Detection of Lung Cancer using SVM Classifier

Dr. D. Nagajyothi[1], Rakshith Addagudi [2], Tejaswini Gunda [3], Sindhu santhoshi Logitla[4]

[1]Associate Professor, Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Hyderabad, Telangana, India,d.nagajyothi@vardhaman.org

[2]UG Student, Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Hyderabad, Telangana, India,rakshith_16ec@vardhaman.org

[3]UG Student, Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Hyderabad, Telangana, India,tejaswini_16ec@vardhaman.org

[4]UG Student, Department of Electronics and Communication Engineering, Vardhaman College of Engineering, Hyderabad, Telangana, India,sindhusanthoshi_16ec@vardhaman.org

ABSTRACT

More number of deaths occurring every year due to a common health issue cancer. It is one of the diseases distributed over large area. Among all, lung cancer is most predominant having high death rate. Hence this is most important and a lot of research is going on and to identify lung cancer CT (Computed Tomography) scanned images are used. They provide a picture detailed with growth of tumor and tracks its growth in body so it is better by providing good results compared to other techniques. Therefore in medical fields most of the image processing techniques are used to detect the cancer in its early stages .In this paper we present an approach to detect lung cancer in scanned images. The steps involved in it is proposed using methods of image preprocessing where median filter is used followed by segmentation here mathematical morphological operations are used. Geometrical features like area, perimeter, and eccentricity are calculated to tumor detected part. At last SVM classifier is used to classify whether it is cancerous (Malignant) or normal (Benign).

Key words: CT, region of interest, SVM, Lung cancer

1. INTRODUCTION

Cancer is a widespread disease where deaths due to it are increasing nowadays. Among all lung cancer is predominant for men and women compared to all other cancers. It occurs because of uncontrollable cell growth in lung tissues. It is caused due to smoking and eating tobacco causing cancer in lung tissues. The cancerous nodules are called malignant tumors [1]. Lung cancer is the leading cause of cancer deaths worldwide. Despite advances in detection and diagnosis and therapies, many people still develop fatal disease. If cancer gets detected early, it's fatality rate decreases. Higher the detection time higher the fatality.

Lung Cancer is classified into two major groups based on the cell size i) cell size is small and ii) cell size is large. Cancer is divided into four stages and this staging has been done according to the size of tumor and node location. Computed Tomography scanned images are more effective to be used in detecting and diagnosing of lung cancer compared to the normal chest X-ray [2]. Cigarette smoking is top in the list for the factors causing lung cancer resulting in 85% and 75% cases of male and female respectively. Lung cancer causes 29% of entire cancer deaths. There is a higher chance of survival rate when the cancer is detected in it's early stages.

It is estimated that males are having 85% of cancer cases and females have 75% which occurs due to cigarette smoking. There are approximately 1,372,910 cancer cases in 2005 and deaths are expected to be 5,70, 280 to occur in United States due to cancer. It is also estimated that among those 1,63,510 deaths are caused due to lung cancer which constitutes 29% of total deaths.

The computed tomography scanned image is a combination of several different X-ray images which are taken from different angles so that using it we can see the internal organ part without cutting the organ i.e; inside the organ without cutting [6] [7]. CT scans provide better information about the images than conventional X-ray imaging, especially about blood vessels and soft tissue such as internal organs and muscles. It is an automated system so there will not be any human errors. It is also preferred over all other modalities as it has less distortion and is more reliable [3-5]. Here in this methodology we proposed a system that includes database, image preprocessing, image segmentation, extraction of features and a classifier.

2. METHODOLOGY

The methodology used for detecting lung cancer is CT image. Where this proposed system is consists of five stages and the detailed explanation about each stage is given in the below figure 1.
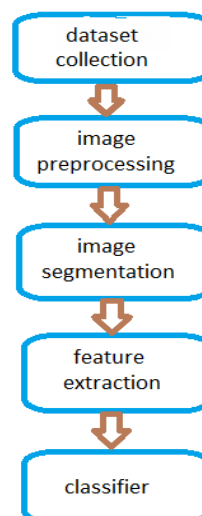

Figure 1: Proposed System

2.1. Data collection

It is the first step, where we store all the CT scanned images. Two varieties of images are taken from the database which contains the computed tomography images of patients having cancer and no cancer.

2.2. Image preprocessing

The main aim of this stage is to remove or suppress the noise or unwanted distortions and to enhance the image. Here we have two steps Image smoothing and Image enhancement. As CT images are taken out from a machine the image is added with some noise [8-10]. Image smoothing is used to eliminate noise. Here median filter is used to eliminate the unwanted distortions or noise. Median filtering is an effective technique as it removes noise while preserving the edges and doesn't blur the image.

To enhance the image we are using contrast adjustment so that affected quality of image caused due to variations in contrast is reduced [14]. It enhances contrast of image by transforming new values to input pixel values. By this data gets saturated default at low and high.

2.3. Image segmentation

In segmentation the required ROI (region of interest) of image will undergo an operation called segmentation which will separate the tumor part from the image [11-13]. Here we are going to use morphological operations. These are very powerful tools used to segment an image. Morphological opening operation is performed over the preprocessed image it firstly erodes the image and then it dilates the eroded image using the same structuring element for both the operations. Here we can also use marker controlled watershed segmentation to segment the image as it also works on morphological operations where we are using opening by reconstruction operation. In this we erode the image and use morphological reconstruction to segment the tumor part.

2.4. Feature extraction

From the segmented image we have to calculate the features of the tumor. Here we are going to calculate the features area, perimeter, eccentricity this is the major step in the whole process. All these features are geometrical features.

1. Area: It gives the no.of pixels that are acquired by value high and registered as 1 in lung nodules. It is a scalar quantity,i.e:taking the values and doing the summation

2. Perimeter: It is also a scalar quantity where it gives the number of pixels that are having high value at the borders of the lung nodule,i.e: having the pixel value as 1 at the border and summating them

3. Eccentricity: It is the ratio of the distance between the foci of the ellipse and its major axis length. Its value is between 0 and 1.

2.5. Classification

This stage involves classifying the cancer type whether it is Malignant or Benign. Here the classifier used is SVM classifier [15]. SVM classifiers are supervised learning models that are identified using their pattern. Support vector machines (SVM) can be used when our data has exactly two classes. The best hyper plane is found to classify the data, it creates two classes with different data points. So that we can find the best hyper plane of classifier with the largest margin between the two classes.

3. RESULTS

The results of proposed system are implemented using MATLAB software. Here we are going to use Arduino Uno and LCD display also in order to print the type of cancer on the screen display i.e, whether it is malignant or benign. The images are stored in the dataset consisting of images with cancer and without cancer and the images in the dataset look like the follows.
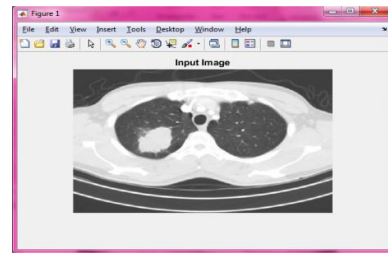

Figure 2: Original image

The median filter will be applied to the input image to remove the noise from it. As the input is a scanned CT image it consists of salt and pepper noise to remove it we are using median filter.
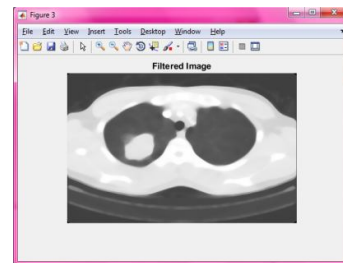

Figure 3: Filtered image

Now to enhance the filtered image we have used contrast adjustment technique.

In segmentation stage the input image is divided into several parts to give the required ROI (region of interest) for the next stage. In order to extract required ROI we are using the technique called morphological operations based on structuring element then it is masked to extract the tumour region from scanned CT image.
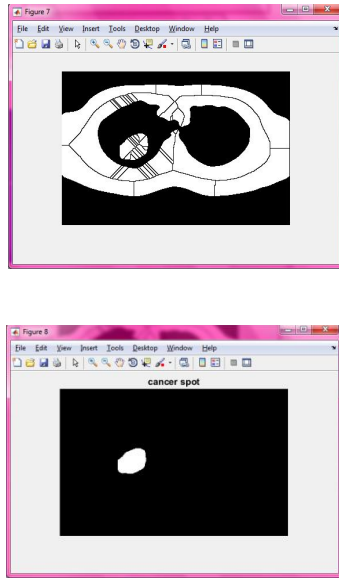
Figure 4: Segmented image

From the acquired data after segmentation we got the tumor region. Where we have to calculate the geometrical features (table 1) like area, perimeter and eccentricity and these features send to classifier as inputs to justify whether the cancer is malignant or benign

Table 1: Geometrical Features

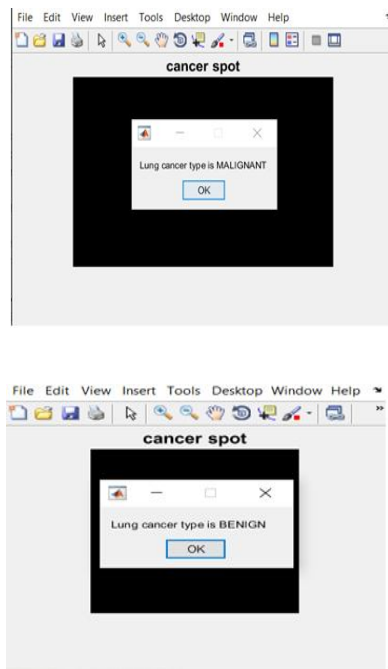| S.No	Features	Values
1	Area	1024
2	Eccentricity	0.78
3	Perimeter	196.3

Figure 5: MATLAB results

After feature extraction, Support Vector Machine (SVM) classifier is used to determine whether the CT scan images are malignant or benign.

After classification the result what we have got is also printed on LCD display using Arduino Uno.

4. CONCLUSION

The most extensive disease in the world is lung cancer. It ordinarily begins in bronchi lined by the cells and other parts of the lung. Detection time plays major role in diagnosing the disease. Detection method and diagnosing methods can be used to decrease the spread to the world. For more accuracy the work that has to be done into three stages 1.Enhancing the image, 2.Segmenting the image and the other is 3.Extracting the features. Scanning gives most effective results. Lung nodule detection in CT scanning plays a major role in research. This technique is emerging, even many enchantment techniques that can be included to make them work more efficiently.

REFERENCES

1. Nidhi S. Nadkarni, Sangam Borkar. "Detection of Lung Cancer in CT Images using Image Processing" , 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019
https://doi.org/10.1109/ICOEI.2019.8862577

2. Anita Chaudhary, Sonit Sukhraj Singh, **Lung Cancer Detection on CTImages using Image Processing**, International conference computing sciences, IEEE, 2012.
https://doi.org/10.1109/ICCS.2012.43

3. Nooshin Hadavi, Md Jan Nordin, Ali Shojaeipour, **Lung Cancer Diagnosis using CT-scan Images based on Cellular Learning Automata**, International conference on Computer and Information Sciences(ICCOINS), IEEE , 2014.

4. K.Punithavathy, M.M.Ramya, Sumathi 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), IEEE, 18-20 February 2015.
https://doi.org/10.1109/RACE.2015.7097244

5. Anjali Kulkarni, Anagha Panditrao, **Classification of Lung Cancer Stages on CT Scan Images Using Image Processing**, International Conference on Advanced Communication Control and Computing Technologies (lCACCCT),IEEE,2014.

6. Gawade Prathamesh Pratap, R.P Chauhan, **Detection of Lung Cancer Cells using Image**

Processing Techniques, 1st IEEE International Conference on Power Electronics, Intelligent control and energy systems, IEEE, 2016.

7. Moffy Crispin Vas, Amita Dessai, **Classification of Cancerous and Non-cancerous Lung Cancer Nodule using Image Processing Technique** International Conference on Academic Research in Engineering and Management(ICAREM-17).

8. Shubhangi Khoragade, Aditya Tiwari, **Automatic Detection of Major Lung diseases using Chest Radiographs and classification by Artificial Neural Network**, 1st IEEE International Conference on Power Electronics, Intelligent control and energy systems, IEEE, 2016.
https://doi.org/10.1109/ICPEICES.2016.7853683

9. Pooja R. Katre, Dr. Anuradha Thakare, **Detection of Lung Cancer Stages using Image Processing and Data Classification Techniques**, 2nd International Conference for Convergence in Technology (I2CT) 2017 .

10. S. Kalaivani, Pramit Chatterjee ,Shikhar Juyal ,Rishi Gupta,**Lung Cancer Detection Using Digital Image Processing and Artificial Neural Network**, International Conference on Electronics, Communication and Aerospace Technology ICECA 2017.

11. K.Gopi, Dr.J.Selvakumar, **Lung Tumor Area Recognition and Classification using EK-Mean Clusstering and SVM**, International Conference on Nextgen Electronic Technologies,IEEE 2017.
https://doi.org/10.1109/ICNETS2.2017.8067906

12. G.Niranjana, Dr.M.Ponnavaikko, **A Review on Image Processing Methods in Detecting Lung Cancer using CT Images**, International Conference on Technical Advancements in Computers and Communications,IEEE 2017. Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019).

13. Sayali Satish Kanitkar, N.D.Thombare,**Detection of Lung Cancer Using Marker-Controlled Watershed Transform**, International Conference on Pervasive Computing(ICPC) 2015
https://doi.org/10.1109/PERVASIVE.2015.7087031

14. Amal Fouad, Hossam M. Moftah, Hesham A. Hefny, **MRI Brain Cancer Diagnosis Approach Using Gabor Filter and Support Vector Machine,** International Journal Of Emerging Trends in Engineering Research (IJTER), Vol. 7, No. 12, pp. 907-914, December 2019.
https://doi.org/10.30534/ijeter/2019/297122019

15. Naveen Kishore Gattim, Subba Reddy Pallerla, Polaiah Bojja, Tallem Pavan kumar Reddy, Vempati Nikhil Chowdary, V Dhiraj, Sk Hasane Ahammad, **Plant Leaf Disease Detection Using SVM Technique**, International Journal Of Emerging Trends in Engineering Research (IJTER), Vol. 7, No. 11, pp.634-637, November 2019.
https://doi.org/10.30534/ijeter/2019/367112019