Australian Journal of Basic and Applied Sciences, 8(9) June 2014, Pages: 422-432

AENSI Journals

Australian Journal of Basic and Applied Sciences

ISSN:1991-8178

Journal home page: www.ajbasweb.com

A Comparative Study on the Automatic Segmentation of Colon in CT Colonography Using K- Means And Fuzzy C-Means Clustering

[1]K. Gayathri Devi, [2]Dr. R. Radhakrishnan

[1]*Anna University, Dr.N G P Institute of Technology, Department of Electronics and Communication, Associate Professor, Coimbatore, India.*
[2] *Anna University ,Sri Sakthi Institute of Engineering and Technology, Principal ,Coimbatore, India.*

ARTICLE INFO

Article history:
Received 2 March 2014
Received in revised form
13 May 2014
Accepted 28 May 2014
Available online 23 June 2014

Keywords:
Computed Tomography, Image Segmentation, Histogram , Colonography, unsupervised learning, k-nearest neighbor, clustering, Histogram, ITK snap ,Sensitivity and Specificity

ABSTRACT

Background: An approach for the automatic segmentation of colon which is filled with air and opacified fluid in CT colonography is presented. The presence of an air and partial volume effect is a challenging factor for the segmentation of colon. Therefore we have developed a new approach for the automatic segmentation of colon.**Objective**: To extract the segment of colon from all the slices of a particular dataset more precisely and to perform 3D rendering of segmented colon structures in order to improve the accuracy and sensitivity. **Methods Used:** In this paper, we use K-means clustering and Fuzzy C-means clustering, an unsupervised methods to identify and group the colon segments as clusters. Automatic colon segmentation from abdomen CT data starts with removal of the air portions that is outside the body of the subject, masking of the lungs should be done because they have the same intensity values as the colon segments which is a very difficult task, segmentation of colon segments from all the slices and finally 3D rendering of all the segmented colon segments. **Conclusion:** The above proposed methods were tested with 5 dataset of abdominal CT images downloaded from TCIA cancer imaging archived. The segmentation was validated by comparing the segmented output with manually segmented output using ITK snap. The three metrics accuracy, sensitivity and specificity were calculated. The proposed method shows an improvement in accuracy and sensitivity compared to the existing graph cut and level set methods.

To Cite This Article: K. Gayathri Devi, Dr. R. Radhakrishnan, A Comparative Study on the Automatic Segmentation of Colon in CT Colonography Using K- Means And Fuzzy C-Means Clustering. ***Aust. J. Basic & Appl. Sci.,*** *8(9): 422-432, 2014*

INTRODUCTION

Colorectal cancer is cancer that occurs either in the colon or rectum. Most colorectal cancers develop first as colorectal polyps, which are growths inside the colon or rectum that may later become cancerous. It affects both men and women mostly above the age of 50. It is the third most common cancer among men after prostate and lung cancer(Cancer facts and figures,2012). For women, colorectal cancer is the third most common cancer after breast and lung cancers (Paul *et al*., 2006).

Computed tomographic colonography (CTC) or virtual colonoscopy is an evolving method for detecting colon polyps which records the 3D image of the patient's abdomen. Virtual colonoscopy combines axial spiral CT data acquisition of the air-filled and cleansed colon with 3-dimensional imaging software to create endoscopic images of the colonic surface (Xiaoyun *et al*., 2008). Two main factors that limit CT colonography are its excessive interpretation time and the variable sensitivity among readers. This paper focuses on the advantages of computer-aided detection (CAD) techniques for the segmentation of colon which will aid the identification of polyps for the detection of colorectal cancer.

Listed below are some of the points which make the colon segmentation a difficult task:

1. The colon is not the only gas filled structure in the abdomen but there are other organs which is having the same intensity values as the colon. Slices adjacent to the colon contains portion of lung which should be removed.
2. The existence of different areas that have the same high CT number (intensity) such as bones, air and contrast enhancement fluid (CEF) should be removed for proper segmentation of colon
3. Obstructions such as stools, small bowel, polyps and residual feces in the colon.
4. Folding, unfolding of the colon structures such as haustral folds on the borders of colon.
5. Length of the colon and the location of the colon vary for different slices.

Corresponding Author: Dr. Radhakrishnan, Anna University, Dr. NGP Institute of Technology, Department of Electronics and Communication, Associate Professor, Coimbatore, India.
E-mail: gayathridevik@yahoo.com

Australian Journal of Basic and Applied Sciences, 8(9) June 2014, Pages: 422-432

These difficulties make the segmentation of colon more complicated specially in situation where an automated algorithm should be implemented.

Related Works:

Research in colon segmentation, has accelerated over the last few years. It is generally difficult to compare different approaches, because of the following factors, preparation of patient varies before image acquisition, the set of datasets used for segmentation and geometric features of colon taken into consideration for segmentation also varies. A common approach is to segment colon by following the three basic steps

- Air surrounding the body should be removed
- Air contained in the lungs is masked.
- Segment the colon on the various slices

Both semi automatic and automatic method for the segmentation of colon is proposed. Alberto Berta *et al*.,2009 proposed the automatic segmentation of colon using 3D seeded region growing algorithm . Are Losnegrad *et al*, 2010 proposed a semi automatic segmentation of sigmoid and descending colon using fast marching method and evaluated using parameter dice coefficient (DC). Dongqing Chen *et al*, 2000 proposed a method which modified vector quantization technique for a low-level image classification and a region-growing strategy for a high-level feature extraction(Dongqing Chen *et al*, 2008) but the vector quantization approach for segmenting the colon were initially proposed by Liang. Z *et al*. in 1999. Liang. Z *et al*,1999 compared the vector quantization approach with Maximum Posteriori Framework (MAP) and Threshold Method approach and proved that for the segmentation of colon the threshold method has consumed less time for the segmentation of colon than the other two methods . Dongqing Chen *et al*,2008 proposed a method which used thresholding to remove the opacified fluid in the colon and then the seed were initialised manually to remove the colon segments which were tested on 10 real CT colonography datasets, and the accuracy achieved was 96.06% and Dongqing Chen *et al*,2009 later extended the above concept for dataset of 22 real CT colongraphy cases and proved that the accuracy is 98.40% where there is a slight improvement in accuracy compared to the previous paper . Lin Lu *et al*, 2012 proposed an improved method for colon segmentation using two widely virtual colon unfolding (VU) techniques, the ray-casting technique and the conformal-mapping technique. Lihong Li *et al*, 2002 and Lihong Li *et al*,2007 developed a method for cleansing the colon using hidden Markov random filed (MRF) model and to identify the enhanced regions of residual stool/fluid which may reside inside the colon maximum a posterior probability (MAP) method . Marek Franaszek *et al*,2006 developed a hybrid algorithm uses modified region growing, fuzzy connectedness and level set segmentation for the segmentation of colon and to identify the polyps ,the algorithm were tested on 160 CT colonography scans . Sebastian Zander *et al*, 2006 developed a method to locate of the colon serosal tissue boundary from a CT colonography (CTC) scan using level set based method. Tarik A. Chowdhury *et al*, 2005 proposed a centroid based method for geometry analysis of the air inflated region inside the CT data for colon segmentation.

Vahid Taimouri *et al*, 2011 uses constrained least-squares filtering (CLSF) technique for pre-processing and both cleansing and stool detection in the colon is done prior to segmentation. Wyatt *et al*, 2000 have developed a method to locate seed points and segment the gas-filled lumen sections without user supervision and tested on 20 datasets.

Automated Segmentation of Colon:

The Proposed Method:

The overall block diagram of the proposed method is as shown in the Figure 1. Abdominal CT image in DICOM format of size 512 X 512 downloaded from TCIA imaging Archive link (cancerimagingarchive.net) is given as the input to our method. The dataset consisting of 312 slices is preprocessed to improve the quality of the input image. The segmentation of colon segments is divided into six phases. Initially grouping the pixels of colon segments into clusters is achieved by using K means clustering or fuzzy C means clustering. The grouped clusters are labelled to find the connectivity in colon segments using 8 neighbourhood connectivity. The labelling is followed by segmenting the air portions that is outside the body of the subject, masking of the lungs and finally segmentation the colon segments. The portions of colon segmented by our algorithm for every slice is compared with the manually segmented colon structures segmented using ITK snap software .The performance of our method were evaluated with three parameters namely accuracy , sensitivity and specificity and compared with the graph cut and level set method.

Segmentation of Colon using K-Means Clustering:

Pre-Processing:

The sample input slice out of 312 slices as shown in fig 4 (a) is preprocessed to improve the quality of the input image of size 512 X 512. The steps for the preprocessing is

- Convert the image to Double precision type (unsigned integer16- bit image).Assign a variable r.
- Calculate the minimum and maximum pixel value

Australian Journal of Basic and Applied Sciences, 8(9) June 2014, Pages: 422-432

- Assignment of two variables q1 and q2.Set q1=0 ; q2=1
- Assignment of two variables p1 and p2 and convert into double format. Set p1=0 ;
  p2 = maximum(Input Image)/(2^{16}-1)
- Create a zero-matrix equivalent to the size of input image to store the preprocessed image
- Replace the pixel (r) which have a value (r $\geq$ minimum & r $\leq$ maximum) pixel with
  P[i,j] *(q2-q1)/(p2-p1)+q1
- Convert the image to 8 bit unsigned integer for viewing.

Clustering Techniques:

The two clustering techniques which have been considered for the segmentation of colon is K-Means clustering and Proposed Fuzzy C-Means clustering as shown in fig.2.These two techniques works on the preprocessed image. The concept of thresholding does not apply as the voxels in the colon, portions of image outside the body of the subject as well as lungs lies in the range of -950 to -1200HU.Hence Clustering techniques was adopted.

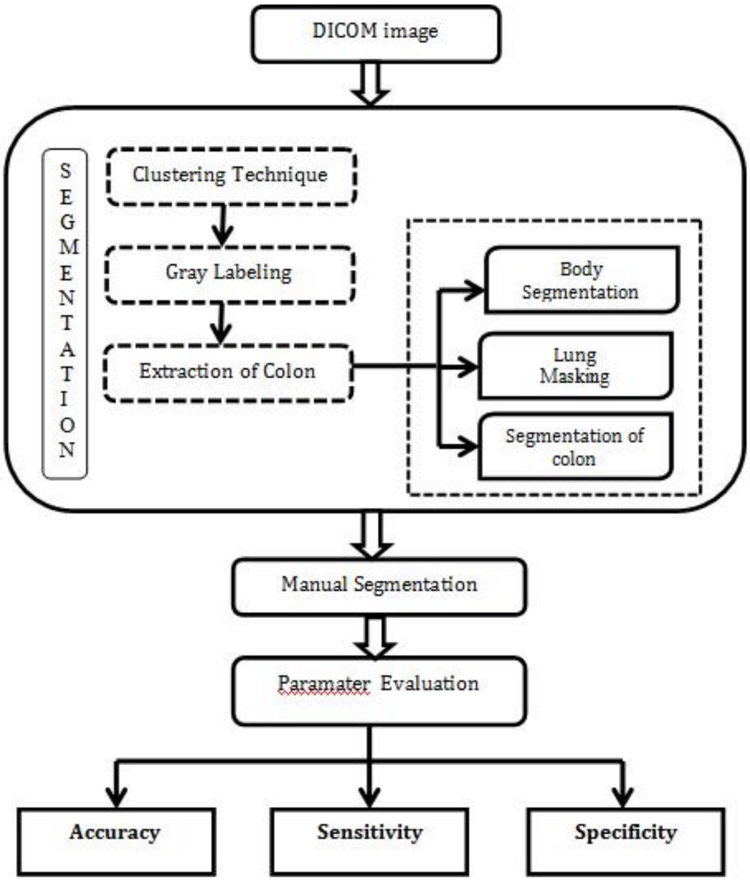


Fig. 1: Framework for the segmentation of colon.

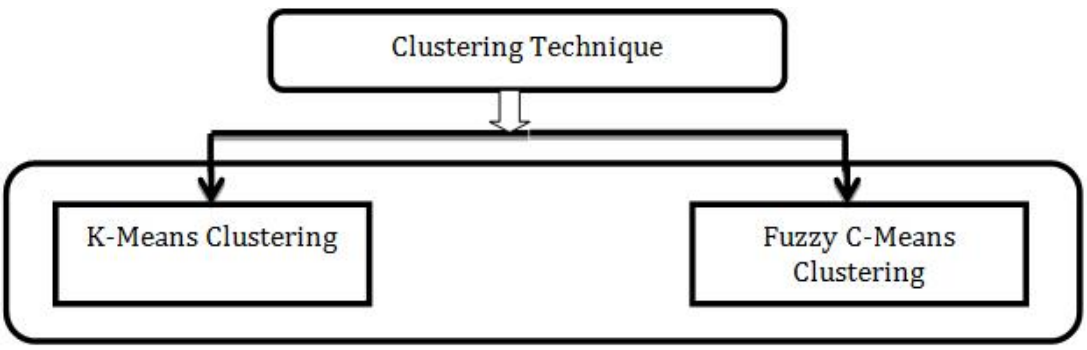


Fig. 2: Clustering techniques for the segmentation of colon.

Australian Journal of Basic and Applied Sciences, 8(9) June 2014, Pages: 422-432

K-means Clustering:

K-Means clustering is applied on the preprocessed image (Pappas, 1992). The performance of this algorithm is determined on the value of K cluster centers which are chosen at random. In this clustering, the set of pixels in the image are represented as $<x_1,x_2\dots x_p>$, these pixels should be grouped into clusters which will exhibit high intra class similarity and low inter class similarity. Initialize cluster centre randomly $< C_1, C_2, \dots\dots\dots C_k >$.The clustering of the pixels is performed based on the following steps (Stephen J Redmond *et al*., 2007)

1. Assign the points in the input space $<x_1,x_2\dots x_p>$ to only one cluster based on the Euclidean distance from the point to the cluster. The element x will assigned to the closest cluster if it satisfies the condition as stated in equation 1, such that the distance between the element x and cluster v should be minimum in comparison with the distance to the other clusters

$$x_i \in C_j \;\text{ if } \left|x_i - v_j\right| < \left|x_i - v_k\right| \; k{=}1,2,3\dots\dots,c \; k \neq j \text{ -----------------} \qquad (1)$$

2. After the assignment of the pixels to the clusters, labelling of the each pixel is done.
3. Moving each cluster centroid C to the mean of the points assigned to it. Arithmetic mean of all the points in the cluster is determined and the random centre which was initially assumed will keep on moving over the eigen values such that the distance of the element x with respect to cluster centroid keeps on reducing after each iteration.
4. The above steps from 1 till 3 are executed till the stopping criterion as stated in equation 2 is satisfied.

$$\sum \left\| C(t+1) - C(t) \right\| < \text{Threshold value -----------------} \qquad (2)$$

It aims at minimizing the objective function as shown in equation 3. The objective function

$$J_m(P,V) = \sum_{j=1}^{c} \sum_{x_i \in C_j} \left|x_i - v_j\right|^2 \text{ -----------------} \qquad (3)$$

where P is the no of Clusters that the image has partioned into

V is the vector of cluster centers to be determined

J measures the sum of the squared distance between each pixel x and the cluster centroid C_j to which it has been assigned. Step 1 minimizes J based on the distance of each pixel with respect to cluster centroid , keeping the cluster centroid constant. In step 3 J is getting minimized with respect to cluster centroid, keeping the distance parameter constant. Therefore J will converge to a minimum value. The number of iteration depends on the random value chosen.

The Pseudo code for the K means clustering is given below.

1. Initialize the center of the clusters	cluster centroid was initialised at [0,150,255]
2. Attribute the closest cluster to each data point 3. Set the position of each cluster to the mean of all data points belonging to that cluster	Create a zero matrix for storing the clustered image. While(stopping criterion) for i=1:number of rows (512) for j=1: number of columns (512) Set the value of the pixel at location (i,j) to current pixel Initialize the variable current pixel vector ,where a unit matrix of the length 9 is generated. Assign a value to current pixel vector by multiplying the value of current pixel to each value in cluster prominent pixels Calculate the Euclidean distance measure The minimum distance is found from the above matrix and the corresponding prominent pixel value is chosen The 1st pixel value is replaced with the selected prominent pixel value in the clustered image (zero matrix), end end
4. Repeat steps 2-3 until convergence	

Proposed Fuzzy C-Means Techniques:

Fuzzy C-Means clustering is applied on the preprocessed image(Bezdek,1981 and Yong Yang *et al*.,2007) . Fuzzy clustering techniques allow the pixel to belong to more than one clusters thus it belongs to the category of soft partitioning of pixels in a dataset, where as in k-means clustering the pixels will belong to only one cluster. The degree of membership of an element belonging to any cluster is inversely proportional to the difference of distance of that element to the clusters which has been initialized (Tolias *et al*., 1998). The larger the distance is less the chance of belonging to that cluster, if the distance of the pixel to that cluster is small than obviously that pixel will belong to that cluster. The objective function as shown in the equation 4 is developed to reduce the Euclidean distance between each point in a cluster and maximise the Euclidean distance between cluster centres.

Australian Journal of Basic and Applied Sciences, 8(9) June 2014, Pages: 422-432

$$J_m(P,V)=\sum_{j=1}^{c}\sum_{x_i\in C_j}\left(\mu_{C_i}(x_k)\right)^{m}\left|x_k-v_i\right|^2 \text{ -----------------} \quad (4)$$

Where P=fuzzy partition of the dataset X formed by $C_1,C_2,\ldots\ldots C_k$
V=vector of cluster centre's that should be determined
v_i= centre of the cluster C_i
m= weight that determines the degree to which partial members of clusters affect the clustering result
$\mu_{C_i}(x_k)$ is the membership degree to which the input x_k belongs to cluster c_i

The steps are
1 .Consider a M-dimensional N data points represented by x_i (i=1,2 ….N) that is to be clustered.
2. Fix the value of cluster centre, the clusters were randomly initialized as 8 and 250.56 .The no of cluster centers should be greater than or equal to 2. Intialize the partition matrix. Each step in this algorithm will be labeled r, where r= 1,2,3,4,5
3. Initialize the matrix of size 512 X 512 to the value of cluster centres. In our approach two matrixes will be initialized as we have considered two clusters.
4. Concatenate the cluster initialised matrix to the dimension of the input image.
5. Initialize the membership matrix to the dimension of the input image. The membership value m' should have a value in the range [0 1].It is also called a weighting parameter.
6. Concatenate the membership initialised matrix U to the dimension equal to cluster center such that the cluster matrix and the membership matrix are of the same size and dimension.
7. Calculate the Euclidean distance between i^{th} data point and j^{th} cluster center with respect to, say m^{th} dimension as shown in equation 5:

$$D_{ijm} = \|(x_{im} - CC_{jm})\|. \qquad \text{-----------------} \quad (5)$$

8. Determine the cluster centres{ $CCjm^{(r)}$ } , for j^{th} cluster and its m^{th} dimension by using the expression 6 given below:

$$CC_{jm}=\frac{\sum_{i=1}^{N}U^{m'}_{ijm}\ x_{im}}{\sum_{i=1}^{N}U^{m'}_{ijm}} \qquad \text{-----------------} \quad (6)$$

9. Update the partition matrix for the rth step as follows in equation 7

$$U_{ijm}^{(r+1)}=\left[\sum_{j=1}^{c}\left(\frac{d_{iim}^{(r)}}{d_{ijm}^{(r)}}\right)^{2/m'-1}\right]^{-1} \qquad \text{-----------------} \quad (7)$$

10. Repeat from Step 3 to Step 9 until the changes in $U\le\varepsilon$, where ε is a pre-specified termination criterion Here the criterion set was 0.0001.In our proposed Fuzzy c means techniques all the portions of the colon will be segmented with 5 iterations as shown in fig 4(b),more no of iterations are not needed.

Labeling of the Clustered Images:

Each connected components are uniquely labeled, which help in locating the colon and non-colon segments. Connected component labeling is used to detect connected regions in binary digital images. Connected component labeling works by scanning an image, pixel by pixel (from top to bottom and left to right) in order to identify the connected pixel region i.e., regions of adjacent pixels which share the same set of intensity values. Connected components are found using 8-connectivity. Thus neighbouring pixels in all eight directions are checked for connectivity.

The gray labeled image is converted into binary image for labeling process. A graph containing vertices and connecting edges is constructed and an algorithm traverses the graph, labeling the vertices based on the connectivity and relative values of the neighbours. After labeling all the connected components, they can be differentiated using rgb colors as shown in fig 4(c).

Fig.3 represents the flow of gray labeling and conversion to binary image respectively.

Australian Journal of Basic and Applied Sciences, 8(9) June 2014, Pages: 422-432

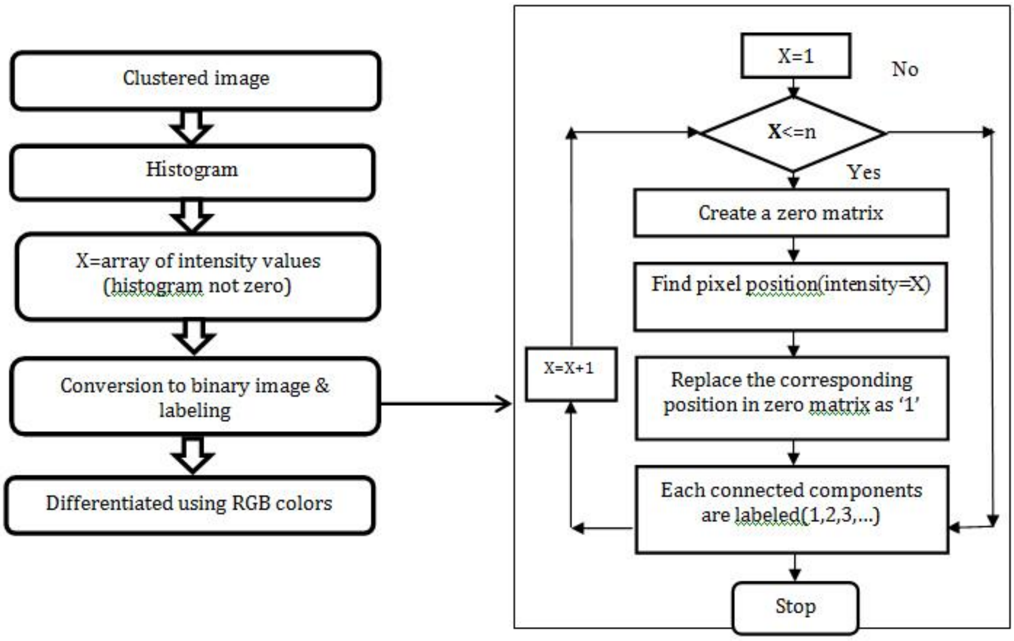


Fig. 3: Framework for labeling of Gray images

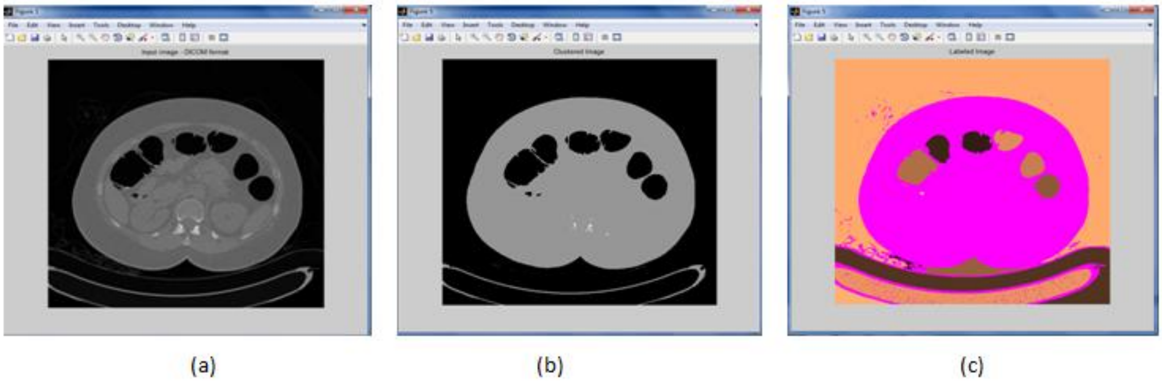


Fig. 4: (a) Input Image b) Clustered Image c) labeling of Gray images

Extraction of Colon:

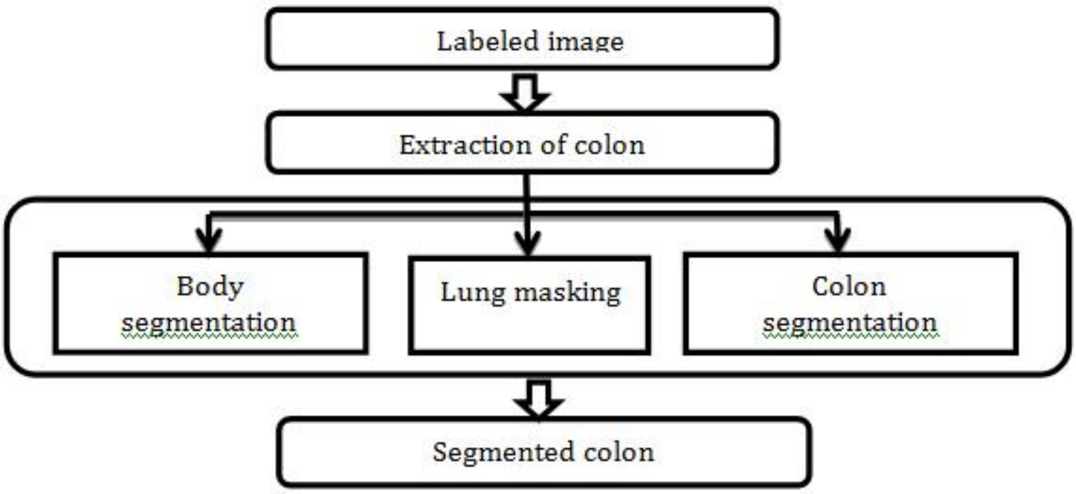


Fig. 5: Framework for Segmentation of colon.

Australian Journal of Basic and Applied Sciences, 8(9) June 2014, Pages: 422-432

Body Segmentation:

The region outside the Body of the subject should be removed as shown if fig 6(a) .The values of the connected component (CC) ranging from 1 to maximum value which have found in previous step is further used for removing the outer region. The process is executed for all the values of connected component one after the another. First the position of the pixels which contain the value 1 is analyzed. The coordinates of the pixel where this connected component exist is stored as a temporary binary image. This is multiplied with the clustered image. The mean value is computed using the formula as shown in equation 8

$$\text{Mean} = \sum_{cc=1}^{\text{Max\_value}} \frac{\sum \text{uint8(clustered\_image).* uint8(temp\_binary\_image)}}{\text{length(pos)}} \quad \text{-----------------} \qquad (8)$$

If the mean value is greater than 50 and less than 200, these cc positions are fixed. After analyzing the position for every CC the corresponding position in Body segmented image is made as ‘1’ and the background is filled with holes using ‘imfill’ function.

Lung Masking:

In a dataset out of 315 slices, 20 slices contain the abdominal organs Kidneys lungs and spleen in additional to colon. Thus a slice which contains lungs and colon is shown in fig 6(b). Thus after the removal of the background outside the body of the subject as shown if fig 6(a), lung masking is performed. Masking of lungs is very important since they have same intensity as colon. Lung reference image is obtained with reference to properties of the CC and if the slices contains both colon and lung portion, with this reference image colon portion can be segmented using K-means clustering as shown in fig 6(c).

(a) **(b)** **(c)**

Fig. 6(a): Removal of the Background outside the body of subject (b)Slice containing both lungs and colon (c) Colon portion segmented by our method.

Colon Segmentation:

After the segmentation of the background and masking of the lungs, the third step involves segmenting the colon portion. The final segmented image will be having ‘1’ in the place where current connected component will have the intensity value less than 50, and area in between 2 and 50,000. This is multiplied with body segmented image and stored in final segmented image. Now the positions of pixel which are ‘0’ in the final segmented image are found and stored in ‘zeros position’. Final segmented image is multiplied with gray image (i.e., contrast enhanced image) and stored in segmented gray image.

Algorithm

Step1: Start the program.

Step2: Find the maximum value in the labeled Image and for loop is executed till this maximum value.

Step3: Create zero matrix for both temporary binary image and temporary gray image.

Step4: Find the position of the connected components in the final label and label it as ‘1’ and the corresponding position in temp binary is made as 1.

Step5: Multiply Gray image (i.e., Contrast enhanced image) with temporary binary image and the result is stored as Temporary Gray image. Length of position is stored in area. Find mean intensity by adding all values in Temporary Gray image and divide it by area.

Step6: Check for area in between 2 and 50,000 and also check for mean intensity.

Step7: If Mean intensity is less than 50, then the corresponding position in final segmented image is made as ‘1’and repeat the process till the maximum value.

Step8: Multiply the final segmented image with body segmented image and stored in final segmented image.

Step9: Find the pixel value which are ‘0’ in the final segmented image and store it in ‘zeros position’.

Australian Journal of Basic and Applied Sciences, 8(9) June 2014, Pages: 422-432

Step10: Multiply the final segmented image with gray image (i.e., contrast enhanced image) and the result is stored in segmented gray image.

Step11: Finally replace the ‘zero position’ in the segmented gray image as 0 and the remaining portion are replaced with 255.

Step12: Stop the program.

Manual Segmentation:

Ground truth is the segmentation of colon structures by radiologist manually by visual inspection. These images are used as reference in predicting how well the algorithm segmented output has detected the colon and the non-colon structures after processing.

Manual segmented output is drawn using ITK-SNAP software (Yushkevich *et al*.,2005) and stored in a separate folder. Then, while processing the slices, ground truth image is called and then matrix values of the ground truth image is compared with the matrix values of the algorithm segmented image. The 3D rendered output of the manually segmented image is shown in fig 9(a).

RESULTS AND DISCUSSION

The previous section considered the segmentation of colon .The figure 7(b) shows the evaluation of the colon segments evaluated by our Fuzzy C means clustering algorithm . Thus most of the colon segments were segmented by our algorithm. The segmentation were evaluated with 5 datasets downloaded from TCIA imaging cancer archive .The segmented output for different slices pertaining to one dataset is shown in the figure 8 by using Fuzzy C-means algorithm and the final 3D rendered segmented output for the dataset consisting of 315 slices is shown in figure 9(b).

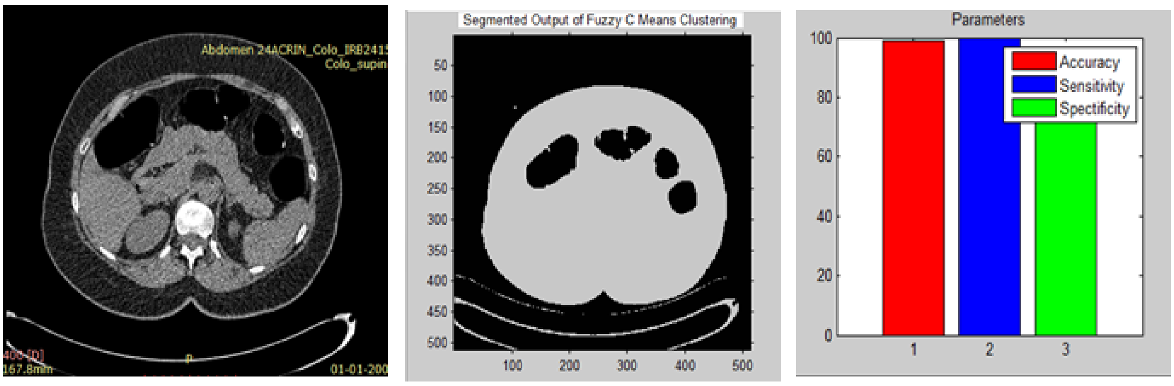


Fig. 7(a): Sample slice after Preprocessing (b) Colon portion segmented by our method (c)Three parameters evaluated by Fuzzy C-means clustering

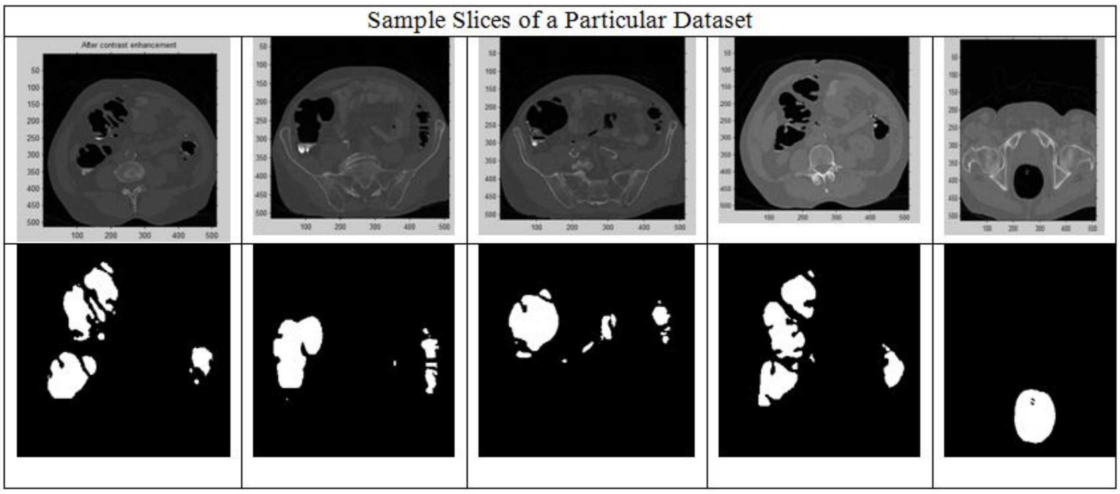


Fig. 8: The Top row shows different slices of a particular dataset and the bottom row shows the corresponding segmented output by Fuzzy C means algorithm

The parameters that were calculated for the evaluation of our algorithm are

- True positive(TP): Colon segments correctly identified as Colon Segments
- True negative(TN):Non-Colon structures correctly identified as Non - Colon parts
- False positive(FP):Non-Colon structures incorrectly identified as Colon
- False negative(FN): Colon incorrectly identified as Non-Colon structures.

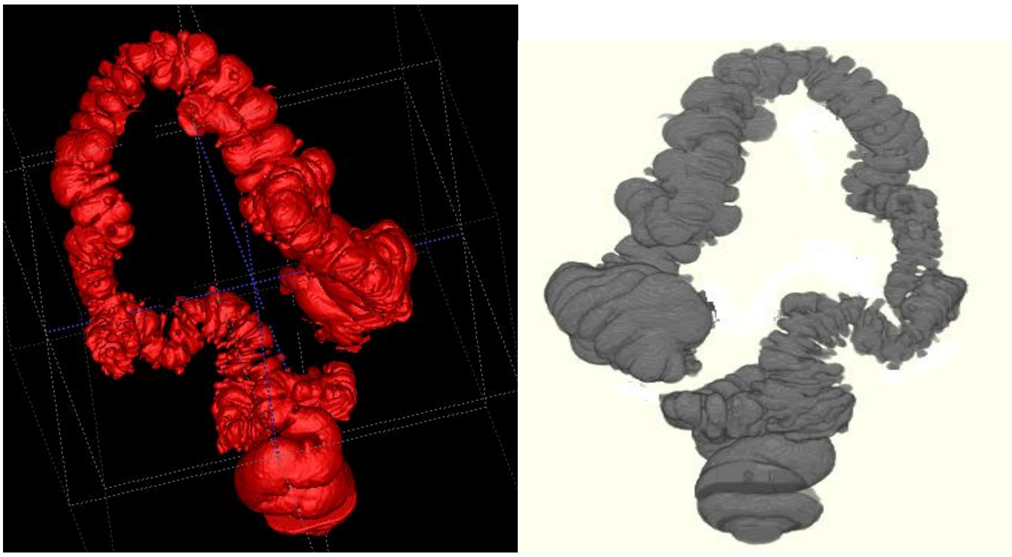
Based on the four parameters the following three parameters were calculated

- Accuracy--is the degree of closeness of measurements of a quantity to that quantity's actual (true) value
- Sensitivity--is the ability to correctly identify colon
- Specificity--is the ability of a test to correctly identify non-colon structures

Table1 lists the parameters considered for the evaluation of our algorithm.

Table1: Tabulation of parameters and their formulae.

PARAMETERS	FORMULA
Accuracy	(TP+TN)/(FP+FN+TP+TN)
Sensitivity	TP/(TP+FN)
Specificity	TN/(FP+TN)



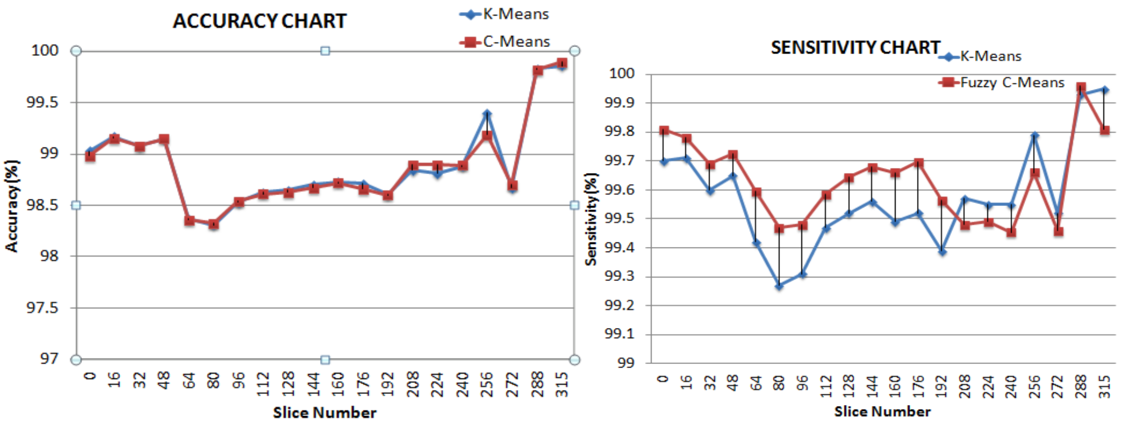
a) Manual Segmentation **b)** Segmentation by the proposed method

Fig. 9: 3D rendering of all the segmented 315 slices.

Table 2 lists the accuracy and sensitivity obtained for sample 20 CT slices of a particular dataset consisting of 315 slices in percentage both for K-means and Fuzzy C-means algorithm. The Accuracy and Sensitivity chart for one dataset consisting of 315 CT Slices is shown in Fig.10. (a) and (b).The charts indicate that both for K-Means and fuzzy C- means the accuracy achieved is 98.8% but there is an improvement in sensitivity for Fuzzy C-means clustering when compared to K-means clustering from 97.5% to 98.5 % . The specificity achieved is nearly 90%. Thus the average value of all the three parameters calculated for one dataset consisting of 315 slices is tabulated in Table.3 .The comparison of the proposed fuzzy C-Means, K-Means method with the previous methods namely Graph-cut and Level set method is tabulated in Table.3. Thus both the accuracy and sensitivity has been improved but the specificity has been decreased because when focusing on sensitivity naturally the specificity will decrease.

Conclusion:

In this paper, an approach for the automatic segmentation of colon was proposed by both K-means and Fuzzy C-means Clustering. The mode of grouping the clusters varies and the time utilized for segmentation of colon using Fuzzy C-means Clustering is lower when compared to K-Means Clustering. The proposed method will segment different portions of colon for each slice using both the clustering techniques and the segmented portions were compared with the ground truth images generated by using ITK Snap. The calculated parameters for one dataset consisting of 315 slices varied in the range of 98±0.86% (accuracy), 99.8±0.15% (sensitivity) and 85±4.0% (specificity). The testing has also been carried out for 5 datasets of abdominal CT images downloaded from TCIA cancer imaging archive and the experimental results showed an enhanced percentage of accuracy, sensitivity and specificity. The proposed Fuzzy C-means shows an improvement in sensitivity when compared to K-Means clustering. In a dataset consisting of 315 slices, Fuzzy C-means shows an improvement in sensitivity up to 210 slices as shown in fig 10(b) which increases the overall sensitivity to 98.5% when compared to k-means clustering whose sensitivity is 97.5% , existing graph cut and level set method whose sensitivity is 94.1% and 96.02%. This method could be successfully used to segment the colon to help diagnose colon cancer detection during virtual colonoscopy.



(a) **(b)**

Fig. 10(a): Accuracy chart (b) Sensitivity chart.

Table 2: Experimental results obtained for 20 CT slices.

Slices	Accuracy (%)		Sensitivity(%)	
	K-Means	C-Means	K-Means	C-Means
0	99	98.9	99.7	99.8
16	99.1	99.1	99.7	99.8
32	99.1	99.1	99.6	99.7
64	99.2	99.2	99.7	99.7
80	98.4	98.4	99.4	99.6
96	98.3	98.3	99.3	99.5
112	98.5	98.5	99.3	99.5
128	98.6	98.6	99.5	99.6
144	98.6	98.6	99.5	99.6
166	98.7	98.7	99.6	99.7
176	98.7	98.7	99.5	99.7
192	98.7	98.7	99.5	99.7
208	98.6	98.6	99.4	99.6
224	98.8	98.6	99.6	99.5
240	98.8	98.6	99.6	99.5
252	98.8	98.6	99.6	99.5
256	99.4	99.1	99.8	99.7
277	98.6	98.4	99.5	99.5
288	99.8	99.8	99.9	100.0
315	99.8	99.7	100.0	99.8

Table 3: Tabulation of comparison of the parameters .

Parameters	GRAPH CUTS (Asem M. Ali *et al.*,2007)	LEVEL SETS (Dongqing Chen.,2008)	K-Means Clustering	Proposed Fuzzy C-Means Clustering
Accuracy	90.8 %	97.6%	98%	98%
Sensitivity	94.1 %	96.02%	97.5 %	98.5%
Specificity	94.3 %	96.8%	92 %	92%

REFERENCES

Alberto Berta, Ivan Dmitriev, Silvano Agliozzoa, Natalia Pietrosemolib, Mark Mandelkernc, Teresa Gallod, Daniele Regged, 2009. "An automatic method for colon segmentation in CT colonography," Computerized Medical Imaging and Graphics, 33:325–331.

Are Losnegrad, Liv Bolstad Hysing, Ludvig Paul Muren, Erlend Hodneland and Arvid Lundervold, 2010. "Semi-automated segmentation of the sigmoid and descending colon for radiotherapy planning using the fast marching method," Physics in Medicine and Biology, 55: 5569-5584.

Asem, M. Ali and A. Aly Farag, 2007.“Graph Cut Based Segmentation of Multimodal Images,” ISSPI,: pp: 1047-1052.

Bezdek, J.C., 1981.” Pattern Recognition with Fuzzy Objective Function Algorithms”. New York: Plenum Press.

“Cancer facts and figures,” Amer. Cancer Soc Online-2012.

Dongqing Chen, Zhengrong Liang, Mark, R. Wax, Lihong Li, Bin Li and E. Arie Kaufman,2000. “A Novel Approach to Extract Colon Lumen from CT Images for Virtual Colonoscopy,” IEEE Transactions On Medical Imaging, 19(12): 1220-1226.

Dongqing Chen, Hossam Abdelmunim, A. Aly Farag, Robert Falk and Gerald Dryden,2008. “Segmentation of colon tissue in CT colonography using adaptive level sets method,” Proc. of MICCAI 2008 Workshop: Computational and Visualization Challenges in the New Era of Virtual Colonoscopy, pp: 108-115.

Dongqing Chen, Rachid Fahmi, A. Aly Farag, L. Robert Falk and Gerald W. Dryden, 2009.“Accurate And Fast 3D Colon Segmentation In CT Colonography,”IEEE, pp: 490-493.

http://cancerimagingarchive.net.

Liang, Z., D. Chen, R. Chiou, Bin Li, A. Kaufman and M. Wax, 1999. “On segmentation of colon lumen for vitual colonscopy,” Proc. of SPIE conference on physiology and function from Multidimensional Images, pp: 270-278.

Lin Lu and Jun Zhoa, 2013. “An Improved method of automatic colon segmentation for virtual colon unfolding” Elsevier Science Ltd Computer Methods and Program in Biomedicine, 09: 1-12.

Lihong Li, Su Wang, Jing Wang, Daria Eremina, Xinzhou Wei and Zhengrong Liang, 2007. “A New Electronic Colon Cleansing Method for Virtual Colonoscopy,” Proc. of SPIE, 6511.

Lihong, Li, Dongqing Chen, Sarang Lakare, Kevin Kreeger, Ingmar Bitter, Arie E. Kaufman, Mark, R. Wac, Peter, M. Djuric and Zhengrong Liang, 2002. “An Image segmentation approach to extract colon limen through colonic material tagging and hidden markov random field model for Virtual Colonoscopy,” Proc. of SPIE Medical Imaging Physiology and Function from Multidimensional Images,4683: 406-411.

Marek Franaszek, M. Ronald Summers, J. Perry Pickhardt and J. Richard Choi Zander, 2006.“Hybrid Segmentation of Colon Filled With Air and Opacified Fluid for CT Colonography,” IEEE Transactions On Medical Imaging, 25(3): 358-368.

Pappas, T., 1992. “An adaptive clustering algorithm for image segmentation,” IEEE Trans. Signal Processing, 40: 901–914.

Sebastian Zander, Thuy Nguyen and Grenville Armitage, 2006. “Detection of colon wall outer boundary and segmentation of colon wall based on level set method,” Proc. of 28th IEEE EMBS Annual International Conf: pp: 108.

Stephen, J. Redmond and Conor Heneghan, 2007. ”A method for initializing the k-means clustering algorithm using kd-trees,”Pattern Recogintion, 28: 965-973.

Tolias, Y.A., S.M. Panas, 1998. Image Segmentation by a Fuzzy Clustering Algorithm Using Adaptive Spatially Constrained Membership Functions. IEEE Trans.Systems, Man, Cybernet, 28: 359–369.

Tarik, Chowdhury, A., F. Paul Whem and Ovidiu Ghita, 2005. “A method for the automatic segmentation of collapsed colon at CT colonography,” Proc. of Second Indian International Conference on Artificial Intelligence: pp: 3517-3532.

Tarik Chowdhury, A. and F. Paul Whelan, “ A Fast and Accurate Method for Automatic Segmentation of Colons at CT Colonography based on Colon Geometrical Features”.

Vahid Taimouri, Xin Liu, Zhaoqiang Lai, Chang Liu, Darshan Pai, and Jing Hua, 2011. “A Colon Segmentation for prepless Virtual Colonoscopy”IEEE Transactions on Information Technology In Biomedicine,15(5): 709-715.

Wyatt, C.L., Y. Ge, and D.J. Vining , 2000. “Automatic segmentation of the colon for virtual colonoscopy” Elsevier Science Ltd Computerized Medical Imaging and Graphics, 24: 1-9.

Xiaoyun Yang, Yalin Zheng, Musib Siddique and Gareth Beddoe, 2008. ”Learning from imbalanced data: a comparative study for Colon CAD,” in Proc. of SPIE Medical imaging, pp: 6915-69150R.

Paul, A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig, 2006. “ User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability,”.Neuroimage 1, 3(3): 1116-28.

Yong Yang, Shuying Huang, 2007. ” Image segmentation by Fuzzy C-Means Clustering algorithm With a novel penalty term,” Computing and Informatics, 26: 17–31.

Yushkevich, P.A., J. Piven, H. Cody, S. Ho, J.C. Gee and G. Gerig, 2005. User-Guided Level Set Segmentation of Anatomical Structures with ITK-SNAP, Insight Journal, Special Issue on ISC/NA-MIC/MICCAI Workshop on Open-Source Software.