

K. Sujatha¹, V. Srividhya², R.S. Ponmagal³, M. Nicholas Ponraj⁴, T. Kalpalatha Reddy⁵ and S. Saranya⁶, N.P.G. Bhavani⁷

 ^{1,4} Department of Biomedical Engineering/EEE, Dr. M.G.R Educational and Research Institute Maduravoyal, Chennai 600095, India.
² Department of EEE, Meenakshi College of Engineering
³Associate Professor, CSE Department, SRM Institute of Science and Technology
⁵Professor, Department of ECE, Dr. M.G.R Educational and Research Institute Maduravoyal, Chennai 600095, India
⁶Department of Biomedical Engineering, Dr. M.G.R Educational and Research Institute ⁷Department of ECE, Saveetha School of Engineering, SIMATS

Abstract

This topic concerned lung analysis and the detection of tumors or cancer. Its general object is to discuss noise removal, binary image, inverted image, segmentation, and circles segmented. The selected image can undergo some filtrations and processing of image segmentation. The results of the study and image can be processed and segmented. Lung cancer is one of the deadliest diseases which cause high death rates throughout the world. Lung cancer is an irregular growth of cells that can be characteristically derived from a single irregular cell which may spread to whole part of the lung. CT scan is one of the sensitive methods used in the medical field for treating the patients as compared to MRI and X-rays. Diagnosis of cancer from the computed tomography (CT) images of lung is very challenging for doctors. Computer aided diagnosis (CAD) is another tool for detection that uses computer-generated output as an assisting tool for a clinician to form a diagnosis. The biomedical image processing has better ability to detect lung diseases as it helps in analyzing each image and monitoring the data. MATLAB has been used through every procedure made in this study. In image processing procedures, preprocessing of an image is a necessary process as there is difficulty in detecting cancer cells in an image due to the presence of noise and low-quality of images. Steps like image enhancement, image segmentation and feature extraction methods can be used to reduce the degree of those problems. Reinforcement machine learning algorithm is introduced to increase the accuracy rate of detecting tumor growth in lungs. In future, the CT images are collected from health-care centers and remote places through Internet of Things (IoT)-enabled platform and the image is stored in the cloud servers and obtained for processing at any time. The main purpose of this paper is to assist the doctors to detect and classify lung cancer by using CT images based on reinforcement learning algorithm through IoT platform.

Section A-Research paper

Keywords: Computed Tomography (CT), Internet of Things (IoT), Computer aided diagnosis (CAD), Pre-processing, Image Segmentation, Reinforcement Learning Neural Networks, lung analysis, detection of tumour, segmentation, processing.

1. Introduction

Cancer is a disease in which cells in the body grow uncontrollably and spread to other parts of the body. Lung cancer begins within the lungs and spread to lymph nodes or other organs within the body like brain. Lung cancers are generally grouped into two main types called small cell and non-small cell. Cancerous cells in a person with small cell cancer appear small and round under a microscope. Non-small cell lung cancer cells are larger. Smoking is a major risk factor for both types. Small cell lung cancer affects 95% of people who have had a history of smoking. People who smoke have the highest risk of lung cancer, though lung cancer can also occur in those who don't smoke. Both small and non-small cell lung cancers have several subtypes. Adenocarcinoma, squamous cell carcinoma, and large cell carcinoma are examples of non-small cell lung cancers. The expression of specific genes determines the type of small cell cancer [14, 15].

A computerized tomography scan combines a series of X-ray images taken from various angles around the body with computer processing to create cross-sectional images of the bones, blood vessels, and soft tissues inside the body. A CT scan is more likely to show lung cancers than routine chest x-rays. A CT scan can reveal small lesions within the lungs which may not be detected on an X-ray. It can also show the size, shape, and position of any lung cancers and can help find enlarged lymph nodes that might contain cancer that has spread. For the year 2020, the projected occurrence of cancer patients in India was 679,421 (94.1 per 100,000) for males and 712,758 (103.6 per 100,000) for females[16,17].



Figure 1. CT Image of Cancer Affected Lung

People at a higher risk of developing lung cancer have been researched in recent years using a procedure called as a low-dose CT (LDCT) scan (primarily because they smoke or were accustomed to smoking). Finding aberrant lung tissue that may be malignant with the use of LDCT scans is possible. According to research, annual LDCT scans to detect adults at higher risk of lung cancer can save lives in contrast chest x-rays. Before symptoms appear, these individuals should get yearly LDCT scans to help reduce their chance of developing lung cancer.

Section A-Research paper

The majority of these photographs have salt and pepper noise, which is a sort of noise in which a portion of the pixels are either black or white. The most popular way to remove this kind of noise is by using a medium filter. "Filtering can eliminate salt and pepper noise without considerably lowering an image sharpness because it is less sensitive than linear approaches to severe changes in pixel values. "Therefore, in this assignment, the method of applying a medium filter was used [18].

2. Literature Survey

To gain a deep understanding of image processing and machine learning, a few research papers were reviewed. This section is a literature review that covers various image processing approaches. Image processing for feature extraction and classification approaches for cancer type and stage recognition are used in computer-aided detection. For the classification of various tumors, Saif Ali et al. [1] has discussed several different algorithms, including SVM, KNN, DT, and others.

A variety of methods, such as texture analysis and image segmentation have been analyzed by Guo Xiuhua et al. [2] to aid radiologists in diagnosing lung cancer in order to increase the accuracy and efficiency of CT screening program for the diagnosis of early-stage lung cancer.

The processing of medical ultrasound images with MATLAB could aid in the identification of Parkinson's disease in the brainstem was proposed by Blahuta et al. [3]. Furthermore, the DICOM standard for medical imaging is introduced, as well as current 3D/4D scanning for a greater quality and accuracy of diagnosis than traditional 2D scanning.

Using a combination of K-Means and support vector machine techniques,D. Suresha et al. [4] presented a system to determine if the brain contains a tumour or is tumour-free from an MR image. The input image is transformed to grey scale using binary thresholding in the first stage, and the spots are detected. To distinguish between the normal and tumour brains, the detected spots are displayed in terms of their intensities.

Manisha et al. [5] in "Tumor region extraction using edge detection method in brain MRI images" has developed a method for determining the threshold value using standard deviation, and the result is an intensity map. This intensity map's standard deviation has now been recalculated and used to calculate the average intensity of pixels that fall outside of this standard deviation. Finally, the threshold value obtained from the computed average intensity will be used to segment the tumour from the original MRI images.

One of the most crucial steps in the detection of irregularities is the appropriate interpretation of edge detection. Image de-noising algorithms are among the oldest in the field of image processing. The core principle of noise reduction through image blurring is shared by Samir Kumar Bandyopadhyay [6].

A measure for image acuity is derived as the ratio of the maximal gradient magnitude and the intensity difference at a boundary, based on a full statistical picture model that includes noise and blur. The object under examination, image acquisition, reconstruction techniques, and any post-processing stages may all affect acuity. The acuity value presented by F. Kruggel [7], easy to comprehend, compute, and integrates well with other conventional picture quality measurements.

Classification techniques such as Naive Bayes, SVM, Decision Tree, and Logistic Regression were used to analyze lung cancer prediction. The primary goal of RadhikaPR et al. [8] in "A Comparative Study of Lung Cancer Detection using Machine Learning Algorithms" is

to investigate the performance of classification algorithms in the early detection of lung cancer [9-13], but it will be a base model for more sophisticated ones, considered in Section 4.

3. Existing System

In the current period, lung cancer can be identified only by sputum cytology or lung biopsy. A lung biopsy is a procedure in which samples of lung tissue are removed to determine if lung disease or cancer is present. Initially, the doctor will advise the patient to take a CT scan, and will examine it to determine if the lung is irregular in form, or has an outgrowth. Further, if the doctor speculates that the patient has some kind of irregularity, they will instruct the patient to undergo sputum cytology or lung biopsy to determine if the patient has cancer cells or tumour growth. This will cause delay in the process of detecting tumour growth which will result in late treatment of cancer. Delayed detection of lung cancer leads to increased mortality rate in people.

4. Proposed System

An image processing system was designed which would detect and classify cancer or tumour growth in the lungs with the help of reinforcement learning in machine learning algorithm. This system can help to increase the efficiency of lung cancer diagnosis by processing the CT image without moving forward to the invasive methods of diagnosis, which will cause patient discomfort, increase wound care, and longer healing periods. As shown in figure 2, the block diagram explains the process of this paper. First, the input image is given as CT image into the program and it undergoes several processes to obtain the final output. The processes include Preprocessing using Median Filter, Segmentation using Gaussian Filter, Feature Extraction using Bilateral Filter and Classification using Machine Learning. The final output of the image is stored in the cloud with the help of IoT.



Figure 2. Block Diagram for Detection of Lung cancer using Image Processing

5. Methodology

5.1. Pre-Processing using Median Filter

The median filter is a non-linear digital filtering technique for removing noise from images and signals. This type of noise reduction is a common pre-processing step used to improve the results of subsequent processing. Median filtering is commonly employed in digital image processing because it preserves edges while reducing noise under specific conditions. The given input image undergoes median filtering technique to remove the noise from the CT image to ensure better results for the detection of lung tumor.

5.2. Segmentation using Gaussian Filter

A Gaussian Filter is a low pass filter that is used to reduce noise and blurring in image regions. To get the desired effect, the filter is built as an Odd sized Symmetric Kernel, which is passed through each pixel of the Region of Interest. The kernel is not resistant to abrupt colour changes because the pixels in the kernel's centre have greater weight in the final value than those in the periphery. The pre-processed CT image after going through median filtering, is segmented using Gaussian filter to again reduce the noises and preserve the edges in the image for easier identification of tumor.



Figure 3. Before and After Pre-Processing with Median Filter



Figure 4. Lung Lesion Segmentation using Gaussian Filter

5.3. Feature Extraction Using Bilateral Filter

A bilateral filter is a non-linear, edge-preserving, and noise-reducing smoothing filter for images. It uses a weighted average of intensity data from surrounding pixels to replace the intensity of each pixel. A Gaussian distribution can be used to calculate this weight. The weights are affected not just by the Euclidean distance between pixels, but also by the radiometric differences. The segmented image is further processed for feature extraction using bilateral filter, which further reduces the noise and smoothens the image, while preserving the edges of the image.



Figure 5. De-noising of CT Image using Bilateral Filter

5.4. Classification Using Machine Learning Based Reinforcement Learning Algorithm

Using Machine Learning, image classification is a method of classifying a picture from data classes that aggregate pixels that are uniform in their specific esteem in different data. Expert efforts to recognize the image based on this information and experience about the space are known as image classifications of interest. The classes could be combined into a single data class. The underlying quantitative data in an image associated with the structure or state of the objective Machine Learning based Reinforcement Learning Algorithms is used in computerized image layout. From the database of previous lung tumor affected images, the system studies the images for the improvement of the detection of lung tumor and to reduce the false rate of detection. The previously processed image is then classified using reinforcement learning algorithm to detect the tumor.

5.5. Internet of Things

The internet of things is a network of interconnected computing devices, mechanical and digital machines, and items with unique identifiers and the ability to transfer data without requiring human-to-human or human-to-computer interaction. The final output is then uploaded into the cloud server for accessing whenever and wherever we need.

6. Results and Discussion

The categorization of lung cancer is identified as primary and secondary stage. Also the CT image of the lungs which are not affected by cancer and is in normal condition is also detected by the Reinforced Learning Neural Networks. Totally Figure 6 shows the final output. It shows the processing of the Lung CT image, and finally, displays the status of the CT image. The CT image of the lung can be given in as an input in this interface. After loading the image, the program starts to process the image for detection of cancer and displays the output image and the result.

6.1. Suppression of entries, gray scale and noise

The issue with the initial image is that there is a significant amount of noise, which must be eliminated before any further procedures can be used to locate the circles within the image. The noise can be seen in the screenshot below and affects the image's brightness and blackness (Bovik, 2009) As a result; the first step in offering a solution to this issue is the elimination of

Section A-Research paper

any noise present in the photos fed into the application. This task's initial step was reading the picture into the programme using the syntax "Selected Image = imread (image jpg);" This makes it possible to read the appropriate input pictures. Once the image has been read, it is transformed from RGB to grayscale using the 'rgb2gray' function. This enables the removal of hue and saturation information while preserving the desired brightness of the image.

Noise can be eliminated after the image has been read and transformed. By applying a medium filter on the grayscale picture, the custom filter function enables for each outputted pixel to include a medium value in the 3-by-3 region surrounding the corresponding pixels inside the image. This distinct function is used to reduce noise. The medium filter uses a similar methodology smoothing approach, and since utilizing comparable pixels improves the capacity to remove noise without degrading an image's clarity, this function is ideal for lowering the salt and pepper noise in the photographs for this project. This is done throughout the project with all the latest pictures implemented inside the project. The output for noise suppression is illustrated in Figure 6.



Figure 6. Noise suppression

6.2. Conversion of grey level images into binary images

The next critical step in resolving the problem is to convert the current gray scale image into a binary image. Making the image into a binary image reduces the number of colours in the image to two, making it easier to discern objects inside it (like circles) (black, and white). The image is converted into a binary image using the im2bw function: "binary picture = im2bw (median_filtering_Image, 0.2);" This function uses the filtered image as an input and, depending on the brightness level, converts each pixel to either the value 1 (white) or the value 0 (black) (black).

6.3. Creation of a morphological structural element (STREL)

Section A-Research paper

In this phase, a structural element in the shape of a rectangle is created by typing "se1 = strel ('disc', 2);" When "2" is used as the radius, a flat disk-shaped structure is produced. In order to create two distinct structures, this must be done twice. These two structures are then applied to the initial binary picture to produce the two distinct images postOpenImage_1 and postOpenImage_2. Once these structures have been applied, the user is shown the first image that has been applied to the subplot, while the other image is kept for additional image processing to aid in the problem's eventual solution.

6.4. Reversing the open image

The binary image must be inverted in this phase, which can be done by simply using the one's function. This function generates an array of all ones and applies it to the binary image as follows: inverted = ones(size(binary picture)); Once this is finished, the code "invertedImage_1 = inverted – postOpenImage_1;" is put into place. This line takes the inverted image created by the array of ones and creates the Open image from it. This inverts the image by switching all prior 0 pixels for 1 pixels and all subsequent 1 pixels for 0 pixels, resulting in the inverted image that is displayed below. This is finished to enable segmentation to be finished in the next stage, enabling the programme to find the circles in the image more effectively as in Figure 7.



Figure 7. Noise suppression

6.5. Initial contour creation and implementation of the segmentation

Create and specify the first contour in the following manner: "mask = zeros (size (invertedImage_1)); mask (50: end-50,50: end-50) = 1;" The mask, which is a binary image that specifies the initial state of the active contour, involves the boundaries of the regions of the image, in this case, the white background and the black regions outside the main part of the image. This allows for contour evolution to happen, which then enables image segmentation to happen. After the image has been segmented, it will be easier to locate the circles inside the lung because the foreground and backdrop will not interfere with the search, as shown in the Figure 8.

Section A-Research paper



Figure 8. Outputs for Contour creation and Segmentation

6.6. Combining images with research circles

Finding the circles within the lung is the next operation that needs to be done. In this software, the image is formed by combining an inverted image with contour segmentation: "mix_Image_1 = invertedImage_1 + bw_1;" This produces an initial image on which the circles can be located. From there, the image is changed from colour to black and white using the "im2bw" programme, and the final image is further filtered using the medium filter "medfilt2" to provide the best possible ability to detect any circles within the lung. The current step's primary procedure is to locate the circles. The 'imfindcircles' function helps to locate the circles.

The function "viscircles (centers, radii,' EdgeColor', 'g');" searches for dark sensitivity in the image below using the centre and radius arguments with a sensitivity specification. The circles are then shown using green circles. The circles are all highlighted in green when this is finished, as seen in the snapshot below of the image after it has been presented. The circles are also tallied using the command "display (size (centres, 1), "Numbers of Circles");" this enables the command window to get the total number of circles that were found and have a centre.

6.7. Segment the circles

The circles must first be located, numbered, and then segmented before being detected. To accomplish this, first create a segmented image using the formula "segment_pic = final_2 – final_1;" This is accomplished by moving from a photo without the circles to one where they are present. The MATLAB code used for the detection of lung cancer is capable to detect small circles within the image and the boundaries and then each circle is filled with green colour to identify they have been found as depicted in Figure 9(a) and (b). Therefore, solving all problems indicated within the brief as the circles have been counted and found and all noise has been removed.

Section A-Research paper



Figure 9(a). Location of the circles after segmentation



Figure 9(b). Location of the circles after segmentation

Finding aberrant lung tissue that may be malignant with the aid of LDCT scans is possible. According to research, annual LDCT scans to detect adults at higher risk of lung cancer can save lives in contrast to chest X-rays. Before symptoms appear, these individuals should undergo yearly LDCT scans to help reduce their risk of developing lung cancer. The chosen image may be subjected to some image filters and segmentation processing after the CT scan. The study's findings and the image can both be analyzed and segmented.

An abnormal tumor or nodule in your lungs could be seen on an X-ray. Small lung lesions may not be seen on an X-ray but can be seen on a CT scan. When sputum is coughed up and examined under a microscope, lung cancer cells can occasionally be seen. This is especially true if you have a cough and are producing sputum. The chosen image is filtered with salt and pepper using MATLAB. The most recent strategies also include the introduction of certain other filters. The entire detection scheme is depicted in Figure 10 and Figure 11, illustrates the stage wise identification of lung cancer.



Figure 10. Detection of Lung cancer



Figure 11. Detection and classification of Lung cancer based Primary and Secondary stages

The performance measurement adhered for this is recall and precision. Both the values of recall and precision are equivalent to 1. Hence the reinforced learning is able to provide an optimum quantity of values. Table 1 represents the behaviour of the classifier for finding out normal, primary and secondary stages from the lung images and their associated values.



Figure 12. Stages of Lung cancer Detection

The features are obtained are classified using RLNN and compared with the existing methods (K-Means algorithm and also by wavelet transform) mentioned in the literature. The precision and recall values for normal, primary and secondary stage abnormal images also equals to 0.83.

On the other hand, if the classification is carried out using RLNN with the selected features its value is equal to 1.

Table 1. Comparison of existing and proposed Algorithms				
Performance	measures for	classification	by K-means	Algorithm,
Wavelet and RLNN based algorithms				
Type of	Abnormal		Normal	
Algorithm	Precision	Recall	Precision	Recall
K-Means	0.83	1	0.83	1
Algorithm	0.85	1	0.05	1
Wavelet	0.83	1	0.83	1
Transform				
RLNN	1	1	1	1
algorithms	L	1	1	T

7. Conclusion

Lung cancer is the leading cause of death worldwide. The medical image is based on various stages in which Computed Tomography samples should be taken throughout the various sections of the lungs. The CT image is used for pre-processing and segmentation of several simultaneous operations. The shape and size of the output image are discovered and analyzed using filtration-based Morphology in this method. Machine learning algorithms remove objects from the image background of test information, also known as preparation information, in order to create forecasts without having to be explicitly configured to deliver out the image. The final stage uses Machine Learning-based Reinforcement Learning Algorithms to classify the picture, and the output image calculation collects data from the significant image output. The final output is then stored into the cloud, which can be accessed from any location.

Thus, using MATLAB, the image is subjected to the removal of salt and pepper noise, binarization, inversion, addition of marks, detection of circles, and counting of circles. As a result, all issues mentioned in the brief must be resolved once the circles have been counted, located, and the noise has been eliminated. The next phase is to more precisely locate the tumour using updated filters and AI masking. The implementation of a fresh deep learning method is planned for the future.

References

- Saif Ali, Aneeqa Tanveer, Azhar Hussain, Saif ur RehmanIdentification of Cancer Disease Using Image Processing Approahes. International Journal of Intelligent Information Systems. Vol. 9, No. 2, 2020, pp. 6-15. doi: 10.11648/j.ijiis.20200902.11
- [2] Guo Xiuhua, Sun Tao, Wang huan and Liang Zhigang (2011). Prediction Models for Malignant Pulmonary Nodules Based-on Texture Features of CT Image, Theory and Applications of CT Imaging and Analysis, Prof. Noriyasu Homma (Ed.), ISBN: 978-953-307-234-0.
- [3] Blahuta, Jiri, Tomás Soukup and Petr Cermak. "Image processing of medical diagnostic neuro sonographical images in MATLAB." (2011).

- [4] D. Suresha, N. Jagadisha, H. S. Shrisha and K. S. Kaushik, "Detection of Brain Tumor Using Image Processing," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 844-848, doi: 10.1109/ICCMC48092.2020.ICCMC-000156.
- [5] Manisha, B. Radhakrishnan and L. P. Suresh, "Tumor region extraction using edge detection method in brain MRI images," 2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT), 2017, pp. 1-5,doi:10.1109/ICCPCT.2017. 8074326.
- [6] Samir Kumar Bandyopadhyay, "Edge Detection from Ct Images of Lung" Published in IJESAT International Journal of Engineering Science & Advanced Technology Volume - 2, Issue - 1, 34 – 37, ISSN: 2250–3676.
- [7] F. Kruggel, "A Simple Measure for Acuity in Medical Images," in IEEE Transactions on Image Processing, vol. 27, no. 11, pp. 5225-5233, Nov. 2018,doi:10.1109/TIP.2018.2851673.
- [8] R. P.R., R. A. S. Nair and V. G., "A Comparative Study of Lung Cancer Detection using Machine Learning Algorithms," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), 2019, pp. 1-4, doi: 10.1109/ICECCT.2019.8869001.
- [9] Suren Makaju, P.W.C. Prasad, Abeer Alsadoon, A.K. Singh, A. Elchouemi, Lung Cancer Detection using CT Scan Images, Procedia Computer Science, Volume 125, 2018, Pages 107-114, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2017.12.016.
- [10] Parveen and A. Singh, "Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM," 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN), 2015, pp. 98-102, doi: 10.1109/SPIN.2015.7095308.
- [11] Ye X, Lin X, Dehmeshki J, Slabaugh G, Beddoe G. Shape-based computer-aided detection of lung nodules in thoracic CT images. IEEE Trans Biomed Eng. 2009 Jul;56(7):1810-20.
- [12] Ribbens A, Hermans J, Maes F, Vandermeulen D, Suetens P. Unsupervised segmentation, clustering, and groupwise registration of heterogeneous populations of brain MR images. IEEE Trans Med Imaging. 2014 Feb;33(2):201-24.
- [13] O. Oktay et al., "Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation," in IEEE Transactions on Medical Imaging, vol. 37, no. 2, pp. 384-395, 2018.
- [14] Sujatha, K. Pappa, N. Senthil, K. Kumar and Siddharth Nambi, U. (2013) Monitoring Power Station Boilers Using ANN and Image Processing, Trans Tech Publications, Switzerland, Advanced Materials Research, Vol. 631-632, pp.1154-1159.
- [15] Sujatha, K. Pappa, N. Senthil, K. Kumar, Siddharth Nambi, U. and Raja Dinakaran, C. R. (2013) Intelligent Parallel Networks for Combustion Quality Monitoring in Power Station Boilers, Trans Tech Publications, Switzerland, Advanced Materials Research, Vol. 699, pp.893-899.
- [16] Sujatha, K. Pappa, N. Senthil, K. Kumar, Siddharth Nambi, U. and Raja Dinakaran, C. R. (2013) Automation of Combustion Monitoring in Boilers using Discriminant Radial Basis Network, *Int. J. Artificial Intelligence and Soft Computing*, Vol. 3, No. 3.

Section A-Research paper

- [17] Sujatha, K. (2012) Flame Monitoring in power station boilers using image processing, *ICTACT Journal on Image and Video Processing*, Dr.M.G.R Educational & Research Institute.
- [18] Sujatha, K. Pappa N. (2011) Combustion Quality Monitoring in PS Boilers Using Discriminant RBF, ISA Transactions, *Elsevier*, Vol.2(7), pp.2623-2631.