 16 in MRI brain image. Clustering algorithm prove their efficiency in many fields. The k-Means algorithm is the widely used algorithm in all domains[19, 20].

3. Experimental Results

In this research work, tumor detection by identifying the pixel values in MRI brain images are taken for analysis. The DICOM Image format of the image is used for preprocessing the image. The source code is written in MATLAB software. The preprocessing work is done by using the outlier method and boundary detection method for the extraction of the input data set. After preprocessing both the images (normal and abnormal), the k-Means algorithm is applied to find the clusters of MRI images by dividing the image into 8 and 16 groups. The normal and abnormal images are given in the first column of the Figure 3 and the preprocessing images are available in the next four columns in the same figure. Also the Figure 3 shows the MRI images before and after the noise removal.

The steps involved in clustering the MRI brain images by k-Means algorithm are given below.

Step 1: Insert the original images as input.
Step 2: Convert the fetched MRI DICOM format file into .JPG
Step 3: Convert the image into preprocessing.
Step 4: Cluster dataset images. (After Noise and Before Noise)
Step 5: Find out the 'k' in image by algorithm itself.
Step 6: Get the clustered objects.

The results of the proposed method by taking the number of clusters as 8 are shown in Figure 5. The various terms used in this clustering are ANOI – Abnormal Original Images, NOI– Normal Original Image, ABNI – Abnormal Before Noise Image, AANI – Abnormal Normal Noise Image, NBNI – Normal Before Image and NANI – Normal After Noise Image. The number of pixels in each category of images in figure 5 is listed in Table 1. It

is easy to identify that some of the clusters are having very less number of pixels and a highest values of pixels are available in some clusters. Wherever the number of pixels are very less, that particular portion of the brain image have affected by tumor. The other areas are not very seriously affected. Figure 6 shows that the results of k-Means algorithm when the number of clusters taken as 8.

The results of normal and abnormal images of taken input are extracted by k-Means algorithm by considering the number of clusters as 16 is shown in figure 7. The number of pixels in each and every figure in the same image is listed in table 2. Figure 8 is the result of the proposed algorithm by splitting both the images by 16 clusters. It is evident that the total number of pixels identified by the algorithm when k = 8 is 262107 and the total number of

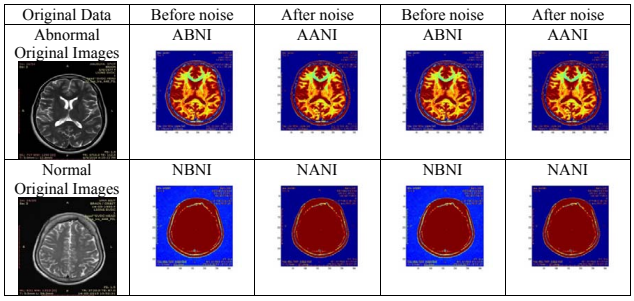

Figure 4. Pre-Processing of MRI Images.

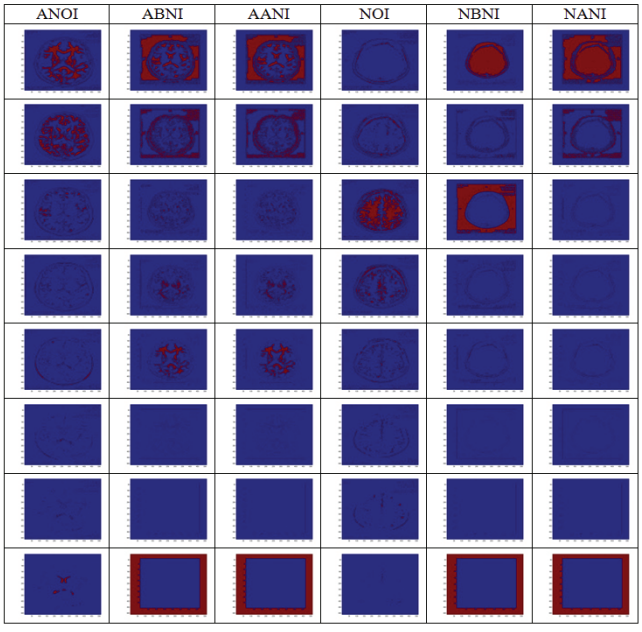

Figure 5. Output of k-Means algorithm when k = 8.

Table 1. Result of k-Means algorithm in pixels when k = 8

Cluster k	ANOI	ABNI	AANI	NOI	NBNI	NANI
1	43663	75438	82217	10595	61643	124552
2	44241	44434	42892	14344	12020	32488
3	15059	13462	10111	57519	80842	3444
4	9127	14769	12309	23220	6404	2768
5	7372	17771	18237	11027	4837	2451
6	3311	2058	2163	5435	2033	2071
7	1943	719	722	3504	661	663
8	2867	93451	93449	282	93670	93670
TOTAL	**127583**	**262102**	**262100**	**125926**	**262110**	**262107**

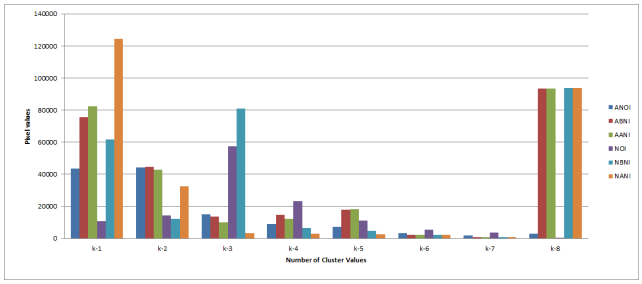

Figure 6. Result of k-Means algorithm when k=8.

pixels when k = 16 is 262124 for normal image. But, for the abnormal image, it is bound that the total number of pixels are 262100 and 262118 for k = 8 and k = 16 respectively. There is a little difference between the total number of pixels identified by the algorithm when the number of clusters taken as 8. It is clearly shown in the respective figures and also shown in figure 9.

4. Conclusions

The boundary detection method is used for the noise removal for preprocessing the input data in this research work. For partitioning the pixels, the simple k-Means algorithm is used. It is proved that the k-Means algorithm perform well in the spherical shaped images by different kind of applications. The approach of k-Means algorithm is very simple for any kind applications. Identifying the calcification in MRI brain image is perfectly done by

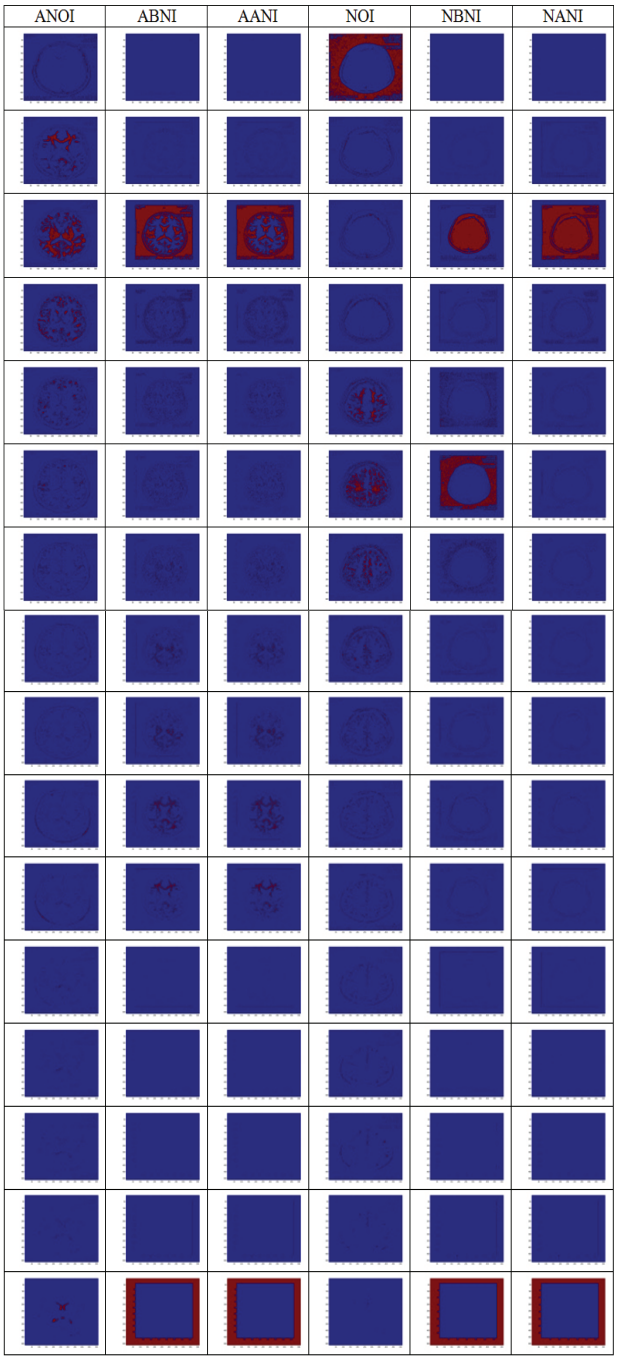

Figure 7. Output of k-Means algorithm when k = 16.

simple k-Means algorithm used in this research work. Also, it is clear that the number of pixels are differ when the numbers of clusters taken as 8 and 16. Therefore, modified versions of the simple k-Means algorithm are requiring for future work to get the better result. In future, some other algorithms and some modified versions of other algorithms are applied to get better results.

Table 2. Result of k-Means algorithm in pixels when k = 16

Cluster k	ANOI	ABNI	AANI	NOI	NBNI	NANI
1	8054	181	138	113536	23	161
2	17196	2320	3093	5228	880	2895
3	40642	100480	109609	4914	66291	147510
4	20861	12923	9722	6131	4537	5517
5	10736	7656	5269	20669	13064	1848
6	7533	6597	4927	31525	62131	1751
7	5446	6664	5248	19078	9931	1632
8	4566	7354	5966	10969	2822	1398
9	3682	7532	6794	7925	2438	1060
10	3894	9332	9739	4994	2567	1294
11	2925	6206	6679	4040	1600	1516
12	1590	640	695	2581	1324	1147
13	1022	99	104	2022	214	89
14	1032	175	178	2009	182	138
15	724	672	652	729	584	624
16	2558	93287	93305	84	93536	93544
TOTAL	132461	262118	262118	236434	262124	262124

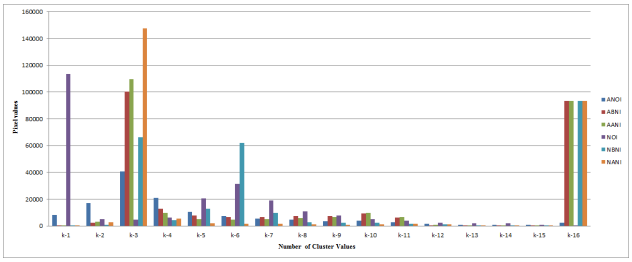

Figure 8. Result of k-Means algorithm when k = 16.

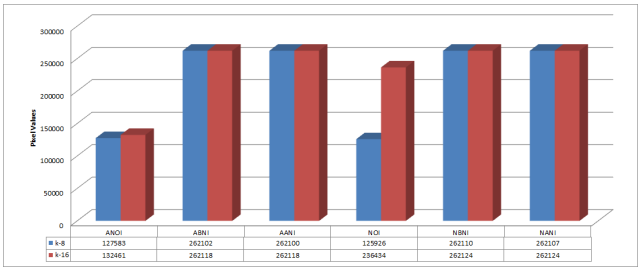

Figure 9. Performance of k-Means Algorithm.

5. References

1. Kumar RS, Karnan M. Review of MRI image classification techniques. International Journal of Research Studies in Computer Science and Engineering. 2014; 1(1):21–8.
2. Deshmukh RD, Jadhav C. Study of different brain tumor MRI image segmentation techniques. International Journal of Computer Science Engineering and Technology. 2014; 4(4):133–6.
3. Kawadiwale RB, Rane ME. Clustering techniques for brain tumor detection. Proc. of Int. Conf. on Recent Trends in Information, Telecommunication and Computing, ITC. Chandigarh, India. 2014 Mar 21. p. 299–305.
4. Sharma Y, Kaur P. Detection and extraction of brain tumor from MRI images using k-Means clustering and watershed algorithms. International Journal of Computer Science Trends and Technology. 2015; 3(2):32–8.
5. Kanade BP, Gumaste PP. Brain tumor detection using MRI images. International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering. 2015; 3(2):146–50.
6. Mathew ST, Nachamai M. Clustering of brain MRI image using data mining algorithm. Proceedings of 19th IRF International Conference; 2015 Feb 1. p. 33–6. ISBN: 978-93-84209-85-8.
7. Godara S, Verma A. Analysis of various clustering algorithms. International Journal of Innovative Technology and Exploring Engineering. 2013; 3(1):186–9.
8. Sharma N, Bajpai A, Litoriya R. Comparison the various clustering algorithms of weka tools. International Journal of Emerging Technology and Advanced Engineering. 2012; 4(5):73–80.
9. Ali SM, Abood LK, Abdoon RS. Brain tumor extraction in MRI images using clustering and morphological operations techniques. International Journal of Geographical Information System Applications and Remote Sensing. 2013; 4(1):12–25.
10. Tamije Selvy T, Palanisamy V, Purusothaman T. Performance analysis of clustering algorithms in brain tumor detection of MR images. European Journal of Scientific Research. 2011; 62(3):321–30.
11. Gulhane A, Paikrao PL, Chaudhari DS. A review of image data clustering techniques. International Journal of Soft Computing and Engineering. 2012; 2(1):212–5
12. Pande SR, Sambare SS, Thakre VM. Data clustering using data mining techniques. International Journal of Advanced Research in Computer and Communication Engineering. 2012; 1(8):494–9.
13. Arockiam L, Baskar SS, Jeyasimman L. Clustering techniques in data mining. Asian Journal of information Technology. 2012; 11(1):40–4.
14. Dharmarajan A, Velmurugan T. Lung cancer data analysis by k-means and farthest first clustering algorithms. Indian Journal of Science and Technology. 2015; 8(15):1–8.
15. Athar H, Raut AB. Extraction of tumour depending cells in brain from rough set theory and magnetic resonance images. International Journal of Advance Research in Computer Science and Management Studies. 2015; 3(1):295–300.
16. Silvia C, Colla V. Data processing for Outliers Detection. Malcolm Gladwell; 2013. p. 1–21.
17. Hou X, Yuille A, Koch C. Boundary detection benchmarking: Beyond F-measures. In: Computer Vision and Pattern Recognition IEEE Conference; 2013 Jun 23-28. p. 2123–30.
18. Gladis D, Rani S. k-Means clustering to identify high active neuron analysis for LSD. International Journal of Innovative Research in Science Engineering and Technology. 2013; 2(9):4817–24.
19. Shah HJ. Detection of tumor in MRI images using image segmentation. International Journal of Advance Research in Computer Science and Management Studies. 2014; 2(6):53–6.
20. Velmurugan T, Santhanam T. Computational complexity between k-Means and k-Medoids clustering algorithms for normal and uniform distributions of data points. Journal of Computer Science. 2010; 6(3):363–8.