Enhanced K-Mean Clustering Algorithm for Liver Image Segmentation to Extract Cyst Region

Ramanjot Kaur
Baba Banda Singh
Engineering College
Fatehgarh Sahib (Pb.), India

Lakhwinder Kaur
University college of Engg.
Punjabi Uni. Patiala (Pb.)
India

Savita Gupta
University Institute of Engg.
Punjab University
Chandigarh, India

ABSTRACT
This paper, first analysis the performance of image segmentation techniques; K-mean clustering algorithm and region growing for cyst area extraction from liver images, then enhances the performance of K-mean by post-processing. The K-mean algorithm makes the clusters effectively. But it could not separate out the desired cluster (cyst) from the image. So, to enhance its performance for cyst region extraction, morphological opening-by-reconstruction is applied on the output of K-mean clustering algorithm. The results are presented both qualitatively and quantitatively, which demonstrate the superiority of enhanced K-mean as compared to standard K-mean and region growing algorithm.

Keywords
Image segmentation, region of interest, k-mean clustering and region growing.

1. INTRODUCTION
Liver diseases are considered seriously as it plays vital role to the life of the human [1]. Traditional methods used to decide whether liver tissue is normal or abnormal are dependent on the experience of the radiologist [1]. Lee et al. [1] classified three type of liver diseases; cyst, hepatoma and cavernous hermangioma. Liver diseases are best identified using grey scale images [2]. Liver cyst is smooth, but the contrast between normal liver tissue and liver cyst tissue is high [1]. The present work segments the grey scale liver CT image to extract cyst region in order to provide a reliable basis for clinical treatment and pathology.

Image segmentation is used to locate objects and boundaries (lines, curves, etc.) in images [3 – 5]. More precisely, it is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. There are a large number of different image segmentation techniques have been proposed in the literature [6-10]. Liver CT image segmentation mainly uses the threshold [11]. Thresholding techniques use brightness constant as threshold value to segment the pixels in the original image. Such methods neglect all the spatial information of the image and do not manage well with noise or blurring at boundaries, which are generally encountered in the ultrasound images [6]. In 2010, Xu et al. [12] used local entropy to segment liver image and morphological method to detect the object regions. In this paper, we have used combination of two approaches; K-mean clustering and morphological opening by reconstruction to segment liver images to extract cyst region.

The paper is organized as follows: section 2 describes cyst, section 3 presents two segmentation approaches used and enhanced K-mean algorithm. The results are presented in section 4 and concluding remarks are given in section 5.

2. CYST
A cyst is a closed, saclike structure that contains fluid, gas, or semisolid material and is not a normal part of the tissue where it is located. Cysts are common and can occur anywhere in the body in people of any age. Cysts vary in size; they may be detectable only under a microscope or they can grow so large that they displace normal organs and tissues. The outer wall of a cyst is called the capsule. Liver cyst, also known as a hepatic cyst, a simple liver cyst is a thin-walled bubble, a fluid-filled cavity in the liver that usually produces no signs or symptoms. Some liver cysts do not contain fluid. They are usually detected by chance during other types of testing, and may be diagnosed through ultrasound, magnetic resonance imaging (MRI) or computerized tomography (CT) scans.

3. SEGMENTATION TECHNIQUES
3.1 Region Growing [3]
It is also classified as a pixel-based image segmentation method since it involves the selection of initial seed points. This approach starts with initial "seed points" and then examines neighboring pixels (using either 4-connectivity or 8-connectivity) to determine whether the pixel neighbors should be added to the region. The process is iterated on, in the same manner as general data clustering algorithms. The region growing algorithm is described in [3] as:

(i) Select a set of seed points. Seed point selection is based on some user criterion (for example, pixels in a certain gray-level range, pixels evenly spaced on a grid, etc.). The initial region begins as the exact location of these seeds.

(ii) The regions are then grown from these seed points to adjacent points depending on a region membership criterion.

The criterion could be pixel intensity, gray level texture or color. Since the regions are grown on the basis of the criterion, the image information itself is important. For example, if the criterion were pixel intensity, examine the adjacent pixels of seed points. If they have the same intensity value with the seed points, classify them into the seed points. It is an iterated process until there is no change in two successive iterative stages. The suitable selection of seed points is an important issue [3].

3.2 K-Means Clustering Algorithm

K-means [3][13-15] is one of the simplest unsupervised learning algorithms that classify a given data set into certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way, because different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early grouping is done. Again re-calculate k new centroids as barycenters of the clusters (resulting from the previous step). After having these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. Repeat the process until centroids do not move any more. In the successive loops, the k centroids change their location step by step.

The K-mean algorithm uses the following distance formula [3] to compute the distance of the *n* data points from their respective *j*th cluster centre.

$$j = \sum_{j=1}^{k} \sum_{i=1}^{x} \left\| x_i^{(j)} - c_j \right\|^2 \qquad (1.1)$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a distance measure between a data point $x_i^{(j)}$ and the cluster centre, c_j.

3.3 ENHANCED K-MEANS CLUSTERING (EK-MEAN)

Using standard algorithm for k mean clustering, cyst region was not extracted properly. To improve its performance, morphological opening-by-reconstruction operation is applied on the output of k- mean clustering algorithm. The idea of opening-by-reconstruction [17] is taken from marker controlled watershed algorithm, which uses both opening and closing-by-reconstruction followed by watershed segmentation [3][16]. Morphological reconstruction processes one image, called the *marker*, based on the characteristics of another image, called the *mask* (e.g. Figure 1).

The high-points, or peaks, in the marker image specify where processing begins. The processing continues until the image values stop changing. Morphological reconstruction can be thought of conceptually as repeated dilations of the marker image until the contour of the marker image fits under the mask image [16]. In this way, the peaks in the marker image "spread out", or dilate as shown in Figure 1. The characteristics of the marker image determine the processing performed in morphological reconstruction. The peaks in the marker image should identify the location of objects in the mask image that we want to emphasize. Opening-by-Reconstruction is more efficient than standard opening and closing at removing small blemishes without affecting the overall shapes of the objects [16].

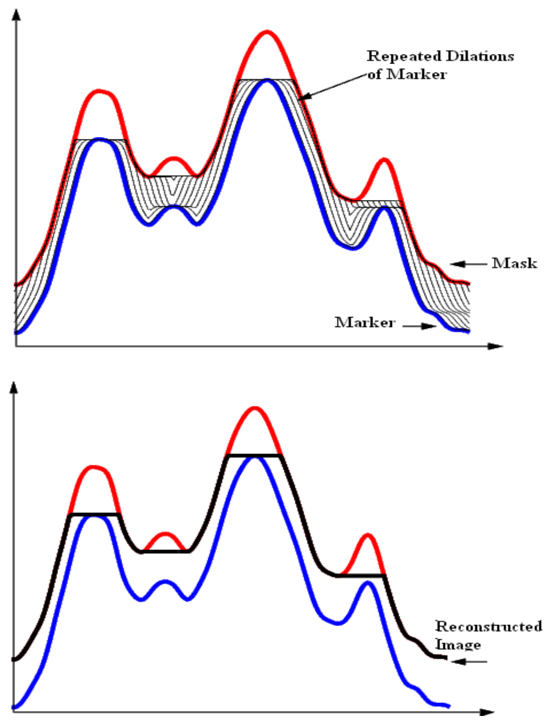


Fig. 1 [16]: Morphological reconstruction: upper figure shows marker, mask and repeated dilations of Marker based on Mask, bottom shows peaks of Marker image spread out.

3.3.1 EK- Mean Algorithm

(i) Apply K-mean algorithm to cluster or classify the image pixels, I. Then mark the desired cluster.

(ii) Apply morphological opening-by-reconstruction operation on desired cluster (cyst) as:

a) Choose suitable structuring element, SE and apply erosion operation to remove background pixels from cyst boundary in Matlab as:

Ie=imerode(cyst, SE)

b) Reconstruct the marker image, Ie, based on mask cyst by imreconstruct command in Matlab as:

RIe =imreconstruct(Ie, cyst);

c) Apply dilation operation on RIe for adding pixels of object boundaries as

DRIe = imdilate(RIe, SE);

d) Apply reconstruction on complement of DRIe based on complement of RIe. as in step (b).

e) By using regional maxima, the cyst area will be highlighted.

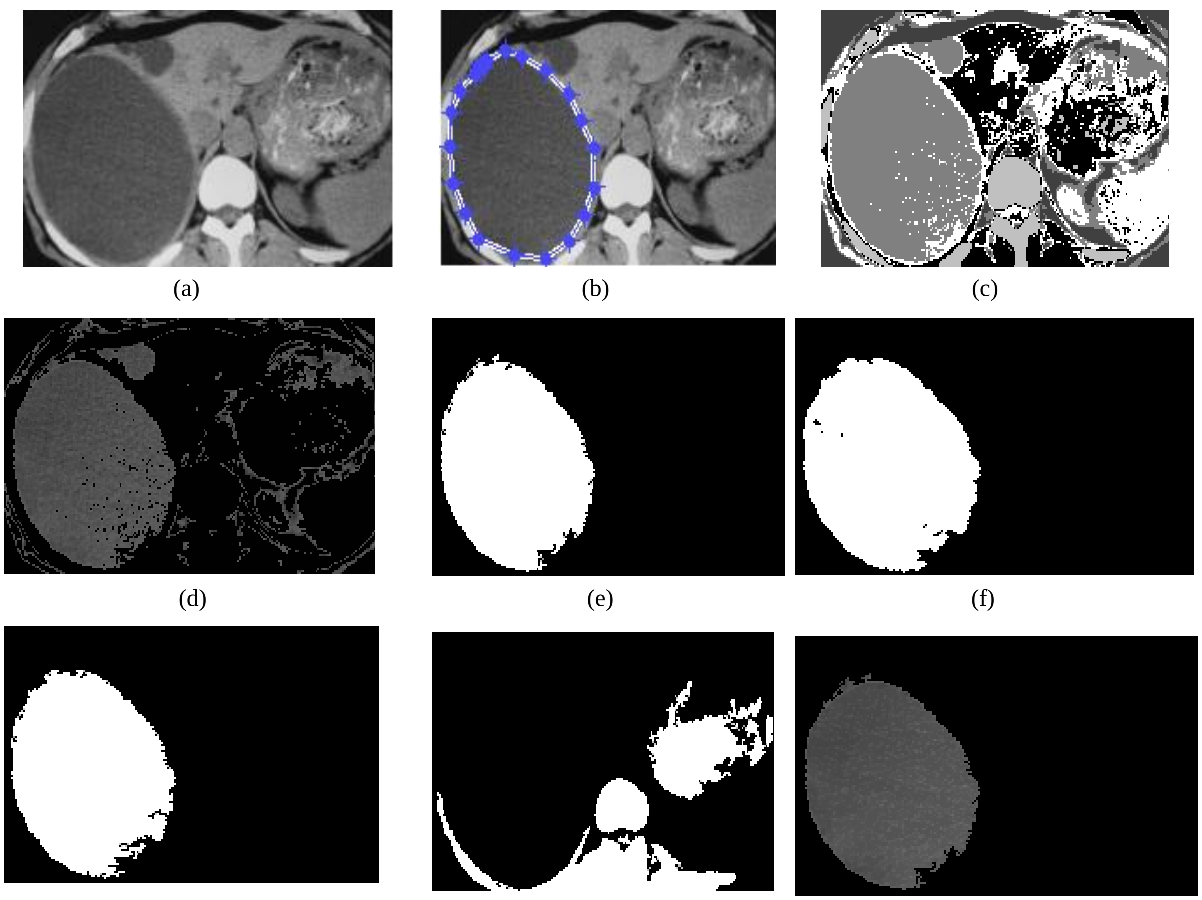
The performance of this algorithm depends on choice of SE.

4. RESULTS AND DISCUSSION

In this paper, the performance of standard k-mean clustering, Region growing and EK algorithm for cyst detection from liver images was analyzed. The liver images are collected from various medical sites [18]. Implementation is done in Matlab 7.5. Each input image is divided into five clusters by using K-mean algorithm. The performance metrics like Mean square error (MSE), Correlation coefficient and Quality index[19 - 20] are computed for quantitative comparison.

For qualitative analysis, output images are shown in fig. 2 to fig. 8. The figures 2(b) – 8(b), show the marked region of interest which we want to segment. The figures 2(c) – 8(c), demonstrate the clusters made by standard k-mean clustering algorithm on original image, but it does not separate out the cyst area from original image. The output of enhanced k-mean (EK) clustering (fig. 2(d) –fig. 9 (d)), shows that it separates the cyst region from original image properly. The fig. 2 (h) to fig. 9 (h), demonstrate that morphological operations if applied directly on the original image, cannot segment cyst area properly. The output images of region growing algorithm and enhanced region growing algorithm (region growing algorithm followed by morphological operations) are shown in fig. 2 (e) – fig 9 (e) and fig.2 (f) – fig. 9 (f) respectively. The visual comparison revealed that EK algorithm has better cyst region segmentation capability.

The quantitative results are presented in table 1 to table 3. To compute parameters, region of interest marked by 'roipoly' command of Matlab is used as reference image. From the results, it is concluded that enhanced k mean clustering have better performance than region growing method. The value of correlation coefficient of both algorithms is very close to 1, which implies that segment region is very close to marked region of the original image. The value of MSE is greater in case of region growing than our proposed work. The value of quality index of EK clustering is higher than that of region growing method. Some discrepancies in the numerical values may be due to marking of region-of-interest. At the end, we concluded that EK clustering method having better performance than region growing for cyst area segmentation in liver images.



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 2: (a) shows Original image, (b) region of interest, (c) Numbers of clusters k=5, and next images show Cyst region segmentation using k mean clustering (d), enhanced k mean clustering (e), region growing (f), enhanced region growing (g), simple morphological operations (h) and Cyst segmented out by EK (i).

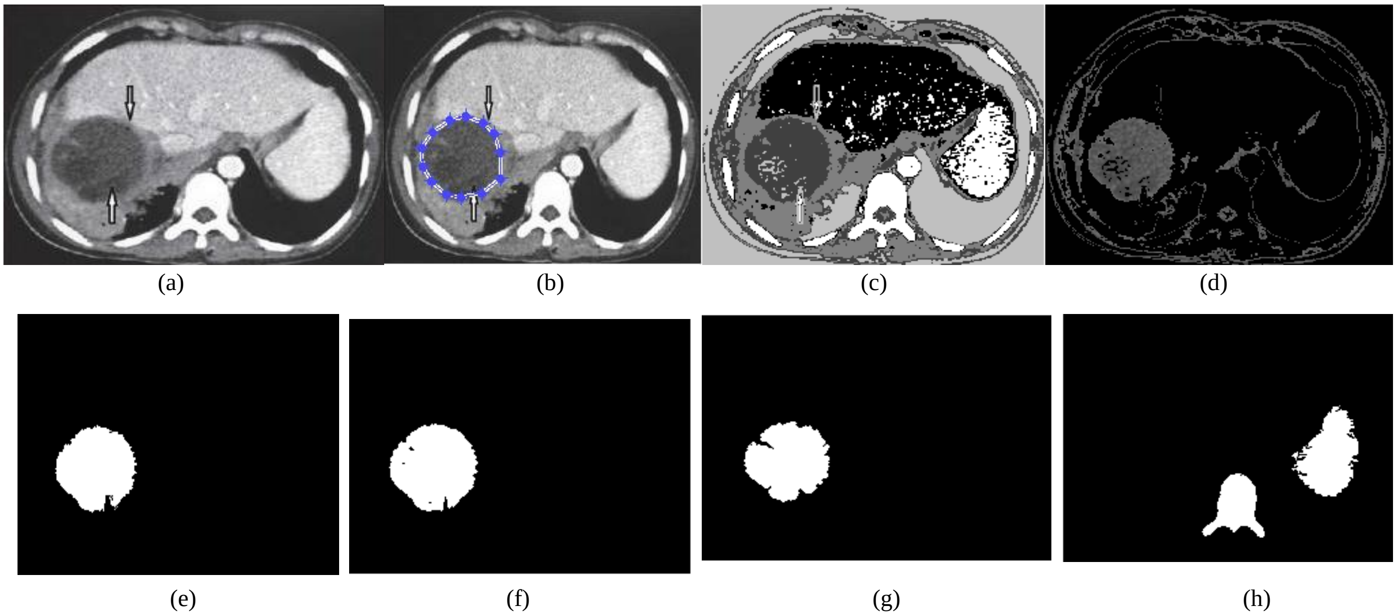


Fig. 3: (a) shows Original image, (b) region of interest, (c) Numbers of clusters k=5, and next images show Cyst region segmentation using k mean clustering (d), enhanced k mean clustering (e), region growing (f), enhanced region growing (g), and morphological operations (h).

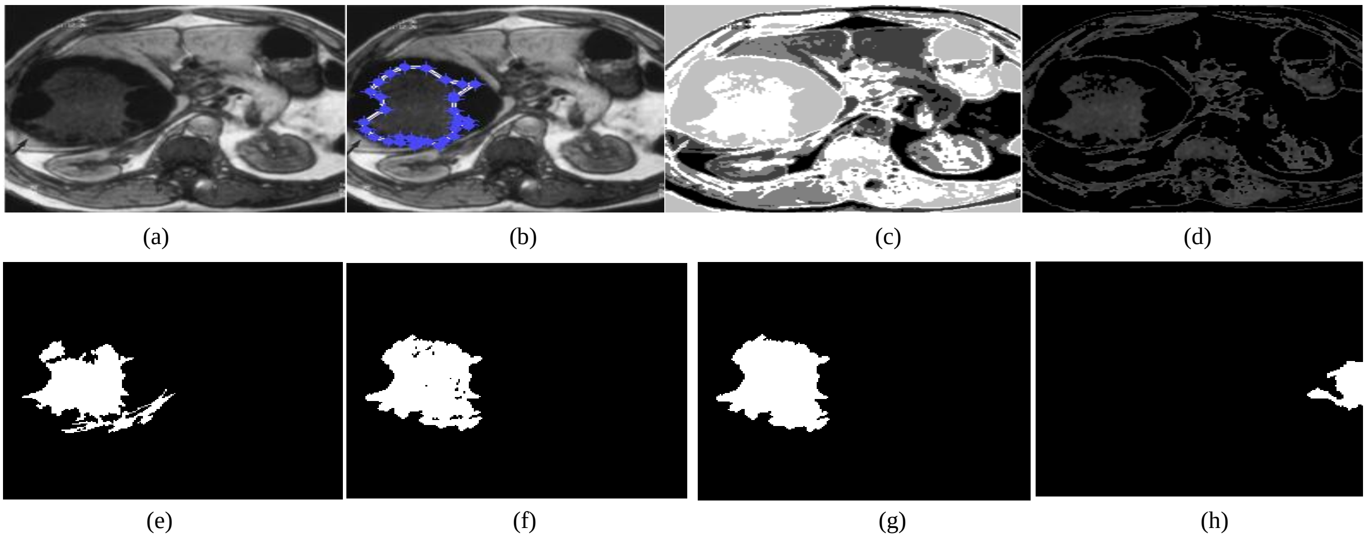


Fig. 4: (a) shows Original image, (b) region of interest, (c) Numbers of clusters k=5, and next images show Cyst region segmentation using k mean clustering (d), enhanced k mean clustering (e), region growing (f), enhanced region growing (g), and morphological operations (h).

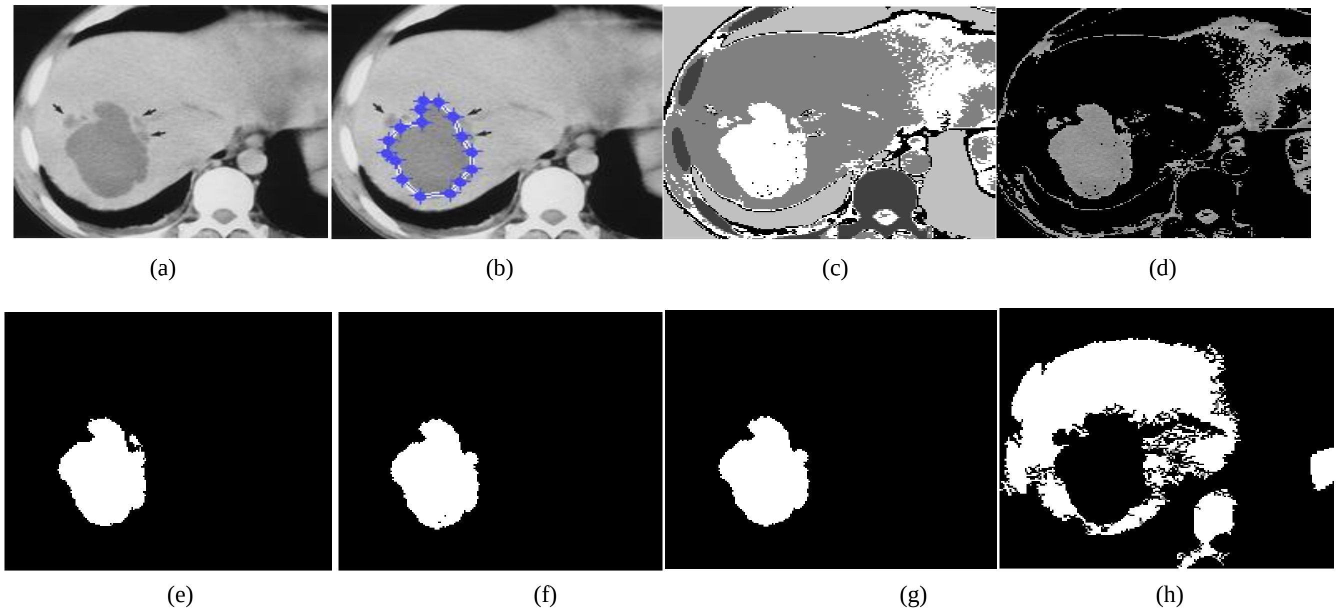


Fig. 5: (a) shows Original image, (b) region of interest, (c) Numbers of clusters k=5, and next images show Cyst region segmentation using; k mean clustering (d), enhanced k mean clustering (e), region growing (f), enhanced region growing (g), and morphological operations (h).

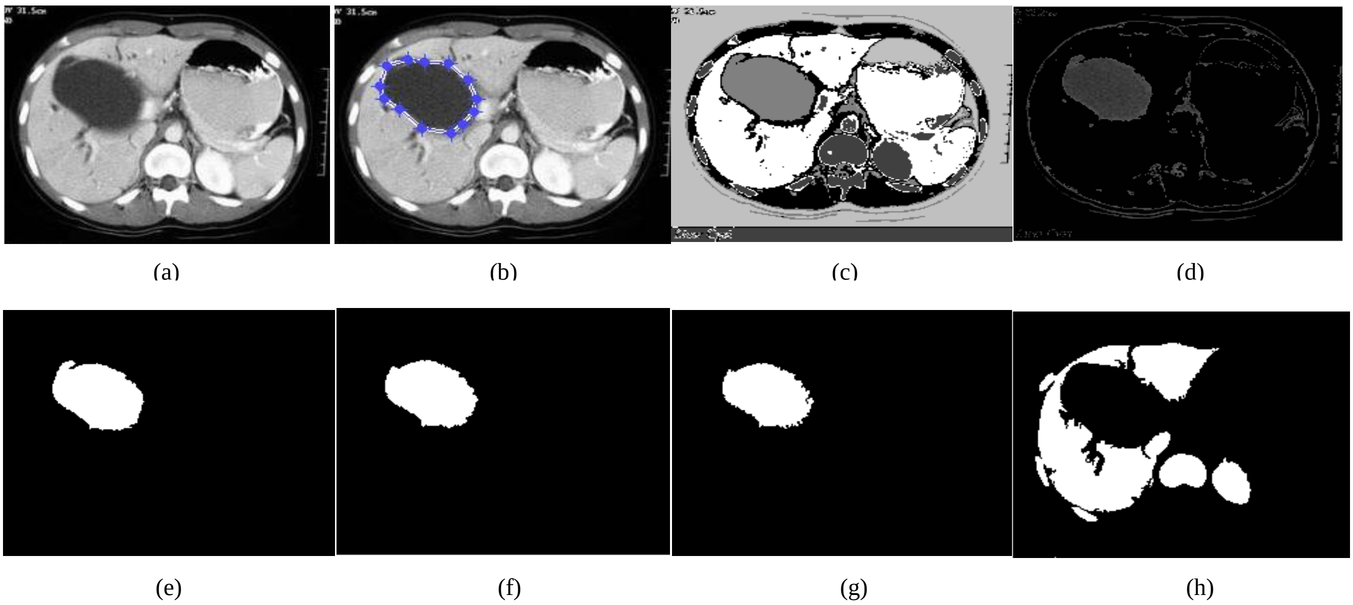


Fig. 6: (a) shows Original image, (b) region of interest, (c) Numbers of clusters k=5, and next images show Cyst region segmentation using k mean clustering (d), enhanced k mean clustering (e), region growing (f), enhanced region growing (g), and morphological operations (h)

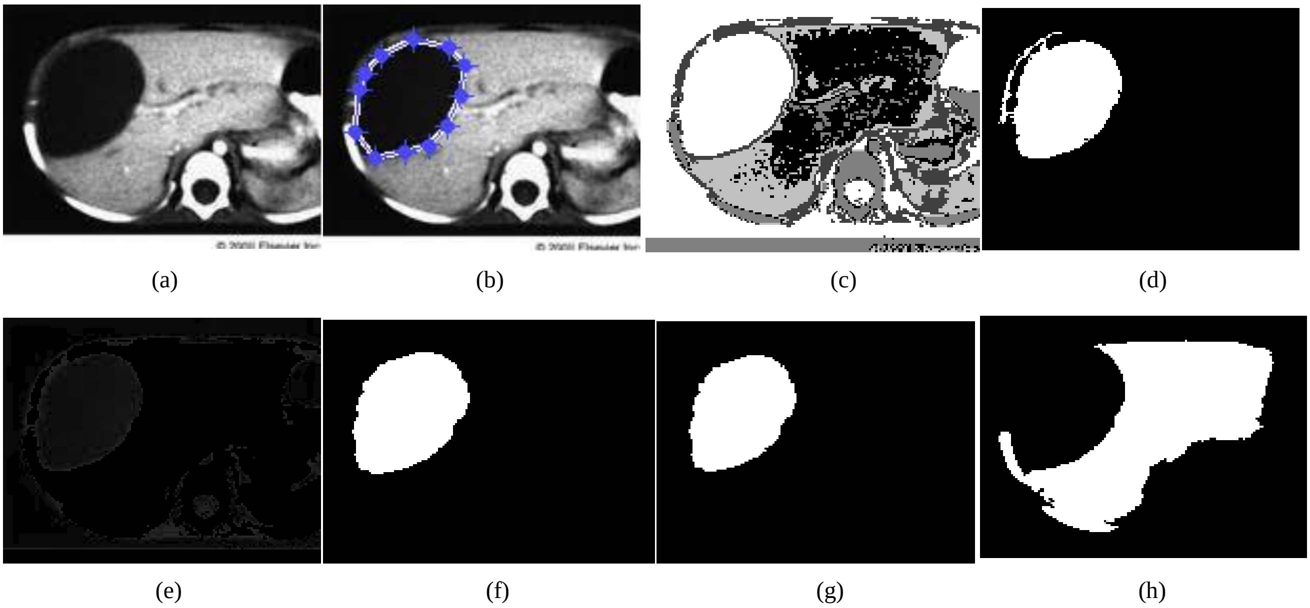


Fig. 7: (a) shows Original image, (b) region of interest, (c) Numbers of clusters k=5 and next images show Cyst region segmentation using k mean clustering (d), enhanced k mean clustering (e), region growing (f), enhanced region growing (g), and morphological operations (h) .

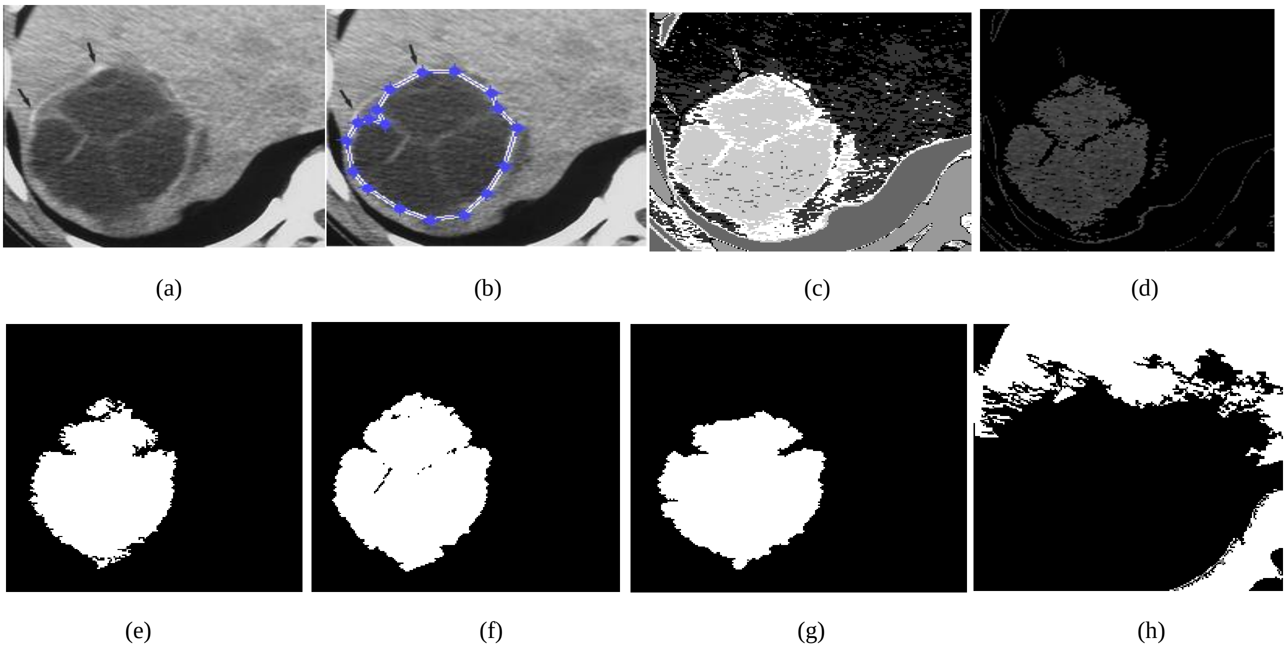
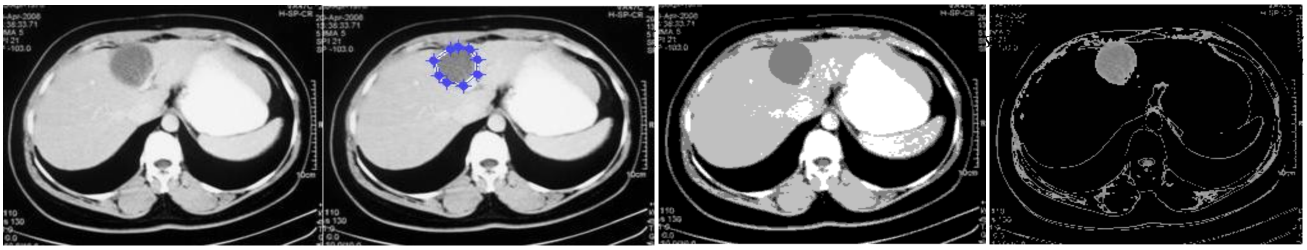


Fig. 8: (a) shows Original image, (b) region of interest, (c) Numbers of clusters k=5, and next images show Cyst region segmentation using k mean clustering (d), enhanced k mean clustering (e), region growing (f), enhanced region growing (g) and morphological operations (h).



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 9: (a) shows Original image, (b) region of interest, (c) Numbers of clusters k=5, and next images show Cyst region segmentation using k mean clustering (d), enhanced k mean clustering (e), region growing (f), enhanced region growing (g), and morphological operations (h) .

Table 1: Values of correlation coefficient obtained from the EK clustering and region growing method.

Image / **Method**	1	2	3	4	5	6	7	8
Enhanced Kmean Clustering	0.9479	0.7741	0.9494	0.9279	0.9442	0.9343	0.9013	0.8533
Region growing	0.9400	0.8614	0.9330	0.9463	0.9513	0.9509	0.8897	0.9040

Table 2: Values of mean square error obtained from the EK clustering and region growing method.

Image / **Method**	1	2	3	4	5	6	7	8
Enhanced Kmean Clustering	0.0056	0.0307	0.0196	0.0096	0.0060	0.0145	0.0329	0.0052
Region growing	0.0721	0.1031	0.2600	0.0797	0.0630	0.1361	0.1771	0.0203

Table 3: **Values of quality index obtained from the EK clustering and region growing.**

Image / **Method**	1	2	3	4	5	6	7	8
Enhanced Kmean Clustering	0.9714	0.9009	0.9134	0.9539	0.9710	0.9278	0.8848	0.9784
Region growing	0.9398	0.8869	0.7012	0.9107	0.9589	0.9081	0.7423	0.9761

5. CONCLUSION

The performance of K-mean, region growing and proposed enhanced K-mean algorithm was evaluated on number of MRI and CT scan liver images to segment cyst region. The novelty of the method is simplicity. Both qualitative as well as quantitative results are in the favour of enhanced k-mean clustering algorithm. Enhanced k-mean clustering can be further used to segment images of other modalities with different diseases, which will helpful in surgeries. It can also be tested on colored images either in 2D or 3D. The performance of proposed algorithm depends on choice of initial centroids as well as on proper selection of structuring element.

6. REFERENCES

[1] Chien-Cheng Lee, S. Chen and Y. Chiang, 2006. Automatic Liver diseases Diagnosis for CT Images using Kernel based classifiers in IEEE world Automation Congress, pp. 1-6.

[2] Poonguzhali, B. Deepalakshmi and G. Ravindran, 2007 Optimal Feature Selection and Automatic Classification of Abnormal Masses in Ultrasound Liver Images in IEEE-ICSCN, pp. 503-506.

[3] Rafael C. Gonzalez and Richard E. Woods, 2009. Digital Image Processing, Prentice-Hall.

[4] Linda G. Shapiro and George C. Stockman (2001): Computer Vision, pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13-030796-3.

[5] C. Bezdek, L. O. Hall and L P. Clarke, 1993. Review of MR image segmentation techniques using pattern recognition in Med. Phys., vol. 20, Issue 4, pp. 1033-1048.

[6] Ashish Thakur and Radhey Shyam Anand, 2010. A Local Statistics Based region Growing Segmentation Method for ultrasound Medical Images in Jr. of World Academy of Science, Engg. and Technology, pp. 321-326.

[7] Madabhushi, A. and Mtaxas, D. N., 2003. Combining low-,high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions in IEEE transactions on medical imaging, VOL. 22, NO. 2.

[8] Eslami, A.; Kasaei, S.; Jahed, M., 2004. Radial multiscale cyst segmentation in ultrasound images of kidney in 4th IEEE International Symposium on Signal Processing and Information Technology, pp. 42-45.

[9] Pavlidis, T., and Liow, Y. T., 1990. Integrating region growing and edge detection", IEEE Transaction on Pattern Analysis and Machine Intelligence, 12, pp. 225-231.

[10] XU D, GUO S. 2009. Research Progress of Liver CT Image Segmentation Techniques in Chinese Medical Equipment Journal•Vol. 30 No. 3.

[11] Yi X., Zhong L. and Lin J., 2010. Liver CT Image Segmentation by Local Entropy Method in International Conf. on computer application and modelling, pp. VIII-591 –VIII-594.

[12] Andrew Moore: K-means and Hierarchical Clustering - Tutorial Slides at http://www.cs.cmu.edu/~awm/tutorials/kmeans.html

[13] Brian T. Luke: K-Means Clustering at http://fconyx.ncifcrf.gov/~lukeb/kmeans.html

[14] Tariq R.: Clustering at http://www.cs.bris.ac.uk/home/tr1690/documentation/fuzzy_clustering_initial_report.

[15] Image Processing Toolbox, available at http://www.caspur.it/risorse/softappl/doc/matlab_help/toolbox/images/morph11.html#34328.

[16] L. Kaur, R.C. Chauhan, and S. C. Saxena, 2006. Wavelet based Compression of Medical Ultrasound images using Vector Quantization in Taylor and Francis JMET, pp. 128-133, vol. 30, no. 3.

[17] S. Gupta, L. Kaur, R.C. Chauhan, and S. C. Saxena, 2007. A versatile Technique for visual enhancement of medical ultrasound images in Elsevier Jr. of Digital Signal Processing, 17, pp. 542-560.