



## Lung Cancer Detection Using Spatially Weighted Fuzzy C-Mean Clustering Algorithm

<sup>1</sup>V. Ramesh Babu\*, <sup>2</sup>A.N. Nandakumar

<sup>1</sup>Ph.D Research Scholar, Department of CSE, St.Peter's University, Avadi, Chennai,TN, India.

<sup>2</sup>Principal, R.L Jalappa Institute of Technology, Doddaballapur, Bangalore, Karnataka, India.

\*Corresponding author's E-mail: [vrbabu252001@yahoo.co.in](mailto:vrbabu252001@yahoo.co.in)

Accepted on: 10-06-2015; Finalized on: 30-06-2015.

### ABSTRACT

Lung cancer seems to be the common cause of death among people throughout the world. Early detection of lung cancer can increase the chance of survival among people. The Computed Tomography (CT) can be more efficient than X-ray. However, problem seemed to merge due to time constraint in detecting the present of lung cancer regarding on the several diagnosing method used. Hence, a lung cancer detection system using image processing is used to classify the present of lung cancer in a CT-images. This paper investigates spatially weighted fuzzy C-mean (SWFCM) techniques that allow detection of lung cancer through analysis of chest computed tomography (CT) images. The overall accuracy, Sensitivity, Predictive values achieved are 76.39%, 88.83% and 85.93% respectively.

**Keywords:** CT lung image, SWFCM, Threshold segmentation, Fundus image.

### INTRODUCTION

Lung cancer is a disease of abnormal multiplication of cells that grows into a tumor. The mortality rate of patients affected by lung cancer is the highest compared to all other types of cancer. Lung cancer is one of the most life threatening diseases in the world, with the less survival rate after the diagnosis, with the maximum rate of deaths.<sup>1</sup> In the earlier days, the identification of lung cancer has been a tedious task in medical image analysis. If the lung cancer can be diagnosed at an early stage, there is possibility to enhance significant percentage of the survival rate of the patients.

nodules using raw chest X-ray images does not provide accurate result since analysis of such medical images is tedious and difficult task. Currently, CT is used in the place of plain chest X-ray in detecting and diagnosing the lung cancer.

The overall 5-year survival rate for lung cancer patients increases from 14 to 49% if the disease is detected in time.<sup>2</sup>

Figure 1 shows the flow chart of the proposed method. In this paper lung CT images are used and the size of images 512\*512 is obtained in Bharath Scan center, India.

### Related Work

Abdullah<sup>3</sup> stated that the segmentation of the lung region due to the limitation regarding on the similarities of the intensity in the x-ray image.

As for lung cancer nodule detection process, it does not seem to be the problem because of the absent of the similar intensity due to the lung segmentation. It can be used in the lung in the lung cancer application and the system can also be used in the application such as the detection and classification of breast tumor in mammography images regarding on the higher variation of intensity present.

Zare<sup>4</sup> declared that the approaches of content-based image retrieval (CBIR) using low level features such as shape and texture are investigated in order to create a framework that classify medical x-ray image automatically. GLCM, canny edge operator, local binary pattern and pixel level information of the images act as image based feature representation.

The performance of image classification offered by combining the promising features stated above investigation. Experimental results using 116 different

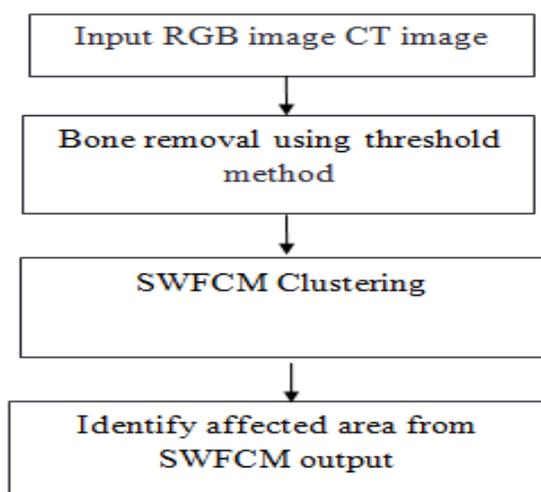


Figure 1: "Flow chart of the proposed method"

Generally, Lung cancer spreads towards the Centre of the chest because the natural flow of lymph out of the lungs is toward the Centre of the chest. In medical terms, Lung cancer is classified into two main category, non-small cell lung cancer and small cell lung cancer. Identifying lung

classes of 11,000 x-ray images 90.7% classification accuracy.

Gomathi<sup>5</sup> expressed that computer aided diagnosing system which uses FPMC algorithm for segmentation to improve the accuracy. Rule based technique is applied to classify the cancer nodule after segmentation. For its better classification, the learning is performed with the help of extreme learning machine.

Jia Tong<sup>6</sup> acknowledged that several steps are followed to detect the cancer like segmentation of lung parenchyma, the detection of suspicious nodule candidates, the feature extraction and classification. Here the author used adaptive threshold segmentation, math morphologic, Gaussian filter, Hessian matrix algorithms.

**Elimination of Bone Region**

The bone region affects the segmentation accuracy so the first step is the removal of bone region from the lung CT image. Separate R-plane, G-plane, and B-plane from RGB Image.<sup>7</sup> In this entire plane, the bone region is detected. Subtract all these images, so that the resultant image is T.

$$T=R-G-B \tag{1}$$

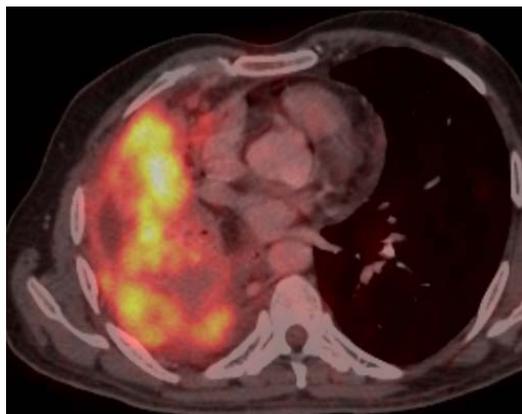


Figure 2: "Input lung CT image"

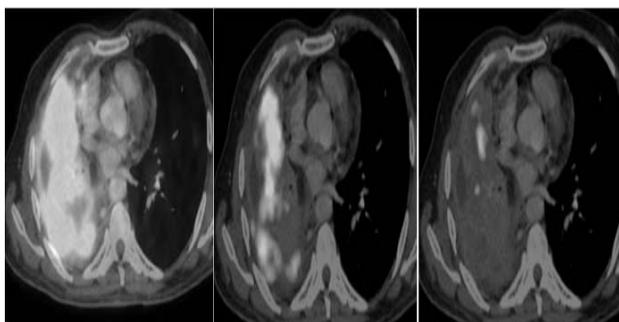


Figure 3: "R-plane, G-plane, and B-plane"

Figure 2 shows the input lung image and Figure 3 shows R-plane, G-plane, B-plane image.

Figure 4 shows the enhanced and bone eliminated image T is added with difference between G-plane and B-plane to obtain the affected area of the disease.

$$S=T+[G-B] \tag{2}$$

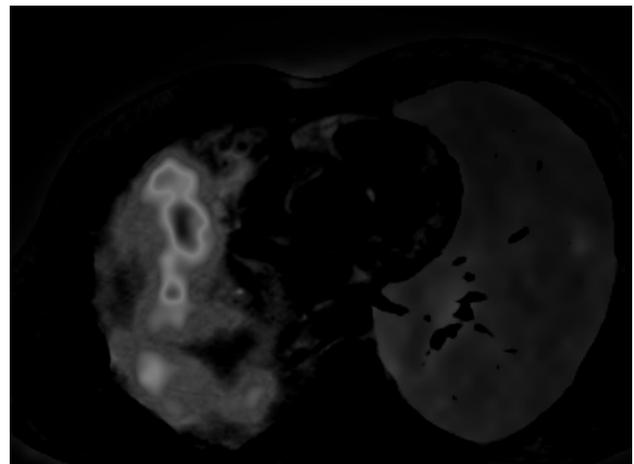


Figure 4: "Enhanced and bone removed image"

Figure 4 shows the cancer affected area image. This image is used to input of the SWFCM clustering.

**Lung Cancer Detection using Spatially Weighted Fuzzy C-Mean Algorithm**

The standard FCM<sup>8</sup> does not consider the spatial information of pixels and in turn, the segmentation result is affected. One of the important characteristics of an image is that neighboring pixels are highly correlated which is considered in spatially weighted fuzzy C-mean (SWFCM) method. The FCM method is utilized by the previous researchers for fundus image segmentation.<sup>9</sup>

Though the quality of CT image will not be as good as the fundus image, SWFCM method is adopted for corresponding region.



Figure 5: "SWFCM image"

**Segmentation of Lung cancer Using SWFCM**

SWFCM is applied to the images in which bones are already removed. One of the important characteristics of an image is that neighboring pixels are highly correlated. The spatial relationship is important in clustering, but it is not utilized in a standard FCM algorithm. In SWFCM, to exploit the spatial information, a spatial function is defined as

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik} \tag{3}$$

Where  $NB(x_j)$  represents a square window centered on pixel  $x_j$  in spatial domain. Larger window size may blur the images and the smaller window size does not remove the noise at high density. Therefore, an optimal window of size 5x5 is used in this work. Just like the membership function, the spatial function  $h_{ij}$  represents the probability that the pixels  $x_j$  belong to the  $i^{th}$  cluster. The spatial function of pixels is large if the majority of its neighborhood belongs to the same clusters. The spatial function is incorporated into membership function as follows

$$u'_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^c u_{kj}^p h_{kj}^q} \tag{4}$$

Where p and q are the controlling parameters of both functions. The spatial functions simply strengthen the original membership in a homogenous region, but it does not change clustering result. However, this formula reduces the weight of a noisy cluster in noisy pixels by the labels to its neighboring pixels. As a result, misclassified pixels from noisy region or spurious blobs can easily be corrected.

The clustering is a two-pass process at each iteration. The first pass is the same as that in standard FCM to calculate the membership function.

In the second pass, the membership information of each pixel is mapped to the spatial domain and the spatial domain function is computed from that.

The FCM iteration proceeds with the new membership that is incorporated with spatial function.

The iteration is stopped when the maximum difference between two cluster centers at two successive iterations is less than 0.00001. After the convergence, defuzzification is applied to assign each pixel to a specific cluster for which the membership is maximal.

**Algorithm**

**Step 1:** Generate the random number with the range from 0 to 1 to be the initial memberships. Let us consider the number of cluster is N then calculate  $V_i$  using (5)

$$V_i = \frac{\sum_{j=1}^N u_{ij}^m x_j}{\sum_{j=1}^N u_{ij}^m} \tag{5}$$

Where,

- $v_i$  =  $i^{th}$  cluster center
- $m$  = fuzziness parameter  $m=2$

Where  $u_{ij}$  is by using Equation (3.21)

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{2/(m-1)}} \tag{6}$$

**Step 2:** Map  $u_{ij}$  into the pixel position and calculate the modified membership  $u'_{ij}$  using (6). Compute objective function J using (7)

$$J = \sum_{j=1}^N \sum_{i=1}^c u_{ij}^m \|x_j - v_i\|^2 \tag{7}$$

**Step 3:** Update the cluster center using (5)

**Step 4:** Repeat steps 2 to step 4 until the following termination criterion is satisfied:

$$\|J_{new} - J_{old}\| < \epsilon \tag{8}$$

Where  $\epsilon=0.00001$  which is same as in the FCM method used previously in this work.

The segmented image has three clusters, namely the backgrounds. The image with these three identified clusters is shown in the Figure 5.

**Detection of Affected Cancer Area**



**Figure 6:** "Identified clustered image"

The clustered image has three forms of cluster namely backgrounds, and affected area. From these three clusters, least cluster size is region of interest. The figure 6 size shows the detected affected area of the cancer. Figure 7 shows the superimposed original RGB image.



**Figure 7:** "Super imposed image"

## RESULTS AND DISCUSSION

Fifteen cancer affected images are used to test the proposed method. All the images are of 8-bit gray scale images and are of size 512x512 pixels. In all the images, areas of the ground truth (or gold standard) regions (T) for cancer affected area is marked manually to assess the accuracy of proposed method. Then the boundaries of affected cancer area are detected. The results of the proposed method are given in fig 7 for cancer affected images.

A simple and effective overlap measure of the match between the ground truth region and detected. Region (R) by the proposed method is used to measure the accuracy (M) as follows:<sup>10</sup>

$$M = \frac{\text{AREA}(T \cap R)}{\text{AREA}(T \cup R)} \quad (9)$$

The other 2 accuracy measures used are

$$\text{Sensitivity (S)} = \frac{TP}{TP+FN} \quad (10)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (11)$$

Where True positive TP= R n T;

False Positive FP=R - (R n T);

False Negative FN=T - (R n T).

The number of true negatives, i.e. the number of pixels that are not classified as exudate pixels, neither by the grader nor by the algorithm is very high. So the specificity is always near 100%.

This is not very meaningful. Therefore, alternative is to calculate the Predictive Value as

$$PV = \frac{TP}{TP+FP} \quad (12)$$

PV is the probability that a pixel that has been classified as exudate is really an exudate.

The overall accuracy, Sensitivity, Predictive value achieved are 76.39%, 88.83% and 85.93% for cancer affected area.

**Table 1:** "Performance of the proposed method on detecting the area (pixels) in abnormal image"

Image	T	R	M	TP	FP	FN	S (%)	SP (%)	Pv (%)
1	631	648	0.8483	587	61	44	0.9303	0.9059	0.9059
2	378	354	0.7387	311	43	67	0.8228	0.8785	0.8785
3	2651	2235	0.8197	2201	34	450	0.8303	0.9848	0.9848
4	3635	3353	0.8066	3120	233	515	0.8583	0.9305	0.9305
5	1885	2133	0.8214	1812	321	73	0.9613	0.8495	0.8495
6	1617	1222	0.6889	1158	64	459	0.7161	0.9476	0.9476
7	1252	1161	0.8448	1105	56	147	0.8826	0.9518	0.9518
8	1954	2289	0.7584	1830	459	124	0.9365	0.7995	0.7995
9	4345	4106	0.8541	3893	213	452	0.8960	0.9841	0.9841
10	1433	1379	0.7686	1112	157	211	0.8528	0.8861	0.8861
11	1860	1762	0.7481	1550	212	310	0.8333	0.8797	0.8797
12	13372	16503	0.7636	12935	3568	437	0.9673	0.7938	0.7938
13	3256	6218	0.5188	3236	2982	20	0.9939	0.5024	0.5024
14	2307	3008	0.7151	2216	792	91	0.9606	0.7367	0.7367
			0.7639				0.8883	0.8593	0.8593

## CONCLUSION

The proposed research work is performed on computed tomography images i.e. CT scan images for the detection of Lung Nodule is cancerous or not. In this paper we show the potential of SWFCM segmentation is used. The overall accuracy, Sensitivity, Predictive values achieved are 76.39%, 88.83% and 85.93% respectively. Besides its use as a potential screening tool for lung cancer, this method can be used to monitor treatment effectiveness, to detect the recurrence of lung cancer, and also to identify patients who may need an invasive diagnostic procedure. In future studies, we also plan to include larger study populations to establish statistical significance.

## REFERENCES

1. American Cancer Society (ACS), "Report on Lung Cancer," 2010
2. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, and M. J. Thun, "Cancer statistics, 2007," CA Cancer J Clin, 57(1), 2007, 43-66.
3. Azian Azamimi Abdullah1, a, Syamimi Mardiah Shaharum1, b Information Engineering Letters, 2, march 2012, 1.
4. Vinod Kumar1, Anil Saini2, Detection system for lung cancer based on neural network: X-Ray validation performance, International Journal of Enhanced Research in Management & Computer Applications, 2(9), 2013, 40-47.



5. M. Gomathi and P. Thangaraj A Computer Aided Diagnosis System for Lung Cancer Detection \ Using Support Vector Machine American Journal of Applied Sciences, 2010.
6. Shamala B. Terdale 1, K.V. Kulhalli 2 CAD System for Lung Cancer Detection using ANN IOSR Journal of Electronics & Communication Engineering (IOSR-JECE), 11-15.
7. Gonzalez. RC & Woods. RE, 'Digital Image Processing', NJ. Prentice Hall, Second Edition, 2005.
8. D. Lin and C. Yan, "Lung nodules identification rules extraction with neural fuzzy network", IEEE, Neural Information Processing, 4, 2002.
9. T.R. Ganeshbabu, R. Satishkumar, and Rengaraj venkatesh 2014, "Segmentation of optic nerve head for glaucoma detection using fundus images", Biomedical and Pharmacology Journal–7(2), Dec 2014, 1-9.
10. Kavitha.D & Shenbaga Devi.S, 'Automatic detection of optic disc and exudates in retinal images', International Conference on Intelligent Sensing and Information Processing, 2005, 501-506.

**Source of Support: Nil, Conflict of Interest: None.**

