International Journal of Electronics Communication and Computer Engineering
Volume 7, Issue 2, ISSN (Online): 2249–071X

Impact of Image Contrast on the Number of Clusters with k-means Algorithm

Lamia AbedNoor Muhammed
University of Al-Qadissiyah
Email ID : Lamia.abed@qu.edu.iq

***Abstract* – Image segmentation is a based step in analysis of image. More techniques are used for segmentation such as clustering. The clustering divide the image into groups of pixels according to its intensity. There are different methods for clustering. K means is one of the clustering methods. This method requires some parameters such as no. of clusters, however this not easy to predetermined. This paper focus on the existence of relationship between the image contrast and the number of clusters. The practical work were applied on medical images according to popularity.**

***Keyword* – Medical Image, Segmentation, k-means, Clustering, Image Contrast.**

I. Introduction

Image segmentation plays a significant role in computer vision. It aims at extracting meaningful objects lying in the image. Generally there is no unique method or approach for image segmentation. Clustering is a powerful technique that has been reached in image segmentation. The cluster analysis is to partition an image data set into a number of disjoint groups or clusters. The clustering methods such as k means, improved k mean, fuzzy c mean (FCM) and improved fuzzy c mean algorithm (IFCM) have been proposed. K means clustering is one of the popular method because of its simplicity and computational efficiency[1].

K-means clustering is the simplest method in supervised classification. It does not require training data. It is an iterative procedure. K-means clustering algorithm clusters the data by iteratively computing a mean intensity for each class and segmenting the image by classifying each pixel in the class with the closest mean. Clustering based on the optimization of an overall measure is a fundamental approach explored since the early days of pattern recognition. The most popular method for pattern recognition is K-means clustering[2].

K-means clustering algorithm, on the other hand, is a simple clustering method with low computational complexity as compared to FCM. The clusters produced by K means clustering do not overlap[3]. However, one of its drawbacks is the requirement for the number of clusters, K, to be specified before the algorithm is applied[4]. There are many methods available to estimate the number of clusters such as statistical indices, variance based method, Information Theoretic, goodness of fit method etc[5].

II. Medical Image Segmentation

An interesting source of images is the medical field. Here, imaging modalities such as CT (Computed Tomography), MRI (Magnetic Resonance Imaging), PET (Positron Emission Tomography) etc. generate a huge amount of image information. Not only does the size and resolution of the images grow with improved technology, also the number of dimensions increase. Previously, medical staff studied two-dimensional images produced by X-ray. Now, three dimensional image volumes are common in everyday practice. Even four dimensional data (three-dimensional images changing over time, i.e. movies) is often used[6].

Image segmentation is an important process for most medical image analysis tasks. Having good segmentations will benefit clinicians and patients as they provide important information for 3-D visualization, surgical planning and early disease detection[7]. Accurate detection of the type of brain abnormality is highly essential for treatment planning in order to minimize diagnostic errors. The accuracy can be improved by using computer aided diagnosis (CAD) systems. The basic concept of CAD is to provide a computer output as a second opinion to assist radiologists' image interpretation and to reduce image reading time. This improves the accuracy and consistency of radiological diagnosis. However, segmentation of the image of brain tumors is a very difficult task[8].

Mainly Medical Image Segmentation relay on five main approaches, namely threshold method, boundary base method, region based method and hybrid method. Clustering algorithm is a region based method, it is use to segment medical image and to find region of interest (ROI) [9].

III. Image Contrast

Contrast is the local change in a brightness and is defined as the ratio between average brightness of an object and the background strictly speaking, we should talk about luminance instead of brightness if our aim is to be physically precise.[10]. The measure of the image contrast(CON) can be expressed by the following equation:-

$$CON = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=0}^{N_g-1} \sum_{\substack{j=0 \\ |i-j|=n}}^{N_g-1} p(i,j) \right\} \quad ..(1)$$

This is a measure of the image contrast- that is a measure of local grey level variations. Indeed, $\sum_i \sum_j p(i,j)$ is the percentage of pixel pairs whose intensity differs by n. The n^2 dependence weighs the big differences more, thus CON takes high values for images of high contrast[11]

IV. K-MEANS CLUSTERING

The most popular partitional algorithm among various clustering algorithms is K-mean clustering. K-means is an exclusive clustering algorithm, which is simple, easy to use and very efficient in dealing with large amount of data with linear time complexity[12]. In 1967 MacQueen first proposed k-Means clustering algorithm. k-Means algorithm is one of the popular partitioning algorithm. The idea is to classify the data into k clusters where k is the input parameter specified in advance through iterative relocation technique which converges to local minimum[13].

K-means algorithm clusters observations into k groups, where k is provided as an input parameter. It then assigns each observation to clusters based upon the observation's proximity to the of the cluster. The cluster's mean is then recomputed and the process begins again[12] as shown in Algorithm(1) .

Algorithm (1): Traditional k-means Algorithm

a) Select k points as initial centroids
b) Repeat
c) From k clusters by assigning each point to its closest centroid
d) Recompute the centroid of each cluster until centroid does not change[13]

The objective function of K-mean algorithm is to minimize the intra-cluster distance and maximize the inter-cluster distance based on the Euclidean distance[12]. So to measure the quality of a clustering , sum square error (SSE) is used i.e. the error for each point is computed. The SSE is formally defined as follows[14]:

$$SSE = \sum_{i=1}^{c} \sum_{x \in C_i} d(x, m_i) \qquad (2)$$

In the above equation, m_i is the center of cluster C_i , while $d(x,mi)$ is the Euclidean distance between a point x and mi. Thus, the criterion function E attempts to minimize the distance of each point from the center of the cluster to which the point belongs [15].

The K-means algorithm requires three user-specified parameters: number of clusters K, cluster initialization, and distance metric. Cluster initialize with one object that is center for each cluster. A cluster is a set of objects such that an object in a cluster is closer (more similar) to the "center" of a cluster, than to the center of any other cluster. The center of a cluster is often a centroid, the average of all the points in the cluster[16].

V. EXPERIMENTAL RESULTS

The practical work was applied on the medical image. Six medical images as shown in Figure(1) were supplied to this work. These images were imported from the site[17] that contains atlas of brain images. The practical procedure of this work would be given. First the image contrast was computed according to equation(1). The results are listed in the table(1).

Table 1. Contrast values of original image

Image No. / Image Contrast	Contrast Value
1	7230.172
2	8584.796
3	11094.054
4	3712.466
5	4700.878
6	4625.713

The contrast values as shown in Table(1) are grouped into two groups; the (image 1,2, and 3) have high values however the second group(image 4,5, and 6) have low values.

Then k-means algorithm that was shown in Algorithm(1) was executed using the previous images with different number of clusters(k); k=2, k=3, k=4, k=5. The resulted images are shown in Figure(2), Figure(3), Figure(4), Figure(5). By human eye, it can be examined for the resulted clustering images(image 1,2, and 3) with k=2 as shown in Figure(2) would be implied more distinctive features than the second group of images(4,5, and 6).

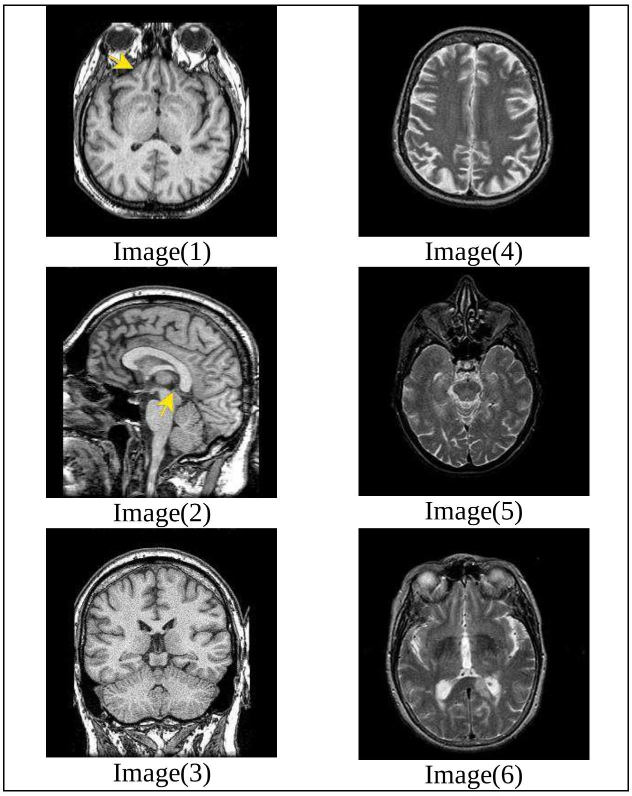

Fig. 1. Original medical images

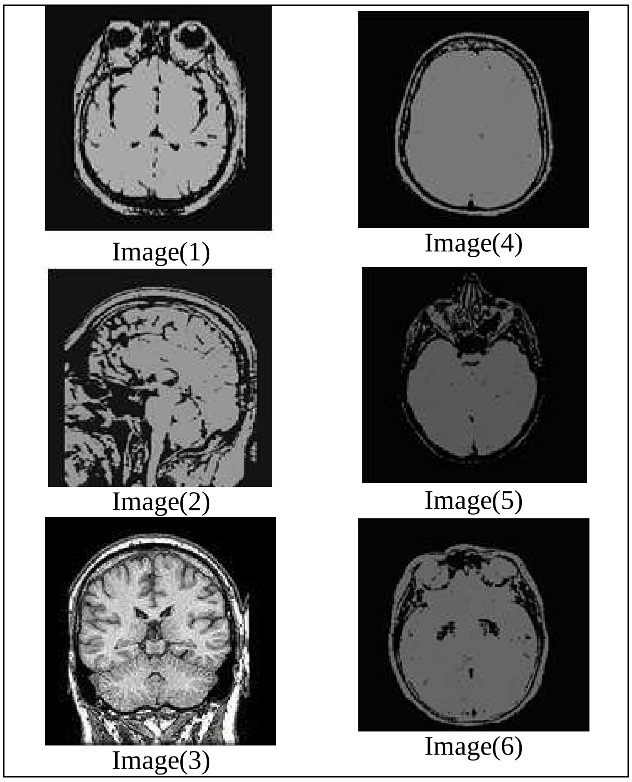

Fig. 2. Resulted images with clusters number(k=2)

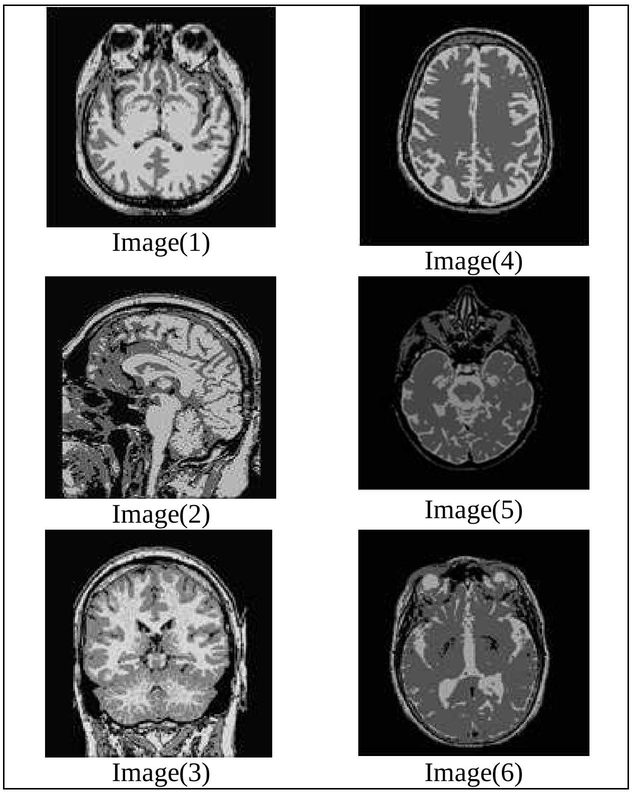

Fig. 3. Resulted images with clusters number(k=3)

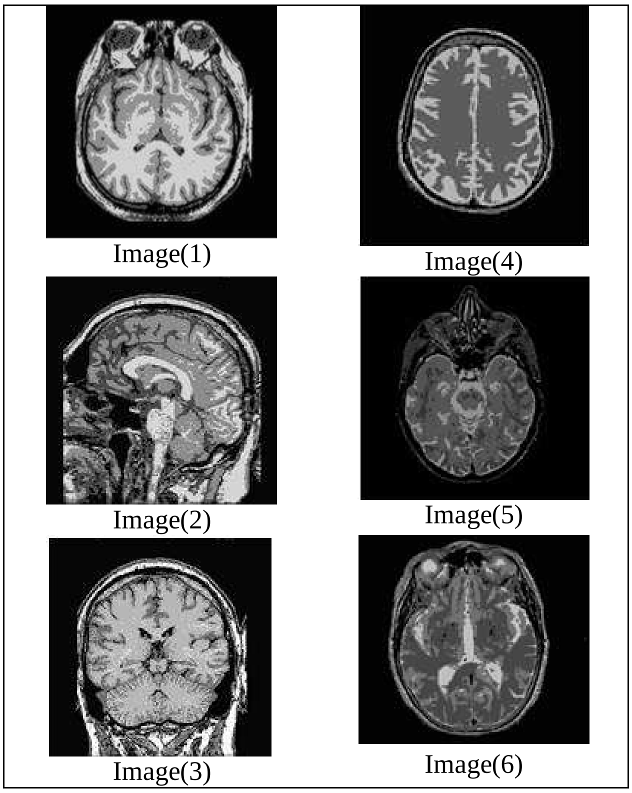

Fig. 4. Resulted images with clusters number(k=4)

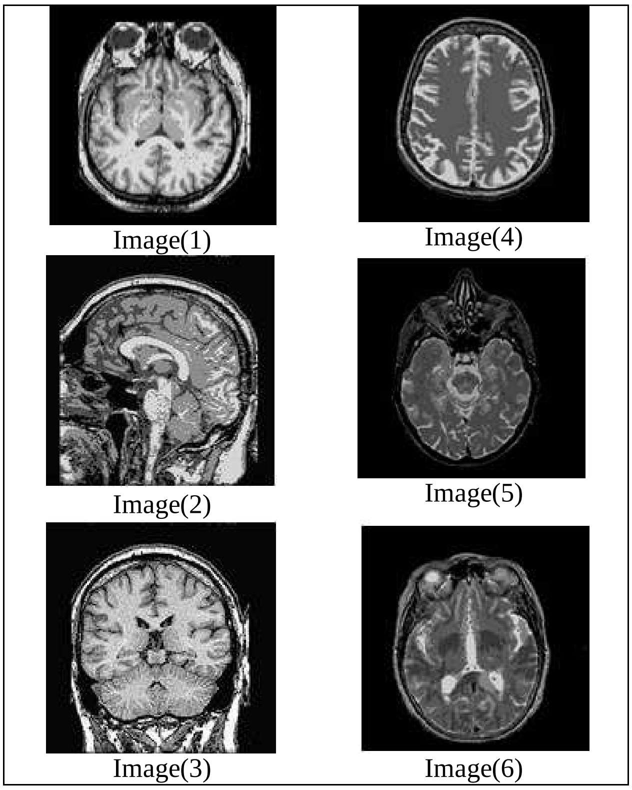

Fig. 5. Resulted images with clusters number(k=5)

Then the clusters number was increased, the quality of images(i.e. more features are become more clear) would improved in the two groups with the superiority of the first group as shown in Figures (3) , (4), and (5) where cluster number k=3,4, and 5.

International Journal of Electronics Communication and Computer Engineering
Volume 7, Issue 2, ISSN (Online): 2249–071X

VI. CONCLUSIONS

Subjective evaluation of the results in this paper give the proof for the role of the image contrast in identifying the number of clusters. The image with high contrast values gives good result with low number of clusters, so it can be chosen in clustering an image. However images with low contrast value didn't give good result with low number of clusters. Therefore, it must be chosen more clusters.

REFERENCES

[1] B.Sathya and R.Manavalan , "Image Segmentation by Clustering Methods: Performance Analysis", International Journal of Computer Applications (0975 – 8887), Volume 29– No.11, September 2011, p.p. 27-32.

[2] Saraswathi, S., y Allirani, A. (2013).Survey on Image Segmentation Via Clustering. Information, Communication and Embedded Systems (ICICES), International Conference, pp. 331-335.

[3] H.P. Ng, S.H. Ong, K.W.C. Foong, P.S. Goh, W.L. Nowinski, " MEDICAL IMAGE SEGMENTATION USING K-MEANS CLUSTERING AND IMPROVED WATERSHED ALGORITHM" , , IEEE Southwest Symposium on Image Analysis and Interpretation2006, Date 26-28 March 2006, p.p. 61-65.

[4] D T Pham_, S S Dimov, and C D Nguyen, "Selection of K in K-means clustering", Proc. IMechE Vol. 219 Part C: J. Mechanical Engineering Science, p.p. 103-110.

[5] K. Trupti M. and M. Prashant R , " Review on determining number of Cluster in K-Means Clustering", Volume 1, Issue 6, November 2013, p.p. 90-95.

[6] Gunnar Lathen, " SEGMENTATION METHODS FOR MEDICAL IMAGE ANALYS I S", Linkoping studies in science and technology, 2010, Thesis No. 1434, p.p. 2.

[7] H.P. Ng, S.H. Ong, K.W.C. Foong, P.S. Goh, W.L. Nowinski, "MEDICAL IMAGE SEGMENTATION USING K-MEANS CLUSTERING AND IMPROVED WATERSHED ALGORITHM", Authorized licensed use limited to: Tamkang University. Downloaded on September 8, 2009 at 01:08 from IEEE Xplore. p.p. 61-65.

[8] MOHAMMED SABBIH HAMOUD AL-TAMIMI, GHAZALI SULONG, " TUMOR BRAIN DETECTION THROUGH MR IMAGES: A REVIEW OF LITERATURE", Journal of Theoretical and Applied Information Technology , 20th April 2014. Vol. 62 No.2, p.p. 387-403

[9] Juilee Anil Katkar, Trupti Baraskar, " A Review: Clustering Techniques for Medical Image Segmentation", International Journal of Advance Foundation and Research in Computer (IJAFRC) Volume 1, Issue 12, December 2014. ISSN 2348 – 4853.

[10] Milan Sonka, Vaclav Halavac and Roger Boyle, " Image Processing, Analysis, and Machine Vision", Third Edition , THOMSON, p.p.25.

[11] T. Sergios and K. Konstantinos, " Pattern Recognition", Fourth Edition, ACADEMIC Press, p.p 416-417.

[12] M. R. Sindhu, S. Rachna, 2013, "Clustering Algorithm Technique", International Journal of Enhanced Research in Management and Computing Applications,vol. 2, Issue 3, March 2013.

[13] M.P.S Bhatia and Deepika Khurana, 2013, " Experimental study of Data clustering using k- Means and modified algorithms "International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.3, No.3, May 2013, p.p. 17-30.

[14] T. Pang-Ning **,** S. Michael**,** and K**.** Vipin, 2006, " Introduction to Data Mining", ©2006 • Addison-Wesley •, p.p487-568.

[15] H. Maria, B. Yannis, and V. Michalis, 2001, " On Clustering Validation Techniques", Journal of Intelligent Information Systems, 17:2/3, 107–145, 2001, Kluwer Academic Publishers. Manufactured in The Netherlands

[16] A.K. Jain, 2010, "Data clustering: 50 years beyond K-means", Pattern Recognition Letters 31 (2010) 651–666.

[17] http://www.med.harvard.edu/aanlib/home.html

AUTHOR'S PROFILE

Lamia AbedNoor Muhammed PhD Computer Science, Iraqi Commission for Computers and Informatics, Informatics Institute for Graduate Studies, Iraq/2008. I have worked in college of Computer Science and Mathematics, Al-Qadissiya university, Iraq 2008 - so far. I have taught the following subjects system analysis and design, database, data mining, image processing.