Lung Image Segmentation using Modified K-Means Algorithm


Abstract: Lung Cancer is also referred as Lung Carcinoma, characterized by unrestrained cell growth in tissues of lung. It has high mortality rate when compared to other cancers. The main reason of Lung Cancer is smoking and exposure to secondhand smoke. A fine Lung Cancer detection system must sense the Lung Cancer in its premature stages. Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are the two tools used to capture the Lung image. The various stages in the Lung Cancer detection include Image Capturing, Image Enhancement, Image Segmentation and Feature Extraction. In this, various image processing techniques are utilized for lung cancer detection and performance of each technique is compared.

Keywords: CT, MRI, K-means clustering.

I. INTRODUCTION

The Lungs are the two spongy organs in our chest that take in oxygen when we inhale and release carbon dioxide when we exhale. The primary types of Lung Cancer are small cell Lung Cancer and non-small cell Lung Cancer. The causes include genetic and non-genetic reasons. Lung Cancer is the major cause of cancer deaths among both men and women. Lung Cancer claims more lives when compared with colon, prostate, ovarian and breast cancers. The United States has the high mortality rate of 165,000 every year. In a survey more than 80% males and more than 70% females Lung Cancer is due to Cigarette Smoking.

Every year in India 63,000 new Lung Cancer has been reported. According to the survey conducted by ‘The Hindu’ in 2010, across the country 5,56,400 people were killed by cancer. In 2010, 8 per cent of the 2.5 million total male deaths and 12 per cent of the 1.6 million total female deaths between 30 to 69 years were accounted by cancer alone.

Lung Cancer takes more lives and the rate rises each year. Only when the disease is found accurately in its early stages, proper medical support will be provided and hence, secure the lives. Its detection in first stage increases the lifespan of the infected persons.

Fig. 1.1 Survival rate chart

From the above chart, the survival rate is increasing every year. By providing proper medical treatment the life span of a patient can be increased up to 5 years.

II. PREVIOUS WORK

Omar, Watson [2] considered on the quality of the section of interest. Fractal examination is utilized to recognize forceful (propelled stage) and nonaggressive (beginning time) tumors by utilizing CE-CT(Contrast Enhanced Computed Tomography). These images were represented by time sequence. The tumor is detected in its early stages by determining the malignant tumor. In this paper, images were acquired by DICOM. The images were modified into a Fractal Dimension images by Differential Box Counting (DBC). After the fractal transformation, the ROI can be chosen manually. The refinement among forceful and non forceful tumor exactness was up to 83%. The data about the forcefulness is additionally examined in this paper.

A.Amutha, Wahidadabanu [5], proposed a set Active contour model to analyse the lung tumor. This works only on 2-D images. This is the main drawback of this system. This paper relies upon the portion work which has the base mean square mistake esteem. The second request highlights were determined. It relies upon the histogram of the clamour free picture. The plan between the normal and abnormal lung image has done dependent on the above highlights.

Anam, Usman, Younus [6], acquainted a technique with evacuate the commotion content. The clamor content is sifted by utilizing middle channel. The foundation was expelled by utilizing inclination mean and change technique. Ideal thresholding is utilized for the division purpose. Then the undesirable data is expelled utilizing diverse morphological tasks.
The segment of intrigue is separated first and after that the five surface based highlights were determined. The vector made by these component were given to the half and half classifier. A blend of self arranging system and multilayer discernment frames a cross breed classifier. The computational time for bigger informational collection was more. This is the principle disadvantage of this framework.

Dasu Vaman Ravi Prasad [8] - In this paper picture quality and exactness are the vital variables of this examination. Picture quality examination just as enhancement relies upon the improvement organize where low pre-preparing procedures are utilized dependent on Gabor channel inside the Gaussian tenets. In light of the division standards, an extreme area of the object of intrigue that is utilized as an essential establishment of highlight extraction is acquired. Contingent upon the general highlights, a typicality correlation is finished. In this, the fundamental highlights detected for precise picture correlation are pixels rate and cover marking.

### III. PROPOSED WORK

In this paper, we have arranged a framework with a few alterations to enhance the exactness and productivity of a picture. The initial step is to evacuate the commotion content utilizing middle channel, a non-direct advanced sifting technique to enhance the highlights of picture. Middle separating replaces the pixel esteems with the middle of pixel esteems. Extraction of lung district is finished by ideal thresholding. At that point morphological capacities, for example, opening, shutting, reproduction, edge identification and gap filling are performed to expel the unimportant data in the lungs. At that point the area of intrigue is separated. Different textural and factual parameters are removed to isolate ROI knob from the other structure of lungs. Highlights arrangement is finished with the assistance of back proliferation organize classifier. In this, we utilized the K-means grouping calculation, it is to bunch the perceptions into a particular number of disjoint perceptions. It isolates a gathering of information into a K number gathering of information. It orders a given arrangement of information into K number of disjoint group. The calculation goes for limiting the separation between the centroid of the bunch and the given perception iteratively to any group and end when the least separation is accomplished. In this, the most ideal approach to ascertain the separation is by the utilization of squared Euclidean distance which squares the separation while the Euclidean separation relates to the base separation between two points.

---

**Fig. 1.2 Flowchart**

Based on the classification results, tumor is determined. The efficiency of the techniques is given as follows.

**Table 1.1 Comparison table**

<table>
<thead>
<tr>
<th>Method</th>
<th>Correlation</th>
<th>Structural similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>K means</td>
<td>0.9838</td>
<td>1.0000</td>
</tr>
<tr>
<td>Global threshold</td>
<td>0.9470</td>
<td>1.0000</td>
</tr>
<tr>
<td>Fuzzy c means</td>
<td>0.9523</td>
<td>1.0000</td>
</tr>
<tr>
<td>Watershed</td>
<td>0.9910</td>
<td>1.0000</td>
</tr>
<tr>
<td>Deformable model</td>
<td>0.7666</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

**Fig. 1.3 Median filter**
IV. CONCLUSION

Various lung cancer detection techniques had been discussed in this paper. The chief aim of all the systems is to achieve the precision as maximum as possible with less false value. It has been concluded that K-means algorithm is the best method for image segmentation. By increasing the accuracy, the lung cancer can be detected in its early stages and we can increase the lifespan of a person.

V. FUTURE IMPLEMENTATION

From the above chart, watershed method is considered as the most preferred method. In watershed algorithm, RGB color space is used. HSV color space is more complex and it has more advantages than RGB. By using HSV color space in watershed algorithm, we can increase the efficiency of the watershed algorithm. Segmentation can be done more efficiently by developing simpler algorithms.

REFERENCES