

# Research Paper on Computer Aided Diagnosis of Lung Cancerous Nodules in X-Ray Image using MATLAB

Mr. Sarvendra Pratap Singh  
M. Tech Scho.  
in Electronics & Communication,  
Department of ECE,  
SKITM, Bahadurgarh,  
Haryana, India.

Ms. Kirandeep  
Assistant professor  
Department of ECE, SKITM, Bahadurgarh,  
Haryana, India.

Mr. Sumit Dalal  
Assistant Professor  
Department of ECE, SKITM, Bahadurgarh,  
Haryana, India.

**Abstract:-** Computer aided diagnosis of cancerous nodules system plays an important role in early detection of cancerous nodules in lung x-rays image. It helps the radiologist in diagnosis of lung cancer. Most of the cancer related deaths are due to lung cancer. So more attention should be given to early detection of lung cancer for improve the survival rate of patients.

We know that cancerous and non- cancerous regions appears slightly differ in x- ray image. For accurate detection of cancerous nodules, it is necessary to differentiate the cancerous nodules from non-cancerous nodules. The artificial neural network is used in this thesis to differentiate them. Now a day, ANN is used in medical imaging for diagnosis the various diseases. They work in similar to the brain and have good decision making criteria when trained appropriately.

## INTRODUCTION:

Most of the cancer related deaths in the world are due to lung cancer. Lung Cancer is the most common cancer related death in both men and women. Early detection and treatment of lung cancer is important for survival of cancer patients. We know that, in a healthy person, cell in the lungs divide and produce at controlled rate to repair worn out or injured tissues and allow for normal growth. When cell inside the lungs developed at an uncontrollable rate it causes lung cancer. These abnormal tissues masses are called tumors. But all tumors are not cancerous (malignant) some are non-cancerous (benign).

## LITERATURE REVIEW:

Medical imaging plays an important role in the early detection and treatment of cancer. It provides physicians with information essential for efficient and effective diagnosis of various diseases. Diagnosis of X-rays can be used as an initial step in nodule detection. The objective of the CAD system developed in this thesis is to help radiologists to improve their accuracy in cancer detection. We developed an Artificial Neural Network (ANN) to differentiate the cancerous nodules from other suspected nodule areas in the X-ray images. To improve the diagnosis accuracy of this ANN we introduced a

pre-processing stage in the CAD system that involves various image enhancement techniques. Our main interest involved the detection of golf-ball tumor type of cancer. We can detect this type of cancer by the presence of round masses (tumors) in the X-ray image.

## Some statistics about Lung Cancer

In Europe in 2004 there were 17,11,000 cancer deaths. Lung cancer deaths were almost 20% of them (3,41,800), followed by colorectal (203700), stomach (137900) and breast (129900). In Italy there are about 35000 to 40000 new cases per year, and the incidence is increasing in women. Tobacco is considered responsible approximately for 85 percent of all cases, while radon exposure for another several percent. Heavy smokers have a much higher risk of dying of lung cancer, about 10 times that of non-smokers. Asbestos, pollution and genetic predisposition are other important risk factors. Only about 10-15% of all people who develop lung cancer survive for 5 years, on average, but individual prognosis depends on the extent of disease at the time of diagnosis. Accurate staging, assessing the extent of local and distant disease, is necessary to determine respectability and overall prognosis. Lung cancer staging encompasses 4 stages, from I to IV, with increasing gravity. The survival rate is about 67% if the cancer is detected at stage I, when it is relatively small (no more than 30 mm in diameter), and drops down to less than 1% for stage IV, when metastases have already developed.

## MOTIVATION OF CAD SYSTEM:

Now a day CAD system has become major research area in the field of medical imaging and diagnostic radiology. CAD is a concept established by taking into account equally the roles of physicians and computers. CAD system works as an assistant to radiologist for final diagnosis of lung cancer. In fact, a large number of CAD systems have been employed for assisting physicians in the early detection of lung cancer in X-ray image. Recently, computer-aided diagnosis (CAD) has become part of the routine clinical work for detection of breast cancer on mammograms and lung cancer in X-rays image at many screening sites and hospitals

in the United States. This seems to indicate that CAD is beginning to be applied widely in the detection and differential diagnosis of many different types of abnormalities in medical images obtained in various examinations by use of different imaging modalities. Even though CAD is well developed clinical tool for diagnosis, however CAD is still in its infancy and much work needs to be done.

#### HISTORICAL BACK GROUND OF CAD SYSTEM:

In the early phase of research on and development of CAD schemes, some computer scientists criticized CAD by saying that it simply would not work, which has been proved to be completely wrong. The reason for this strong criticism at that time might have been related to an unsuccessful attempt in previous research efforts toward the development of automated computer diagnosis. Systematic investigation on CAD began in the 1980s with a fundamental change in the concept for utilization of the computer output, from automated computer diagnosis to computer-aided diagnosis.

In the early 1980s many research and development work on CAD system were begun at the Kurt Rossamann Laboratories for the Radiological image research in the department of Radiology at the University of Chicago. Since that time many research and development works on CAD system have been done for better diagnosis and early detection of disease like lung cancer breast cancer etc. In fact, CAD has become one of the major research subjects in medical imaging and diagnostic radiology.

#### DETECTION OF LUNG CANCER:-

We know that lung cancer is leading cause of cancer death in the world. Many efforts are being undertaken by scientific community to early detection of cancer. So, we can improve the survival chances of patients. There are many technologies used for earlier detection of lung cancer like X-rays, CT scans, FDG-PET scans, bronchoscope, sputum cytology and fluorescent bronchoscope. But now days most commonly used technologies are X-ray, CT, MRI and PET scans.

#### DIAGNOSIS PROCEDURE:

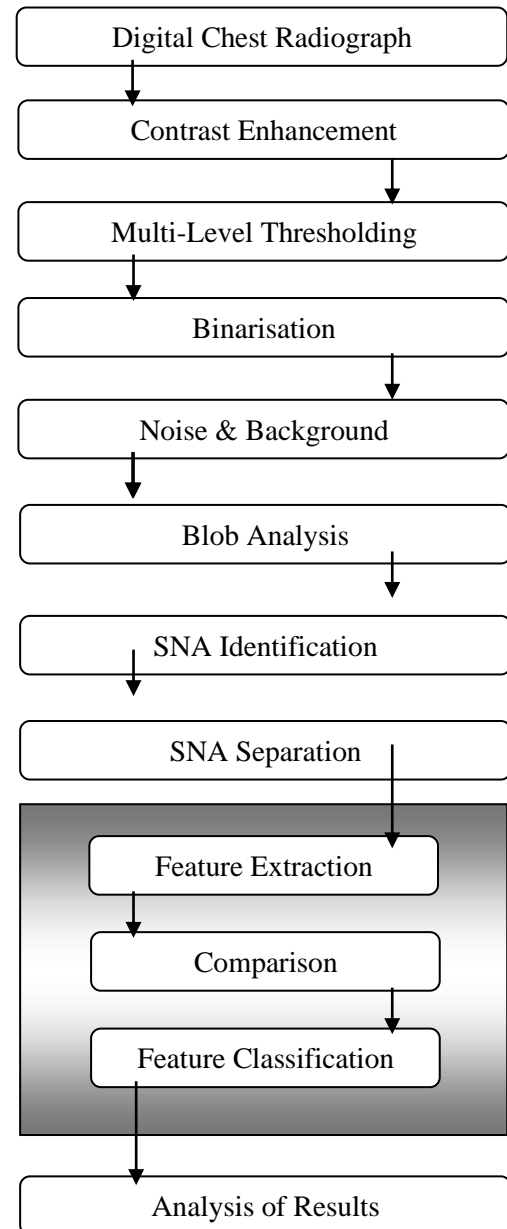
##### Computer Aided Diagnosis:

To identify the early stage tumors is an important goal of physicians, because then it is easy for successful treatment for patient. Computer aided diagnosis (CAD) involves the use of computers to brings suspicious areas on The medical image to radiologist' sattention. It assists the radiologist for better diagnosis. CAD system scans the digital image and marks the suspicious looking areas in the image. Radiologists focus on those areas and decide if a biopsy or further evaluation is needed.

##### Block Diagram Representation:

This CAD system contains two stages for lung cancer detection. In the first stage I processed the image to detect a set of potential nodules. It not only enhances the image but also removes the unwanted background information. The second stage classifies the suspicious region into positive and negative

regions. A positive region is a region that the radiologist feels should go to follow up for additional information. Various morphological features that can be taken into consideration while developing a CAD system include area, size, shape, and moments. These parameters show some variation between cancerous and non-cancerous nodules.



#### Evaluation of CAD

#### RECEIVER OPERATION CHARACTERISTICS:

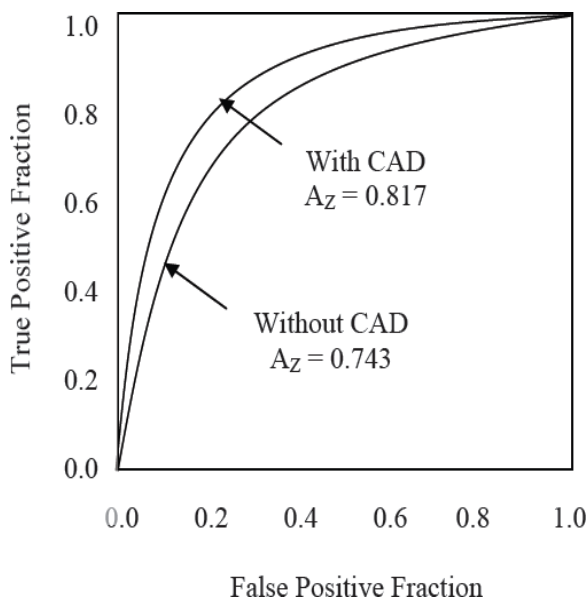
Relative (or receiver) operation characteristic (ROC) analysis is an analytical procedure for measuring the accuracy of a system. ROC curves show a relationship between the true- positive probability and the false-positive probability. The evaluation factor is the area under the curve  $A_z$ . (Figure 2.2 is an example.) ROC

provides a desirable index of accuracy and the appropriate basis for an index of efficacy.

It has the following three unique features.

1. The index of accuracy in ROC analysis is independent of the criterion adopted in the system for making a particular decision.
2. It supplies estimates of the probabilities of decision outcomes of various kinds for any criterion used by the system.
3. The ROC analysis supplies an index of the decision criterion, which reflects together the subjective probabilities and utilities that usually determine this criterion.

Researchers have used ROC analysis to compare the diagnostic performance of radiologists with and without the use of CAD.



#### Median Filter:

Image after multi-level thresholding contains salt and pepper noise. This noise is due to the presence of minute gray scale variations in the image. Median filter is a common image enhancement method for removing pepper and salt noise without significant reduction of the sharpness of the image.

This filter are most popular because, for certain types of random noise, they provide excellent noise reduction capabilities, with considerably less blurring than linear smoothing filters of similar size. Median filter are particularly effective in the presence of impulse noise, also called salt and pepper noise because of its appearance as white and black dots superimpose on an image.

In this filtering process, we take the any pixel and its neighbor \*pixels, determine their median, and assign this median value to that pixel. We calculate the median by first sorting all the pixels values from the neighborhood into numerical order and then replace the pixel being considered with the middle pixel value.

When number of pixels for calculating the median are even then average of the two middle pixel values is the median value.

The medial has two main advantages.

1. The median is a more robust average than the mean and so a single very unrepresentative pixel in a neighborhood will not affect the median value significantly.
2. The median value is value of one of the pixel sintheneighborhood,sothe median filter does not create new unrealistic pixels values when the filter straddles an edge. For this reason the median filter is much better at preserving sharp edges.

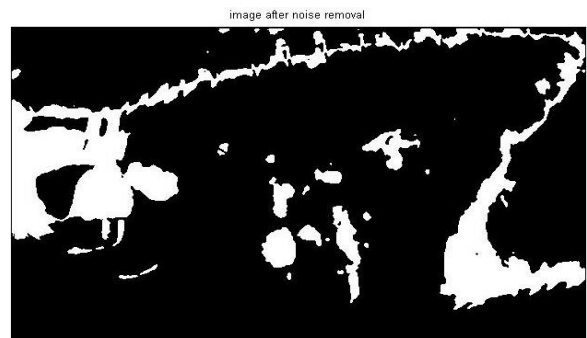


Figure : Image after salt and pepper noise removal

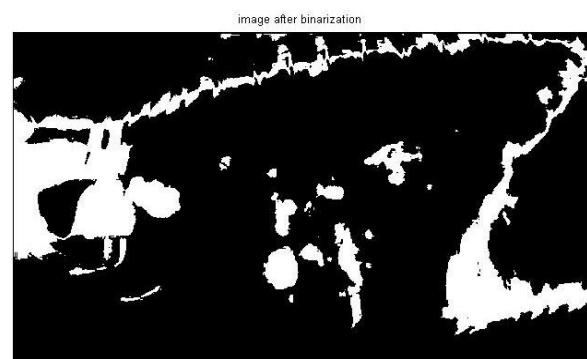


Figure : Image before salt and pepper noise removal

### Method 1:-

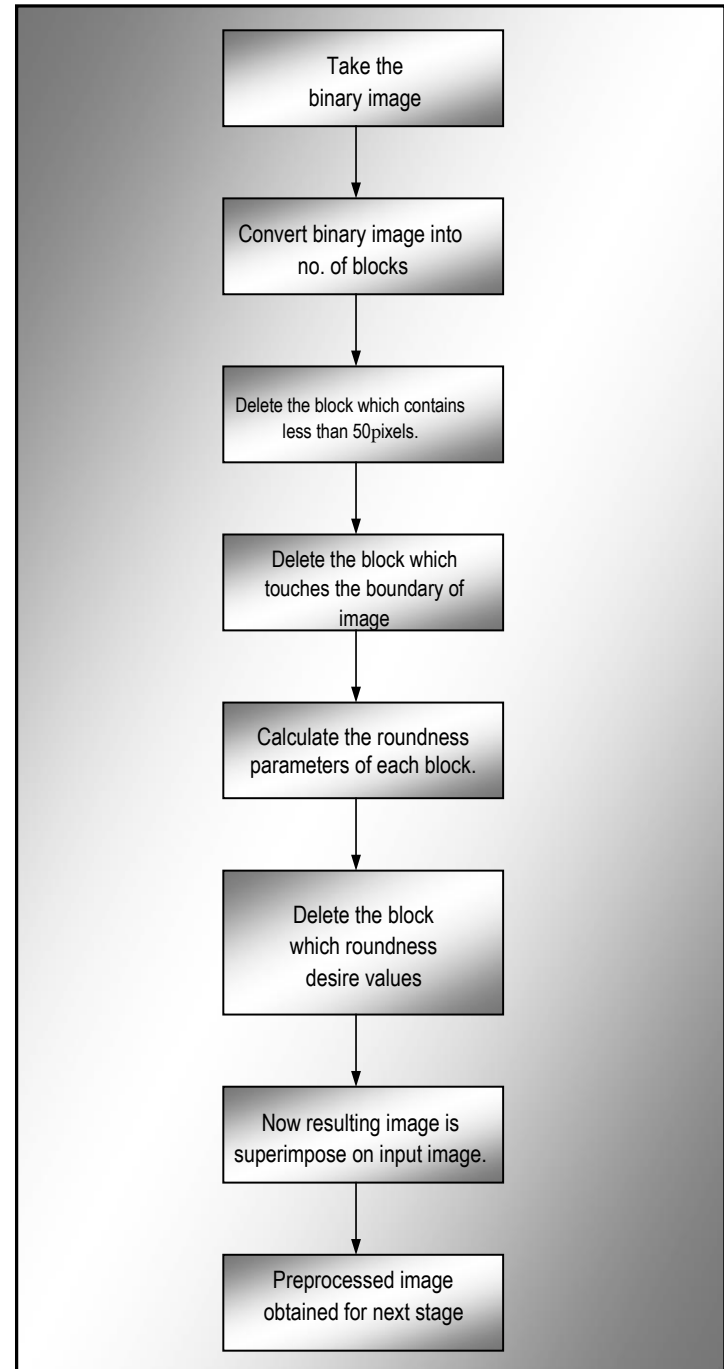
In this method we calculate the area of each blob and obtain a circle with radius, „r“ and area equal to that blob. This circle act as template to locate the round objects. After that we obtain the boundary co-ordinates of each blob. For each boundary co-ordinate say p of each blob, we center a circle with radius r at p and calculate the following ratio value.

$$\text{Ratio} = \frac{\text{Area of intersec. of circle \& blob}}{\text{Area of circle blob}}$$

### Method 2:

In this method we calculate the area and perimeter of each blob and then calculate the following parameter. If any blob which both parameters (ratio and compactness) are less than 0.5, we eliminate that blob by setting all pixels of that blob to the background value(zero).The image obtained after blob analysis act as a mask. Which is projected onto original input image appears relatively high

$$\text{Compactness} = 4\pi \text{ area} / \text{perimeter}^2$$



Block diagram of Blob Analysis of Binary image:

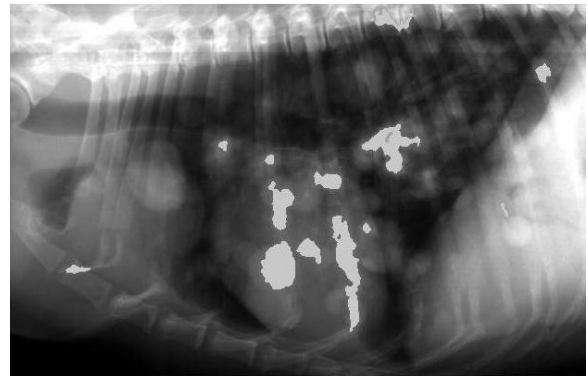
Figure: Block diagram of Blob analysis

## Summary

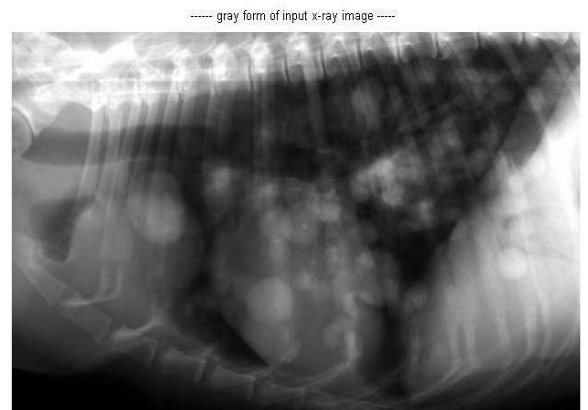
The following produces the brief summary of entire preprocessing stage of for cancer detection. First increase the relative contrast in the x-ray image by histogram equalization. Then we used multi-level (two level) thresholding to remove the unwanted background information. After that binarization method is used to make background information pixels value to zero and foreground pixels value to 1. During this some salt and pepper noise induces in image which is reduced by use median filter. Blob analysis method involves a sequence of steps for removal of unwanted foreground information and check the roundness of blob. Final image obtained after blob analysis are projected onto the original x-ray image.

A manual approach is used to separate the SNA from the final output image. We use the SNA regions separated from this image as input to the next stage for further classification as cancerous or non-cancerous. Some of the SNA regions sent to the next stage contain un-enhanced image information. This happens because the enhancement procedures used in this

stage do not enhance all the cancerous nodule regions.



**Figure 4.14 Final output image**







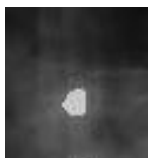
**Figure : Input Image**

## Testing

Table 6.1 shows a few sample images, their input target values, and output results for both processed and unprocessed images in the training sets. These output values lie in a range between 0 to 1.

We selected a threshold value of .5 to differentiate the value of cancerous and non cancerous region.

Table 6.1 Table showing a few sample results of the ANN used in this thesis.

	Target Value=0.91 (non-cancerous)	Processed Image Output=0.71	Unprocessed Image Output=0.59
	Target Value=0.91 (non-cancerous)	Processed Image Output=0.75	Unprocessed Image Output=0.7
	Target value=0.1 (cancerous)	Processed Image Output=0.18	Unprocessed Image Output=0.18
	Target Value=0.91 (non-cancerous)	Processed Image Output=0.69	Unprocessed Image Output=0.69
	Target value=0.1 (cancerous)	Processed Image Output=0.21	Unprocessed Image Output=0.33



## CONCLUSION:-

In this thesis we developed the Computer Aided Detection method for lung cancer nodules in X-ray image. This works as assistant of radiologist for better detection of lung cancer in X-Ray image. It contains two stages, preprocessing stage and classification stage. In preprocessing stage we implement various image enhancement techniques, like contrast enhancement, multi-level thresholding, binarization, noise removal and lastly blob analysis for better visualization of lung nodules in X-ray image. Now SNAs region is separated from preprocessed image and applied to the input of Artificial Neural Network for classification of SNA as cancerous and non- cancerous regions. We have used back-propagation algorithm for training the neural network for better classification of SNAs. The detection accuracy is improved and mainly depends on the correct classification. By including the pre-processing stage in the CAD system, we observed a significant improvement in the detection accuracy. The accuracy values increased from 70% to 83.3%.

## BIBLIOGRAPHY

- [1] Automated cancer diagnosis based on histopathological images: a systematic survey -- CigdemDemir and BulentYener
- [2] Development of Neural Networks For System Identification .... Siobhan Murphy 98276506
- [3] Neural Network Model of the Backpropagation Algorithm ..... Rudolf Jak'sa, MiroslavKatr'ak, Department of Cybernetics and Artificial Intelligence, Technical University of Ko'sice, Letn'a 9, 041 20 Ko'sice Slovakia
- [4] "Salt and Pepper Noise Removal by Median Type Noise Detectors and Edge Preserving Regularization,"H. Chan, C. W. Ho, and M.Nikolova, Department of Mathematics-Hong Kong, pp.1- 7, 2003.
- [5] *Digital Image Processing*, Second Edition, Pearson Education, India, 2003. R. C. Gonzalez and R. E. Woods
- [6] A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database .....Arnold M.R. Schilham \*, Bram van Ginneken, Marco Loog Image Sciences Institute, University Medical Center Utrecht, The Netherlands
- [7]. Knowledge-based method for segmentation and analysis of lung boundaries in chest X-ray images  
Matthew S. Brown, 1,\*, Laurence S. Wilson, Bruce D. Doughty, Robert W. Gill, Changming Sund
- [8]. Notes on Artificial Neural Networks, ....John Larchev'eque
- [9]. Neural Network Based System Identification TOOLBOX, Version 1.1 For Use with MATLAB
- [10]. Computer Aided Diagnosis Tool for the detection of cancerous Nodules in X-ray image bybomma
- [11]. Computer-aided Diagnosis for Improved Detection of Lung Nodules by Use of PA and Lateral Chest Radiographs..... JunjiShiraishi, PhD, Feng Li, MD, PhD, and KunioDoi, PhD
- [12]. Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, 5841 S. Maryland Avenue, Chicago Illinois 60637
- [13]. Lung Cancer Diagnosis - "http://www.iressa-us.com/content/patient/nsclc/diagnosis.asp"
- [14]. Fundamentals of Neural Networks ..... Ali Zilouchian
- [15]. Computer-Aided Diagnosis: A Neural-Network-Based Approach to Lung Nodule Detection ..... Manuel G. Penedo,\*, Mar'ia J. Carreira, Antonio Mosquera, and Diego Cabello, Associate Member, IEEE
- [16]. A New Computer Aided System for the Detection of Nodules in Lung CT Exam ..... Alessandro Riccardi