ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 10, April 2014

Segmentation and Analysis of Lung Cancer Images Using Optimization Technique

Joel George R, Anitha Jeba Kumari D
Department of Applied Electronics, Sri Lakshmi Ammaal Institute of Technology, Chennai, TN, India
Department of Instrumentation and Control, St.Joseph's College of Eng, Chennai, TN, India

Abstract: In this paper an optimization method is proposed to segment the lung cancer image. The image is acquired for which Otsu's thresholding approach is used and then histogram equalization of the image is identified. It is a computational procedure that sort images into groups according to their similarities. Two methods such as Fuzzy K- means clustering method and Particle Swarm optimization method are implemented for segmenting an image. The various parameters like PSNR, MSE are calculated and compared. The proposed optimization method gives optimal solution.

***Keywords*–Thresholding based segmentation, Fuzzy K-means (FKM), PSO.**

I. INTRODUCTION

To detect lung cancer at an early stage is quite difficult [1] with the existing methods of thresholding [3] and segmentation. Here the lung CT scan images are taken for the work. So to get information from the scan images, it has to be processed using certain techniques like making the gray level of the image to be black and white which is done by fixing the thresholding level. After the image is done with thresholding, the image is segmented using a good segmentation technique that divides the image according to the region of interest. Medical image segmentation is a vital component of a large number of applications such as to study anatomical structure, identify regions of interest ie., locate tumor lesion, measure tissue volume to assess growth or decrease in the size of tumor, helps in treatment planning prior to radiation therapy and in radiation dose calculation. In segmentation process, contour detection is still a challenging problem in medical imaging because contour delineation error above 10% may lead to an unacceptable risk to irradiate healthy tissues instead of affected ones. So a method based on computational intelligence is proposed to deploy a more effective segmentation procedure. In this work, the theory of fuzzy K means is implemented as it is an interesting and useful tool that provides a good theoretical basis to represent imprecision of the information. The Fuzzy K-Means method is designed to classify the images. Another optimization algorithm named PSO is used for making the segmentation more successful.

II. METHODOLOGY

This system is fully implemented (in matlab) and tested with real CT scan images. [6]

A. Image Acquisition

An attempt has been made to collect few lung cancer images from a private hospital (APOLLO SPECIALITY HOSPITALS, CHENNAI). The digitized information is stored in the DICOM format with a resolution of 8 bits per plane. CT images have low noise, better clarity and less distortion so it is been taken for studying the segmentation methods. The Peak Signal to Noise Ratio (PSNR) values and Mean Squared Error (MSE) Values are calculated for the various images processed using different segmentation methods.

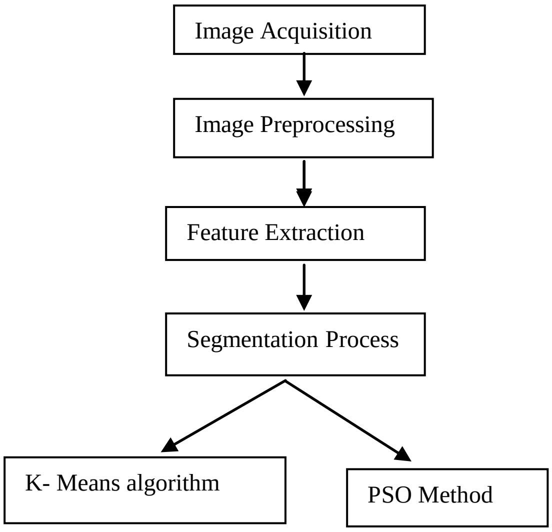

Fig 1. Methodology of work

B. Preprocessing of Image

The quality of the image is affected by different artifacts like non- uniform intensity, variations in motion, shift and noise. So the image is processed by certain methods like thresholding, histogram equalization etc., to remove redundancy present in the scanned images without affecting the features of the image. In this work thresholding method is used for preprocessing. Thresholding is a non-linear operation that converts a gray-scale image into a binary image where the two levels are assigned to pixels that are below or above the specified threshold value. Here Otsu's method [9] is used (gray thresh) for computing global image threshold. Otsu's method is based on threshold selection by statistical criteria. Otsu suggested minimizing the weighted sum of within-class variances of the object and background pixels to establish an optimum threshold. Recall that minimization of within-class variances is equivalent to maximization of

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 10, April 2014

between-class variance. This method gives satisfactory results for bimodal histogram images. Threshold value based on this method will be between 0 and 1, after achieve this value we can segment an image based on it.

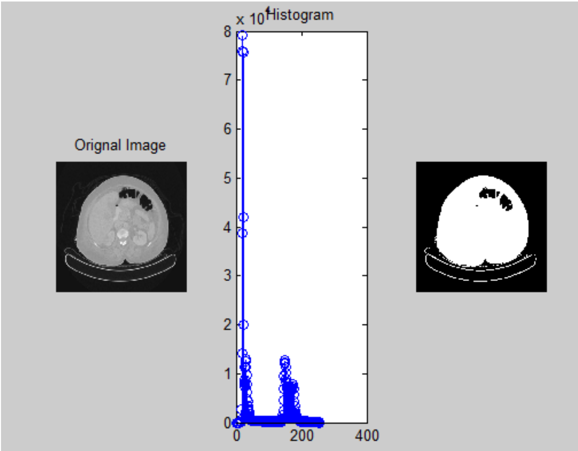

Fig 2 Otsu's Gray threshold method on CT scan image

C. Feature Extraction

The various features of the image is extracted using different techniques like binarization [6],[10] and Gray Level Co-Occurrence Matrix (GLCM) [4] where both methods are based on facts that strongly related to lung anatomy and information of lung CT imaging. The features are extracted to detect and isolate various desired portions or shapes (features) of the image.

D. Gray Level Co-Occurrence Matrix Method

GLCM is a tabulation of how often different combination of pixel brightness value (gray level) occur in an image. Here the matrix is formed from the image using gray co-matrix function in MATLAB. Then the matrix is normalize d using the following formula

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}}$$

Where, i is the row number and j is the column number. From this we calculate texture measures from the GLCM. The following features are extracted using this method

- Contrast
- Correlation
- Energy
- Homogeneity

III. K-MEANS ALGORITHM

K-Means [5] is a rather simple but well known algorithm for grouping objects, clustering. The K-Means method is numerical, unsupervised, non-deterministic and iterative Again all objects need to be represented as a set of numerical features. In addition the user has to specify the number of groups (referred to as k) he wishes to identify. Each object can be thought of as being represented by some feature vector in an n dimensional space, n being the number of all features used to describe the objects to cluster. The algorithm then randomly chooses k points in that vector space, these points serve as the initial centers of the clusters. Afterwards all objects are each assigned to center they are closest to. Usually the distance measure is chosen by the user and determined by the learning task. After that, for each cluster a new center is computed by averaging the feature vectors of all objects assigned to it. The process of assigning objects and recomputing centers is repeated until the process converges. The algorithm can be proven to converge after a finite number of iterations. Several tweaks concerning distance measure, initial center choice and computation of new average centers have been explored, as well as the estimation of the number of clusters k. Yet the main principle always remains the same do that for you.

A. K-means algorithm properties

- There are always K clusters.
- There is always at least one item in each cluster.
- The clusters are non-hierarchical and they do not overlap.
- Every member of a cluster is closer to its cluster than any other cluster because closeness does not always involve the 'center' of clusters.

B. K-means algorithm process

- The dataset is partitioned into K clusters and the data points are randomly assigned to the clusters resulting in clusters that have roughly the same number of data points.
- For each data point:
- Calculate the distance from the data point to each cluster.
- If the data point is closest to its own cluster, leave it where it is. If the data point is not closest to its own cluster, move it into the closest cluster.
- Repeat the above step until a complete pass

Through all the data points results in no data point moving from one cluster to another. At this point the clusters are stable and the clustering process ends. The choice of initial partition can greatly affect the final clusters that result, in terms of inter-cluster and intra cluster distances and cohesion.

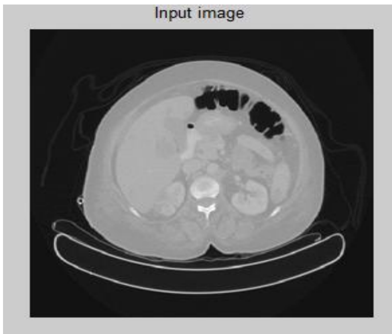

Fig 3 Original Image

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 10, April 2014

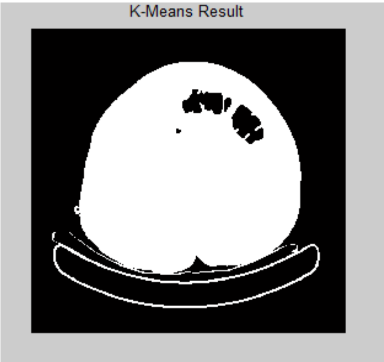

Fig 4 Segmented image using K- means algorithm

IV. PARTICLE SWARM OPTIMIZATION METHOD

Here, a multilevel thresholding method segmenting images based on particle swarm optimization (PSO) is proposed [7]. In the proposed method, the thresholding problem is treated as an optimization problem, and solved by using the principle of PSO. The algorithm of PSO is used to find the best values of thresholds that can give us an appropriate partition for a target image according to a fitness function. The proposed method has been tested on different images, and the experimental results demonstrate its effectiveness. Particle swarm optimization (PSO) [7] is a population-based optimization algorithm modeled after the simulation of social behavior of birds in a flock [11], [12]. The algorithm of PSO is initialized with a group of random particles and then searches for optima by updating generations. Each particle is flown through the search space having its position adjusted based on its distance from its own personal best position and the distance from the best particle of the swarm. The performance of each particle, i.e. how close the particle is from the global optimum, is measured using a fitness function which depends on the optimization problem. Each particle, i, flies through an n–dimensional search space, Rn, and maintains the following information: xi, the current position of ith particle (x - vector), pi, the personal best position of ith particle (p - vector), and vi, the current velocity of ith particle i (v - vector). The personal best position associated with a particle, i, is the best position that the particle has visited so far. If *f* denotes the fitness function, then the personal best of i at a time step t is updated as:

$$P_i(t+1) = \{P_i(t) iff(x_i(t+1) \geq f(P_i(t)) \quad (1)$$

$$\{X_i(t+1) iff(X_i(t+1) < f(P_i(t+1) \quad (2)$$

If the position of the global best particle is denoted by *gbest*, then:

$$gbest \in \{ p1(t), p2(t), \ldots, pm(t) \}$$
$$= \min\{ f(p1(t)), f(p2(t)), \ldots, f(pm(t)) \} \quad (3)$$

The velocity updates are calculated as a linear combination of position and velocity vectors. Thus, the velocity of particle i is updated using equation (4) [13] and the position of particle i is updated using equation (5) [14]

$$v_i(t+1) = w\, v_i(t) + c_1 r_1(p_i(t) - x_i(t)) + c_2 r_2 (gbest - x_i(t)) \quad (4)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (5)$$

In the formula, w is the inertia weight [16] , c1 and c2 are the acceleration constants, r1 and r2 are random numbers in the range [0,1] and Vi must be in the range [-Vmax, Vmax], where Vmax is the maximum velocity.

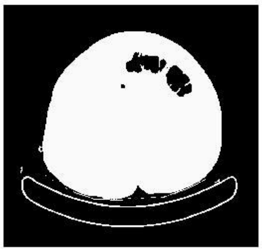

Fig 5 Segmented image using PSO Algorithm

The values chosen for PSO processing n=256; c1=2.1; c2=2.1; wmax=0.8; wmin=0.4; G=30; M=2; Where,n=no.of pixels, M = centroids

V. RESULTS AND DISCUSSION

The GLCM method is implemented on the sample image and the various values are tabulated below

Parameters	Min	Max
Contrast	0.1404	4.5162
Correlation	0.3220	0.9781
Energy	0.1018	0.2994
Homogeneity	0.5899	0.9558

Table 1 Feature extracted using GLCM method

The obtained values are used for further segmentation of the images using K- means algorithm. Additional to these values the PSNR value and MSE value is calculated using the formula [17]

$$PSNR = 10 \log_{10} [R^2/MSE] \quad (6)$$

$$MSE = \sum_{M,N} [I_1(m,n) - I_2(m,n)]^2 \quad (7)$$

The PSNR value and MSE for the original image is given below

Image	PSNR	MSE
Original Image	30.2570	61.5146

Table 2 PSNR and MSE for original image

Now the image is segmented using the two methods, K-Means Algorithm and PSO algorithm. The PSNR and MSE values are calculated using Equation (6) & (7) for the segmented image and the values are tabulated below

Name of the Patient	PSO		K-Means	
	PSNR	MSE	PSNR	MSE
Sample 1	32.8618	33.9082	32.8156	34.2706
Sample 2	32.8460	34.0314	32.8210	34.2275
Sample 3	32.8807	33.716	32.8237	34.2065
Sample 4	32.9102	33.5319	32.8612	33.9125
Sample 5	32.8270	34.1802	32.7860	34.5046
Sample 6	32.9297	33.3818	32.8250	34.1916
Sample 7	32.8842	33.7336	32.8515	33.9886
Sample 8	32.8413	34.0134	32.7799	34.5534

In general, PSNR must increase and MSE must decrease for a good segmented image, so here the PSNR and MSE values are compared between original image and the segmented image. Here the segmentation done with PSO algorithm proves to be optimal solution when compared to that segmentation done with K- Means Algorithm.

VI. CONCLUSION

In this work, the CT images are acquired and the series of operations are performed to enhance the image quality. Here the image is first converted to gray scale image by the otsu's thresholding method. Then the segmentation of the image is done by two methods such as K means algorithm and PSO algorithm [11]. The PSNR value and MSE value is calculated for both the segmented images and compared. Here the PSO method is proved to best in obtaining the PSNR and MSE value. This process can be improved by implementing any other evaluation algorithm to obtain the best segmentation of the image.

VII. ACKNOWLEDGEMENT

The author would like to thank **Apollo Specialty Hospitals, Chennai** for providing the lung CT scan images.

REFERENCES

[1] Yongjun WU, Na Wang, Hongsheng ZHANG, Lijuan Qin, Zhen YAN, Yiming WU, "Application of Artificial Neural Networks in the Diagnosis of Lung Cancer by Computed Tomography", 2010 Sixth International Conference on Natural Computation (ICNC 2010).

[2] M. Gomathi and P. Thangaraj, "A Computer Aided Diagnosis System for Lung Cancer Detection \Using Support Vector Machine", American Journal of Applied Sciences 7 (12): 1532-1538, 2010 ISSN 1546 9239.

[3] Mokhled S. AL-TARAWNEH, "Lung Cancer Detection Using Image Processing Techniques", Leonardo Electronic Journal of Practices and Technologies ISSN 20, January-June 2012 p. 147-158.

[4] P. Mohanaiah, P. Sathyanarayana, L. Guru Kumar, "Image Texture Feature Extraction using GLCM Approach", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013 1 ISSN 2250-3153.

[5] Almas Pathan, Bairu.K.saptalkar, "Detection and Classification of Lung Cancer Using Artificial Neural Network", International Journal on Advanced Computer Engineering and Communication Technology Vol-1 Issue: 1: ISSN 2278 – 5140.

[6] Ada, Rajneet Kaur, "Feature Extraction and Principal Component Analysis for Lung Cancer Detection in CT Scan Images", International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 3, March 2013 ISSN: 2277128X.

[7] Fahd M. A. Mohsen, Mohiy M. Hadhoud, Khalid Amin, "A new Optimization-Based Image Segmentation method By Particle Swarm Optimization", (IJACSA) International Journal of Advanced Computer Science and Applications, Special Issue on Image Processing and Analysis.

[8] Disha Sharma, Gagandeep Jindal, "Identifying Lung Cancer Using Image Processing Techniques", International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2011).

[9] Anita Chaudhary, Sonit Sukhraj Singh, "Lung cancer Detection using Digital Image Processing", International Journal of Research in Engineering & Applied Sciences(IJREAS), Volume 2, Issue 2, (Feb 2012),ISSN:2249-3905.

[10] Jose, Enrique, Francisco, Antonio, Paula, Gisela, Marfa, Femat, Moises, "Development of An optimized multi-biomarker panel for the detection of lung cancer based on principal component analysis and artificial neural network modeling", ELSEVIER Expert Systems with Applications 39 (2012) 10851-10856.

[11] J. Marcello, F. Marques and F. Eugenio, "Evaluation of thresholding techniques applied to oceanographic remote sensing imagery," SPIE, 5573, pp. 96-103, 2004.

[12] J. Kennedy, and R. Eberhart, "Swarm Intelligence", San Francisco: Morgan Kaufmann Publishers, 2001

[13] R. O. Duda and P. E. Hart, "Pattern Classification and Scene Analysis", John Wiley & Sons, New-York, 1973.

[14] A. A. Younes, I. Truck, and H. Akdaj, "Color Image Profiling Using Fuzzy Sets," Turk J Elec. Engin., Vol.13, No.3, 2005.

[15] Fatma Taher1,*, Naoufel Werghi1, Hussain Al-Ahmad1, Rachid Sammouda"Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering Methods" American Journal of Biomedical Engineering 2012, 2(3): 136-142 DOI: 10.5923/j.ajbe.20120203.08.

[16] Ai-Qin Mu, De-Xin Cao, Xiao-Hua Wang, "A Modified Particle Swarm Optimization Algorithm" Vol.1, No.2, 151-155 (2009)

[17] Niveditta Thakur, Swapna Devi "A New Method for Color Image Quality Assessment". International Journal of Computer Applications (0975 – 8887) Volume 15– No.2, February 2011.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 10, April 2014

AUHTOR'S PROFILE

Joel George R is a PG student in the Department of Aplied Electronics in Sri Lakshmi Ammaal Institute of Technology, Anna University affiliated. He has completed his UG in Electronics and Communication in Karunya Institute of Technology, Coimbatore.

Anitha Jeba Kumari D has completed her PG in Medical Electronics, Anna University, CEG campus and UG in Department of Instrumentation and Control engineering in Sri Sairam Engineering College. She is currently working as Assistant Professor in St. Joseph's College of Engineering, Chennai.