International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August 2019

Automatic Lung Cancer Detection using Sobel & Morphological Operations

Akanksha Soni, Avinash Rai, Ekta Shivhare

Abstract: ***Cancer is the most dangerous disease that may cause death and lung cancer is one of them which is more common among all. There are various imaging techniques through which organs can be scanned for diagnosis. Lung cancer is a disease that may be caused by unrestrained cell growth in lung. Lung cancer is the most common and most dangerous cancer. CT scan can obtain the lung images, but still it has been recognized manually. Manual lung cancer detection is a challenging task because false error rate may lead you to compromise with human's life. There are lots of researches that has been done in this field but still failed to obtain high precision with minimal error rate. Here the system proposes automatic lung cancer detection using Sobel & Morphological operations that can acquire good precision along with cancer area detection. Sobel is a gradient edge detection technique through which absolute gradient magnitude is computed in the reference of 2D input lung image that is later dilated with morphological operator. The obtained result is liable to attain high precision with less false alarm rate.***

Keywords: ***Lung Cancer Detection, Sobel, Dilation, Gradient Magnitude, Computed Tomography, Morphological Operations.***

I. INTRODUCTION

CT-Scan is firstly used for scanning lung cancer images, it is a method of medical imaging for tomography to produce three-dimensional images in one axis, it works like a digital x-ray where hard tissues looks lighter whereas soft tissue as darker [1]. Nowadays image processing technology is the most widely used area in medical science for detecting various cancers or tumor. Most of the lung cancer disease occur because of tobacco smoke up to the ratio of 85 to 90% and rest 10 to 15% due to other reasons. These cases are now regularly formed by a mix of hereditary components and presentation to asbestos and different types of air contamination [2]. Detecting lung cancer is not an easy task, it is generally handled by doctors and diagnosis is done manually.

Revised Manuscript Received on August 30, 2019.
* Correspondence Author
Akanksha Soni*, Electronics & Communication, University Institute of Technology, RGPV Bhopal, Madhya Pradesh, India. Email: akansha.sonivits@gmail.com
Dr. Avinash Rai, Electronics & Communication, University Institute of Technology, RGPV Bhopal, Madhya Pradesh, India. Email: avinashrai@rgtu.net
Ekta Shivhare, Electronics & Communication, University Institute of Technology, RGPV Bhopal, Madhya Pradesh, India. Email: ektashivhare@gmail.com

Here the system has been proposed which is able to detect cancer automatically from an input image with less processing time.

System uses Sobel edge detection technique through which edges are extracted and smoothen area can be determined for masking unwanted area and only focusing over region of interest (ROI). Dilation is also used to fill small holes and wider the lighter area for extracting ROI or to obtain cancer area from lung image. Lung image is already in black and white color mode, nevertheless it will be further converted into grayscale image and binary image for image enhancement using various filtration techniques. There are various intermediate states for processing a lung image to obtain cancerous spots. If a spot is obtained then the contrast gets increased automatically, that results into cancerous lung otherwise it is a healthy lung. System possesses higher precision rate with less false alarm rate.

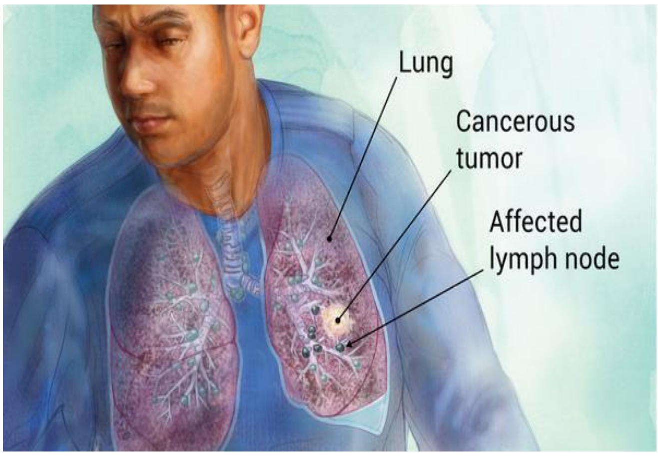


Fig. 1. Human Lung [3]

Lungs are a pair of spongy organs in chest which inhale oxygen when breathe and exhale carbon dioxide. Lung cancer is a disease of abnormal cells multiplying and growing as a tumor. The main cause of cancer deaths in both men and women in the India is lung cancer. Lung cancer claims more life each year than colon, prostate, ovarian and breast cancer. People who smoke are at the greatest risk of lung cancer. With increase in the consumption of cigarettes, the chances of lung cancer also increase. If a smoker quit smoking after many years, he can greatly reduce the chances of developing lung cancer [4]. Qing Wu et al. proposed a system which is based on supervised machine learning algorithm and the result has been obtained from entropy degradation method (EDM).

An EDM along with histogram features to detect small cell lung cancer (SCLC) for early cancer prediction. According to the authors EDM is reasonably good for prediction accuracy. EDM has been designed with the concept of shallow neural network, which transforms the vectorized histogram of each training set into a score.
The ultimate goal of this study is to develop a clinical based decision-making system for radiologist to better predict lung cancer with computed tomography (CT) imaging [10].

Fig. 2. Healthy Lung [11]

Fig. 3. Cancer Affected Lung [11]

Fig.2 and Fig.3. Represents healthy and cancer affected lung respectively.

II. PROBLEM IDENTIFICATION

The system has been trained only with 12 lung CT scans images where 6 images have been taken from healthy lungs and 6 from SCLC. There are 5 scans for each image set for training and the algorithm achieved 77.8 percent of accuracy which would be bit higher. The SCLC images are labeled as cluster-1 and the other as cluster 0. Because not all CT scans images reveals cancer cells. For testing, they choose two additional CT images, one for a patient with SCLC and other without SCLC. Here the problem has been identified with the accuracy and the dealing technique that jounce system towards low accuracy [10].

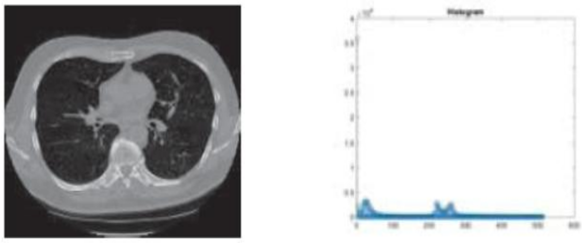


Fig.4. Cluster 0 & Histogram of Healthy Lung [10]

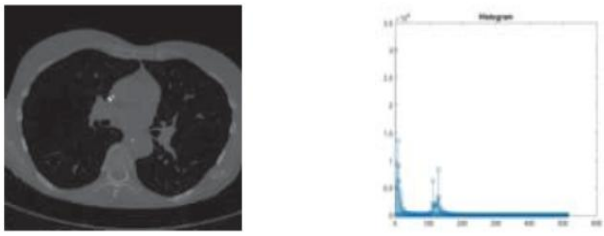


Fig.5. Cluster 1 & Histogram of Lung Cancer [10]

Fig 4 and Fig 5. Represents the histogram of healthy lung image and cancerous image respectively. But it seems, there is a bit difference between both the images and their histograms. A spot in the image can be a part of nerve or any blood vessel. Vessels and nerves carrying fluid inside lung might be mistaken as cancer, if it is over there. This may cause false alarm rate along with moderate precision and recall in the reference of true positive and false positive.

III. PROPOSED WORK & IMPLEMENTATION

Proposed system is able to classify cancerous images as well as non-cancerous images. System is based on Sobel Edge Detection and dilation technique that enhances cancerous spot extraction for better prediction. There are various phases including histogram equalization, segmentation and thresholding, sobel edge classification with inverse matrix i.e. smoothening, dilation, image filling and cancer area extraction. System is capable enough to classify the nerves or background area of lung and provide better level of precision.

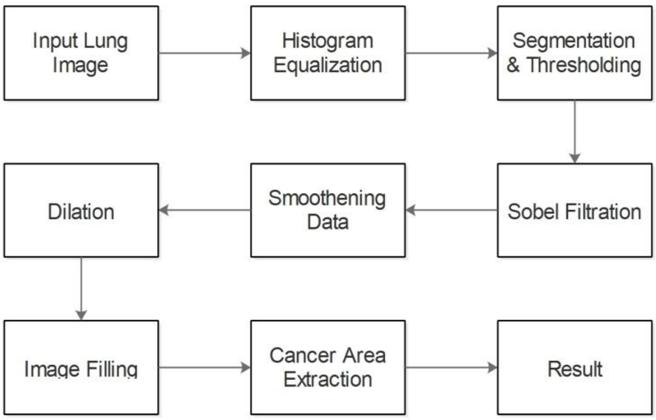


Fig.6. Block Diagram

A. Histogram Equalization

Histogram Equalization is an image processing technique for adjusting contrast in images. The best example where histogram equalization is very useful is x-ray that leads to better representation of the bone structure by increasing the intensity values of bone photographs. Histogram equalization can be done by effectively spreading out the frequent intensity values. In this paper, histogram equalization is used for adjusting contrast by reduction that resulted classifications of lung nerves background and image background that need to be masked in segmentation.

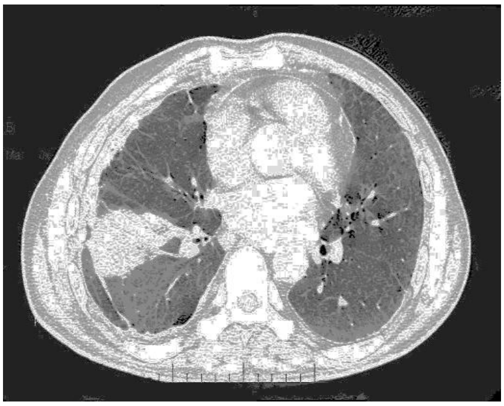


Fig.7. Histogram Equalization

B. Segmentation by Thresholding

Thresholding is a method of image segmentation. The thresholding may replace each pixel in an image with a black pixel if the image intensity is less than the threshold T or a white pixel if the image intensity is greater than T. Automatic thresholding is a great way to extract useful information encoded into pixels with minimizing background noise. So, T can be selected automatically and it can be either background pixel or foreground pixel. It can be generated by computing mean of background pixels or foreground pixels. Here in this system, it has been applied for masking nerves and image background where all low pixel intensities that are less than the threshold pixel should be masked for extracting region of interest.

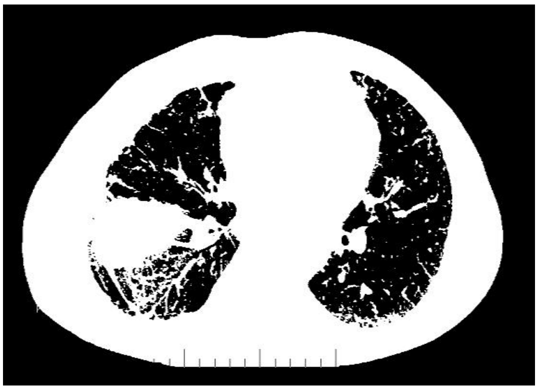


Fig.8. Segmentation by Thresholding

C. Sobel Edge Detection

Sobel is a filter in image processing for extracting edges or slight changes in pixel intensities. Sobel is a digital filter which helps to extract edges on the basis of angular matrix. Lesser the value of gray level, darker the area and greater the value of gray level, lighter the area. The color range lies between 0-255. Typically, it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. By the help of Sobel edge detection an object can be detected from an image with high level of accuracy. Sobel is better than all other edge detection technique due to its edge detailing. Consider A as an input matrix-

A =

a11	a12	a13	...	
a21	a22	a23	...	
a31	a32	a33	...	
...	...	...	...	

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A, \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

Where A as an input 2D image array or matrix, G_x and G_y are the gradient kernel that will be multiplied with input image array A. Where G_x is the horizontal gradient and G_y is the vertical gradient. Negative gradients appear darker, and positive gradients appear brighter. Computing the value at each pixel by shifting the row towards right till the end row has been reached. The example below shows the calculation of a value of G_x:

a11	a12	a13	…	
a21	a22	a23	…	
a31	a32	a33	…	
…	…	…	…	

Input Matrix

b11	b12	b13	…	
b21	b22	b23	…	
b31	b32	b33	…	
…	…	…	…	

Output Matrix

Kernel =

+1	0	-1
+2	0	-2
+1	0	-1

$$b_{11} = a_{11}*1+a_{12}*0+a_{13}*(-1)+a_{21}*2+a_{22}*0+ a_{23}*(-2)+ a_{31}*1+ a_{32}*0+ a_{33}*(-1)$$

Similarly, each pixel will be calculated according to the kernel matrix and finally G_x has been computed.

The example below shows the calculation of a value of G_y:

Kernel =

+1	+2	+1
0	0	0
-1	-2	-1

$$b_{11} =a_{11}*1+a_{21}*0+a_{31}*(-1)+a_{12}*2+a_{22}*0+ a_{32}*(-2)+ a_{13}*1+ a_{23}*0+ a_{33}*(-1)$$

Retrieval Number F8834088619/2019©BEIESP
DOI: 10.35940/ijeat.F8834.088619
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

At each pixel in the image, the gradient approximations given by Gx and Gy are combined to give the gradient magnitude, using:

$$G = \sqrt{G_x^2 + G_y^2}$$

Fig.9. Sobel Edge Detection

Once the sobel has been obtained, compute its inverse matrix for extracting smoothen data G^{-1}.

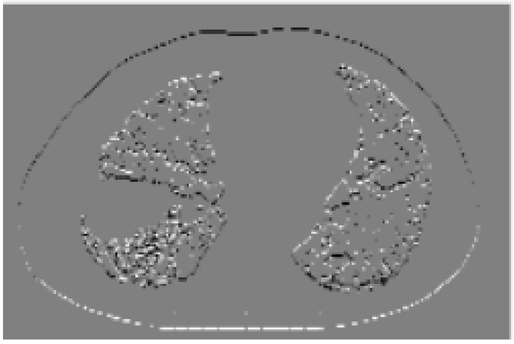


Fig.10. Smoothen Data

D. Morphological operation

Morphology is a set of image processing operations that process the images on the basis of their shape. In a morphological operation, each pixel value in the output image is based on a comparison of the corresponding pixel in the input image with its neighbors. The most common morphological operations are erosion and dilation. Erosion removes pixels to the boundaries of objects in an image, while dilation adds pixels on object boundaries. In this paper, dilation has been applied for filling small holes that may contain information.

1	1	0	1	1
0	0	0	1	0
1	1	1	1	0

1	1	1	1	1
0	0	0	1	0
1	1	1	1	1

System does not fill those cells that contain consecutive 0. It only fills those 0 where consecutive 1 is present in a row or column.

E. Filling Regions & Holes

It works like dilation iteration where whole image is dilated again for filling holes and region but not using dilation function, instead of that imfill function has been used for this operation. It is a flood fill operation where a region is filled if neighboring pixels intensities are same.

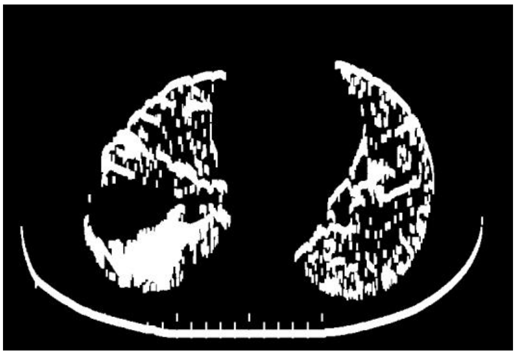


Fig.11. Filling Regions & Holes

F. Cancer Area Extraction

Cancer area can be extracted using a comparison between the output image of the proposed system and healthy lung image. If image contain big uncertainties (holes), that means cancerous area is present in the output image. This is done by foreground extraction in the field of image processing.

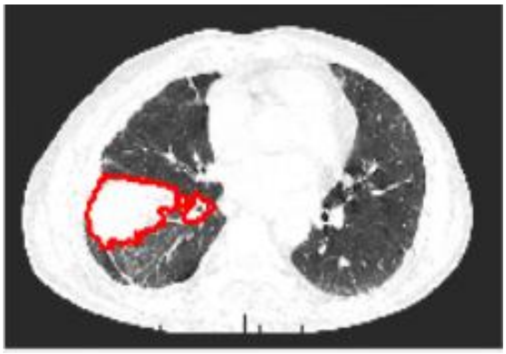


Fig.12. Cancer Area Extraction

The flow chart explains the same procedure which is explained above. First of all; an input image will be preprocessed with histogram equalization where contrast is adjusted for classifying nerve background and unwanted image background. Image background should not affect the accuracy of the system, that is why it is required to mask before analyzing final result. After that segmentation and thresholding will finally classify the unwanted background by deleting small pixels which are less than the threshold value, so the darker region gets more darker and lighter region gets lighter. Then sobel edge detection will be applied for extracting edges and inverse of the result is obtained.

Retrieval Number F8834088619/2019©BEIESP
DOI: 10.35940/ijeat.F8834.088619
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

So, smoothen data has been achieved, that equalize cancer area with background. Then dilation process fills the small holes where consecutive 1 appears. Finally, image filling smoothens the background and highlight the cancer spot area in the final resulting image. If contrast is greater than the threshold value, it means that image contains cancer part, otherwise no cancer will be detected.

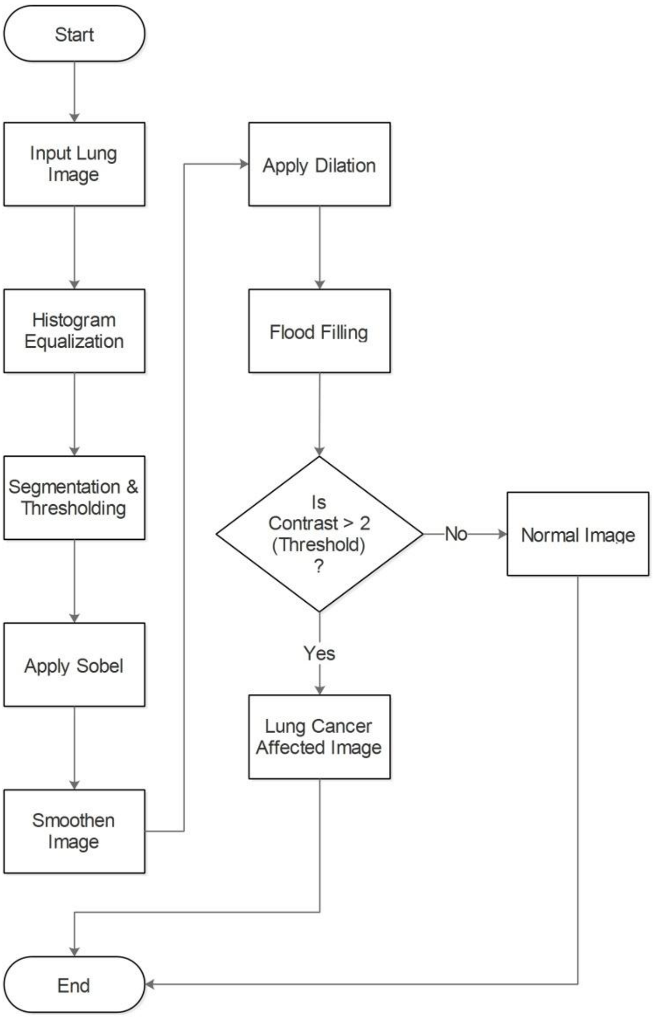


Fig.13. Flow Chart

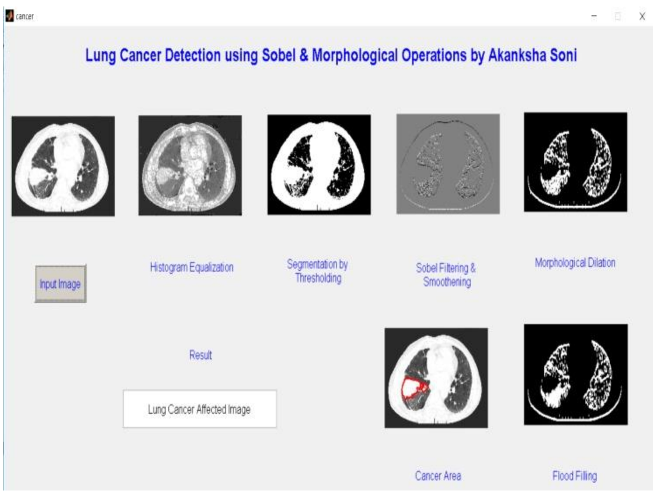


Fig. 14. Graphical User Interface – Proposed system

IV. RESULT ANALYSIS

The result has been analyzed on the basis of true positive, true negative, false positive and false negative. Total number of images are 72 that have been tested. Here 30 are true positive, which means that 30 lung images are cancerous images; recognized as cancerous image by the system. Similarly, 32 false negative where cancer does not exist and indeed it has not been detected by the system. Also 6 are true negative where cancer exist but not detected by the system and 4 are false positive where cancer does not exist but cancer is detected by the system. So, on the basis of the above-mentioned data, over all accuracy has been computed as 86.11 % of the proposed system.

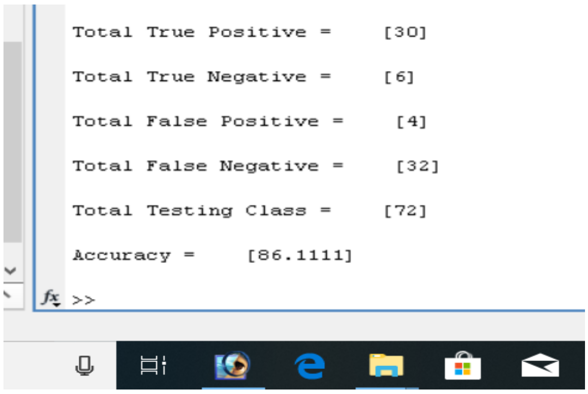
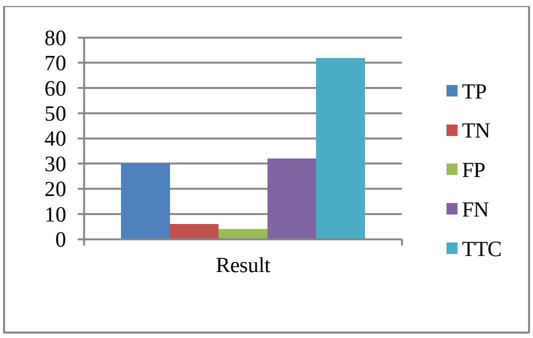


Fig.15. Flow Chart

Table- I: Result Analysis

Terms	Proposed
Total Testing Class (TTC)	72
True Positive (TP)	30
True Negative (TN)	6
False Positive (FP)	4
False Negative (FN)	32

$$Accuracy = \frac{TTC - (TN + FP)}{TTC} \times 100\ \%$$



Graph 1. Result Analysis Graph

Graph 1 shows the result outcomes on the basis of 72 testing calsses. The accuracy is computed accordingly.

Table- II: Result Comparison

Terms	Qing Wu [10]	Proposed
TTC	72	**72**
TP	30	**30**
TN	26	**6**
FP	10	**4**
FN	6	**32**
Accuracy	77.8 %	**86.11 %**

V. CONCLUSION

The current proposed system is able to detect lung cancer from CT-Scan images with less processing time and acquire high level of accuracy as compare to the Entropy Degradation Method (EDM). System is also able to classify the cancer area with minimal error rate. System uses Sobel Edge Detection and Dilation techniques for acquiring such accuracy which is higher than the EDM. The system proposed is developed in MATLAB and accuracy of the proposed system is 86.11 %.

VI. FUTURE SCOPE

In the field of medical science, diagnosing diseases through image processing is now become in trend that indeed saves time as well as human's life. The above proposed system can be enhanced in future by implementing it with different techniques and filters, which may acquire good accuracy and minimal false alarm rate. Because as per the ideal system, accuracy is an important parameter in detection of cancer, that is why accuracy of system can be enhanced in future with different techniques or filters.

REFERENCES

1. R. Wulandari, R. Sigit and S. Wardhana, "Automatic lung cancer detection using color histogram calculation," *2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC)*, Surabaya, 2017, pp. 120-126.
2. P. B. Sangamithraa and S. Govindaraju, "Lung tumour detection and classification using EK-Mean clustering," 2016 *International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Chennai, 2016, pp. 2201-2206.
3. Apollo Hospital Sources, "Lung Cancer", [Online]. Available: https://g.co/kgs/pxVUUg, [Accessed: 28-July-2019]
4. Mayoclinic, "Lung cancer", April 02, 2019. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/lung-cancer/symptoms -causes/syc-20374620, [Accessed: 28-July-2019].
5. M. B. A. Miah and M. A. Yousuf, "Detection of lung cancer from CT image using image processing and neural network," 2015 *International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Dhaka, 2015, pp. 1-6.
6. G. P. Pratap and R. P. Chauhan, "Detection of Lung cancer cells using image processing techniques," 2016 *IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, Delhi, 2016, pp. 1-6.
7. S. Kalaivani, P. Chatterjee, S. Juyal and R. Gupta, "Lung cancer detection using digital image processing and artificial neural networks," *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, 2017, pp. 100-103.
8. Y. Yin *et al.*, "Tumor Cell Load and Heterogeneity Estimation From Diffusion-Weighted MRI Calibrated With Histological Data: an Example From Lung Cancer," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 1, pp. 35-46, Jan. 2018.
9. Selin Uzelaltinbulat, Buse Ugur, "Lung tumor segmentation algorithm", Procedia Computer Science, Volume 120, 2017, Pages 140-147, Springer.
10. Q. Wu and W. Zhao, "Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm," 2017 *International Symposium on Computer Science and Intelligent Controls (ISCSIC)*, Budapest, 2017, pp. 88-91.
11. Dr. Sandeep Bansal, Lung Images Data Collection, Nobel Imaging & Diagnostic Center, Bhopal, [Collection Date: 23-May-2019].

AUTHORS PROFILE

Akanksha Soni has done B.E. (Electronics & Communication) and is currently an M.E. scholar (Digital Communication) in the Department of Electronics & Communication Engineering at University Institute of Technology,Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (M.P). Akanksha Soni has undertaken research work presented in this paper under the guidance of Dr. Avinash Rai which is required in the partial fulfillment for the award of M.E. in digital communication.

Dr. Avinash Rai has done B.E. (Electronics), M.E. (VLSI Design), Ph.D. (Wireless Sensor Network) and is currently working as Assistant Professor in Department of Electronics & Communication Engineering at University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (M.P). Dr. Avinash Rai does researches in Wireless Sensor Network, Telecommunication Engineering and Electronics Engineering. His most recent publication is "Enhancing Energy of Sensor Node to Increase sEfficiency in Wireless Sensor Network"

Mrs. Ekta Shivhare has done B.E. (E.I), M.E. (ECE) and is currently pursuing Ph.D. in the field of Histopathological image processing. Mrs. Ekta Shivhare is currently working as Assistant Professor in Department of Electronics & Communication Engineering at University institute of technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya Bhopal (M.P.)