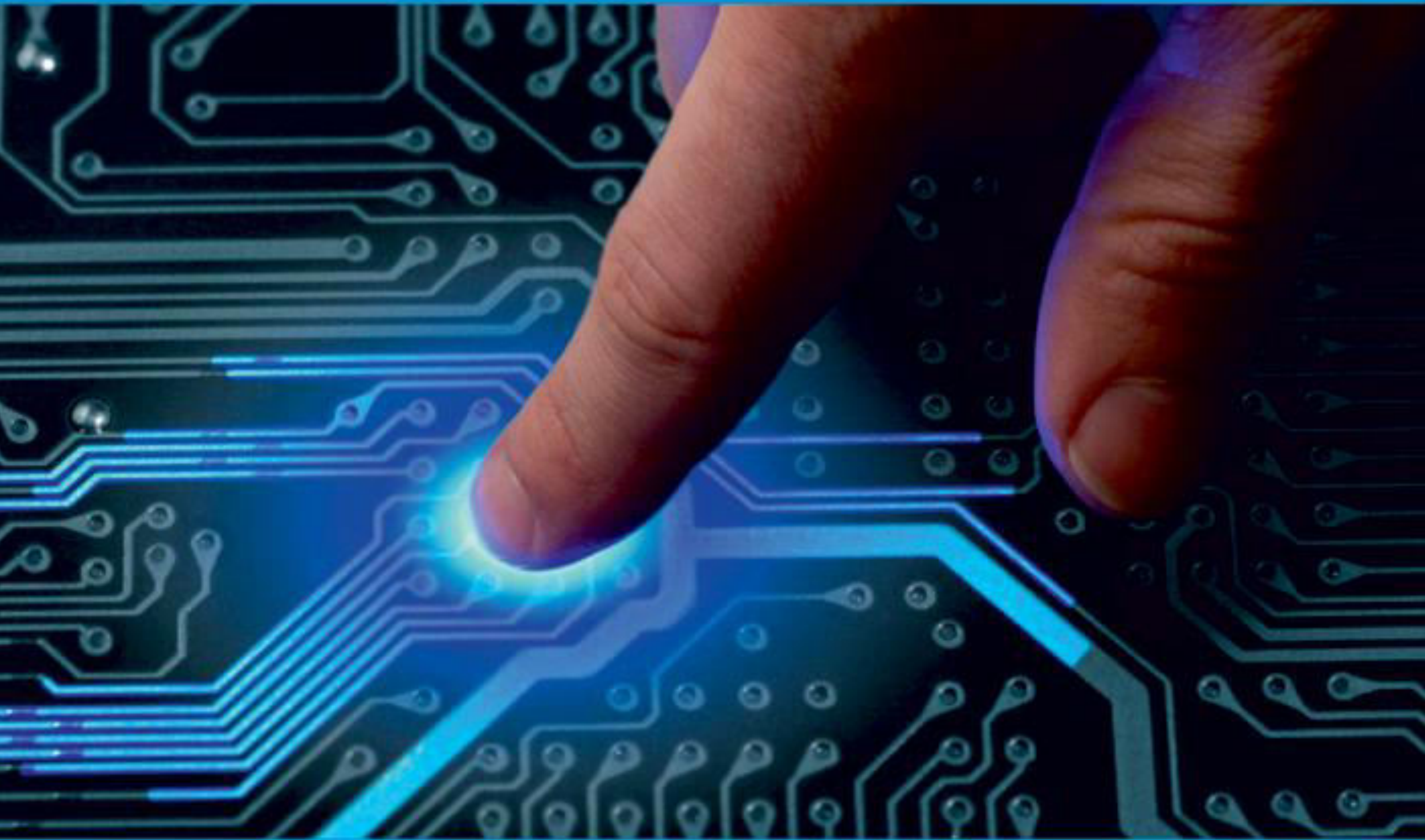
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 9, September 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 9, September 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1009032 |

Lung Cancer Detection using Chest CT-Scan Images for Non-Smokers by using Watershed Transform with SVM Classifier

B. Mohamed Faize Basha[1], Dr. M. Mohamed Surputheen[2]

Assistant Professor, PG & Research Dept. of Computer Science, Jamal Mohamed College (Autonomous)

[Affiliated to Bharathidasan University, Tiruchirappalli], TN, India[1]

Associate Professor, PG & Research Dept. of Computer Science, Jamal Mohamed College (Autonomous)

[Affiliated to Bharathidasan University, Tiruchirappalli], TN, India[2]

ABSTRACT: Lung Cancer is a kind of disease that spreads quickly and it is the main source of getting more mortality rate in worldwide. The Computer Aided Detection (CAD) framework incorporates preprocessing methods like; division, lung knob identification, and bogus positive decrease, which includes extraction. It utilizes a dataset, to find a relevant frame works in developing an exactness and accuracy in location. The Computed Tomography (CT) chest images are used, which are more efficient than X-ray. "MATrix LABoratory" (MATLAB) is widely used software in the study of lung cancer detection from chest CT scan images. The process includes image pre-processing, image segmentation, feature extraction and classification techniques. Lung illness can debilitate the respiratory framework, is a premier reason for human sickness in the worldwide space, which influences on its mortality rate. To extract the feature of cancer affected lung image, Support Vector Machine (SVM) classifier is proposed to detect normal and abnormal lung cells after implementing the features. Our experimental evaluation via MATLAB, which demonstrates the efficient performance of the proposed system using the method of watershed transform. This present work proposes a method to detect cancerous types based on benign, malignant and normal cells from the lung chest CT scan images. It will minimize the detection error made with the help of image technique, which gives the high performance of accuracy as 98.99%.

KEYWORDS: Lung Cancer, Computed Tomography (CT), Computer-Aided Diagnosis (CAD), Image Detection, Sobel Gradient, Support Vector Machine (SVM), Watershed Transform, MATLAB.

I. INTRODUCTION

Lung cancer is a disease of abnormal cells multiplying and increasing into a tumour. Cancer cells continue to increase and form new, abnormal cells. The major cause of cancer death is lung cancer. Detection of cancer [1] in the early phase can provide more treatment options, less invasive surgery and increases the survival rate. For lung cancer, if the disease [2] is detected in time, the survival rate of patient increases from 14 to 49% in recent 5 years. It is the most dangerous and widespread disease in the world. The cancer cells present in lung causes lung cancer is ease. This cells detection is very important issue for medical researchers. The chances of an effective treatment will significantly increases with early detection.

Recently, the image processing mechanisms are used widely in different medical areas for increasing earlier detection and treatment stages. The time is very significant factor to discover the disease in the patient as possible as fast. Its early detection increases the chances of an effective treatment. In 2005, approximately 1,372,910 new cancer cases are predictable and about 570,280 cancer deaths are too occurred. It is expected that there will be 163,510 deaths from lung cancer, which forms29% of all cancer deaths. When cells start to grow out of control, cancer begins in a part of the body. The cancer cell starts because of out of control expansion of abnormal cells.

1.1 Lung Cancer

The cancer cells of each type grow and spread in different ways, and they are treated differently. Like other cancers, lung cancer develops when normal processes of cell division and growth are disrupted, giving way to abnormal, uncontrollable growth. The cells grow into a mass, or tumor. Any abnormal growth in the body that directly invades surrounding tissues and organs, spreads to other parts of the body, or has the potential to grow back after being

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 9, September 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1009032 |

removed is called "malignant," or cancerous.

1.2 To who gets Lung Cancer?

Lung cancer can take several years to develop. Cigarette smoking is the most common risk factor for developing lung cancer. Many people exposed to cigarette smoke – or some of its components – will end up with permanent abnormal changes in their lungs. These changes can cause a cancerous tumor to develop within the lung. Twenty-five percent of all cases of lung cancer worldwide are diagnosed in people who have never smoked. The underlying cause in these cases is not well understood. Two out of three people diagnosed with lung cancer are over age 65.The most common age at diagnosis is 70 years.

Lung cancer is the second most common malignancy in the country, but it is the deadliest as it causes the highest number of deaths of all cancers. Staging allows the physician to fully understand the extent of the patient's cancer to determine treatment decisions and predict expected outcomes. Doctors use specific terms to describe the stages of cancer, but a straightforward way of describing staging might be as follows:

- **Localized**: The cancer is confined to the lung.
- **Regional**: The cancer has spread to lymph nodes (or glands) within the chest.
- **Distant**: The cancer has spread (or metastasized) to other parts of the body.

So, finding of lung cancer earlier is most important role for successful [3] treatment with Lung Image Database Consortium (LIDC), diagnosis is mainly based on chest CT scan images. Cancerous tumour starts in the part of lung is called primary lung cancer. Following are the types of lung cancer and these are divided into two main types:

- **Small Cell Lung Cancer (SCLC)**
- **Non-Small Cell Lung Cancer (NSCLC)**

In this, current work focuses on finding tumour and its stages. In this Marker-controlled Watershed segmentation is used to isolate a lung of a chest CT scan images.

SCLC: It is less common than non-small cell lung cancer accounting for about 15 percent of all lung cancers. This type of lung cancer grows rapidly, is likely to be advanced by the time of diagnosis and spreads to other parts of the body quickly.

NSCLC: It is much more common, and usually grows and spreads more slowly than small cell lung cancer. There are three main types of non-small cell lung cancer, named for the type of cells from which the cancer develops:

1. **Adenocarcinoma:** Often starts growing near the periphery of the lung and may vary in both size and growth rate. This is the most common type of lung cancer in both smokers and those who have never smoked.
2. **Squamous cell carcinoma:** Usually starts in one of the larger breathing tubes near the center of the chest. The size of these lung tumors can range from very small to large.
3. **Large cell carcinoma:** Often starts near the periphery of the lung, grows rapidly and is usually quite extensive when diagnosed.

1.3 CT-Scan

Computed Tomography (CT) is one of the imaging techniques most often used in the diagnosis of lung cancer. Lung cancer with multiple pathological residues can be seen on CT. Lung cancer is categorized into benign and malignant cancer. During diagnosis, cancer of a certain density can be categorized as benign cancer in some cases. However, in many cases that occur, the lung which tends to be congested is usually classified as malignant cancer. It is very important to make a diagnosis of lung cancer at an early stage to improve the treatment and treatment process. Systems designed for medical applications provide a variety of benefits to successfully optimize lung cancer detection.

II. LITERATURE REVIEW

Several researchers have proposed and implemented detection of lung cancer using different approaches of image processing and machine learning. SurenMakaj et.al.,[4] have proposed with the detection using CT Scan images in the form of image processing with LIDC and developed in MATLAB R2016a. They took 16- different DICOM

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 9, September 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1009032 |

images 23 are detected and get the accuracy of 92.0%, sensitivity of 100% and specificity of 50% only using SVM classifier.

Ignatious and Joseph [5] developed a system using watershed segmentation. In preprocessing, it uses Gabor filter to enhance the image quality. It compares the accuracy with neural fuzzy model and region growing method. Accuracy of the proposed is 90.1%, which is comparatively higher than the model with segmentation using neural fuzzy model and region growing method. The advantage of this model uses a marker controlled watershed segmentation, which solves over segmentation problem. As a limitation, it does not classify the cancer as benign or malignant and accuracy is high but still not satisfactory.

Roy, Sirohi, and Patle[6] developed a system to detect lung cancer nodule using fuzzy interference system and active contour model. This system uses gray transformation for image contrast enhancement. Image binarization is performed before segmentation and resulted image is segmented using active contour model. Cancer classification is performed using fuzzy inference method. Features like area, mean, entropy, correlation, major axis length, minor axis length are extracted to train the classifier. Overall, accuracy of the system is 94.12%. Counting its limitation it does not classify the cancer as benign or malignant which is future scope of this proposed model.

Radiologists have interior relied on examining images from chest radiography and PET scans to detect lung cancer [7]. Fatema et al. [8] proposed a technique based on CT technology, the proposed system consists of a few steps like photo acquisition, preprocessing, feature extraction and categorization. At the preprocessing stage, RGB images are converted to grayscale images, used for reducing the Gabor filter, the adaptive thresholding method transforms CT scan images into binary images, and the REGIONPROPS function is used for the right body, from the binary image to the region.

Sangamithraa and Govindaraju[9] uses K- means unsupervised learning algorithm for clustering or segmentation. It groups the pixel dataset according to certain characteristics. For classification this model implements back propagation network. Features like entropy, correlation, homogeneity, PSNR, SSIM are extracted using gray-level co-occurrence matrix (GLCM) method. The system has accuracy of about 90.7%. Image preprocessing median filter is used for noise removal which can be useful for our new model to remove the noise and improve the accuracy.

Jin, Zhang and [10] Jin used convolution neural network as classifier in his CAD system to detect the lung cancer. The system has 84.6% of accuracy, 82.5% of sensitivity and 86.7% of specificity. The advantage of this model is that it uses circular filter in Region of interest (ROI) extraction phase which reduces the cost of training and recognition steps. Although, implementation cost is reduced, it has still unsatisfactory accuracy.

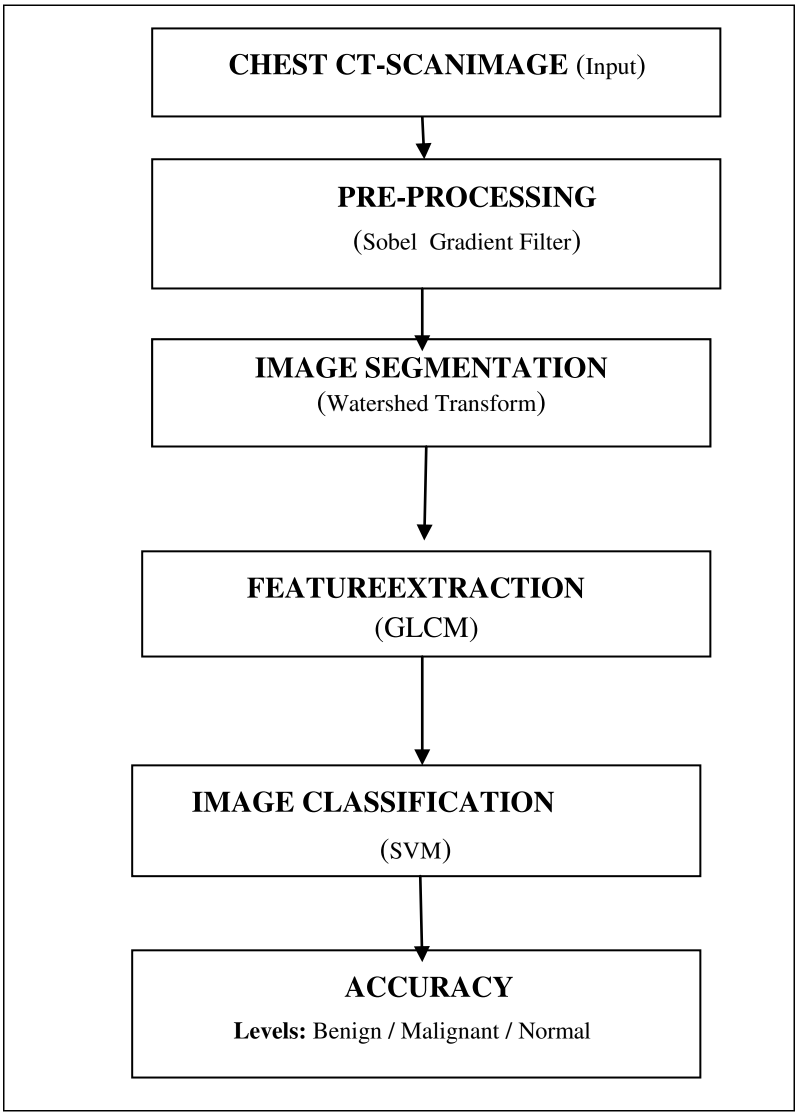
Aggarwal, Furquan and Kalra[11] proposed a model that provides classification between nodules and normal lung anatomy structure. The method extracts geometrical, statistical and gray level characteristics. LDA is used as classifier and optimal thresholding for segmentation. The system has 84% accuracy, 97.14% sensitivity and 53.33% specificity.

Rendon-Gonzalez and Ponomaryvo[12] proposed a system that classifies lung cancer as benign or malignant. The system uses the priori information and HousefieldUnit(HU) to calculate the Region of Interest(ROI). Shape features like area, eccentricity, circularity, fractal dimension and textural features like mean, variance, energy, entropy, skewness, contrast, and smoothness are extracted to train and classify by using the support vector machine to identify whether the nodule is benign or malignant. The advantage of this model is that, it classifies a cancer as benign or malignant.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 9, September 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1009032 |



III. PROPOSED MODEL

A new model has been proposed as illustrated in the Figure 1. Instead of Gabor Filter, Median filter and Gaussian filter, Sobel Gradient filter have been implemented in pre-processing stage. After preprocessing, the processed image is segmented using watershed Transform segmentation. It gives an image with cancer nodules marked. The intensity has been extracted in feature extraction stage for the detected cancer nodules. Classification of cancer nodule has been performed by using Support Vector Machine (SVM) in the next stage Extracted features are used as training features and trained model is generated.

3.1 Pre-processing

The pre-processing of image aims for selective elimination of the redundancy in scanned images without affecting the original image. It plays a vital role in the diagnosis of lung cancer. The image Pre-Processing stage in this system begins with chest CT- image, which aims to improve the interpretability or sensitivity of information included in them to provide better output from the image processing techniques.

3.1.1 Sobel Gradient Filter

The Sobel Gradient filter is used for edge detection. It works by calculating the gradient of image intensity at each pixel within the image. It finds the direction of the largest increase from light to dark and the rate of change in that direction. This study uses a watermarked image segmentation approach, where a result is a group of segments covering several contours that are completely explained. In the segmentation process, images that have been segmented and differentiated with their background aim to facilitate feature extraction using the Sobel Gradiant process. In this study, a cancer detection system procedure has been implemented and the lungs based on the image on the chest CT-Scan.

3.1.2 Image Segmentation

Image segmentation is a crucial process for most image analysis consequent tasks. Especially, most of the existing techniques for image description and recognition are highly depend on the segmentation results. Segmentation splits the image into its constituent regions or objects, where the objects used are distinguished from one another. The process of giving a label to each pixel in an image aims to distinguish each characteristic it has. Image segmentation produces a set of contours separated from the background. Each pixel in an image has a different color, intensity, and texture value. This segmentation process takes pictures from the pre-processing results and then used to take the area detected by cancer from the original image.

3.1.3 Watershed Transform

In the study of image processing, a watershed is a transformation, which is defined on a grayscale image. In an image by treating it as surface, where light pixels represents a high elevations and a dark pixel represents a low elevations. The watershed transform can be used to segment contiguous regions of interest into distinct objects.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 9, September 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1009032 |

3.1.4 Feature Extraction

Image feature extraction stage is a crucial stage that uses algorithms and methods to detect and separate various preferred portions or shapes of an input image approach has been applied for detection of cancer. In extract the number of white pixels and check them against some threshold to check the normal and abnormal lung cells. After this process, the condition is whether number of white pixels of a new image is less than the threshold or not. The feature extraction stage based on the shape is done by calculating the area of value in the segmented cancer object, which is then reconstructed to the original image color.

3.1.5 Image Classification

Image classification is perhaps the most important part of digital image analysis. Image classification is the process of categorizing and labeling groups of pixels or vectors within a chest CT image based on specific rules. The objective of image classification is to identify and portray, as a unique gray level (or color), the features occurring in an image in terms of the object or type of land cover.

3.1.6 Support Vector Machine (SVM)

The main advantage of SVM is that it can be used for both classification and regression problems. SVM draws a decision boundary, which is a hyper plane between any two classes in order to separate them or classify them. SVM also used in Object Detection and image classification. SVM are a set of supervised learning methods used for classification, regression and outlier's detection.

IV. DATASET SOURCE

This section describes the experiments developed to evaluate the performance of SVM as feature extractors in the lung cancer malignancy classification task using chest CT scan images from the LIDC/IDRI database. Here, LIDC/IDRI database is filtered, extracted and formatted by using 320 Images, where 240 images are trained and 90 images are tested. In which, each of the implemented extractors is applied to the LIDC/IDRI database. Finally, evaluation metrics are computed for each SVM classifier applied. Model's classification of benign or malignant or normal using support vector machine can be useful in the proposed new model.

V. IMPLEMENTATION

MATALAB Code for Watershed with images:

```
imshow(bwImage) title('cancer trained images  '); NewImage=bwImage;
J = -bwdist(~NewImage); figure
imshow(J,[])
Ld = watershed(J); figure
imshow(label2rgb(Ld)) title('watershed image'); bw2 = NewImage; bw2(Ld == 0) = 0;
figure imshow(bw2)
mask = imextendedmin(J,330);
D2 = imimposemin(J,mask); Ld2 = watershed(D2);
bw3 = NewImage; bw3(Ld2 == 0) = 0;
bwImage=bw3;
bigMask = bwareaopen(testsing Image, 240); finalImage = bwImage; finalImage(bigMask) = false;
bwImage=bwareaopen(finalImage,90); labeledImage = bwlabel(bwImage, );
RegionMeasurements = regionprops(labeledImage, initImage, 'all'); Ecc = [RegionMeasurements.Eccentricity];
RegionNo = size(RegionMeasurements, 1); allowableAccIndexes = (Acc< 0.98); keeperIndexes =
    find(allowableEccIndexes);
RegionImage = ismember(labeledImage, keeperIndexes); bwImage=RegionImage;
figure imshow(bwImage);
title('cancer spot', 'FontSize', FontSize);
 %%%%%clear labeledImage;

clear RegionMeasurements;
```

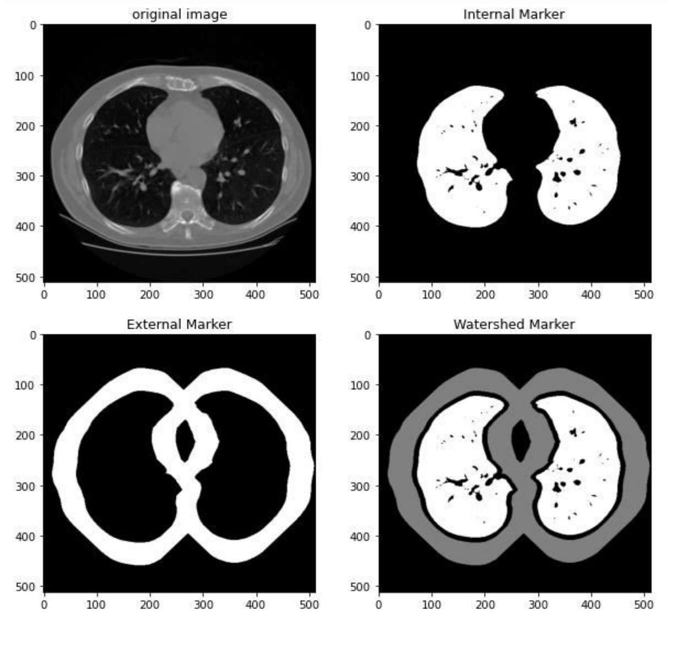


Figure 2: Different markers extracted from Lung chest CT scan images using watershed Transform

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 9, September 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1009032 |

VI. RESULTS AND DISCUSSION

Output results are interpreted in our model, are detailed and compared to a few other high performing techniques. The proposed algorithm is applied to lung chest CT images to calculate tumor detection results. This model extracts the background part of the image. Proposed method based performance analyses such as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

In the below Figures 3 and 4 are different markers in original image with internal, external and watershed marker to the chest CT images. Based on these the results of all CT scan images are explained.

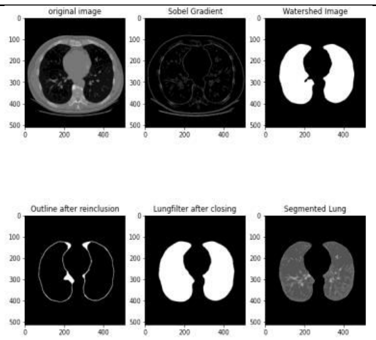
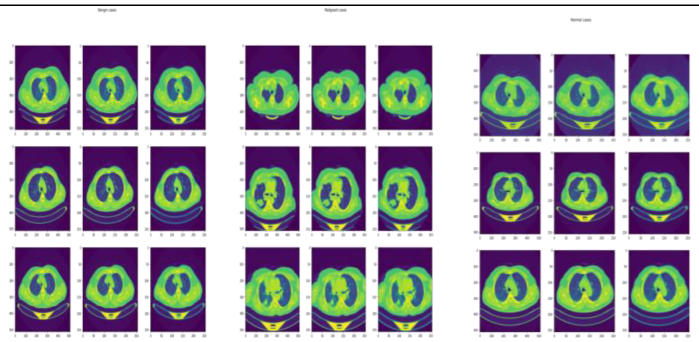


Figure 3: Image segmentation process visualization



a) Benign Images b) Malignant Images c) Normal Images

Figure 4: Lung Cancer Detected Images via Watershed (a,b,c)

Performance Evaluation Metrics

Performance measurement is an ongoing process that monitors and reports on a program's progress and accomplishments by using pre-selected performance measures. In evaluation process, it uses the measurement and analysis to answer the specific questions about how well a program is achieving its outcomes.

Accuracy **TP+(TN/TP)+(TN+FP+FN)**	Number of correct predictions / Total number of Predictions made
Precision **TP/(TP+FP)**	It is the number of correct positive results divided by the number of positive results predicted by the classifier.
Sensitivity (or) **Recall** **TP/(TP+FN)**	It is the number of correct positive results divided by the number of all relevant samples. In mathematical form. True Positive / (True Positives + False Negatives)
Specificity **TN/(TN+FP)**	True Negatives / (True Negatives + False Positives)

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 9, September 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1009032 |

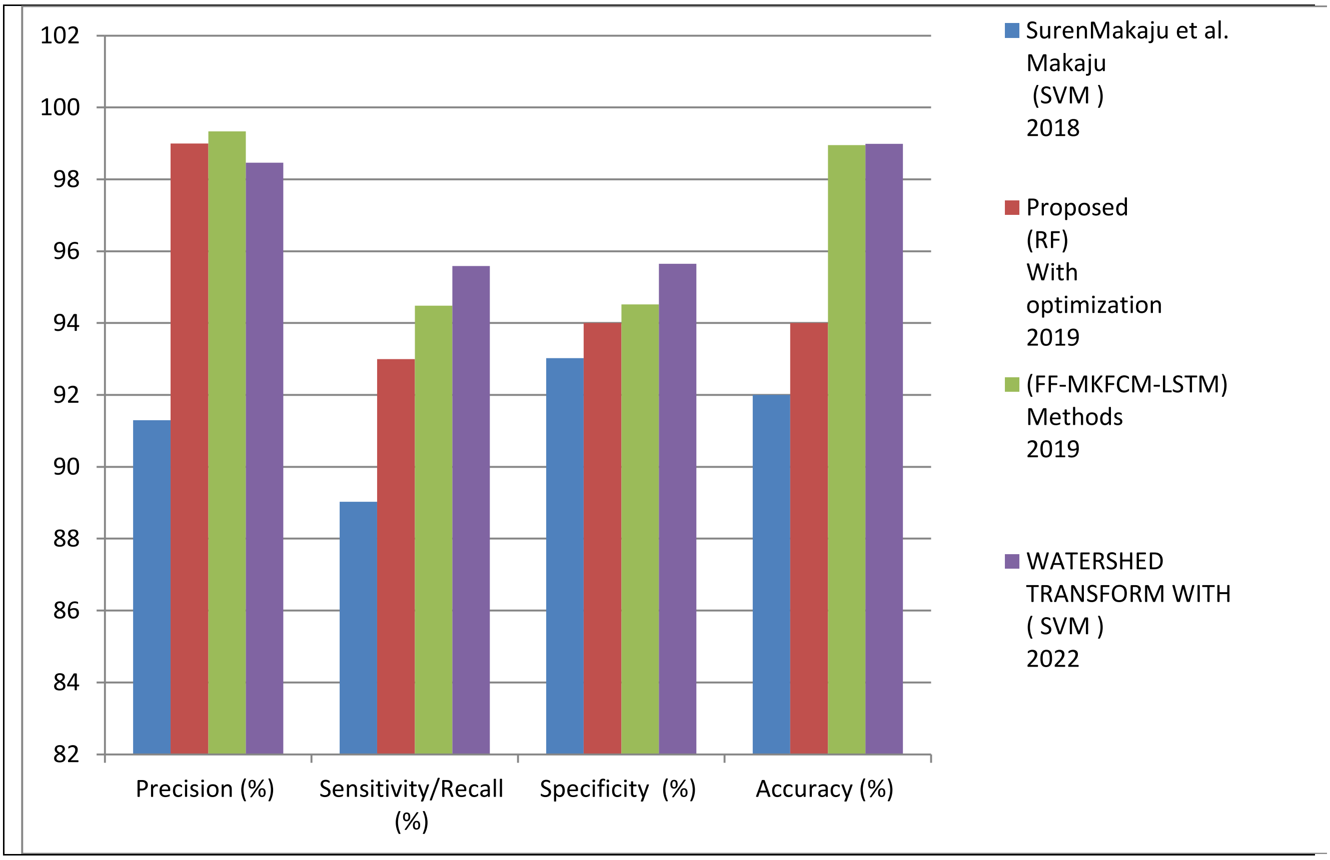


Figure 7: Graphical Representation -Accuracy level of Watershed with SVM

VII. CONCLUSION AND FUTURE DIRECTIONS

In this paper, an image processing techniques used to detect early stage of lung cancer in chest CT scan images among non-smokers. The chest CT scan image has been pre-processed followed by the process of segmentation of Watershed Transform of the lung. The SVM classifier has been evaluated and achieves an accuracy of 98.99%, sensitivity of 95.39%, specificity of 95.65%, and precision as 98.46%. As a feature work, these proposed techniques can be used for other medical problems such as breast cancer, blood cancer, throat cancer, brain cancer, tongue cancer, mouth cancer and skin cancer among non-smokers too.

REFERENCES

[1] Miah, M.B.A., &Yousuf, M.A. “Detection of lung cancer from CT image using image processing and neural network.” *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT):*1-6, 2015.

[2] Khobragade, S., Tiwari, A., Patil, C., &Narke, V. “Automatic detection of major lung diseases using Chest Radiographs and classification by feed-forward artificial neural network.” *IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES):*1-5, 2016.

[3] Armato, I., Samuel McLennan, G., McNitt-Gray, F. R., Michael, Charles, Reeves, Anthony P., Clarke, Laurenc, “*Data from LIDC-IDRI. The Cancer Imaging,* 2015.

[4] S. Makaju, P. Prasad, A. Alsadoon, A. Singh and A. Elchouemi, "Lung Cancer Detection using CT Scan Images", *Procedia Computer Science,* Vol. 125, pp. 107-114, 2018.

[5] Ignatious, S., & Joseph, R. (2015) “Computer aided lung cancer detection system.” *Global Conference On Communication Technologies (GCCT),* DOI:10.1109/GCCT.2015.7342723, 2015.

[6] Roy, T., Sirohi, N., &Patle, A. “Classification of lung image and nodule detection using fuzzy inference system.” *International Conference On Computing, Communication & Automation.* DOI:10.1109/CCAA.2015.7148560, 2015.

[7] P. Reid and S. Walter, “Yield from a fast track referal system for radiologists suspecting lung cancer*”, Lung*

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 9, September 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1009032 |

Cancer, Vol. 115, p. S16, 2018.

[8] F. TujJohora, M. Jony and P. Khatun, "A New Strategy to Detect Lung Cancer on CT Images", *International Research Journal of Engineering and Technology (IRJET)*, Vol. 05, No. 12, pp. 27-32, 2018.

[9] Sangamithraa, P., &Govindaraju, S. "Lung tumour detection and classification using EK-Mean clustering." *International Conference on Wireless Communications, Signal Processing and Networking (Wispnet).* DOI: 10.1109/WiSPNET.2016.7566533, 2016.

[10] Jin, X., Zhang, Y., & Jin, Q. "Pulmonary Nodule Detection Based on CT Images Using Convolution Neural Network." *9Th International Symposium on Computational Intelligence and Design (ISCID).* DOI: 10.1109/ISCID.2016.1053, 2016.

[11] Aggarwal, T., Furqan, A., &Kalra, K. "Feature extraction and LDA based classification of lung nodules in chest CT scan images." *International Conference On Advances In Computing,CommunicationsAndInformatics(ICACCI),*DOI:10.1109/ICACCI.2015.7275773, 2015.

[12] Rendon-Gonzalez, E., &Ponomaryov, V. "Automatic Lung nodule segmentation and classification in CT images based on SVM." *9Th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter andSubmillimeter Waves (MSMW).* DOI:10.1109/MSMW.2016.7537995.13, 2016.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

doi crossref

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING