Pulmonary CT Image analysis for Nodule Detection using Inspired FCM Clustering

Dr. M. Mohamed Sathik
Principal,
Sadakathullah Appa College,
Tirunelveli,India

S Piramu Kailasam,
Assistant Professor ,
PG & Research Department of Computer Science ,
Sadakathullah Appa College, Tirunelveli,
India

Abstract—**The segmentation of white blob nodule area from pulmonary CT images are tedious and time consuming task. The accuracy of nodules as tumor depends on the radiologist experience only. Hence the CAD still necessary to overcome these limitations. In this study to improve the performance and reduce the complexity of segmentation, adaptive histogram equalization method and Fuzzy C means applied with high level iterations. Further to improve the accuracy local and global threshold also used. The resultant parameters mean , median showed good accuracy results.**

Keyword— *Pre-processing; Global Threshold; Fuzzy C Means(FCM); Segmentation*

I. INTRODUCTION

Image Segmentation is the process that segment an image into many disjoint small area based on the certain principle, and it is the one of the most basic area of research in the image processing. Image segmentation is used in many applications; with Image Retrieval it is possible to analyse the images in Geology Information Systems, Medical Image Analysis, Film and Photo Analysis etc.

II. RELATED WORKS

Many image processing techniques for the detection of lung cancer by using CT images are reviewed in [1], lung cancer detection is carried out by splitting the review in different aspects such as pre-processing, nodule segmentation and segmentation, lung nodule detection are presented in [2], the performance of the recent lung nodule detection techniques are compared. A lung cancer classification system[3] is proposed on the basis of wavelet recurrent neural network also employs wavelet to remove the noise from the input image and the recurrent neural network is utilized for classification. Perhaps, this author could not achieve better specificity rates and this implies that the false positive rates of the work are reasonably greater. In paper [4], FCM is applied to achieve segmentation and the GLCM (Gray Level Co - occurence Matrix) features are extracted. The author Manasee[5] says a lung cancer detection technique that relies on genetic approach is proposed, this work involves more time complexity and the number of connected objects have been calculated by assigning 1 to inside and 0 to outside of the object that shows the medical image based on threshold technique to analyze the performance.

A. *Segmentation Techniques*

Some of the image segmentation methods are Local entropy image segmentation [6]. A Discrete cosine texture feature is applied by [7].The parallel algorithm for grey scale image was clearly declared[8]. Clustering of spatial patterns and watershed algorithm[9] is done. Medical image segmentation performed [5] in the year 2010. Iterative regularized likelihood learning algorithm [11]is performed well by Zhi Lu . The Automatic model selection and unsupervised image segmentation by Zhi Wu Lu , finite mixtures and entropy regularization by Zhi wu Lu [12] is well delt in the paper.

As described by the Sun Hee Kim[13] , acquisition of the image, extraction of the region , intensity and noises of the image affect the accuracy. Resize the image to process in matlab. also another measurable factor. The image processing is applied to separate the region of nodules and backgrounds. Then the smaller size is obtained by processing it in matlab12a. The array editor convert the image to values. Identify the cluster where nodules are present by iterations. The place where the nodules are there shows higher value because of brightness. The conversion of the image to array editor values done by Harvey A et al., [14].

Among the large number of lung segmentation methods bidirectional differential chain code combined with machine learning framework is able to correctly include the just a pleura nodules into the lung tissue on minimizing over segmentation. This method identifying low-dose, concave/convex regions [15]. In the Hessian-based matrix approach, 3D lung nodule segmented in the multiscale process through the combination of Shape and Curvedness methods. Image characteristics, Nodule position, nodule size, and nodule characteristics are included also Eigenvalues are computed from the 3 x 3 Hessian Matrix .

B. *CT Image Pre-processing by AHE*

The contrast of the CT images is enhanced by employing adaptive histogram equalization technique, which is versatile for both greyscale and colour images. Kernelized Fuzzy C Means (FCM) is utilized to segment the CT images, which shows better performance than the standard FCM. Unlike Histogram Equalization, Adoptive Histogram Equalization works with region of image. FCM

algorithm shows the group of pixels which are having same characters.

III. MATERIALS AND METHODS

This section shows the materials , the source of pulmonary nodule image dataset and the methods used on pulmonary image. LIDC dataset taken as source bench mark dataset available in pulmonary image database online.

A. *Pre-processing*

Noise is a variation of image intensity in random manner and visible as dots or grains in the image. Noisy pixels in the image show different intensity values instead of true pixel values. The
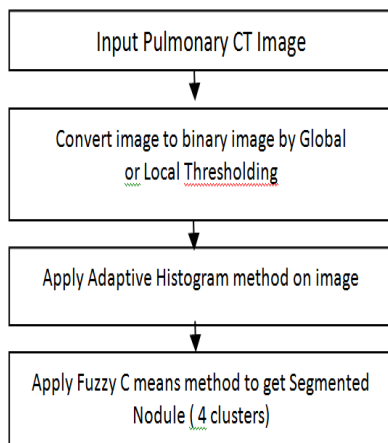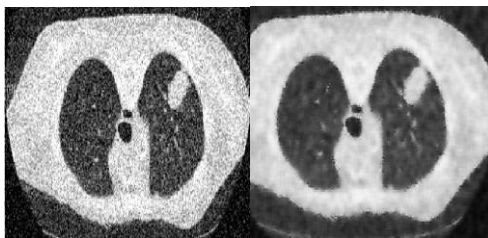

Fig1.Flow Diagram

Fig 2 a. Gaussian Noise   b .Wiener Filter

common types of noise arises in the image area. The imaging sensor may be affected by climatic conditions during image acquisition period. It may affected by various hardware devices. Insufficient light and sensor temperature may introduce the noise in the image. Interference in the transmission channel may also corrupt the image. The dust particles on scanner screen scan also introduce noise in the image. The fig 1. shows the flow diagram of the work and fig 2 shows the Gaussian noise,wiener filter applied CT pulmonary images.

B. *Denoising in Medical Images*

The main properties of a good image denoising model are that it will remove noise while preserving edges. Subband decomposition of an image is performed by a mathematical transform. Some of the familiar transforms are discrete wavelet transform (DWT), Bandelet transform,

contourlet transform, curvelet transform and ridgelet transform. Curvelet transform and ridgelet take the form of basic elements which exhibit very high anisotropic. Therefore the curvelet transform represents edges better than wavelets, and is well suited for multiscale edge enhancement. Curvelet have a variable width and length so that it has a variable anisotropy.

Most of the medical images suffer from poor contrast issues and so the details of the images cannot be distinguished easily. At this juncture, the role of contrast enhancement algorithm comes into picture.

The CT images are pre-processed by employing the Adaptive Histogram Equilization (AHE) technique and the major reason for the exploitation of AHE is that this technique computes several histograms by considering separate areas of the CT image. This means that the contrast of every particular area of the image is enhanced. This method consider pixel (pi) and alters into nearest neighbourhood pixel intensity. This focus on the region of image unlike Histogram Equivalization method. This computes several histogram for each region of the CT images. This means that contrast of every particular region is enhanced. This ensures the balanced distribution of values.

The AHE method improves the edges and regions. The computed histogram is adaptive, which ensures the balanced distribution of values. The AHE technique improves the local contrast of the CT image. Some of the sample pre-processed images are presented in fig 2 (a) & (b) by Gaussian Noised and Wiener filter methods.

Consider a CT image with $n \times n$ pixels and is applied with the contrast enhancement process. The AHE processes each and every pixel p_i and alters p_i by considering the intensity immediate neighbourhood pixels.

Unlike standard HE technique, AHE processes the regions of an image separately. This results in the contrast enhancement of each and every region, which in turn enhances the contrast of a complete image. By this way, the contrast of the CT image is enhanced and the contrast enhanced images are segmented as follows.

C. *Types of Image Segmentation Techniques*

Image segmentation by different methods are done by many people. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Segmentation is a crucial task in medical image processing Some of them are Local entropy image segmentation is done by Ali Shojaee Bakhiari[6], Discrete cosine texture feature by Chi-man pun, parallel algorithm for grey scale image by Harvey, Clustering of spatial patterns and watershed algorithm by Kai-Jian,Medical image segmentation by Prasantha, Finite bivariate doubly truncated Gaussian mixture model by Rajkumar[16], Random set of view of texture by Ramana reddy[17], Sobel operator technique, Prewitt technique, Kiresh technique, Laplacian technique, Canny technique , Roberts technique and Edge Maximization Technique by Salem Saleh[18],

Mathematical image processing by Sun hee kim, Simple algorithm for image denoising by Vijaya[19], Iterative regularized likelihood learning algorithm by Zhiwu Lu, Automatic model selection and unsupervised image segmentation by Zhi Wu Lu, Finite mixtures and entropy regularization by Zhi wu Lu[11].

The various algorithms used for image segmentation such as watershed, Sobel operator technique, Roberts technique, Edge Maximization Technique (EMT), Bivariate Finite doubly Truncated Gaussian Mixture Model, Expectation Maximization (EM) algorithm, K-Means algorithm, Iterative Entropy Regularized Likelihood Learning Algorithm, Partial Differential Equation (PDE), Mean shift (MS) segmentation. Wavelet based segmentation, Genetic algorithm based segmentation, Parallel Algorithm, Discrete Cosine Texture Feature are explained in the following.
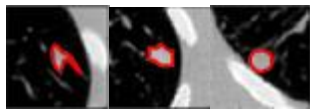

Fig 3. Manual Segmentation

In fig 3, the red color shows the manual segmentation of lung nodules. The traditional water shed algorithm is a mathematical morphology segmentation algorithm based on the topology theory. The basic idea is that the image is regarded as the topological landform in the geodesy, gray-scale value of each pixel in the image represents the Altitude of this point. Every local minimum and influence area is called the catchments basin, while its boundary forms a watershed. The Watershed concept and process of the formation can be illustrated by simulating immersion. In every local minimum surface, pierce a small hole, and then slowly immerse the model into the water, with the immersing deeper and deeper, the influence area of every local minimum value gradually expands outward, and construct the dam basin at the confluence of the two catchments basin. It is just the watershed.

The computation of the partial derivation in gradient may be approximated in digital images by using the Sobel operators, Prewitt operator and Canny operator which are shown in the masks below fig 4.

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	1

Fig 4 The Sobel Masks

The prewitt operator uses the same equations as the Sobel operator, except the constant value one. The Prewitt operator (fig.5) measures two components, the vertical edge component (kernel Gx) and the horizontal edge component with (kernel Gy). |Gx| + |Gy| gives e current pixel's intensity gradient.

-1	0	1
-1	0	1
-1	0	1

1	1	1
0	0	0
-1	-1	-1

Fig 5. Gx   Fig 6. Gy

Prewitt Technique

The digital implementation of the Laplacian function is usually made through the mask (Fig 7) below :

0	-1	0
-1	4	1
0	-1	0

Fig 7. The Laplacian mask

The Laplacian is used to establish whether a pixel is on the dark or light side of an edge. The performance of edge detection of nodules can also done by canny operator.

The Canny method applies a best edge detector based on a set of conditions which finding the most edges by low error rate.

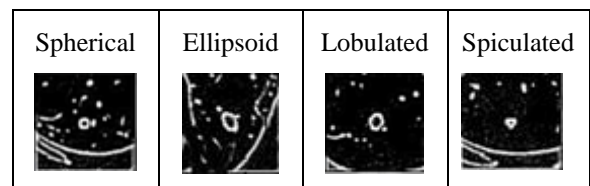| Spherical	Ellipsoid	Lobulated	Spiculated

Fig 8. Canny Edge Detection of Nodules

The Roberts Cross operator performs a simple, fast compute method. It measures the 2D spatial gradient of an image. Hence highlights the regions of high level spatial frequency which correspond to edges.

In images when there is more than one homogenous region or different gray levels. In this situation, some part of the objects may be merged with the background. An automatic threshold selection technique's performance becomes much better in images with large homogenous splitted regions.
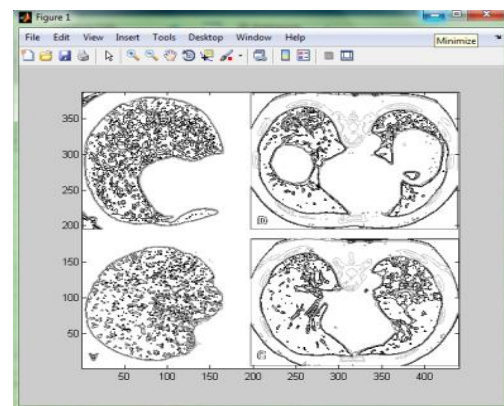

Fig 9. Edge Detection using prewitt

In the image the vertical edge component and the horizontal edge component is calculated using Prewitt operator. Hence edge of lung is detected in above Fig .9.

Image is a matrix in which each element is a pixel also the pixel value is a number that shows intensity or color of the image. Based on the probability values the pixel intensities of their corresponding segments based on GMM (Gaussian Mixure Model).

EM (Expectation Maximum) algorithm is highly dependent on initialization of variable. A common approach is using multiple random starts and choosing the final estimate with the highest likelihood. This will greatly increase the computation complexity. Using the K-Means algorithm, the number of image regions is identified. The number k of clusters in the data set is usually assumed to be pre-known. However, since image databases in the context of image retrieval are often huge, the prior setting of cluster number for each image is no longer feasible.

The segmentation consists of two steps. The first step is mixture estimation. The EM plus MML method is used to estimate the mixture parameters and determine the number of Gaussian components Minimum Message Length. In the second step, pixel clustering, the segmentation is carried out by assigning each pixel into a proper cluster according to the Maximum Likelihood (ML) estimation.

As for unsupervised image segmentation, one important application is content based image retrieval. In this context, the key problem is to automatically determine the number of regions (ie. clusters) for each image so that we can then perform a query on the region of interest. An iterative Entropy Regularized Likelihood (ERL) learning algorithm for cluster analysis based on a mixture model is used to solve this problem.

Segment the image by using Gibou-Fedkiw algorithm, considered 2-means clustering method by using level set function.

The MS algorithm is a robust feature – space analysis approach which can be applied to discontinuity preserving smoothing and image segmentation problems. It can significantly reduce the number of basic image entities, and due to the good discontinuity preserving filtering characteristic, the salient features of overall image are retained. The later property is particularly important in the portioning natural images, in which only several distinct regions are used in representing different scenes such as sky, lake, sand beach, person and animal, whereas other information within a region is often less important and can be neglected. MS algorithm is an unsupervised clustering – based segmentation method, where the number and the shape of the data cluster are unknown, a priori. Moreover, the termination of the segmentation process is based on some region-merging strategy applied to the filtered image result, and the number of regions in the segmented image is mainly determined by the number of pixels in a region

which is denoted as M pixels. Here M will be eliminated and merged into its neighboring region.

For a z1 X z2 color image to be segmented, we consider an 8 – dimensional vector consisting of color, texture, and position features for each pixel. The three color features are the coordinates in the Lab color space, and we smooth these features of the image to avoid over-segmentation arising from local color variations due to texture. The three texture features are contrast, anisotropy, and polarity, which are extracted at an automatically selected scale. The position features are simply the [x, y] position of the pixel, and including the position generally decreased over segmentation and leads to smoother regions. We can get a sample set S of $z = z1 . z2$ samples for each image in the database.

In case of images where in all of the boundaries arise from neighbouring patches of different texture, we compute pair wise texture similarities by comparing windowed texton histograms.

D. Threshold based Image Segmentation

The thresholding transformation sets each gray level that is less than or equal to some prescribed value T called the threshold value to 0, and each gray level greater than T is changed to K-1. Thresholding is useful when one wants to separate bright objects of interest from a darker background or vice versa. The thresholding does not produce edges that correspond to the edges of the objects in the scene. The global threshold image is better than manual threshold (fig.10 a,b,c). Segmentation for nodule is recommended with global thresholding. But in the case of nonnodule adaptive threshold is better. On the whole adaptive threshold is best for nodule and nonnodule images.
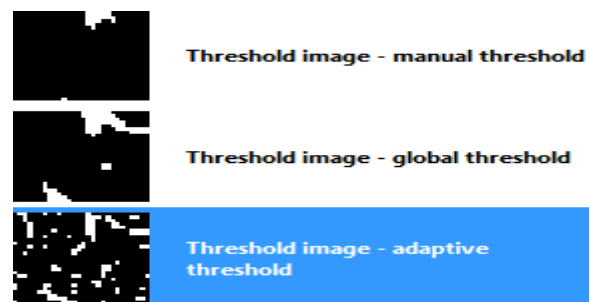

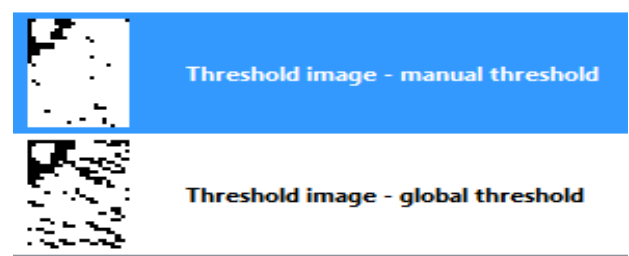Fig 10 (a) Threshold based Image Segmentation for Nodule

Fig 10 (b) Threshold based Image Segmentation for Non-Nodule 1

Published by :

http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 07, July-2019

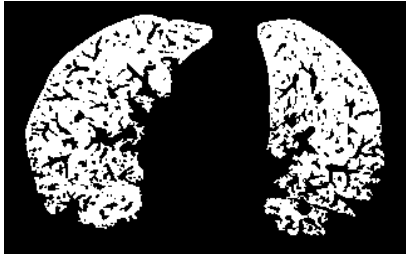Fig 10 (c) Global threshold and manual threshold samples of nodules
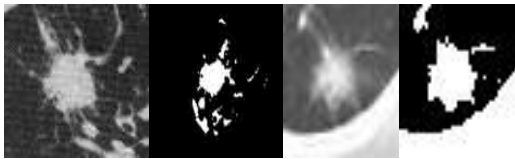

Fig 11. Paranchyma with threshold value 128

Fig 12. Region of Interest "Spericity" by Threshold Value 128 using Fuzzy Edge detection method

Edge detection is a basic step in image segmentation, image enhancement, image extraction, image selection, restoration, image registration and image compression.

The aim of edge detection is to identify the physical phenomena and localize the variations in the intensity of an image. Edge detection is an easy way of extracting most important features in all type of binary images (fig 12).
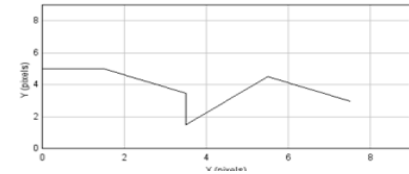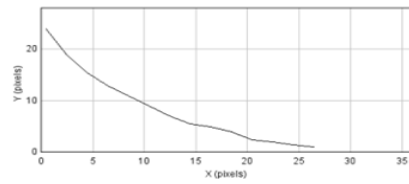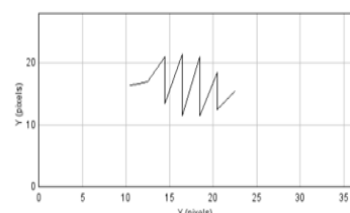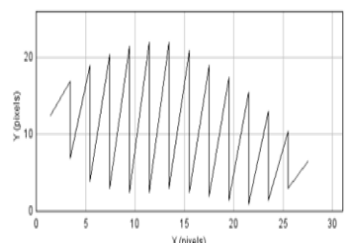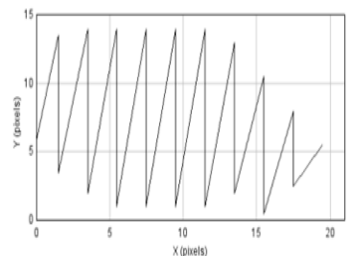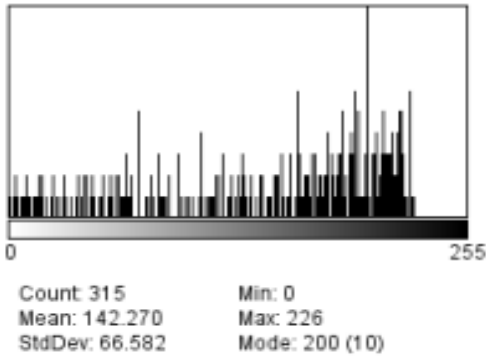

Fig 13. Line Graphs of Nodule Edges

Edge detection should be efficient because the edge detection should provide all significant information about the image. However, differentiation of an image is sensitive to electronic, semantic and discreatization or quantification effects. So it is tough work to design a general edge detection algorithm which performs well in many aspects.Fig 13, shows the line graph of edges in (x,y) pixels for nodules. Types of edges are step edge, ramp edge, line edge and roof edge. Because of low frequency elements or smoothing applied by sensing devices, sharp discontinuities like step and line edges, rarely exist in real signals/images. Step edges become ramp edges and line edges become roof edges. Intensity changes are not instantaneous by occur over a fixed distance.

Table I. Statistical Result of Nodules

NODULE	Image 3m-75-1.bmp	Image 3-44-1.bmp	Image 3-115-1.bmp	Image 3-n-71.bmp	Image 240-10-1.bmp	Image 24-69-1.bmp	Image 476-217-1.bmp	Image 11-115-1.bmp
MEAN	103	54	60	62	120	115	115	119
STANDARD DEVIATION	16	32	28	15	68	62	76	68

The mean of an image is calculated by sum up the pixel values of nodule image divided by the total number of pixels. In the above table mean value 11-115-1.bmp is 119. This is the maximum mean value. The probability distribution and intensity value of image 476-217-1.bmp indicated the better intensity level and high contrast of edges. Hence the standard deviation is the measure of inhomogenity of nodule.

Count: 315   Min: 0
Mean: 142.270   Max: 226
StdDev: 66.582   Mode: 200 (10)

Fig 14. Histogram of trained nodule

Histogram is a graphical representation of pixels in grey or level. The fig 14 shows the pixel falls in bins. The min value is zero and max value is 226 showed perfectly in graph. The mean and standard deviation also calculated for segmentation.
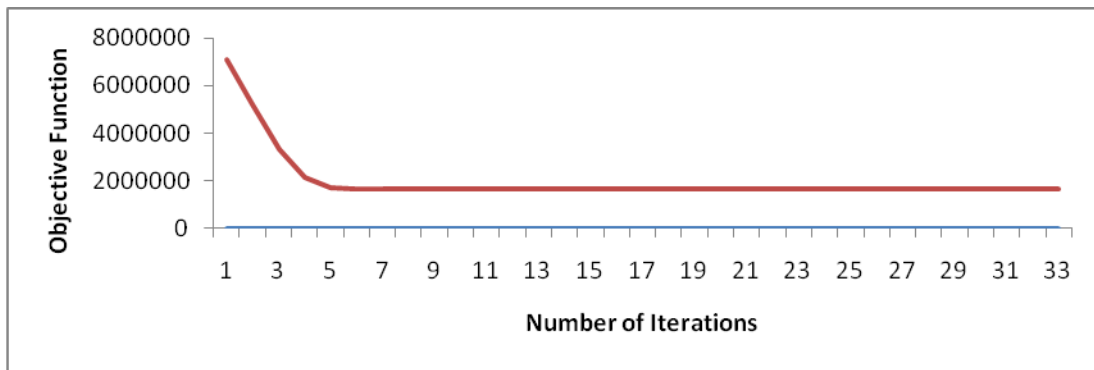

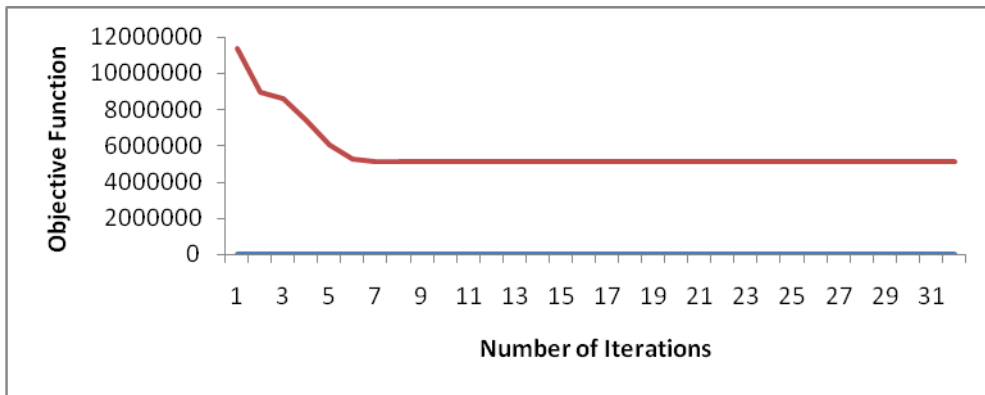Fig 15. (a) FCM Result from Cluster 1

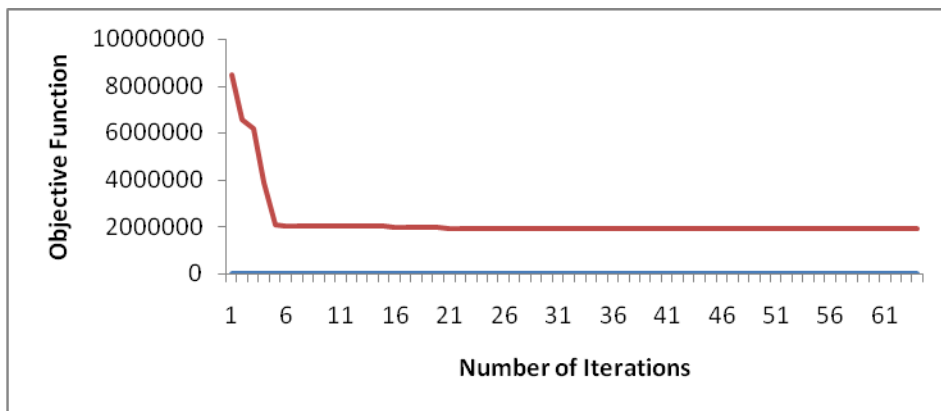Fig 15. (b) FCM Result from Cluster 2

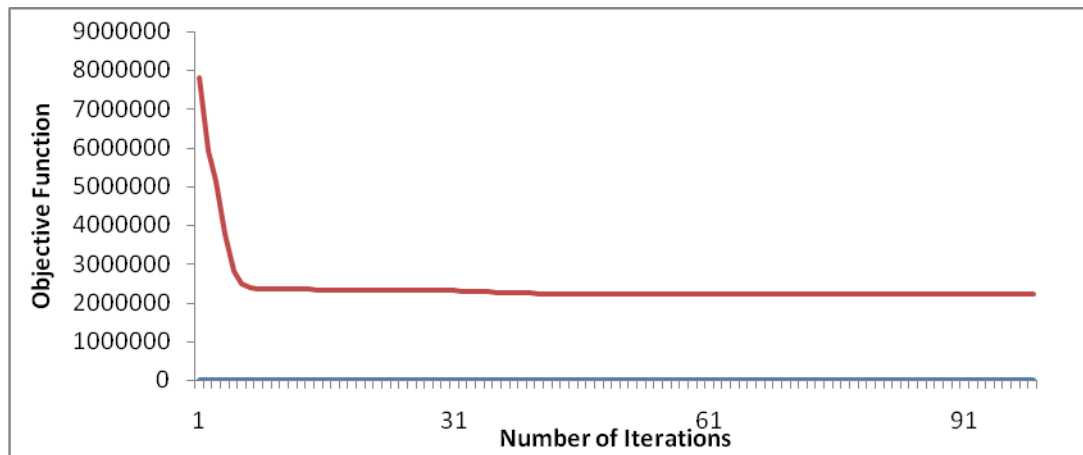Fig 15. (c) FCM Result from Cluster 3

Fig 15. (d) FCM Result from Cluster 4

The FCM result with cluster2 showed the less number of iterations for clustering the pixels of preprocessed nodule. Compare to other cluster this showed the best. Clustering algorithms achieve region segmentation by partitioning the image into sets or clusters of pixels that have strong similarity in the feature space. In hard clustering, data are divided into distinct clusters, where each data element belongs to exactly one cluster. In fuzzy clustering, data elements can belong to more than one cluster. with Each element is associated with set of membership levels. These indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a inspired process of assigning membership levels. Then assign data elements to one or more clusters. One of the most widely used fuzzy clustering algorithms is the FCM Algorithm. This can be used in many fields such as pattern recognition, fuzzy identification, and feature extraction.

IV. CONCLUSION

Segmentation types are discussed with edge based, threshold based, region based, soft computing based methods. Gaussian Radial Basis Function (GRBF) using Fuzzy Method works well for CT images. Computer-aided diagnosis of lung cancer is used to segment pathologically changed tissues fast and accurately. The proposed method HE+FCM successfully segment the pulmonary nodule from the CT image with accuracy value 91.6%. For lesion the average volume error obtained is 0.968%. The coefficients of similarity, spatial overlap, RMSE, average over and under of segmentation ratio are 0.914, 0.584, 0.10, 0.63% and 0.015% respectively. The experimental results indicate that the newly proposed method could segment blood vessel adhesion, pleura adhesion fast and exactly performs better than traditional methods.

REFERENCES

[1] Niranjana, G.; M. Ponnavaikko, "A Review on Image Processing Methods in Detecting Lung Cancer Using CT Images", International Conference on Technical Advancements in Computers and Communications, 10-11 April, Melmaruvathur

[2] Rabia Naseem ; Khurram Saleem Alimgeer ; Tariq Bashir, "Recent trends in Computer Aided diagnosis of lung nodules in thorax CT scans", International Conference on Innovations in Electrical Engineering and Computational Technologies, 5-7 April, Karachi, Pakistan, 2017

[3] Devi Nurtiyasari ; Dedi Rosadi ; Abdurakhman, "The application of Wavelet Recurrent Neural Network for lung cancer classification", International Conference on Science and Technology - Computer, 11-12 July, Yogyakarta, Indonesia, 2017.

[4] Bhagyarekha U. Dhaware ; Anjali C. Pise, "Lung cancer detection using Bayasein classifier and FCM segmentation", International Conference on Automatic Control and Dynamic Optimization Techniques, 9-10 Sept, Pune, India, 2016.

[5] Manasee Kurkure ; Anuradha Thakare, "Lung cancer detection using Genetic approach", International Conference on Computing Communication Control and automation, 12-13 August, Pune, India, 2016.

[6] Ali Shojaee Bakhtiari, Ali Asghar Beheshti Shirazi and Amir Sepasi Zahmati, "An Efficient segmentation Method Based Local Entropy Characteristics of Iris Biometrices", World Academy of Science , Engineering and Technology, 28, 2007, 64-68.

[7] Chi-Man Pun and Hong-Min Zhu, "Image segmentation using Discrete Cosine Texture Feature, International Journal of computers Issue1,Volume 4, 2010, 19-26

[8] Harvey A, Cohen computer science and computer engineering La Trobe University Bundoora 3083 Parallel algorithm for gray scale image segmentation Harvey A. Cohen, Parallel algorithm for gray –scale image segmentation, proc. Australian and new Zealand conf. Intelligent information systems, ANZIS- 96,Adelaide, Nov 18-20, 1996, pp 143-146.

[9] Prasantha, H.S., Shashidhra. H., K.N.B.Murthy, Madhavi Lata, Medical Image Segmentation. S.Prasantha et al. (IJCSE) international journal on computer science and engineering, 02(04), 2010, 1209-1218

[10] Zhiwu Lu Unsupervised Image Segmentation using an Iterative Entropy Regularized Likelihood Learning Algorithm J. Wang et al.(Eds.): ISNN 2006, LNCS 3972 , pp. 492-497, 2006. C Springer- Verlag Berlin Heidelberg 2006.

[11] Zhiwu Lu, Xiaoqing Lu, and Zhiyuan Yc, entropy Regularization, Automatic V Model Selection V, and Unsupervised Image Segmentation Z. –H. Zhou, H.Li, and Q. Yang (Eds.):PAKDD 2007. LNAI 4426, pp. 173-182, 2007._c Springer –Verlag Berlin Heidelberg 2007.

[12] Zhiwu Lu, Yuxin Peng and Jianguo xiao Institute of Computer Science and Technology, Peking University, Beijing, China, "Unsupervised learning of Finite Mixtures Using Entropy Regularization and its application to Image Segmentation", 978-1-4244-2243-2/08/$25.00, 2008 IEEE.

[13] Sunhee kim, Seugmi oh and Myungjoo Kangy, "Mathematical Image Processing for Automatic number plate Recognition System", 14(1), 57-66, 2010.

[14] Harvey A, Cohen computer science and computer engineering La Trobe University Bundoora 3083 Parallel algorithm for gray scale image segmentation Harvey A. Cohen, Parallel algorithm for gray –scale image segmentation, proc. Australian and new Zealand conf. Intelligent information systems, ANZIS- 96,Adelaide, Nov 18-20, 1996, pp 143-146

[15] Raj Kumar, GVS, Srinivas Yarramalle, A.Tejaswi, Naveen Kumar K., Chhajar Sunanda. "Image Segmentation by using finite bivariate doubly Truncated Gausian Mixture model", International Journal of Engineering Science and Technology, 2(4), 2010, 694-703.

[16] Ramana Reddy, B.V., M. Radhika Mani, V. Vijaya Kumar, "A Random Set View of Texture Segmentation International Journal of Signal and Image processing", (Vol.1-2010/Iss.1), A Random Set view of Texture Segmentation / pp.24-30.

[17] Salem Saleh Al-amri, N. V.Kalyankar and Khamitkar S. D, "Image Segmentation by using edge Detection", (IJCSE) International Journal on Computer Science and Engineering, 02(03), 2010, 804-807

[18] Vijaya, G., V. Vasudevan, "A Simple Algorithem for Image Denoising Based on MS Segmentation", International Journal of Computer Applications (0975 - 8887). 2(6), 2010.