

An Implementation of Lung Cancer Detection and Classification System from Medical Images using Matlab

¹Swati Sendre, ²Shailendra Kumar Singh

¹MTech Scholar, ²Asst. Professor

¹Eleetronics & Telecommunication, ²Eleetronics & Telecommunication

¹GD-Rungta College of Engineering & Technology, Bhilai, India

²GD-Rungta College of Engineering & Technology, Bhilai, India

Abstract: In previous years, image processing had covered a large area over biomedical applications in diagnosis of several of diseases in medical images. Lung Cancer detection and classification is one of the most widely used applications by many researchers. In review paper we already described various techniques and algorithm. This study of this paper is aims to highlight the significance of data analytics and machine learning in prognosis in health sciences, particularly in detecting life threatening and terminal diseases like cancer. Here, we consider lung cancer for our study. For this purpose, preexisting lung cancer patient's data are collected to get the desired results. This paper presents the better Computer Aided Diagnosing (CAD) system for automatic detection of lung cancer. The initial process is lung region detection by applying basic image processing techniques such as Bit-Plane Slicing, Erosion, Median Filter, Dilation, Outlining, Lung Border Extraction and Flood-Fill algorithms to the CT scan images. After the lung region is detected, the segmentation is carried out with the help of Mean Shift clustering algorithm. With these, the features are extracted and the diagnosis rules are generated. These rules are then used for learning with the help of Random Forest.

Keywords – Lung cancer detection, CT scan images, Mean Shift clustering algorithm, Computer Aided Diagnosing, Lung Border Extraction.

I. INTRODUCTION

The term cancer is a technical term refers to uncontrolled cell growth in tissues leading to malfunctioning of body organs at extreme state of influence cause major suffering and even death. Cancer leads to excessive multiplication of abnormal cells without control and are able to affect other tissues. Cancer cells infect neighboring part of human body through connective tissue. Most of the cancers are named for the organ where cancer starts – rapid growth of abnormal cancer cells in the colon is called colon cancer; that of skin is called basal carcinoma.

Lung cancer is a disease of abnormal cells multiplying and growing into a tumor. Cancer cells can be carried away from the lungs in blood, or lymph fluid that surrounds lung tissue. It is the second most common disease in men and women. When symptoms are seen in person, lung cancer tests are done to determine the type of lung cancer and if it has spread. Lymph flows through lymphatic vessels, which drain into lymph nodes located in the lungs and in the center of the chest. Lung cancer often spreads toward the center of the chest because the natural flow of lymph out of the lungs is toward the center of the chest. Metastasis occurs when a cancer cell leaves the site where it began and moves into a lymph node or to another part of the body through the blood stream [1]. Cancer that starts in the lung is called primary lung cancer. There are several different types of lung cancer, and these are divided into two main groups: Small cell lung cancer and non-small cell lung cancer which has three subtypes: Carcinoma, adenocarcinoma and Squamous cell carcinomas.

Figure 1 shows a general description of lung cancer detection system that contains four basic stages. The first stage starts with taking a collection of CT images (normal and abnormal) from the available Database from IMBA Home (VIA-ELCAP Public Access). The second stage applies several techniques of image enhancement, to get best level of quality and clearness. The third stage applies image segmentation algorithms which play an effective rule in image processing stages, and the fourth stage obtains the general features from enhanced segmented image which gives indicators of normality or abnormality of images.

Objective

We presented a novel method to classify a given CT scan images as normal or abnormal. The aim of this research was to detect features for accurate images comparison as pixels percentage and mask-labelling.

II. PROBLEM IDENTIFICATION

Lung cancer is the second most common cancer in both men and women (not counting skin cancer), and is by far the leading cause of cancer death among both men and women. Each year, more people die of lung cancer than of colon, breast, and prostate cancers combined. Most lung cancers could be prevented, because they are related to smoking (or secondhand smoke), or less often to exposure to radon or other environmental factors. But some lung cancers occur in people without any known risk factors for the disease. It is not yet clear if these cancers can be prevented. Most lung cancers have already spread widely and are at an advanced stage when they are first found. These cancers are very hard to cure. But in recent years, doctors have found a test that can be used

to screen for lung cancer in people at high risk of the disease. This test can help find some of these cancers early, which can lower the risk of dying from this disease.

III. SYSTEM ALGORITHM

As in the figure 1 we have some of the steps to detect Lung Cancer

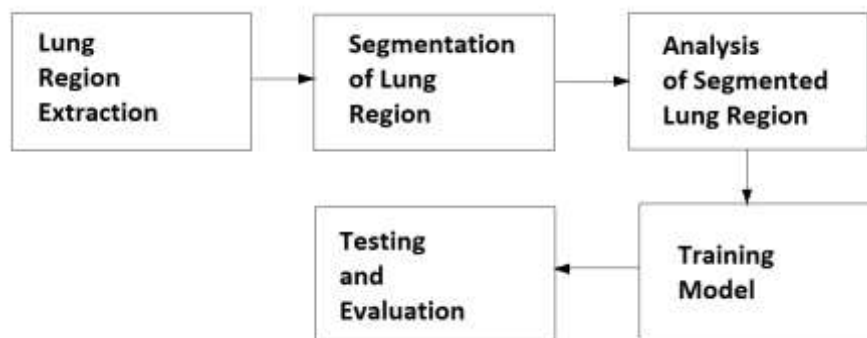


Figure 1: Lung Cancer Detection System

CT images acquisition

We load the CT image data which is store in our system. We collect the number of sample data from the radiologist and test this data in our system. We collect the high resolution structural and functional CT image data in normal and abnormal patient with different types of brain tumors.

Lung region Extraction

The goal of the lung extraction step is to separate the voxels corresponding to lung tissue from those belonging to the surrounding anatomical structures. Since the gray level of the background is close to the gray level of the lung tissues, it is advantageous to remove the background (all pixels outside the chest cavity) from each slice before doing extraction of the lung regions, in order to increase the accuracy of segmentation. Background is simply removed as follows: starting from the edges of the CT slices, all pixels that have similar gray levels, within a certain range, are removed. This process is applied, on each slice, from the left to right comers (i.e., horizontally or row-wise) as well as vertically (column-wise), in order to cover the whole slice. This simple approach has been shown to be quite adequate for removal of the background pixels outside the chest cavity region. After removing the background from each slice, we assume that each CT slice (image) contains only two types of pixel: 1) Lungs - pixels within very dense regions in the CT scan. These pixels have an average gray level g_b ; and 2) other tissues organs – low density pixels within the lungs and the surrounding regions (e.g., ribs, heart, liver, and other parts in the chest cavity).

Detection of Abnormal Tissue in Lung Cancer

We assume that the segmented lung volume is formed of two types of voxels: normal, which describes the healthy lung tissue, and abnormal, which describes all abnormalities in the lungs, the bronchi and bronchioles. The first step to detect the abnormalities is to remove the normal tissues. The gray level histogram shows the distribution of the intensity of lung, which is usually a homogenous region. Lung abnormalities show in the histogram of the CT slices as bright (or homogeneous) areas with distinct gray levels from the surrounding lung tissues. Therefore, it is logical to assume that the abnormalities will correspond to the peaks (usually one) in the gray level histogram. Therefore, as a first step in isolating the abnormities, we threshold all pixels having a gray level below the gray level value of the peak of the histogram. These pixels will correspond to normal lung tissues.

The abnormalities, in general, appear as elliptical-shaped regions in the lungs. These elliptical shapes can be isolated by the analysis of the distribution of the gradient maxima in the neighborhood of each pixel. Ellipses and rings are patterns symmetrical relatively to their centroids. In the case of the polar co-ordinates, the expression of the intensity and spatial distribution of the most significant edges has characteristics properties of symmetry and uniformity. In the case of CT images, the gradient should show local maxima at the point on the border of the bronchi, bronchioles, and abnormal tissues.

Training Model using Random Forest algorithm

The random forest starts with a standard machine learning technique called ‘Decision Tree’ which, in ensemble terms, corresponds to the weak learners. In a decision tree, an input is entered at the top and as it traverses down the tree the data gets bucketed into smaller and smaller sets. The random forest takes this notion to the next level by combining trees with the notion of ensemble. Thus, the ensemble terms, the trees are weak learners and random forest is a strong learner.

Random forest runtimes are quite fast, and they are able to deal with unbalanced and missing data. Random Forest weaknesses are that when used for regression they cannot predict beyond the range in the training data, and that they may over-fit data sets that are particularly noisy.

Testing and Evaluation

(i) Cross-validation

We test our classifier using a technique called “cross-validation”: train the classifier on all projects except for one. Here the projects mean the partition of the data set. Of course, we also know the ground truth for this held-out project, so we can see how well the classifier does on it, without cheating by training the classifier on the hold-out.

(ii) Mean Precision

The effectiveness of the classifier is the distance between the two means, which does not vary as threshold changes. One way of measuring the effectiveness of the classifier is the “precision”. Precision is the number of truly correct items (“hits”) divided by the number of items that the classifier says are correct (hits + false alarms). “Mean precision” takes into account the issues with choosing a threshold, noted above, by performing this calculation at a range of thresholds and taking the mean.

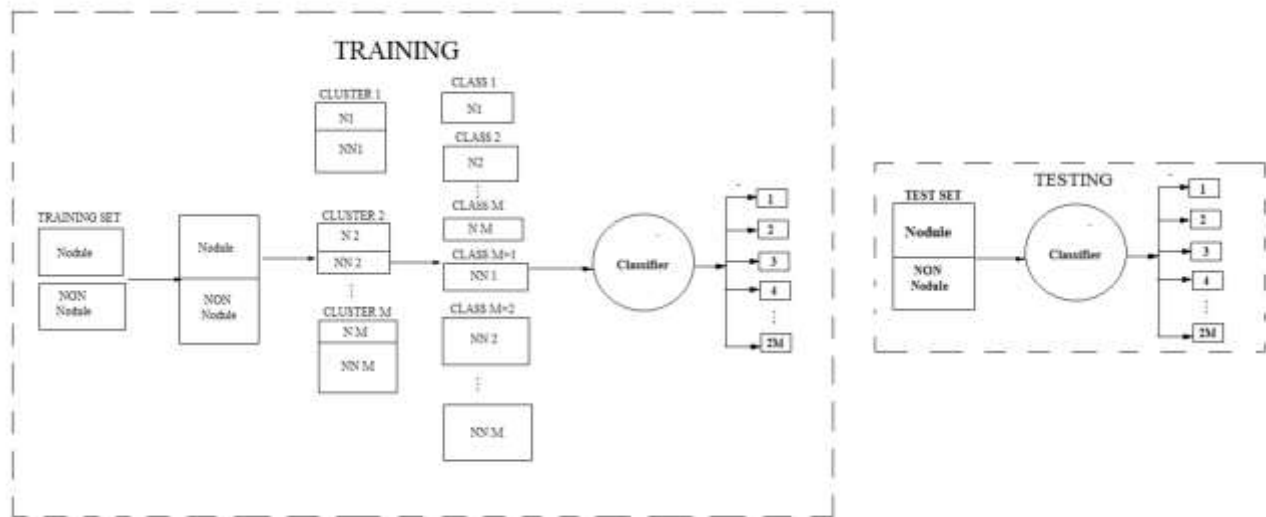


Figure 2: Architecture of the proposed classification aided by clustering method using Random Forest Algorithm.

The random forests [21] refer to a collection of tree structures classifiers $\{h(x, \theta_k), k=1\}$ where $\{\theta_k\}$ are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input x . For each observation, each individual tree votes for one class and the forests predicts the class that has most votes. The user has to specify the number of randomly selected variables m try to be searched through for the best split at each node. The node is split using the best among a subset of predictors randomly chosen at that node. The largest tree possible is grown and is unpruned. The root node of each tree in the forest contains a bootstrap sample from the original data as the training set. The observations that are not in the training set are referred to as “out-of-bag” observations. The out-of-bag instances for classifier $h_k(x)$ are represented as $O_k(x)$. $Q(x, y_j)$ is voted as the out-of-bag proportion for class y_j based on input x and the approximation of

$$P(h(x)=y_j):$$

$$Q_{(x,y_j)} = \frac{\sum_{k=1}^K I(h_k(x) = y_j; (x, y) \in O_k)}{\sum_{k=1}^K I(h_k(x); (x, y) \in O_k)}$$

Where $I(-)$ is the indicator function.

The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them.

Since an individual tree is unpruned, the terminal nodes can contain only a small number of observations. The training data are run down each tree. If observations i and j both end up in the same terminal node, the similarity between i and j is increased by one. At the end of the forest construction, the similarities are symmetrized and divided by the number of trees. The similarity between an observation and itself is set to one. The similarities between objects form a matrix which is symmetric, and each entry lies in the unit interval $[0, 1]$.

IV. TEST AND IMPLEMENTATION

Data Pre-processing

For data pre-processing, after reading the dcm format lung scan image into Matlab I count the dark intensity in the specific lung region and compute a percentage threshold. With the threshold, I can filter a number of useless CT slices and finally pick 10 slices for each patient. As Figure 3 shown, we do not want lung scans like CT3 but prefer lung scans link CT1, CT2 and CT3. After shrinking the data set, I obtain a smaller data set that can be processed in the next stage.

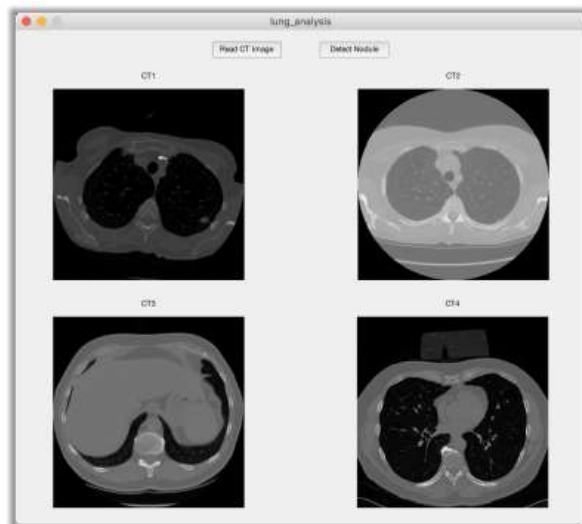


Figure 3: Some scan slices are useless

*Detection Algorithm
Lung Region Extraction*

The initial stage of the proposed Computer Aided Diagnosing (CAD) [2, 7] techniques is the extraction of lung region from the CT scan image. The basic image processing techniques are utilized for this purpose. The methods and steps involved in the extraction of lung region from CT image are shown in figure 3. The image processing techniques applied in the proposed technique are Bit-Plane Slicing, Erosion, Median Filter, Dilation, Outlining, Lung Border Extraction and Flood-Fill algorithms. Usually, the CT chest image not only contains the lung region, it also contains background, heart, liver and other organs areas. The main aim of this lung region extraction process is to detect the lung region and regions of interest (ROIs) from the CT scan image.

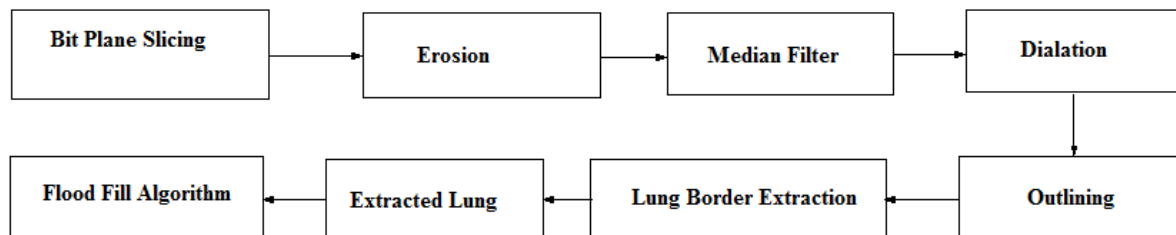


Figure 4: Lung Region Extraction Process

Bit-Plane Slicing

The databases comprise a set of 2D-CT images of normal and abnormal subjects. Each image is stored as a double subscripted array of type unsigned integer (8 bits) and of size 512 x 512 pixels, and each element of the array contains the CT-intensity value of a single pixel in a plane resolution 0.78 x 0.78, slice thickness 2 mm. the 512x512 pixel slices were acquired in the upper, middle, and lower part of the lung of each subject. Figure 4 shows sample of the acquired data from the same subject at different levels from top to bottom. The extraction of lung region starts by applying the bit-plane slicing algorithm to the slimed data from step one. The resulted binary slices are then analyze to choose from among then, the best bit-plane image that may help in extracting automatically the lung region from raw data with a certain degree of accuracy and sharpness.

Lung Border Extraction

The lung border extraction is based on the result of applying outlining algorithm images. In Matlab, I use **bwboundaries** function to outline the objects in the input images. The **bwboundaries** function implements the Moore-Neighbor tracing algorithm modified by Jacob's stopping criteria. This function is based on the boundaries function presented in the first edition of Digital Image Processing Using MATLAB.

As Figure 5 shown, I need to delete useless outlines and finally get a result like the right one. Based on the return value of **bwboundaries** function, I know the number of points of each object in the outlining image. Thus I can set a threshold to filter the small circles region. Plus, I can filter the outer ring encircling the two lungs by designed detecting algorithm.

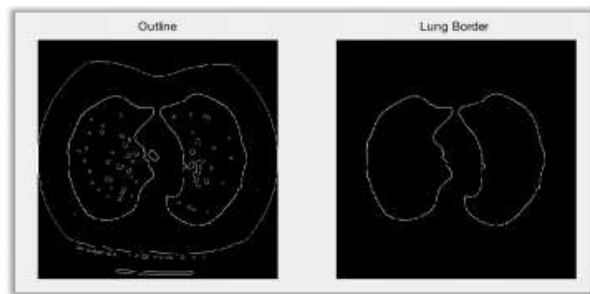


Figure 5: The result of lung border extraction

Flood Fill algorithm and extracted lung

Finally, flood fill algorithm is applied to fill the obtained lung border with the lung region. After applying these algorithms, the lung region is extracted from the CT scan image. This obtained lung region is further used for segmentation in order to detect the cancer nodule.

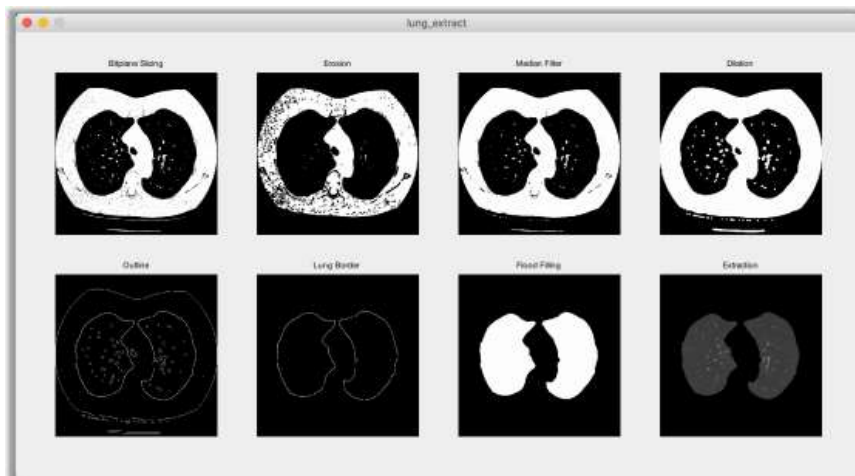


Figure 6: Lung region extraction algorithm

Segmentation of Lung Region

After the lung region is detected, the next process is segmentation of lung region in order to find the cancer nodules. This step will identify the region of interest (ROIs) which helps in determining the cancer region. In this project, Mean Shift is implemented for segmentation.

Intuitive Idea

Mean shift considers feature space as an empirical probability density function. If the input is a set of points then Mean shift considers them as sampled from the underlying probability density function. If dense regions (or clusters) are present in the feature space, then they correspond to the mode (or local maxima) of the probability density function. We can also identify clusters associated with the given mode using Mean Shift. For each data point, Mean shift associates it with the nearby peak of the dataset's probability density function. For each data point, mean shift defines a window around it and computes the mean of the data point. Then it shifts the center of the window to the mean and repeats the algorithm till it converges. After each iteration, we can consider that the window shifts to a denser region of the dataset.

At the high level, we can specify Mean Shift as follows:

1. Fix a window around each data point.
2. Compute the mean of data within the window.
3. Shift the window to the mean and repeat till convergence.

Kernel Density Estimation

Kernel density estimation is a non-parametric way to estimate the density function of a random variable. This is usually called as the Parzen window technique. Given a kernel K , bandwidth parameter h , Kernel density estimator for a given set of d -dimensional points is

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)$$

Generate nodule candidates

With the input of lung region extraction images, Mean Shift is applied to cluster nodule candidates and the result is shown in Figure 7. After the segmentation is performed to the lung region, the feature extraction and cancer diagnosis can be performed with the segmented image.

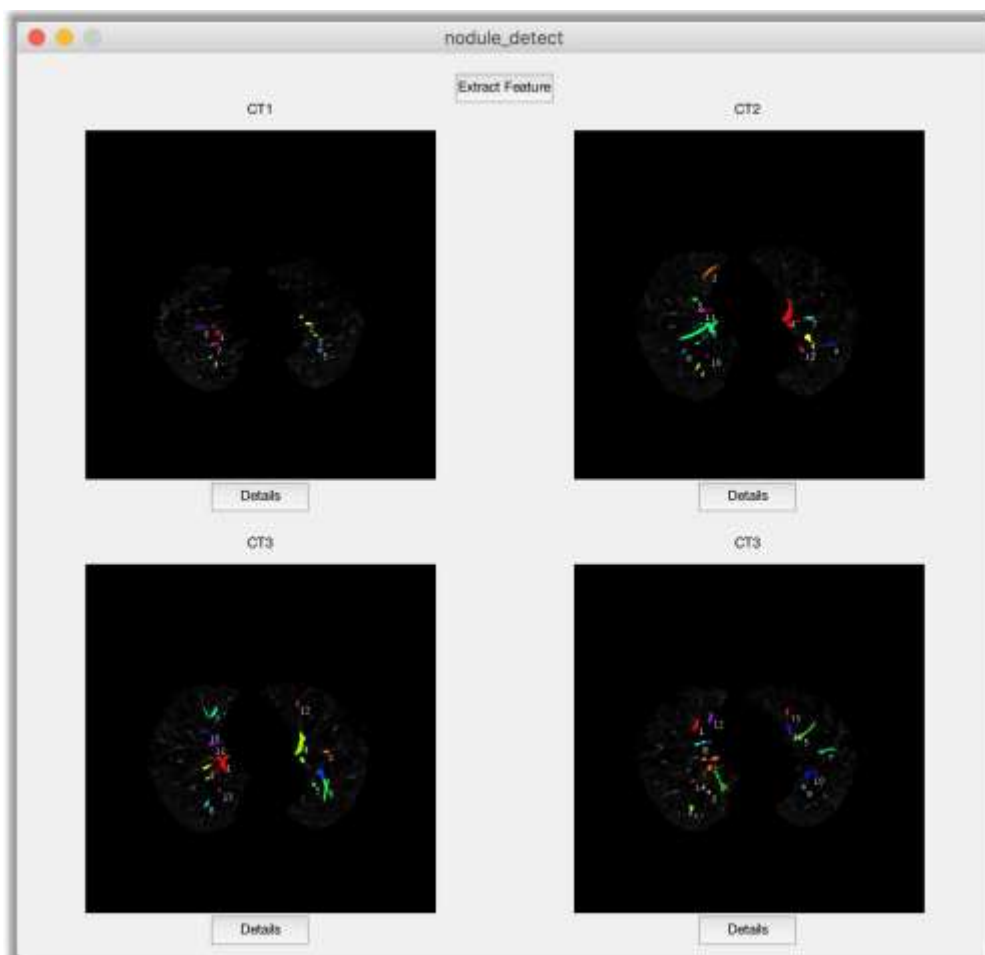


Figure 7: Nodule candidates generated result

As Figure 7 shown, each lung region extraction image can generate from ten to thirty nodule candidates. Consequently, each patient generates more than one hundred nodules candidates, which they needed to be filter since not all of them are useful.

Analysis of Lung Region

After the segmentation is performed to the lung region, the features can be obtained from it and the diagnosis can be designed to exactly detect the cancer nodules in the lungs. The diagnosis rules can eliminate the false detection of cancer nodules resulted in segmentation and provides better diagnosis.

Feature Extraction

The features that are used in this project in order to generate diagnosis rules are:

- Area of the candidate region
- The Maximum Drawable Circle (MDC) inside the candidate region
- Mean intensity percentage of candidate region

Area of the candidate region

- This feature can be used here in order to
- Eliminate isolated pixels
- Eliminate very small candidate object

With the help of this feature, the detected region that do not have the chance to form cancer nodule are detected and can be eliminated. This helps in reducing the processing in the further steps and also reducing times taken by further steps.

Maximum Drawable Circle (MDC)

This feature is used to indicate the candidate regions with its maximum drawable circle (MDC). All the pixels inside the candidate region is considered as center point for drawing the circle. The obtained circle within the region is taken for consideration. Initially radius of the circle is chosen as one pixel and then the radius is incremented by one pixel every time until no circle can be drawn with that radius. Maximum drawable circle helps in the diagnostic procedure to remove more and more false positive cancerous candidates.

Mean intensity percentage of the candidate region

In this feature, the mean intensity percentage for the candidate region is calculated which helps in rejecting the further regions which does not indicate cancer nodule. The mean intensity percentage indicates the average intensity percentage of all the pixels that belong to the same region and is calculated using the formula:

$$\text{MeanPtg}(j) = \frac{\sum_{i=1}^n \text{Intensity}(i)/n}{\text{Max}}$$

where j characterizes the region index and range from the number of candidate regions in the whole image, which we specify as N. Intensity (i) indicates the CT intensity value of pixel i, and i ranges from 1 to n, where n is the total number of pixels belonging to region j. Max indicates the maximum intensity values of all of the candidate regions.

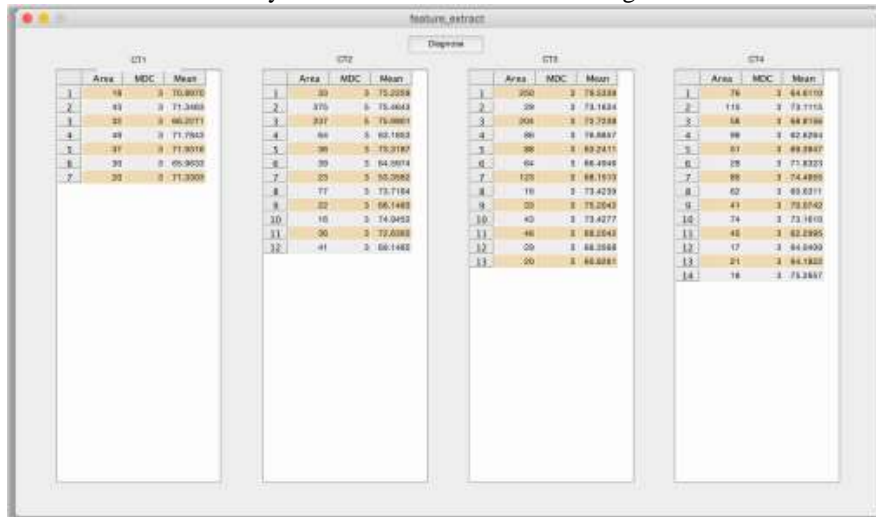


Figure 8: Feature extraction result

Formulation of Diagnosis rules

After the necessary features are extracted, the following diagnosis rules can be applied to detect the occurrence of cancer nodule. There are three rules which are involved are as follows:

Rule 1: Initially the threshold value T1 is set for area of region. If the area of candidate region is lower than the threshold value, then it is eliminated for further consideration. This rule will help in reducing the steps and time necessary for the upcoming steps.

Rule 2: In this rule maximum drawable circle (MDC) is considered. The threshold T2 is defined for value of maximum drawable circle (MDC). If the radius of the drawable circle for the candidate region is less than the threshold T2, then that region is considered as non-cancerous nodule and is eliminated for further consideration. Applying this rule has the effect of rejecting large number of vessels, which in general have a thin oblong, or line shape.

Rule 3: In this, the range of value T3 and T4 are set as threshold for the mean intensity percentage of candidate region. Then the mean intensity percentage for the candidate regions are calculated. If the mean intensity percentage of candidate region goes below minimum threshold or goes beyond maximum threshold, then that region is assumed as non-cancerous region.

By implementing all the above rules, the maximum of regions which does not considered as cancerous nodules are eliminated. The remaining candidate regions are considered as cancerous regions. This CAD system helps in neglecting all the false positive cancer regions and helps in detecting the cancer regions more accurately. These rules can be passed to the Random Forest Classifier in order to detect the cancer nodules for the supplied lung image.

Training of Model

In this training section, I use Random Forest model to train and test the data.

Trees and Forests

The random forest starts with a standard machine learning technique called 'Decision Tree' which, in ensemble terms, corresponds to the weak learners. In a decision tree, an input is entered at the top and as it traverses down the tree the data gets bucketed into smaller and smaller sets. The random forest takes this notion to the next level by combining trees with the notion of ensemble. Thus, the ensemble terms, the trees are weak learners and random forest is a strong learner.

V. CONCLUSION

Since lung diseases are dynamic and evolutionary in nature, their detection, treatment will also progress based on the dynamic nature of the disease. The treatment detection technique in lung cancer keeps on changing with time. Therefore, the image processing technique must also progress in a direction of finding cancer as early as possible. The Mean shift clustering algorithms in cancer detection must be reliable, robust and must have high-level of diagnostic value. For this purpose, right from the image acquisition to the detection of cancer we must be very careful enough to identify the real indicators of cancer nodules. For this purpose, we have used algorithm inspired by the analogy of filling water in an uneven topology. In our case, also cancerous part of the lung becomes very high or a region having global maxima and non-cancerous part having less intensity as compared to cancerous part or simply the complete lung tissues having typical intensity

value with respect to their nature. This process is called Random Forest algorithm with distance transformation. This algorithm goes highly successful in segmenting or identifying the cancerous nodule after the removal of thorax or background of lungs.

VI. EXPECTED OUTCOME

This project presents the better Computer Aided Diagnosing (CAD) system for automatic detection of lung cancer. The initial process is lung region detection by applying basic image processing techniques such as Bit-Plane Slicing, Erosion, Median Filter, Dilation, Outlining, Lung Border Extraction and Flood-Fill algorithms to the CT scan images. After the lung region is detected, the segmentation is carried out with the help of Mean Shift clustering algorithm. With these, the features are extracted and the diagnosis rules are generated. These rules are then used for learning with the help of Random Forest. The experimental result shows that the proposed CAD system can able to tell the posterior probability of lung cancer for a patient based on the detection algorithm. Also the usage of Random Forest will increase the accuracy of detecting the cancer nodules.

VII. ACKNOWLEDGMENT

Expression of giving thanks is just a part of those feeling which are too large for words, but shall remain as memories of wonderful people with whom I have got the pleasure of working during the completion of this work. I am grateful to GD-Rungta College of engineering and technology, which helped me to complete my work by giving encouraging environment. I would like to express my deep and sincere gratitude to my supervisor, Assistant Professor, Mr. Shailendra Kumar Singh. His wide knowledge and his logical way of thinking have been of great value for me. His understanding, encouraging and personal guidance have provided a good basis for the present work.

REFERENCES

- [1] Ch. Hima Bindu , and K. Satya Prasad, "An Efficient Medical Image Segmentation Using Conventional OTSU Method," Int. Journal of Advanced Science and Technology, January 2012 Vol. 38.
- [2] B. Wu and Y. Yang, [2012], "Local-and global-statistics-based active contour model for image segmentation," Mathematical Problems in Engineering, vol. 2012.
- [3] M. Erdt, S. Steger, G. Sakas, [2012], "Regimentation: A New View of Image Segmentation and Registration," Journal of Radiation Oncology Informatics ISSN: 1663-618X, J Radiat On col Inform.
- [4] A. Nyma, M. Kang, Y. KeunKwon, C. HongKim, and J. MyonKim, [2012] "A Hybrid Technique for Medical Image Segmentation," Hindawi Publishing Corporation Journal of Biomedicine and Biotechnology, Article ID 830252, 7 pages doi:10.1155/2012/830252.
- [5] N. T. N. Anh, J. Cai, J. Zhang, and J. Zheng, [2012], "Constrained active contours for boundary refinement in interactive image segmentation," IEEE Int. Symposium on Circuits and Systems (ISCAS), pp. 870-873.
- [6] D. Kaushik, S. Utkarsha , P. Singha, and V. Singh, [2013]"Medical Image Segmentation using Genetic Algorithm," Int. Journal of Computer Applications (0975 – 8887) , November, Volume 81.
- [7] Ashfaq Ahmed K and Syed NaimatullahHussain [2013], "Comparative Prediction Performance with Support Vector Machine and Random Forest Classification Techniques", International Journal of Computer Applications, Volume 69– No.11, 0975 – 8887.
- [8] A. Norouzi1, M. Rahim, A. Altameem, T. Saba, E. Abdolvahab, A. Rehman and M. Uddin,[2014], "Medical Image Segmentation Methods, Algorithms, and Applications," IETE Technical review ,vol 31.
- [9] BingquanHuo, Guoxin Li, and Fengling Yin ,[2015] "Medical and Natural Image Segmentation Algorithm using M-F based Optimization Model and Modified Fuzzy Clustering: A Novel Approach," Int. Journal of Signal Processing, Image Processing and Pattern Recognition, Vol.8.
- [10] Christopher T and J.Jamerabanu [2015], "A Study on Mining Lung Cancer Data for Increasing or Decreasing Disease Prediction Value by Using Ant Colony Optimization Techniques", Proceedings of the UGC Sponsored National Conference on Advanced Networking and Applications, Special Issue Published in International Journal of Advanced Networking and Applications (IJANA).
- [11] AiswaryaIyer, et. al. [2015], "Diagnosis of Diabetes Using Classification Mining Techniques", International Journal of Data Mining & Knowledge Management Process (IJDMP) Vol.5, No.1.
- [12] Gawade Prathamesh Pratap, R.P. Chauhan [2016], "Detection of Lung cancer cells using image processing techniques", IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), ISBN: 978-1-4673-8587-9.
- [13] Ping Liu, Jianqi Sun, Yijing Guan[2016], "Detection of Lung Cancer with Phase-Contrast X-ray Imaging Using Synchrotron Radiation", Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE, ISBN: 1-4244-0032-5.
- [14] Durgalakshmi R and MannarMannan.J [2016], "Prognosis of Blood Carcinoma using Data Mining Techniques", International Journal of Contemporary Research in Computer Science and Technology (IJCRCT), Volume2, Issue 4, e-ISSN: 2395-5325.
- [15] S Avinash, K. Manjunath, S. Senthilkumar[2017], "Analysis and comparison of image enhancement techniques for the prediction of lung cancer", 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), ISBN: 978-1-5090-3704-9.
- [16] Pradeep Singh,Namrata Singh,[2017], "Intelligent approaches for prognosticating post-operative life expectancy in the lung cancer patients", IEEE International Conference on Inventive Computing and Informatics (ICICI), ISBN: 978-1-5386-4031-9.

- [17]Anjali Kulkarni, Anagha Panditrao [2016],"Classification of lung cancer stages on CT scan images using image processing",IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), ISBN: 978-1-4799-3914-5.
- [18]Lilik Anifah,Haryanto,Rina Harimurti ,[2017],"Cancer lungs detection on CT scan image using artificial neural network backpropagation based gray level cocurrence matrices feature", IEEE International Conference on Advanced Computer Science and Information Systems (ICACSIS), ISBN: 978-1-5386-3172-0.
- [19]Snehal Dabade, Shubhangi Chaudhari, Sneha Jadhav [2017],"A review paper on computer aided system for lung cancer detection ", IEEE International Conference on Big Data, IoT and Data Science, ISBN: 978-1-5090-6593-6.

