



# Fast and Adaptive Detection of Pulmonary Nodules in Thoracic CT Images Using a Contextual Clustering Based Region Growing

Bhakkialakshmi R<sup>1\*</sup>, Ponnammal P<sup>2</sup>, Srilekha M.K<sup>3</sup>, Abhishikt Sai.K.<sup>4</sup>

<sup>1,2,3</sup>Assistant Professor, <sup>4</sup>B.Tech Student,

SRM Institute of Science and Technology, Chennai, Tamil Nadu, India

\*Corresponding Author E-mail: [bhakkialakshmi.r@ktr.srmuniv.ac.in](mailto:bhakkialakshmi.r@ktr.srmuniv.ac.in)

## Abstract

For segmenting the Region of interest and for analyzing each area separately to locate whether pathologies present in it or not, we use segmentation process as the first step to diagnose lung image using ComputerAided Diagnosis. In this paper, ROI is segmented by using supervised Contextual Clustering in addition to the Region growing algorithm. Accurate segmentation of the lungs from the chest volume is obtained from the Contextual clustering which is better than all other thresholding approaches that are simple. Initial Nodule Candidates can be detected and segmented effectively by contextual clustering which is considered to be the most effective approach when compared to the remaining approaches present. We combine rule-based filtering and a feature based support vector machine using which we can reduce the False-positives (FP), custom CNN, Alex net, neuro-fuzzy classifier.

**Keywords:** Contextual Clustering, Pulmonary Lung Image, Initial Nodule Candidates support vector machine classifier, False-positive (FP) reduction.

## 1. Introduction

The representation of objects present in image is done by partitioning it in to multiple segments by segmentation method. [9] For the location of objects, boundaries, lines, curves and for assigning a label to each pixel in an image, segmentation process is used such that certain visual characteristics are shared by pixels with the same label.

The various regions obtained after image segmentation are used for further analysis and interpretations. Hence, image segmentation involves extracting important features and to derive the relevant metrics to segregate regions of homogeneous intensities. A selective region of interest should be chosen to do segmentation by taking the application requirements into consideration. However many image segmentation methods were proposed by various researchers to perform successive image analysis. Moreover, many researchers have used the existing thresholding techniques to segment various regions of interest. The most frequently used techniques used for segmentation are statistical, geometrical, structural, model-based, signal processing, spatial domain filters, Fourier domain filtering, Gabor and wavelet models [8].

A method in which it combines CC and region growing is used to segment the CT lung images effectively The main objective of our work is to increase the efficiency of detection with less computational cost than the previous techniques. The hierarchical INCs detection method proposed here is fast and uses contextual clustering and rule-based filtering. SVM classifier is further used to reduce false positives in the detected INCs.

## 2. Contextual Clustering with Region Growing

The contextual clustering is an algorithm uses an overlapping window of 5x5 pixels for the segmented image. The Threshold value for the segmented image is  $T=140$  and control parameter  $\beta$  ranges from 0 to 1 [1,2]. All the steps of contextual clustering algorithm is adopted to segment the lung region from CT lung image .

In images, Region growing is a method used in the identification of connected regions of interest that makes use of the threshold values with respect to a discrete connectivity. The region growing is getting started by selecting an initial pixel which is considered to be the seed point that is there in the initial region to be grown. Then the other pixels are included as neighbors in the region based on the threshold value. Each included pixel now becomes the seed point for further iterations. The above process continues until all pixels got added in a growing region based on the set of rules and the threshold value. For setting the threshold value in segmenting the region of interest that are present in the CT lung image, the combination of region growing and clustering is used. The implementation of region growing based contextual clustering is done as follows [1,2]:

1. Read image.
2. Sort block of pixels vales and generate pattern.
3. Find Pattern Median.
4. Find the count of values greater than the Median Value.
5. Using equation 1 calculate CC and assign it to values segmented
6. Compare threshold of Region growing with segmented values.

The contextual value  $V_{cc}$  is calculated as below [1]

$$V_{cc} = \text{Median value} + \frac{\beta}{\text{threshold}} \times \left\{ u - \frac{\text{Window size}}{2} \right\} \quad (1)$$

Input lung image was taken from the LIDC-Lung image database consortium which is further processed for the segmentation. In the contextual clustering splitting of the entire image is done as 5x5 windows on the basis of the region-based growing method which is an iterative method.

### 3. Border Corrected Mask and Segmentation

Spatial filtering is also called masking which are done by either a kernel or convolution matrix or a mask for blurring, sharpening, embossing, and edge-detection. This is done by convolving a kernel and an image. The mask is created to find the exact operations in an image so that the problems or features to be found in an image can be identified. The border corrected mask is one in which the edges are closed to find all the features of an image.

### 4. Rule Based Filtering

This process was done for Initial Nodule Candidates which are got in every initial level and also for the other Initial Nodule Candidates in which no opening operation is done in preserving small solitary nodules. Finally obtained Initial Nodule Candidates after doing filtering operations were formed by a logical union operation of the filtered INCs Initial Nodule Candidates at all levels.

### 5. Classifiers

In this work, the comparison of four main classifiers namely SVM, Alex Net, CNN (Convolution neural network) and neuro-fuzzy classifiers are done using MATLAB. A dataset was derived from the LIDC-IDRI dataset (lid (2015)). To reduce the further False Positives SVM classifier is used to carry out supervised learning strategy. The feature-based SVM classifier is relying on a series of features that are got from each and every Initial Nodule candidates that are obtained after performing rule-based filtering operations. Features are extracted according to four categories like geometry or shape, intensity, gradient features and Hessian eigen value features. Extraction of two dimensional and three-dimensional features is done are for the first three categories. The difference in the image resolution is taken for the calculation of all related features [3, 4]. Figure 1 and 2 shows the CC based region growing segmentation result and Initial Nodule candidates after rule-based filtering respectively.

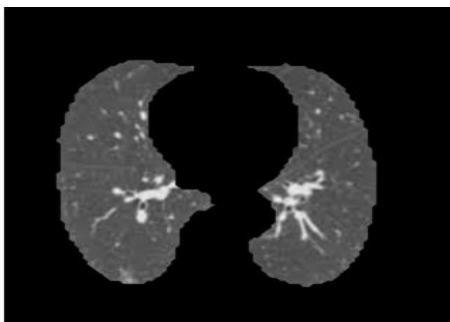


Fig.1: Segmented image

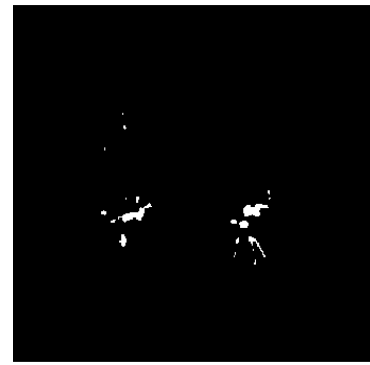


Fig. 2: Initial Nodule Candidates

### Custom Convolutional Neural Networks (CNN)

A CNN based segmentation method based on Convolutional Neural Networks(CNN) was adapted from deep CNN method for lung image segmentation( Pereira et al. (2016); Simonyan & Zisserman (2014)). This design uses multiple layers of 5X5 filters to give a deeper architecture and reduces over-fitting. It is found that using several of these convolutional layers decreases the False Positive Rate significantly.[5]

#### AlexNet

AlexNet is one of the most famous of the CNN architectures(Krizhevsky et al. (2012)). It has a multi-layered architecture that consists of five convolutional layers, with three max-pooling layers, and three fully connected layers. This highly flexible network can identify 1000 different objects within images. The network can also have certain layers retrained in order to identify objects within images, not from the original training set.[5]

#### Neuro-fuzzy

It is the combination of artificial neural networks and fuzzy logic. In this classifier, for classification, the feature vector is fed to a hybrid classifier based on neural network and fuzzy known as neuro-fuzzy. It contains two sub networks i.e. fuzzy self-organizing network and Multilayer Perceptron (MLP) which are connected together.[6,7]For the test sample classification, a fuzzy layer takes feature vector as input for generation of pre-classification vector which in turn is given to MLP. Fuzzy self -organizing layer is used for grouping and clustering of nodule pixels in regions based on their membership. This layer would help in differentiating in nodule and non nodule clusters[6,8]. By performing different experiments, it is found that the proposed fuzzy neural network is more tolerant to the noisy data related to nodules. Sensitivity is calculated by using TP and FN of classifiers using the formula [10]

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

Where TP represents the number of true positives and FN number of false negatives

Table 1: Comparison of proposed method results

Classification technique	Accuracy	Sensitivity	Running Time (seconds)
SVM	93.33%	98%	9.2
CNN	99.51%	99.4%	6.1
Alex Net	99.56%	99.5%	6.3
Neuro fuzzy	96%	96.5%	11

The accuracy of different classifier techniques like SVM, CNN, AlexNet and Neuro fuzzy compared in Table 1. , in which CNN and AlexNet shows better accuracy for CT images in 5X5 pixel.

## 6. Conclusion

In this work, lung segmentation is done and lung nodules are detected by using the classifiers like SVM, CNN, AlexNet and neuro fuzzy for 5x5 pixels. When compared to existing classifiers the accuracy of SVM, CNN, AlexNet and neuro fuzzy classifiers have improved and the values are 93.33%, 99.51%, 99.56% and 96% respectively. While analyzing SVM and neuro fuzzy classifier the accuracies are not as much as good compared to CNN and AlexNet. A strong classifier has good accuracy, short running time and low false positive. AlexNet does better than CNN and it shows high false positive, Compared to CNN though the accuracy and sensitivity is very good. CNN is the stronger classifier among all even though accuracy is less compared to Alexnet, due to low false positive and short running time.

## References

- [1] S.Santhosh Baboo, E.Iyyapparaj ,”Analysis and classification Methods for Diagnosis of Pulmonary Nodules in CT Images”, International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST) Vol.3 Issue.5 May 2017.
- [2] A.PRABIN , DR. J.VEERAPPAN, “Automatic Segmentation Of Lung Ct Images By Cc Based Region Growing”, Journal of Theoretical and Applied Information Technology, 10th October 2014. Vol. 68 No.1
- [3] Aydin Kaya, Ahmet Burak Can, “A weighted rule based method for predicting malignancy of pulmonary nodules by nodule characteristics”, Journal of Biomedical Informatics, Volume 56, August 2015,Pages 69-79.
- [4] Wook-Jin Choi and Tae-Sun Choi, “Automated Pulmonary Nodule Detection System in Computed Tomography Images: A Hierarchical Block Classification Approach”, Entropy 2013, 15, 507-523
- [5] Anonymous authors, “Lung Tumor Location And Identification With Alexnet And A Custom CNN”, Under review as a conference paper at ICLR 2018.
- [6] Rakesh Kumar Khare G. R. Sinha Sushil Kumar, “Cancer Detection Using Neuro Fuzzy Classifier in CT Images”, International Journal on Future Revolution in Computer Science & Communication Engineering, Volume: 3 Issue: 12, December 2017.
- [7] Anam Tariq , M. Usman Akram “Lung Nodule Detection In Ct Images Using Neuro Fuzzy Classifier” TELKOMNIKA Vol. 11, No. 2, June 2013: 331 – 336.
- [8] Hamid bagherieh,Atiyeh Hashemi, Abdol Hamid Pilevar, “Mass Detection in Lung CT Images using Region Growing Segmentation and Decision Making based on Fuzzy Systems”, I.J. Image, Graphics and Signal Processing, 2014, 1, 1-8
- [9] S.L.A. Lee, A.Z. Kouzani, and G. Nasierding, E.J. Hu, “Pulmonary Nodule Classification Aided by Clustering”, Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, USA - October 2009.
- [10] Hao Han, Lihong Li, “Fast and Adaptive Detection of Pulmonary Nodules in Thoracic CT Images Using a Hierarchical Vector Quantization Scheme” IEEE J Biomed Health Inform. 2015 Mar; 19(2): 648–659.
- [11] T. Padmapriya and V. Saminadan, “Improving Throughput for Downlink Multi user MIMO-LTE Advanced Networks using SINR approximation and Hierarchical CSI feedback”, International Journal of Mobile Design Network and Innovation- Inderscience Publisher, ISSN : 1744-2850 vol. 6, no.1, pp. 14-23, May 20 15.
- [12] S.V.Manikanthan and T.Padmapriya “Recent Trends In M2m Communications In 4g Networks And Evolution Towards 5g”, International Journal of Pure and Applied Mathematics, ISSN NO: 1314-3395, Vol-115, Issue -8, Sep 2017.