Lung Cancer Nodule Detection

(Pranav vadanere [1], Vrushali Shinde [2], Karan bendigiri [3])
(Prof Rakhi bharadwaj [4])

[1, 2, 3](Students, Trinity College of engineering, Pune, Maharashtra).
[4] (Professor, Trinity College of engineering, Pune, Maharashtra).

Abstract- *Lung cancer is a third leading type of cancer that may cause death of a person. Early detection and diagnosis of the disease is essential to prolong the life of the patients affected with this scourge. For determining the cancer cells from medical images, various image processing and soft computing techniques can be used. CT-images are most commonly used for image processing. CT-image has properties like high resolution, better clarity, low noise and low distortion and etc. It is the best technique of image for detection of small nodules. This paper focuses on different techniques that have been proposed to provide detection of lung cancer nodules.*

Keywords: - ***Cancer, Nodule Detection, Lung Nodule, CT Images, Preprocessing, Segmentation.***

I. INTRODUCTION

About 80% of patients are diagnosed correctly in the middle or advanced stages of cancer [3]. If this had been diagnosed at initial stages, then the chance of survival rate would be high. An utmost important but difficult task for any radiologist is to detect and examine the tumorous nodules from the chest radiographs. Few nodules may not be detected as they may be masked by the anatomical structure or due to the inferior quality of the images. Also, it is influenced by the subjective and variable decision criteria used by radiologists. Hence, there is a great need for the upcoming technologies to detect the lung cancer in its primary stages.

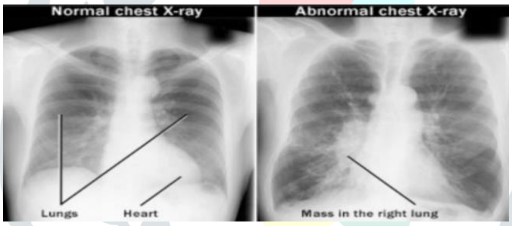

Figure: - Comparison between Normal and cancer Nodule.

STAGES OF CANCER:-

Cancer leads to excessive multiplication of normal cells without control and is able to affect other issues. Cancer spreads stage by stage; Cancer is the rapid multiplication of abnormal cells with cluster formation the figure below shows the images with different stages of cancer.

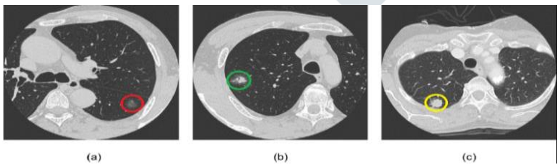

Figure: - Nodule with smaller to increasing size.

II. LITERATURE REVIEW:

REVIEW ON LUNG CANCER DETECTION SYSTEMS

Image Smoothing using Median Filtering

A. Kulkarni et al. [2] proposed a system for lung nodule cancer detection using CT images in DCOM format. Image smoothing was done by Median filter to reduce blurring of edges. The advantage of using a median filter is that it is not affected by individual noise spike, to eliminate impulsive noise quite well and it does not blur edges much and can be applied iteratively. Gabor filter is used for enhancement as it gives a better result compared to Fast Fourier Transform and auto enhancement.

Advantage: -

The median filter gives more accurate results compared to Gaussian, mean and Wiener filters.

Genetic Algorithm in Lung Cancer Detection

El-Baz et al (2013) developed a new algorithm for lung nodule detection using GATM. Their proposed algorithm was based on three steps: (i) Isolation of nodules, arteries, veins, bronchi and bronchioles from other anatomical structures (ii) Isolation of the nodules using deformable 3-D and 2-D templates and (iii) Elimination of the FPs. The algorithm was validated using private database which yielded 82.3% of sensitivity and 9.2% of FPs.

Tan et al (2011) described a novel computer-aided lung nodule detection system for CT images, which used three classifiers; genetic algorithms, artificial neural network and fixed-topology neural network. Their lung nodule detection was based on the filters, which highlighted the nodules, vessels, and divergence features. After the detection of candidate nodules, invariant features were extracted and applied to the three classifiers. The results obtained with the genetic algorithm had the sensitivity of 87.5%, with 4 FPs/patient for nodules with diameter larger than or equal to 3 mm.

Linear Discriminant Analysis in Lung Cancer Detection

Messay et al. (2010) described a new computationally efficient CAD system for pulmonary nodule detection by combining simple image processing techniques, such as intensity thresholding and morphological processing, to segment and detect structures that are lung nodule candidates. The lung nodule candidates are determined by extracting 245 features from the segmented lung CT image. Then, significant features are selected and fed to the LDA classifier. This method was able to detect 92.8% of the structures, which are nodule candidates.

Suarez et al (2011) demonstrated an automated detection of pulmonary nodules in CT for the FP reduction by combining multiple classifiers. Experimented classifiers are LDA, Quadratic Discriminant Analysis (QDL), ANN and SVM. They are applied independently and combined to the Lung Image Database Consortium (LIDC) using 85 images which has 110 cancerous nodules. The reported sensitivity is 80%, and the number of FPs/patient for each of the six classifiers is 6.1 for LDA, 19.9 for QDA, 8.6 for ANN and 17.0 for SVM. When the classifiers are used in combination, the number of FPs per patient is greatly reduced.

Artificial Neural Networks in Lung Cancer Detection

Kruvilla and Gunavathi (2014) described a system for the detection of lung cancer in CT scan images using ANN. In their approach, the CT scan images which were in gray scale was firstly converted to binary image using grey level thresholding. The lungs were segmented using morphological operation. Then the statistical parameters such as mean, standard deviation, skew ness, kurtosis, fifth and sixth central moment were calculated. The classifications were done by feed forward and feed forward-back propagation networks. Their study concluded that the feed forward-back propagation network provided better classification results. The implemented ANN yielded the sensitivity of 82% and specificity of 90% with 0.5 FP/patient.

Rule based Classifier in Lung Cancer Detection

Sharma et al (2011) described an automated CAD system for early detection of lung cancer for CT images using several steps. First, lung a region was extracted by image processing techniques, including bit image slicing, erosion, and Weiner filter. The CT image was converted into a binary image during the extraction process by bit plane slicing technique. After extraction of lung region, lungs and nodules were segmented by region growing segmentation. Then the set of features were extracted and fed to the rule based classifier for the final decision on the segmented nodules. The proposed system achieved the accuracy of 80%.

Lung Cancer Detection from the Observed Symptoms

A detailed study on lung cancer diagnosis based on fuzzy rules was conducted by Durai and Iyengar (2010). The efficiency of their system was low, because of their simple algorithm. Later (2010), they developed a diagnostic model for the stage-wise lung cancer detection using improved fuzzy rules. In addition, it also suggests the type of treatment for the patients. The key characteristic of their latter system was easier modification and updating of database. Their later system efficiency was better than the former, could be used as medical diagnosis model for finding the stages of lung cancer patients.

Lung Cancer Detection from 2-D CT Scan Images

Commonly lung nodule detection using CAD systems for CT scan images involves four important steps: lung region segmentation, nodule candidate detection, feature extraction and classification. To extract the lung region from the CT scan images several methods were found in the literatures. The methods such as multiple thresholding, optimal thresholding, and global thresholding were successfully implemented for lung region segmentation (Suzuki et al, 2003; Ye et al, 2009; Suarez-Cuenca et al, 2009).
However, in these studies the threshold values were calculated manually by considering the pixel intensity of the CT scan images. Also morphological processing was performed on CT images to remove fat, bone and background noise of lung parenchyma in threshold based segmentation techniques. Since the threshold value is chosen based on the CT technology and X-ray dose, it varies from one CT machine to another CT machine. Hence this technique will not segment the lung region universally. Auto thresholding based algorithms were implemented in many literatures to overcome the difficulty of hard threshold segmentation.

III. PROPOSED SYSTEMS

A new automatic method has been proposed based upon black circular neighborhood algorithm and image processing techniques to extract the nodules. Feature extraction is an important step in algorithms. These separate the area which is then analyzed for detection of nodules to diagnose the disease. Computed tomography (CT), allows effective mapping .CT Image decreases the time complexity. We have used the GLCM Features Which Helps in detection of the nodule .The use of Otsu's Algorithm helps in detection of size and stage of the tumor.

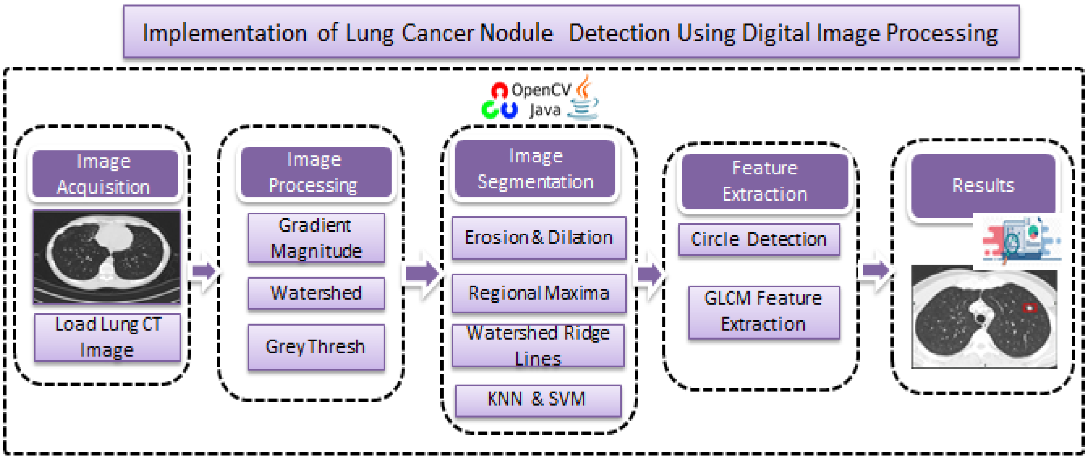

Figure: - System Architecture.

1. Image Preprocessing

Apply Processing on input Image (Open CV image)

a. Gradient Magnitude: we measure's the intensity change in the image to find difference between the nodule and other objects.

b. Then we apply the watershed technique to perform image segmentation to separate the objects present in lung CT images.

c. Gray thresh will now clusterises the image using the intensity value of pixels. It uses Otsu's Algorithm for performing this operation that checks the brightness of each pixel from the CT image to find its height map the lines that are similar to the given height.

2. Erosion & Dilation(Opening and Closing)

In image Segmentation, pixels are allocated to categories according to the range of values in which a pixel lies. Erosion will removes the objects that are smaller than the structure of the nodule and reconstructs the shape of remaining objects. After that the dilation removes the small holes from the large objects from the image.

3. Object Segmentation

In segmentation, regional maxima applied to the image, pixels are categorized as their intensity value. Watershed based segmentation is a fact that computes the pixels which belongs to the same object and separate's the image from foreground with background as well as objects also segmented with each other with it. This step is very necessary that derives the actual differentiation between the nodule present in image and other image

4. Nodule detection

Now we have segmented image and we can now detect the number of circles with their radios that specifies the nodule in CT image.

5. Feature Extraction

Following GLCM feature will be computed from the detected lung nodule in CT image.
a. Contrast
b. Correlation
c. Correlation
d. Entropy
e. Homogeneity

We can apply KNN algorithm for identification of the complete Lung Nodule.

IV. ALGORITHM USED

1. KNN

a. In machine learning, a KNN is a type of feed-forward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex, whose individual neurons are arranged in such a way that they respond to overlapping regions tiling the visual field.
b. KNN was inspired by biological processes and is variations of multilayer perceptron's designed to use minimal amounts of preprocessing.

a) **Initialization:** it is important to achieve convergence. Xavier initialization is used. With this, the activations and the gradients are maintained in controlled levels; otherwise back-propagated gradients could vanish or explode.
b) Activation Function: is responsible for non-linearly transforming the data. Rectifier linear units (ReLU), defined as

$$f(x) = (0, x),$$

Were found to achieve better results than the more classical sigmoid, or hyperbolic tangent functions, and speed up training. However, imposing a constant 0 can impair the gradient flowing and consequent adjustment of the weights. These limitations are coped using a variant called leaky rectifier linear unit (LReLU) that introduces a small slope on the negative part of the function. This function is defined as

$$f(x) = (0, x) + \alpha\, min(0, x)$$

Where α is the leakiness parameter. In the last FC layer, we use softmax.

2. Otsu's Algorithm

$$T(x, y) = m(x, y).\,[1 + k.\left(\frac{s(x, y)}{R}\right) - 1)]$$

3. Image Thresholding algorithm

The simplest implementation of thresholding is to choose an intensity value as a threshold level and the values below this threshold become 0 (black) and the values above this threshold become 1 (white). If T is the global threshold of image f(x, y) and the g(x,y) is the thresholding image, then:

$$g(x, y) = \{\frac{1 \quad ,if\ f(x,y) \geq T}{0, \quad Otherwise}\}$$

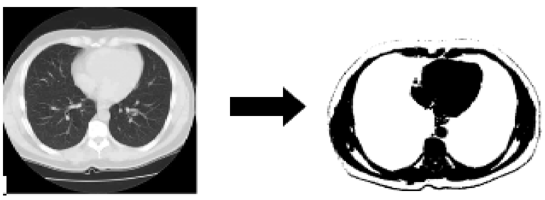

Figure: - Image Before and After Thresholding

V. RESULT AND DISCUSSIONS

Results

Following table represent the normal and abnormal images and it accuracy.

$$ACCURACY = \frac{TRUE\ POSITIVE}{TOTAL\ NO.OF\ IMAGES} * 100$$

IMAGES	ACCURACY
Abnormal Images	86.16
Normal Images	91.35

Figure: - Accuracy of Normal and Abnormal Images

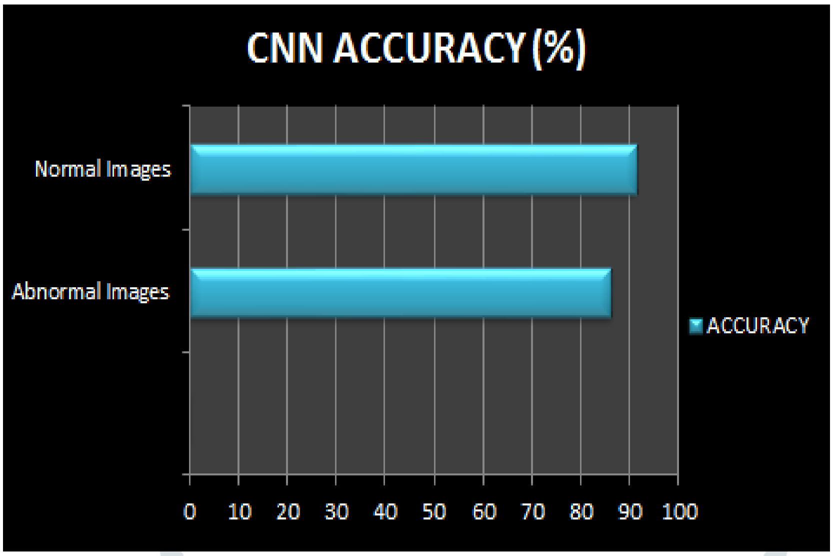

Figure: - Accuracy of Normal and Abnormal Images

ADVANTAGES:-

• The whole system can be implemented in very low cost and provides better accuracy.

CONCLUSION

A new method for robust detection of pulmonary lung nodule in CT images. Initially the images are preprocessed where the contrast levels are adjusted. Then segmentation is done and nodules are identified. The features are extracted from the nodules and it is given as an input to KNN for classification. Here the normal and abnormal CT images are classified using KNN and accuracy is calculated.

REFERENCES:

[1] C. T. Henschke, et al. Early lung cancer Action Project: Overall Design and Findings from baseline Screening,Cancer.2000 (Suppl), Vol.89 (11), pp. 2474-2482, 2000.

[2] Kulkarni ,A. Panditrao, Classification of lung cancer stages on CT scan images using image processing, IEEE,International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 2014

[3] K. Punithavathy, M. Ramya, S. Poobal Analysis of Statistical Texture Features for Automatic Lung Cancer Detection in PET/CT Images, International Conference on Robotics, Automation, Control and

Embedded Systems RACE 2015, 18-20,February 2015.

[4] Christoph Georg Eichkitz n, Johannes Amtmann, Marcellus Gregor Schreilechner Calculation of grey level co-occurrence matrix-based seismic attributes in three dimensions 2013 Elsevier.

[5] A.R. Talebpour, H.R. Hemmati, M. Zarif Hosseinian Automatic Lung Nodules Detection In Computed Tomography Images Using Nodule Filtering And Neural Networks 2014 IEEE

[6] N. Hadavi Md. Jan Nordin, Ali Shojaeipour, Lung Cancer Diagnosis Using CT-Scan Images Based on Cellular Learning Automata , IEEE International Conference on Computer and Information Sciences (ICCOINS), 3-5 June 2014.

[7] S. Kanitkar, N. Thombare , S. Lokhande Detection of Lung Cancer Using Marker-Controlled Watershed Transform, International Conference on Pervasive Computing (ICPC), 2015.

[8] Lakshmi N arayanan A and Prof. Jeeva J.B A Computer Aided Diagnosis for detection and classification of lung nodules IEEE Sponsored 9th International Conference on Intelligent Systems and Control (lSCO) 2015.

[9] Lei Fan, Zhaoqiang Xia, Xiaobiao Zhang, Xiaoyi Feng Lung Nodule Detection Based on 3D Convolutional Neural Networks 2017 IEEE.

[10] B. Patil, Sanjeev N. Jain, Cancer Cells Detection Using Digital Image Processing Methods, International Journal of Latest Trends in Engineering and Technology (IJLTET), Vol. 3 Issue 4 March 2014.