International Journal of Electrical and Electronics Engineers
Volume 16, Issue No. 01, Jan-June 2024
ISSN (O) 2321-2055
ISSN (P) 2321 -2045

LUNG CANCER DETECTION USING DIFFERENT LEARNING ALGORITHMS

LIKHITHA VARMA[1], M SUPRIYA[2], K. PRAVEEN CHOWDARI[3], SK. JANI BASHA[4], L. ASHOK REDDY[5], MR. G. MALLIKARJUNA RAO[6]

[1,2,3,4,5]UG student, Department of ECE, Tirumala Engineering College

[6]Associate professor, Department of ECE, Tirumala Engineering College

Abstract:

As per the technical evolution and latest trend taken into consideration, we have decided to make research in biomedical field particularly Lung cancer detection. Recently, image processing techniques are widely used in several medical areas for image improvement in earlier detection and treatment stages. There are various types of cancers i.e. lung cancer, Breast cancer, blood cancer, throat cancer, brain cancer, tongs cancer, mouth cancer etc. Lung cancer is a disease of abnormal cells multiplying and growing into a tumor. Cancer cells can be carried away from the lungs in blood, or lymph fluid that surrounds lung tissue. In this project we collect cancer image from different hospitals where present work is going on and this available image was color image we have taken the image and followed by conversion into a suitable form.

Image quality and accuracy is the core factors of this research, image quality assessment as well as improvement are depending on the enhancement stage where low pre- processing techniques are usedbased on Gabor filter. The segmentation and enhancement techniques are used to obtain the feature extraction of normal and abnormal image. Relying on general features, a normality comparison is made. In this paper the main detected features for accurate images comparison are pixels percentage and mask-labelling.

Keyword: Lung Cancer Detection, CT Scan Image, Cancer, Image Processing

1. Introduction

1. Lung cancer is of disease of abnormalcells multiplying and growing into a Tumor. Cancer cells can be carriedaway from the lungs in blood, or lymph fluid that surrounds lung tissue. Lymph flows through lymphatic vessels, which drain into lymph nodes located in the lungs and in the center of the chest. Lung cancer often spreadstoward thecenter of the chest

International Journal of Electrical and Electronics Engineers
Volume 16, Issue No. 01, Jan-June 2024
ISSN (O) 2321-2055
ISSN (P) 2321 -2045

because the natural flow of lymph out of the lung is towards the center of the chest.

2. Metastasis occurs when a cancer cell leaves the site where it began and moves into a lymph node or to another part of the body through the blood streams. Cancer that starts in the lung is called a primary lung cancer. There are various types of lung cancer such as Carcinoma, Adenocarcinoma and Squamous cell carcinomas. The rank order of cancers for both males and females among Jordanians in 2008 indicated that were 356 cases of lungcancer accounting for (7.7%) of all newly diagnosed cancercases in 2008.
3. Lung cancer affected 297 (13.1%) males and 59 (2.5%) females with a male to female ratio of 5:1 with alung cancer ranked second among males and 10th among females. It consists of few stages. The first stage starts with taking a collection of CT images (normal and abnormal) from the available database from IMBA home. The second stage applies the several techniques of image enhancement, to get a best level of quality of clearness. The third stage applies image segmentation algorithms which play a effective role in image processing stages, and thefourth stage obtains the general features from enhanced segmented image which gives indicator of normality or abnormality of images.
4. Lung cancer is the most dangerous and widespread cancer in the world according to stage of discovery ofthe cancer cells in the lungs, so the process

early detection of the disease plays a very important and essential role to avoid the serious advance stages to reduce its percentage of distribution.

2. Literature Review

Several researchers has proposed and implemented detection of lung cancer using different approaches of image processing and machine learning. Aggarwal, Furquan and Kalra

[4] 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. Although the system detects the cancer nodule, its accuracy is still unacceptable. No any machine learning techniques has been used to classify and simple segmentation techniques is used. Therefore, combination of any of its steps in our new model does not provide probability of improvement. Jin, Zhang and Jin [5] 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

International Journal of Electrical and Electronics Engineers
Volume 16, Issue No. 01, Jan-June 2024
ISSN (O) 2321-2055
ISSN (P) 2321 -2045

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.

Sangamithraa and Govindaraju [6] uses K mean 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 pre processing median filter is used for noise removal which can be useful for our new model to remove the noise and improve the accuracy.Roy, Sirohi, and Patle [7] developed a system to detect lung cancer nodule using fuzzy interference system and activecontour 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 isperformed using fuzzy inference method. Features like area, mean, entropy, correlation, major axis length, minor axislength are extracted to train the classifier. Overall, accuracy of the system is 94.12%. Counting its limitation it doesnot classify the cancer as benign or malignant which is future scope of this proposed model.

Ignatious and Joseph [8] developed a system using watershed segmentation. In pre processing it uses Gabor filter toenhance the image quality. It compares the accuracy with neural fuzzy model and region growing method. Accuracyof 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 is that it uses 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. Some changes and contribution in this model has probability of increasingthe accuracy to satisfactory level.

Gonzalez and Ponomaryvo [9] 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.

3. Existing model

Eccentricity, circularity, fractal dimension and textural features like mean, variance, energy, entropy, skewness, contrast, and smoothness are extracted to train and classify the support vector machine to identify whether the nodule is benign or malignant. The advantage of this model is that it classifies cancer as benign or malignant, however the limitation of it is that prior

International Journal of Electrical and Electronics Engineers
Volume 16, Issue No. 01, Jan-June 2024
ISSN (O) 2321-2055
ISSN (P) 2321 -2045

information is required about region of interest. Model's classification of benign or malignant using support vector machine can beuseful in our new model.

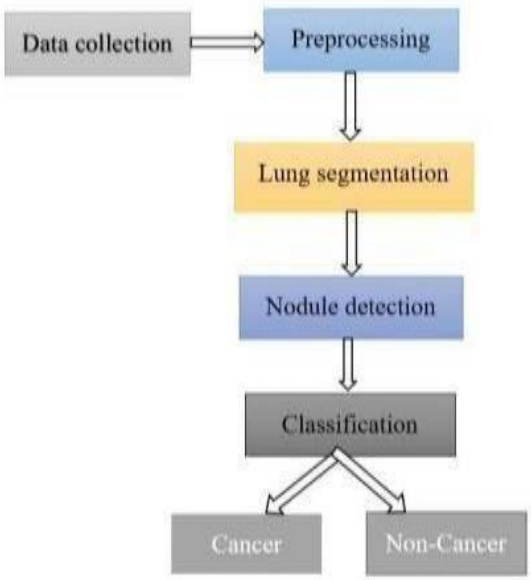


Fig. 1. Proposed model

Analyzing the literature reviews, on the basis of accuracy and advantages of the steps used, the system proposed by Ignatius and Joseph [8] is current best solution. In image pre processing it uses Gabor filter to enhance the image and uses marker controlled watershed method for segmentation and detects the cancer nodule. This model also extracts thefeatures like area, perimeter, and eccentricity only of the cancer nodules. It shows the comparison with other previously proposed models and highlights its accuracy 90.1% which is higher than of those. Even the system is current best solution (fig. 1), it has some limitations. They are highlightedbelow.

Only few features has been extracted for c cancer nodules

- No preprocessing like noise removal, image smoothing which can probably assists in increasing the detection of nodules accurately has been implemented
- No classification as benign or malignant of extracted cancer has been performed

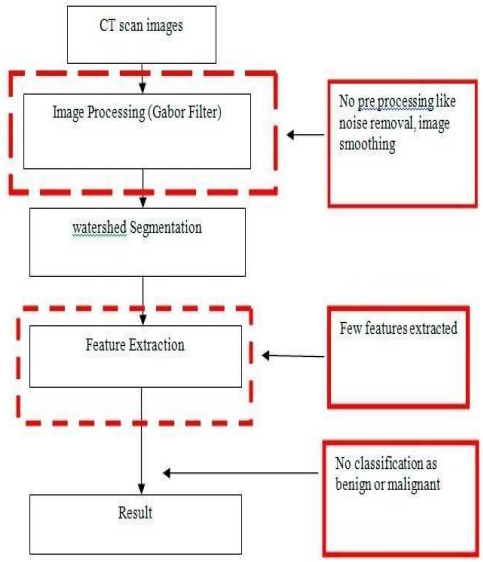
4. Proposed Model

Deep Learning has been proposed as promising tool to classify malignant nodules. Our aim was to retrospectively validate our Lung Cancer Prediction Convolutional Neural Network (LCPCNN), which was trained on US screeningdata, on an independent dataset of indeterminate nodules in an European multi centered trial, to rule out benign nodules maintaining a high lung cancer sensitivity

Main strengths of the proposed model are pointed as below:

- Increase in accuracy of cancer nodule detection than the best current model.
- The classification is the major portion where the cancerous and non- cancerous is identified with the pretrained model
- Together with strength, model has some weakness too. They are pointed as below:
- There is increase in the accuracy but stillit has not reached to best level i.e. nearertowards 100%
- It classifies the cancer as justmalignant or benign but does not classify into different stages like stageI, II,III, IV.

International Journal of Electrical and Electronics Engineers
Volume 16, Issue No. 01, Jan-June 2024
ISSN (O) 2321-2055
ISSN (P) 2321 -2045



5. Implementation

For implementation, real patient CT scan images are obtained from Lung Image Database Consortium (LIDC) archive[12]. It is the database of lung cancer screening CT images for development, training, and evaluation of computer assisted diagnostic methods for lung cancer detection and diagnosis. It was initiated by National Cancer Institute. It consists of 1018 cases of dataset contributed by seven academic center and eightmedical imaging companies. Images are in DICOM format with size 512*512 pixel. DICOM format is difficult to process; therefore, those images are converted to JPEG Gray scale image using software MicroDicomsoftware. MicroDicom opens the DICOM CT scanimages and can also convert to appropriateJPEG format.

Python libraries are designed to solve common problems, such as handling data, performing mathematical operations, interacting with databases, working with files, implementing networking protocols, creating graphical user interfaces (GUIs), and much more. They provide ready-to-use functions, classes, and methods that simplify complex operationsand save development time.

6. Result and Evaluation of Implementation

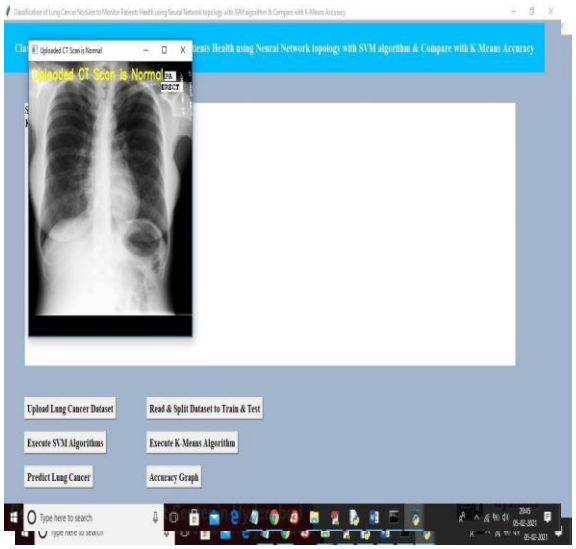


Fig 2: The ct scan describes the result as Normal and abnormal

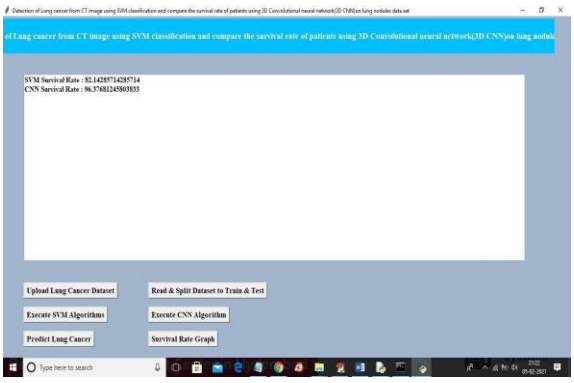


Fig 3: Describes about survival rate of compared algorithms.

International Journal of Electrical and Electronics Engineers
Volume 16, Issue No. 01, Jan-June 2024
ISSN (O) 2321-2055
ISSN (P) 2321 -2045

Comparing the accuracy of proposed model with current model it can be seen that there is progressive increase in accuracy from 88.4% to 92%. Sensitivity remained same. Specificity

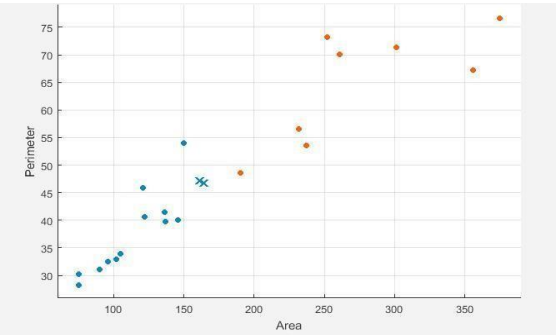


Fig 4. Comparison analysis of Proposed and Current model increased from 40% to 50%

From the detected cancer nodes, features like Area, Perimeter, Centroid, Diameter, Eccentricity and Mean Intensity of the Pixels were extracted. Extracted features were used to Train Support vector machine and trained model was developed. Training time for classification learner app was 5.93 seconds. Classification learner app evaluates the prediction time for the developed trained model to be 310 observations per second. Scatter plot of trained model are as below.

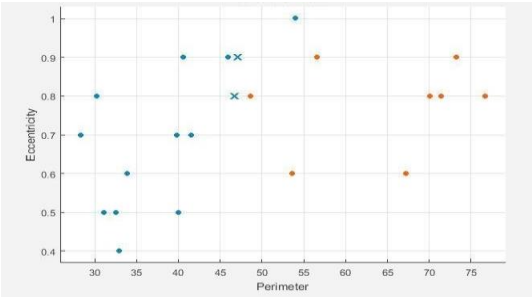


Fig. 5. Scatter plot for Perimeter vs Eccentricity of Trained Model

Conclusion

The current best model has no satisfactory result of accuracy and does not classify degree of cancer of detected nodules. Therefore new system is proposed. The proposed system is used to detect the cancerous nodule from the lung CT scan image using watershed segmentation for detection and SVM for classification of nodule as Malignant or benign. Proposed model detects the cancer with 92% accuracy which is higher than current model and classifier has accuracy of 86.6%. Overall, we can see improvement in the proposed system in comparison to current bestmodel However, this proposed does not classifies into different stages as stage I, II, III, IV of cancer. Therefore, as future scope improvement in this can be done by implementing classification in different stages. Also, further accuracy can be increased by proper pre-processing and eliminations of falseobjects.

References

[1] Gindi,A. M., Al Attiatalla, T. A., & Sami, M.M. (2014) "A Comparative Study for Comparing Two Feature Extraction Methods and Two Classifiers in Classification of Early stage Lung Cancer Diagnosis of chest x-ray images." Journal of American Science, **10(6):** 13-22.

[2] Suzuki, K., Kusumoto, M., Watanabe, S. I., Tsuchiya, R., & Asamura, H. (2006) "Radiologic classification of small

adenocarcinoma of the lung: radiologic-pathologic correlation and its prognostic impact," The Annals of Thoracic Surgery. **81(2):** 413-419.

[3] Xiuhua,G., Tao, S., & Zhigang, L.(2011) "Prediction Models for Malignant Pulmonary Nodules Based-on Texture Features of CT Image." In Theory and Applications of CT Imaging and Analysis. DOI: 10.5772/14766.

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

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

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

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

[8] Ignatius, S., & Joseph, R. (2015) "Computer aided lung cancer detection system." 2015 Global Conference On Communication Technologies (GCCT), DOI: 10.1109/GCCT.2015.7342723.

[9] Rendon-Gonzalez, E., & Ponomaryov, V. (2016) "Automatic Lung nodule segmentation and classification in CT images based on SVM." 2016 9Th International Kharkiv Symposium On Physics And Engineering Of Microwaves, Millimeter And Submillimeter Waves (MSMW). DOI: 10.1109/MSMW.2016.7537995.

[10] Miah, M.B.A., & Yousuf, M.A. (2015) "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.

Author profile

Itte Likhitha Varma currently Pursuing B.Tech (Electronics & Communication Engineering) in Tirumala EngineeringCollege, Affiliated to Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh in the year 2024.

Maddarapu Supriya currently Pursuing B.Tech (Electronics & Communication Engineering) in Tirumala Engineering College, Affiliated to Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh in the year 2024.

Kalluri Praveen Chowdaricurrently Pursuing B.Tech (Electronics & Communication Engineering) inTirumala Engineering College, Affiliated to Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh in the year 2024.

Shaik Jani Basha currently Pursuing B.Tech (Electronics & Communication Engineering) in Tirumala EngineeringCollege, Affiliated to Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh in the year 2024.

Lakku Ashok Kumar currently Pursuing B.Tech (Electronics & Communication Engineering) in Tirumala EngineeringCollege, Affiliated to Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh in the year 2024.