Optimal Diagnosis of Lung Cancer using CT Images

Karthikeyan S, Navinkumar M, Aishwarya M R, Deva Subramanian S, Gurulakshana M

Abstract— According to the American Cancer Society, lung cancer is the second most widespread cancer and the leading cause of cancer deaths in both men and women. The death rate of lung cancer every year is greater than that of colon, breast, and prostate cancers combined. CT scan is a non-invasive method for diagnosis of any ailment, and can be used to detect lung cancer as well. The proposed project involves cell detection using image processing techniques. Because the time is a very important factor in cancer treatment, especially in cancers such as the lung, imaging techniques are used to accelerate diagnosis. The image processing paired with data analysis techniques helps us diagnose the particular type of cancer by comparing the output of the CT scan to an available database of images. This improves accuracy and reduces the time required for the diagnosis. Features of the image under test are extracted and analysed, and the decision regarding the morphological characteristics of the image are made. This helps us arrive at a decision regarding the nature of the image.

Keywords—CT scan, image processing, segmentation, feature extraction.

I. INTRODUCTION

Cancer is a condition that is signified by an uncontrollable proliferation of cells in the body. Cancer is dangerous as it disrupts the normal metabolism of the body. There are various causes of cancer, such as alcohol consumption, smoking, exposure to aerosols, and at times even overweight can be a potential cause of cancer. This paper deals with the detection of lung cancer using image analysis because according to the American Cancer Society, lung cancer is the second most commonplace cancer in both genders, leading to the most number cancer deaths. The death rate of lung cancer every year is greater than that of colon, breast, and prostate cancers combined. Lung cancer spreads toward the centre of the chest corresponding to the natural flow of lymph out of the lungs is toward the centre of the chest. There are four stages of lung cancer. Lung nodule is an abnormality that leads to lung cancer, it is a small round or oval shaped growth on the lung which appears as a white shadow in the CT scan. These uncontrolled proliferation of cells restrict the growth

of healthy lung tissues. If not treated, this growth can spread beyond the lung in the nearby tissue called metastasis and, form tumours. Long-term exposure to tobacco smoke, which causes 80-90% of lung cancers, is one of the primary causes of lung cancer. Chemotherapy, radiotherapy and immunotherapy are general methods of cancer treatment. Image processing can be used to detect lung cancer from CT scan images. This paper aims at reducing the cases of misdiagnosis associated with manual examination due to interpretation errors as well as variation between observers, and improves the speed and efficiency by image analysis. By using an image system, a radiologist considers the output from a computer analysis of the interpreted image, though the final diagnosis is made by the radiologist.

Few models center their investigation for the most part on the order of lung pictures as influenced by malignancy malady or not. In their proposed strategy lung CT examine picture is utilized as input image then they extract the particular interested region using active contour model and then some important features are extracted to classify the image. To characterize the picture Fuzzy Interface System is utilized. This framework utilized both neural arrange and fuzzy logic system. The exactness of the proposed framework is 94.12% (Tanushree Sinha Roy et al., 2015).

Some papers studied the concept of medical content based image retrieval. As per author numerous procedures are accessible to decide the presence of nodule in the lung in its initial stage yet none of the procedure give the best outcome. To give the more accuracy and best outcome author proposes a framework called as Computer Aided Diagnosing System (CAD) for recognizing the presence of nodule in lung (Ritika Agarwal et al., 2015).

Some systems proposes a method which uses a median filter and wiener to remove unwanted noise from the image, K-mean clustering approach is used for segmentation and Gray Level Co-occurrence matrix is applied for feature extraction. After applying this entire image processing techniques the back propagation network is used for the classification (P.B.Sangamithraa et al., 2016).

Since lung cancer is a very prevalent type of cancer, developing a system that can detect cancer from the input image of a CT scan can be a very beneficial contribution to the medical community. This system is accurate to a very high percentage, the working of the system is very simple and consists of simple algorithms. The system is efficient and fast and seamlessly integrates with the system it works on.

Dr. Karthikeyan S, Associate Professor / Dept. of ECE, Malla Reddy Engineering College and Management Sciences, Kistapur Village, Medchal Mandal, Ranga Reddy, Telangana, India. (Email: meckarthikkct@gmail.com)

Mr.Navinkumar M, Assistant Professor / Dept. of ECE, Sri Krishna College of Technology, Coimbatore, Tamilnadu, India. (Email: navinkumar.m@skct.edu.in)

Ms.Aishwarya M R, Student / Dept. of ECE, Sri Krishna College of Technology, Coimbatore, Tamilnadu, India.

Mr.Deva Subramanian S, Student / Dept. of ECE, Sri Krishna College of Technology, Coimbatore, Tamilnadu, India.

Ms.Gurulakshana M, Student / Dept. of ECE, Sri Krishna College of Technology, Coimbatore, Tamilnadu, India.

II. EXISTING SYSTEMS

There are various systems that perform analysis on images and detect cancer, some of them are listed below.

A. Detection of Lobular Carcinoma In Situ(LCIS) by Image Analysis

Quantitative image analysis method used. Glandular segmentation and feature extraction are major steps in the process. Computationally different features in LCIS diagnostic groups and normal image groups and the detection accuracy were reported (Sujin Kim et al., 2011).

B. Cancer Detection based on Cell Counting, Cell Area Measurement and Clump Detection

Processes images in four major steps: Pre-processing, Classification, Separating Bound Nucleus and Cell Counting. From the results, the Error Ratio and Standard Deviation of the proposed method are calculated to be 6.75% and 6.39% respectively (Monirul Islam Pavel et al., 2017).

C. Thermographic Image Analysis Method in Detection of Cancer

Uses Thermographic Images to detect cancer. A thermographic image (thermogram) is a colour image in which each colour represents a specific temperature. The thermograms used in this study use an 18-shade colour map. The colour map describes warmer temperatures as white and red and cooler temperatures as blue and black (Maryamsadat Amini et al., 2012).

III. PROPOSED SYSTEM

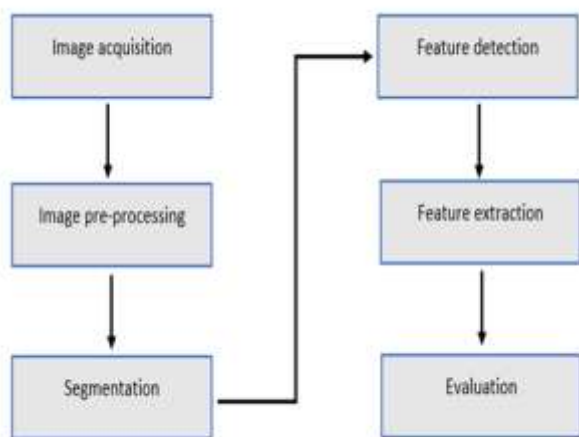The proposed system has the following stages.

Fig. 1. Flow of the proposed system

The system proposed in the paper, in accordance with Fig. 1. has the following stages or module-image acquisition, image pre-processing, image segmentation, feature detection, feature extraction, evaluation. Each stage is very important, and helps in the seamless and easy detection of the tumour in the given CT scan image. To analyse a CT scan image of the lung, we need to know the small differences and nuances between a lung having cancer, and one not having cancer, as illustrated in Fig. 2. This is a very important step on the side of the user, as we need to still verify the outputs of the system. An example of a lung having and not having cancer is given below.
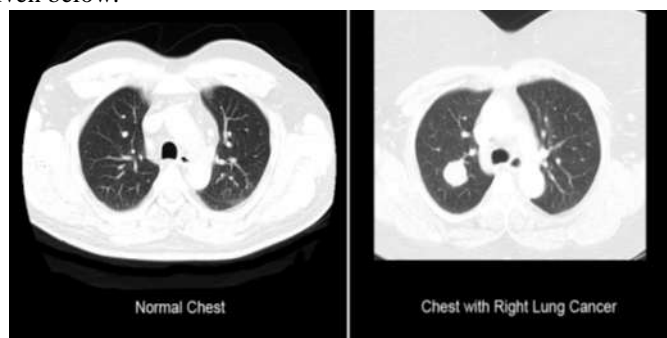

Fig. 2. CT scans of lungs with and without cancer

The various modules of the system are explained in detail below.

A. Image Acquisition

Different images either malignant or benign lung cancer are acquired from different medical sources such as atlases, pathological labs, internet websites & hospitals. The images that are acquired are completely unprocessed and are likely to possess noise. The images can be in the JPEG format. The acquired images have a lot of noise, thus to improve the contrast, clarity and to suppress the background noise, it is required to pre-process the images. Thus, different methods like smoothing, enhancement, segmentation are applied to bring the image to a form apt for processing.

B. Image Pre-processing

Pre-processing technique consists of various steps for noise removal or noise suppression. Several filters are used for noise removal. Few of them are Wiener Filter, Averaging Filter & Median Filter. This stage is divided into Image Smoothing and Image Enhancement.

Image Smoothing

Image smoothing is the process of subduing noise. The most common filters used for image smoothing are Averaging filter and Median filter. To apply filters on an image, we first induce noise in the image, as shown in Fig. 3. ourselves and see how it reacts to each filter. Then we decide which filter is to be used, based on which filter removes most noise without disturbing the basic nature of the image. We also check if all the essential features are still present in the image such as sharp edges and boundaries. This is important as the image should not lose its features as a result of filtering.
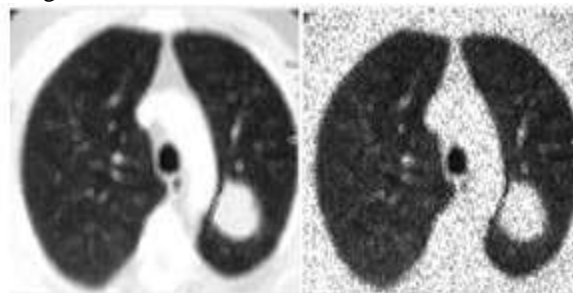

Fig. 3. Original image (left) and noisy image (right)

Averaging filter

Produces smoothing in the spatial domain. This produces an average of different images. But higher order of mask can lead to over smoothing leading to loss of important boundaries, thus making the image lose its value.
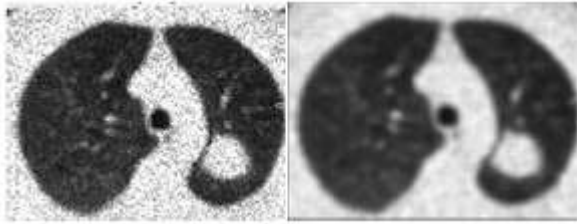

Fig. 4. Noisy image (left) and output of averaging filter (right)

Median Filter

This filter is commonly used for image enhancement and restoration. In this filter, the mask consists of pixels replaced by the median of their corresponding neighbours. This filter effectively removes salt and pepper noise without diminishing the high detail boundaries and features of the image.
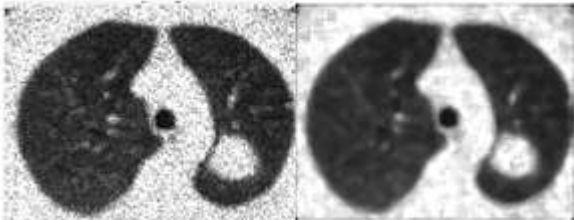

Fig. 5. Noisy image (left) and output of median filter (right)

As we can see from the Fig. 4. And Fig. 5., median filter does a better job at retaining the edges and boundaries in the noisy image than the averaging filter.

- *Image Enhancement*

Image enhancement is the process of improving the image for processing. We use two filters her for image enhancement namely Gabor filter and Fast Fourier transform.

Gabor Filter

This filter is used to recognize any pattern or disparity in the frequency content of the image, as illustrated in Fig. 6. This makes Gabor Filter a perfect candidate for our application. Gabor filters are orientation sensitive filters, and are used for texture analysis. This filter had an extremely useful property. This property is that it is very useful in feature extraction. Since feature extraction is one of the most important steps in our system, using of Gabor filter is essential to obtain our goal of segmenting the image and extracting necessary features from it. This is the reason we use Gabor filter.
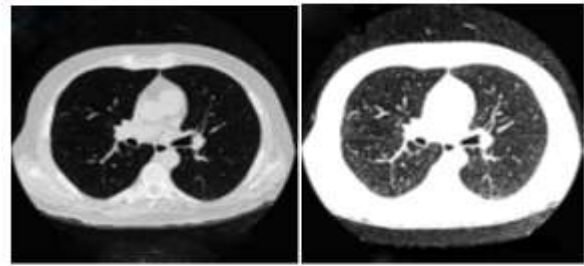

Fig. 6. Original image (left) and output of Gabor filter (right)

Fast Fourier Transform

Fast Fourier transform acts on the frequency domain of the image. It generally arbitrarily divides the image into blocks and acts on them as shown below in Fig. 7.
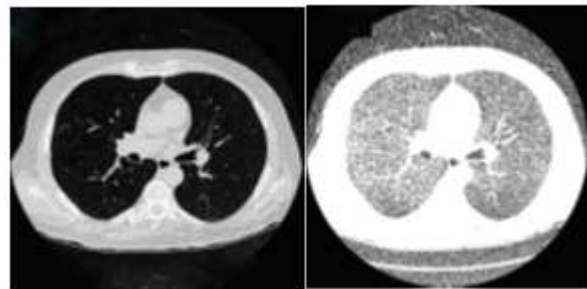

Fig. 7. Original image (left) and output of FFT (right)

After we have performed image enhancement, we move towards image segmentation.

C. Image Segmentation

Segmentation is the process in which a particular region of interest or the ROI is highlighted better than the rest of the image. Image segmentation is most important procedure for image analysis. It divides the image into multiple regions of constituent segments. Segmentation involves grouping of similar pixels, as segmentation is generally based on detecting disparities such as edges and boundaries. Segmentation techniques are generally based on two properties of the intensities of the pixels in the image-discontinuity and similarity.

The discontinuity based approach separates parts of the image at abrupt changes in the grayscale level of the pixels, such as at edges and boundaries. The similarity based approach groups together pixels of similar features such as similar or comparable grayscale values.

- *Thresholding*

This method of segmentation is a similarity based approach. The 256 levels of the grayscale value are converted into binary, which is either one or zero, represented as white and black respectively. Here a threshold value is defined. If the gray level value is greater than the threshold, it is mapped as binary one or white. If the gray level value is lower than the threshold, it is mapped as binary zero or black, as clearly shown in Fig. 8. This method produces a clear difference between the region of interest and neighbouring regions. We

use thresholding to highlight the region of interest, in our case, the tumour.
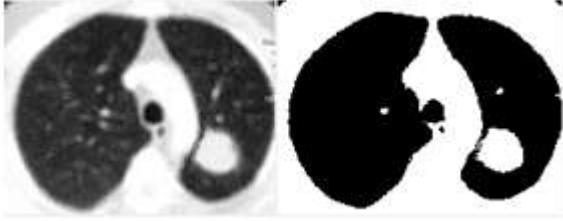
Fig. 8. Original image (left) and output of Thresholding (right)

D. Feature detection and extraction

This stage helps us to detection and isolate desired features of the region of interest. Feature extraction is an important step as it reduces bulky sets of data into simpler sets of features corresponding to each image. The most common features that are extracted to analyse a cancerous tissue are perimeter, area and eccentricity.

- *Area*

It is the actual number of all pixels present in the tumour portion, or the overall area of the ROI.

- *Perimeter*

It is the actual number of all pixels which are interconnected on the edges of the tumour, or the total length of the boundary of the ROI.

- *Eccentricity*

The roundness or matric value or irregularity index or circularity is one for circular shape and is less than one for other shape.

Depending on the area, perimeter and the eccentricity of the region of interest (ROI) or the tumour, the various stages of lung cancer can also be determined. The stages of lung cancer are I, II, III and IV. This is determined by the size and spread of the image. Thus area, and eccentricity can determine the stage of lung cancer, and thus its threat level. This is the reason why feature extraction is one of the most important stages.

E. Evaluation

After the image is segmented and the features are extracted, they are compared with an already existing database, to determine if the given image of lung has cancer or not. The process of evaluation is performed using fuzzy logic and deep learning techniques.
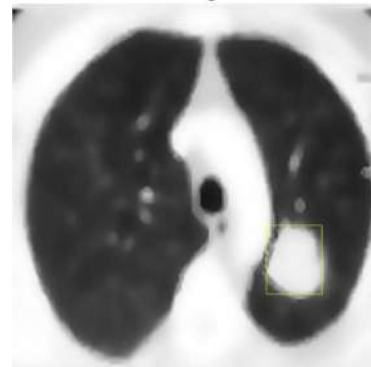
IV. RESULTS

Various results are obtained at the end of the program. Some of the results are given below.

Fig. 9. Original image
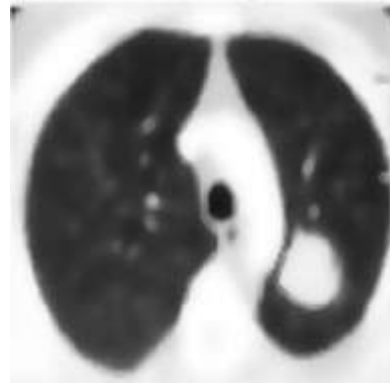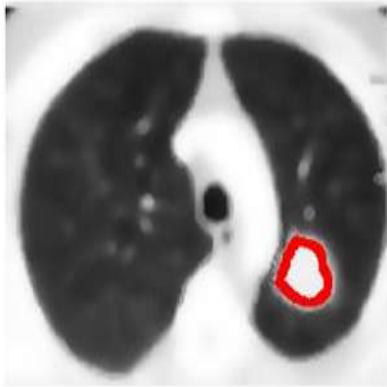

Fig. 10. Filtered image

Fig. 11. Bounding box

Fig. 12. Tumour alone

Fig. 13. Outline of tumour

Fig. 14.   Detected tumour

The input image, shown as Fig. 9, is filtered first as depicted in Fig. 10. The a bounding box as illustrated in Fig. 11. is applied to produce the outline of the tumour(Fig. 13) and the tumour alone(Fig. 12). Then the detected tumour is highlighted as illustrated in Fig. 14.

V. LIMITATIONS

May give false positive if the section of bones are too thick, or a false negative if the tumour image is hollow, i.e., density of tumour is low. False positives representing cells classified as cancerous result in putting a patient through unnecessary radiation and surgical operations. This can be quite problematic.

VI. CONCLUSION AND FUTURE SCOPE

Lung cancer is a widely prevalent type of cancer. Developing a system to identify cancer easily can be very beneficial in the medical field. The mentioned process has been completed till segmentation. Evaluation using fuzzy logic has been reserved for future work. The process can be extended to 3D images to overcome the above mentioned limitations. The proper anatomical position of the tumors can also be detected. Fixed thresholds were used for segmentation. Machine Learning may be implemented to train the system dynamically change the thresholds.

REFERENCES

1. Sujin Kim, Hyun-Joo Choi, Desok Kim, Hee Jae Joo, "Detection of Lobular Carcinoma In Situ(LCIS) by Image Analysis",2011, IEEE International Conference on Bioinformatics and Biomedicine Workshops
2. Tanushree Sinha Roy,Neeraj Sirohi, Arti Patle , "Classi_cation of Lung Image and Nodule Detection Using Fuzzy Inference System," International Conference on Computing, Communication and Automation (ICCCA2015)
3. Ritika Agarwal, Ankit Shankhadhar, Raj Kumar Sagar , "DETECTION OF LUNG CANCER USING CONTENT BASED MEDICAL IMAGE RETRIEVAL," 2015 Fifth International Conference on Advanced Computing and Communication Technologies
4. P.B.Sangamithraa, S.Govindaraju, "Lung Tumour Detection and Classification using EK-Mean Clustering," IEEE WiSPNET 2016 conference.
5. Shraddha G. Kulkarni, Sahebrao B. Bagal lung cancer tumor detection using image processing and soft computing techniques 2016, International journal of science
6. Mokhled S. AL-TARAWNEH, "Lung Cancer Detection Using Image Processing Techniques", Leonardo Electronic Journal of Practices and Technologies.
7. Maryamsadat Amini, Peng Liu,Scott E Umbaugh, Dominic J. Marino Catherine A. Loughin, "Thermographic Image Analysis Method in Detection of Canine Bone Cancer (Osteosarcoma)", 2012, 5th International Congress on Image and Signal Processing.
8. Monirul Islam Pavel, Ashique Mohammad Sadique, Rifat Ahmed Ritul, SynthiaKhan SnigdhaNath, "Cancer Detection using Image Processing Techniques Based on Cell Counting, Cell Area Measurement and Clump Detection", 2017, Department of Computer Science and Engineering, BRAC University.