Early Detection Of Lung Cancer Using Digital Image Processing

[1]Alla Padmashri, [2]Nalli Sravya, [3]Palegari Pavan Kumar, [4]Kontham Satish, [5]Dr.S.V.R.K.Rao
[1,2,3,4]B.Tech Student, [5]Professor and Head Of the Department,
[1]Department Of Electronics and Communication,
[1]Godavari Institute Of Engineering and Technology, Rajamahendravaram, Andhra Pradesh,India.

Abstract : Apart from other types of cancer, Lung Cancer is one of the most common causes of death throughout the world.In this type,the cancer cells are increased at an abnormal condition. It is necessary to get proper and efficient treatment in time, to reduce death rate of population in the world.In the Lung cancer detection system, we have detected various stages of cancer by using Support Vector Machine classifier (SVM**)**.This is done by using virtual machine where Jupyter notebook is used as a console for python implementation.

***IndexTerms* – Lung Cancer,stages,support vector machine,jupyter notebook.**

I. INTRODUCTION

Cancer is now the biggest cause of death now -a-days. Lung cancer is one of the most common cancers . Due to the lifestyle of people there is a gradual increase in cancer patient. Pain, breathlessness, cough, weight loss and fatigue are some of the common symptoms of cancer. Survival from the disease is difficult if it is not detected at the early stage. Only 15% of lung cancer is detected at the early stage. So it is necessary to detect the cancer in the early stage. This has not changed in the past three decades [1].

Mostly lung cancer occurs in males and females, it is caused by cigarette smoking, alcohol consumption etc. An estimated 85% of lung Cancer cases in males and 75% in females are caused by cigarette smoking [2]. The purpose of this paper is to find the early stages of lung cancer and more accurate result by using different techniques like fusion, enhancement and segmentation process.It performs above stages such as,

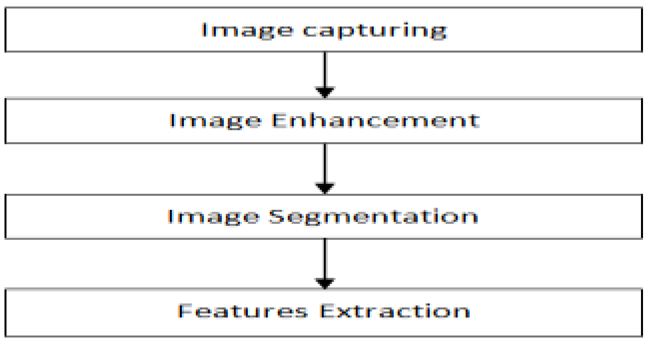

figure-1 :cancer stages in image processing

II. LUNG CANCER APPLICATION

A web model has been developed to demonstrate a proof of the concept. The application requires a user to upload CT Scan image. The application then processes the file and displays the images to the user. The user then chooses which scan he or she wants to predict and then the application pre-processes the CT Scan and infers the image to the predictive model. The output of the model is then displayed to the user. The user has the choice to view the images via a carousel or a gallery mode.

.

III. RESEARCH ELEMENT

This report contains many aspects of research that support deep learning's ability to find lung cancer within CT Scans. The data used to train the model was gathered from a deep learning competition which are CT Scans that have been annotated by radiologists. The radiologists determine whether a nodule is lung cancer or not and this location has been specified in 3D coordinates.

3.1 Deep Learning

Deep Learning research has been conducted to ensure that the correct paradigm would be used to to train the model. The architecture used is called U-Net. It acts as a deep learning segmentation of images. Using our evaluation methods we could see that the deep learning architecture is able to segment regions of interest relative to it's accuracy but found that the model creates many false positives which means that the model required more regularity and training time, although it was trained for many hours.

IV. METHODOLGY

Considering the fact that lung nodules have relatively higher densities than those of lung parenchyma, density-based detection methods employ techniques such as multiple thresholding, region-growing, locally adaptive thresholding in combination with region growing, opening and closing, using the histogram, the top 20% gray values considered as initial cancerous candidate regions, using the histogram the normal tissues are removed, then elliptical-shaped regions, which is in general represent abnormalities, are detected, and fuzzy clustering used to identify nodule candidates in the lungs. False-positive results can then be reduced from the detected nodule candidates by employing a priori knowledge of small lung nodules. For the model-based detection approaches, the relatively compact shape of a small lung nodule is taken into account while establishing the models to identify nodules in the lungs. Techniques such as Morphological filter and the anatomy based generic model have been proposed to identify sphere shaped small nodules in the lung. Support vector machines are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification. The basic SVM takes a set of input data and for each given input, predicts which of two classes forms the input, making it a non-probabilistic binary linear classifier. SVM uses a kernel function which maps the given data into a different space and the separations can be made even with very complex boundaries.

The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image Segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics [3]. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture

V. RESULT

The metrics given at TABLE 5.1 give an indication that the model is able to a high percentage of accuracy when detecting lung cancer nodules.

TPR, TNR, FPR, FNR and Mean Dice Coefficient for the Dev Set and Test Set including results for post processing. An interesting observation is that the model is able to gain predict the true positive cancer nodule within the scan. The data consists of about 18 positive labels (with cancer) on each dev and training set and about 20 negative images (without cancer) which approximates to about 28 images per set. The positive samples were taken from the subsets of the training set and the negative ones were taken from random CT scans.

table 5.1: model evaluation

Data	Post Processing	TPR	TNR	FPR	FNR	Dice Coefficient
Dev Set	None	0.833%	0.256%	0.743%	0.166%	0.0003%
Test Set	None	0.928%	0.312%	0.687%	0.0714%	0.00062%
Dev Set	Rounding	0.388%	0.833%	0.166%	0.61%	0.183%
Test Set	Rounding	0.529%	0.833%	0.166%	0.470%	0.105%

Staging involves evaluation of a cancer size and its penetration into surrounding tissues as well as presence or absence of metastasis in the lymph nodes or other organs [4]

VI. CONCLUSION

Lung cancer is one of the most dangerous diseases in the world. Correct Diagnosis and early detection of lung cancer can increase the survival rate. The present techniques include study of X-ray, CT scan, MRI, PET images. The expert physicians diagnose the disease and identify the stage of cancer by experience. The treatment includes surgery, chemotherapy, radiation therapy and targeted therapy. These treatments are lengthy, costly and painful. Hence, an attempt is made to atomize this procedure to detect the lung cancer using image processing techniques. CT scan images are acquired from various hospitals. These images include less noise as compared to X-ray and MRI images

Also for classification purpose, Support Vector Machines are an attractive approach to data modeling. They combine generalization control with a technique to address the curse of dimensionality. The kernel mapping provides a unifying framework for most of the commonly employed model architectures, enabling comparisons to be performed. In classification problems generalization control is obtained by maximizing the margin, which corresponds to minimization of the weight vector in a canonical framework. For future work, we can implement this technique on some more images. Increasing the number of images used for the process, can improve the accuracy. Also MRI, X-ray, PET images can be considered for this technique.

V. ACKNOWLEDGMENT

I am grateful to my guide **Dr. S. V. R. K. RAO ,**Professor and Head Of The Department of ECE for having given me the opportunity to carry out this project work. I take this opportunity to express my profound and whole heartful thanks to my guide, who with his patience support and sincere guidance helped me in successful completion of the project

REFERENCES

[1] American Cancer Society, "Cancer Statistics, 2005", CA: A Cancer Journal for Clinicians, 55: 10-30, 2005, "http://caonline.amcancersoc.org/cgi/content/full/55/1/10".

[2] Singapore Cancer Society, http://www. singaporecancersociety.org.sg.

[3] Shapiro L.G., Stockman G.C., Computer Vision: Theory and Applications, Prentice Hall,2001.

[4] D. Lin and C. Yan, "Lung nodules identification rules extraction with neural fuzzy network", IEEE, Neural Information Processing, vol. 4, (2002)