Advances in Computational Sciences and Technology
ISSN 0973-6107 Volume 12, Number 1 (2019) pp. 45-51

http://www.ripublication.com

Analysis of CT Images and Localization of Lung Tumor using K-Means Clustering

Arpitha C G
M.Tech (Information Technology)
Department of ISE, RVCE
Bengaluru, India.

Dr. Rajashekara Murthy S
Associate Professor
Department of ISE, RVCE
Bengaluru, India.

Abstract

Lung cancer is the abnormal growth of cells in the lung causing severe risk to human health because lung has a connected network of blood vessels and lymphatic channels vulnerable to metastasis. There are two type of lung cancer: small cell lung cancer and non-small cell lung cancer. This paper aims to implement how K-Means clustering algorithm is applied on CT (Computed Tomography) images of lung to detect the lung cancer affected region and then to calculate the area of that lung affected region.

Keywords: Lung Cancer, CT image, K-means clustering algorithm.

I. INTRODUCTION

Computer vision is an interdisciplinary field make use of digital videos and digital images to gain deep and high-level understanding using computers. Here, understanding is nothing but the transformation of images caught by eyes into depictions of the world that can interface with other thought processes and bring out appropriate action. Clinical diagnosis and treatment planning is done by radiologist depending on the observation made using Computed Tomographic (CT) images or Magnetic Resonance Images (MRI) and doctors use computer to assist them. [10]

In medical imaging, segmentation plays a vital role. The accurate detection and diagnosis depends on the segmentation. [9] There are different segmentation techniques like thresholding, k-means clustering algorithm, histogram based method, edge detection, dual-clustering method, region growing method and many more. In this paper, the technique used for segmenting lung image is by k-means clustering algorithm. MATLAB is used as the developing tool. MATLAB permits framework

controls, plotting of capacities and information, execution of calculations, formation of UIs, and interfacing with projects written in different dialects, including C, C++, C#, Java, Fortran and Python.

II. RELATED WORK

Computer aided method for detecting and segmenting the cancer affected area from the lung MR image is done using two algorithms and they are k-means clustering and fuzzy c-mean algorithm. The lung affected area is calculated and also the stage of tumor is identified. [1]

Features derived from wavelet domain and intensity enhancement method are used to detect and classify stroke automatically using two-level classification scheme. [2]

Dzulkifli Bin Mohammad and M. Masroor Ahmed explains how tumor is extracted and segmented using k-means clustering algorithm from the affected images combining with Perona-Malik Anisotropic Diffusion Model. [3]

S.Mary Praveena and Dr. I1aVennila explains how image clustering and segmentation is done base on image pixels using K-means algorithm and image segmentation is done using the Optimization Fusion Approach. [4]

Threshold value will be selected. If the pixel value in the image is less than the threshold then it is replaced with the black pixel or with the white pixel. Segmentation of the lung affected region from MR images is done by thresholding technique. [7]

Thresholding can also be used to create binary image from gray scale image. This method some time ignores the lung affected cells also, so thresholding is not recommended. [8]

III. PROPOSED METHOD

The proposed technique comprises of five phases: acquisition of images, preprocessing the input image by removing noises, segmentation of affected region, extracting features from segmented region and finally calculating the area of affected region. Image acquisition is a phase in which the images which are fed as input the system are collected from various sources, in our case hospitals and scan centers are the sources. The image collected will not be in the required form so preprocessing phase is needed. Then the region of interest is identified and segmented from the background. In this paper k-means clustering algorithm is used for segmentation process. The required features are extracted in feature extraction phase. In the last phase i.e., area calculation phase the area of the lung affected part is calculated. Block diagram of proposed system is shown in below figure 1.

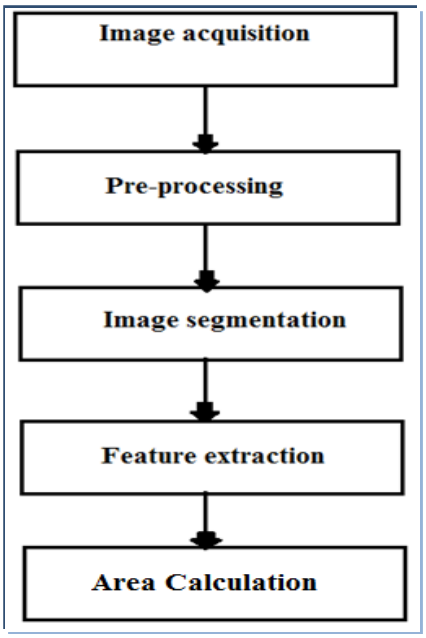

Figure 1: Block diagram of proposed method.

A. Image Acquisition

Image acquisition as the name says, it is collection of images fromdifferent sources like hospitals and scan centers. These images are fed as input to the system. Many problems were encountered while collecting the images. Hospitals maintain secrecy of their patient's data. So hospitals will not agree to share the data with outsiders. Figure 1 shows the sample input images collected.

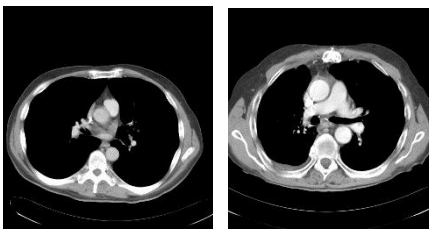

Figure 2: Sample input images.

B. Preprocessing

The collected images will not be in the required format, so the preprocessing of those images becomes necessary. The images are converted from RGB to gray scale using rgb2gray().

C. Image Segmentation

Image Segmentation plays a very important role in medical imaging. By using image segmentation process on medical images, the diagnosis of the affected region is made

easy and this also helps in treatment planning. [3]

There are many techniques for segmenting the region of interest (RoI) and they are: thresholding, k-means clustering algorithm, histogram based method, edge detection, dual-clustering method, region growing method and many more.

In this proposed paper k-means clustering algorithm is used for segmentation. It is the least difficult unsupervised learning method. Cluster value k is given in random. Trial and error method is used while choosing k till proper clustering is done. Depending on the type of image whether it is CT image or MR image k value varies. [5][6]

a. AGORITHM
   1. Give the number of cluster value as k.
   2. Pick the centers of k clusters arbitrarily.
   3. Calculate midpoint (center) of the cluster.
   4. The distance between every pixel to center of every cluster is calculated.
   5. The pixel is moved to the cluster to which calculated distance is minimum.
   6. Or else move the pixel to next cluster.
   7. Re-calculate the center.
   8. Repeat the whole process until the center of the cluster do no change.

Figure 3 shows the flowchart of k-means clustering algorithm. It shows how the number of clusters is chosen, the distance of each pixel to every cluster is calculated and the pixel is added to that cluster to which distance is minimum. Repeat this process until the cluster's center doesn't move.

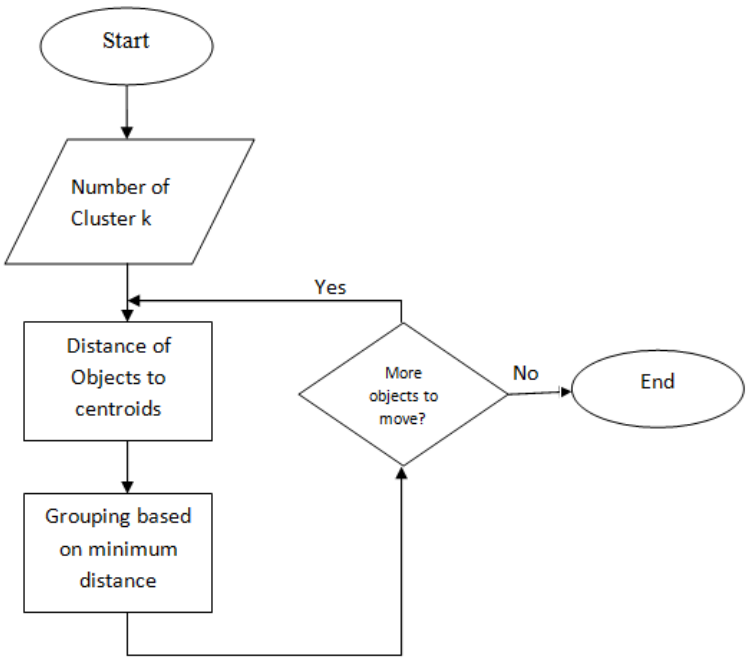

Figure 3: Flowchart of k-means clustering algorithm.

Figure 4 shows the subplots after segmenting the clustered image. Each subplot represents four clusters.

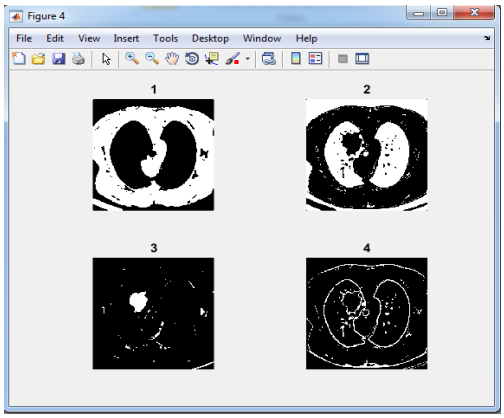

Figure 4: Output image of k-means.

D. Feature Extraction And Area Calculation.

In feature extraction phase, the features which are required for further phases are extracted. Here pixel values are needed to calculate the area, so the pixel values are extracted, then the area of the lung affected region is calculated and the same is displayed in command window along with number of the subplot to which the affected region belong. The lung affected region is shown in figure 5.

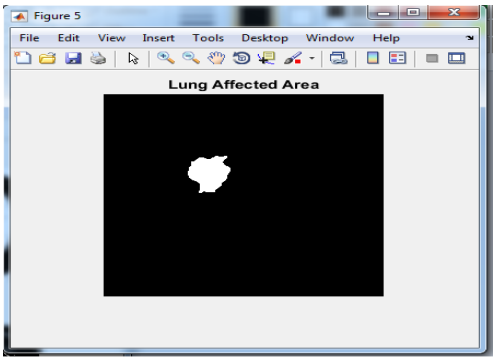

Figure 5: The stroke affected region.

Mathematical Representation:

$$\text{Image, } I=\sum_{w=0}^{511}\sum_{h=0}^{511}[f(0)+f(1)]$$

Pixel=W x H =512 x 512

W= Width h=Height

$f(0)$ = white pixel $f(1)$ = black pixel

$$\sum_{w=0}^{511}\sum_{h=0}^{511}$$

Number of white pixels $P=\Sigma w=0\Sigma h=0\ [f\ (0)]$

Where,

1 pixel = 0.264mm

P = Number of white pixels

Area Calculation of stroke region:

$$S=[(\sqrt{P})*0.264]\ mm^2$$

Where,

S = Stroke size

IV. RESULTS AND DISCUSSION

The acquired input image is fed to the system. That image is pre-processed, and k-means algorithm is used. In this method, first the preprocessed image is clustered into k (predefined value) clusters. Later image segmentation takes place according to the clusters which are done in previous step and segmented images are displayed in subplots. The subplot number in which lung affected region is present is displayed in command window along with the area of that lung affected region which is the final output as shown in below figure 6.

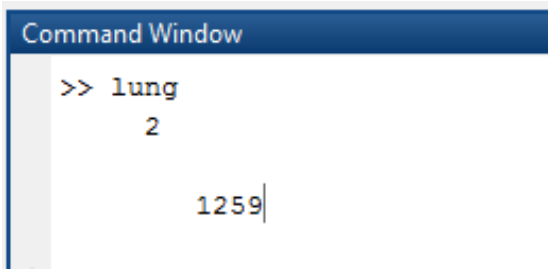

Figure 6: This image shows the final output displayed in command window.

The numbers of input images given for testing are 42. Among them area of 38 images are calculated correctly and 4 images shows the wrong output. So the accuracy of the project is 90.47% which is shown in Table 1.

Table 1: This table shows the number of input images, correct detection and wrong detection.

Images	Total number of input images	Correct Detection	Wrong Detection
Lung CT Images	42	38	4

V. CONCLUSION

The method used in this paper, detects the lung affected region and the area of that affected region is calculated. The input image is pre-processed and then K- Means clustering algorithm, which is unsupervised learning method, is used to cluster the input image and segment the affected region. The area of the affected region is then calculated and displayed on the command window. The accuracy of the project is 90.47%.

REFERENCES

[1] Prionjit Sarker1, Md. Maruf Hossain Shuvo2, Zakir Hossain3, and Sabbir Hasan4. "Segmentation and Classification of Lung Tumor 3D CT Image using K-means Clustering Algorithm" 4th International Conference on Advances in Electrical Engineering (ICAEE)28-30,September-2017,Dhaka, Bangladeshakshmi.

[2] R. Krishnapuram, J. M. Keller (1993), "A possibilistic approach to clustering" IEEE Transactionson Fuzzy SystemsI..

[3] Tse-Wei Chen , Yi-Ling Chen , Shao-Yi Chien (2010), "Fast Image Segmentation Based on K-Means Clustering with Histograms in HSV Color Space" Journal of Scientific Research ISSN I452-2I6X Vol. 44 No.2, pp.337-35I.

[4] Anil Z Chitade (2010j , " Colour based image segmentation using kmeans clustering," International Journal of Engineering Science and Technology Vol. 2(10), 5319-5325.

[5] Shengdong Nie, Lihong Li, Yuanjun Wang,."A Segmentation Method for Sub-solid Pulmonary Nodules Based on Fuzzy C-means Clustering" 2012 5th International Conference on BioMedical Engineering and Informatics (BMEI 2012)

[6] Kan Chen, Bin Li, Lian-Fang Tian, Jing Zhang "Segmentation of Pulmonary Nodules Using Fuzzy Clustering Based on Coefficient of Curvature" 2013 Seventh International Conference on Image and Graphics

[7] Nameirakpam Dhanachandra, Khumanthem Manglem and Yambem Jina Chanu, "Image Segmentation using K-means Clustering Algorithm and Subtractive Clustering Algorithm", Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015).