LUNG CANCER DETECTION USING CONVOLUTIONAL NEURAL NETWORKS

Akash P Keladi[1], Ananthsai Raghava K[2], Chennasamudram Harsha[3], Manoj Krishna D[4], Dr. M. Vinoth Kumar[5]

[1] BE Student, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Karnataka, India
[2] BE Student, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Karnataka, India
[3] BE Student, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Karnataka, India
[4] BE Student, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Karnataka, India
[5] Associate Professor, Department of Information Science and Engineering, Dayananda Sagar Academy of Technology and Management, Karnataka, India

ABSTRACT

Cancer-related deaths are on the rise, and one of the leading causes is lung cancer. Detecting lung cancer in its early stages reduces the number of patients who die and dramatically increases the patient's chances of survival. The goal of this study is to use the Convolutional Neural Network (CNN) algorithm to distinguish between malignant and non-malignant lung nodule developments. The CNN technique, as a gradually mechanized methodology, uses digital images as input information and can be easily classified as yield. When observing the disruption in the standard representation of the lung nodule due to its radiological complexity, the machine will detect the image of the lung nodule present in characteristics with various targets and dimensions, radiological complexity and fluctuation of sizes and shapes, thereby doing the constructive side of the classification function and enhancing the precision of classification steps. Many different methods of detecting lung cancer nodules exist but we will be focusing on Convolutional neural networks that utilizes deep learning techniques.

Keyword*: Lung Cancer Detection, CNN, Convolutional neural networks, Deep CNN, Detection, Deep Learning.*

I. INTRODUCTION

The cancer known as lung cancer is responsible for nearly 1.8 million deaths [1]. Despite the fact that many cancer-curing methods, such as chemotherapy and targeted therapeutics, have been shown to be effective, a poor prognosis persisting in the detection of advanced lung cancer nodules results in a survival time of less than one year. As a result, detecting a lung lesion at an early stage is critical for increasing a patient's lifespan. CT (chest tomography) scans are widely used in the medical field to detect lung tumors [2]. A protuberance is a small round shaped protuberance with a diameter of 15/16 mm in the lung (nodule). The chest tomography scanner produces images that provide a detailed picture of the internal organs that form a nodule. This also takes into account the round-shaped protuberance structure, which is not always cancerous. Radiologists dislike these tasks because they require them to analyze numerous small protuberances, which takes more time. This task is made easier by a computer-aided diagnosis (CAD) system. After processing the digital image, the computer-aided diagnosis system highlights sections that are clearly affected by lung cancer. Radiologists and doctors' benefit from improved perception and interpretation of medical images [3]. In the approach we are proposing a method using CNN and deep learning aptly called Deep CNN.

II.METHODOLOGY

we propose a method that uses a Convolutional Neural Network to detect cancerous nodules supported by Chest Tomography images of the lungs. In the first step, we extract lung regions from the Chest tomography image, and each of the slices in this region is segmented to isolate tumors. By inputting the segment in the previous step, the Convolutional Neural Network architecture is trained. Following that, CNN is used to check the patient scans, and the softmax layer determines the final outcome. The study's ultimate goal is to determine whether the patient's lung nodule is malignant or benign [4].

2.1 DATASET

Lung X-Ray images from JSRT dataset. LIDC-IDRI is where we get our dataset, from which we train our model. LIDC-IDRI contains 1000s of CT scans or X-ray images that consists of large and small nodules saved in DICOM format and converted them into png format.

Fig -1: CT Scan slice of a normal lung.

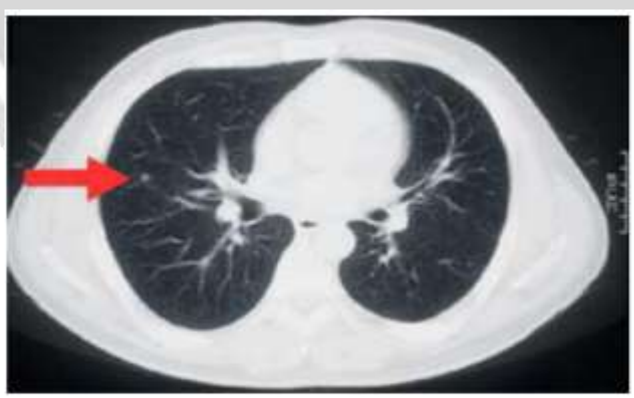

Fig -2: CT Scan slice of an early-stage of lung cancer.

Fig -3: X-Ray image of a normal lung.

Fig -4: X-Ray image of Cancerous lung.

2.2 FLOW DIAGRAM

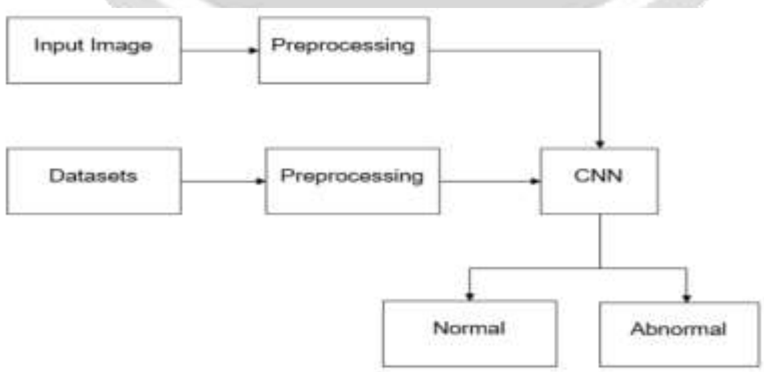

Fig -5: Flow diagram for the CNN architecture.

III.PROPOSED METHODOLOGY

3.1 PRE-PROCESSING

A median filter is used in the pre-processing stage to revive test beneath test by minimizing the effects of degradations during acquisition. The median filter simply replaces each pixel value with the norm of its adjacent pixels, which includes itself. As a result, pixel values that differ from their adjacent pixels will be removed during this process [6].

3.2 DEEP LEARNING

It is made up of multiple layers that are non-linear nodes, coupled with a computer file, and a collection of weights for assigning importance to inputs for the interrelated task that the algorithm is attempting to tell in supervised and/or unsupervised behavior. The sum of that input's product and the weights is passed through the activation function of nodes. Every layer of the output is given as input to the next layer ranging from the input layer at the same time. Learning is frequently performed in a variety of forms that correspond to a variety of levels of abstraction [7].

3.3 CONVOLUTION NEURAL NETWORKS (CNN)

Convolutional layer: By applying a filter that scans the entire image, a few pixels at a time, a feature map is created to predict the class probabilities for each feature.
Pooling layer (down-sampling): reduces the amount of information generated by the convolutional layer for each feature while retaining the most important information (the process of the convolutional and pooling layers usually repeats several times).
Fully connected input layer: compresses the outputs of previous layers into a single vector that can be used as an input to the next layer Fully connected layer: To predict an accurate label, weights are applied to the input generated by the feature analysis.
Fully connected output layer: Produces the final probabilities for classifying the image.

CNN is a Deep-NN that is made up of many hidden layers, including a convolutional layer, a rectified linear unit layer, a pooling layer, and a completely connected normalized layer. To reduce memory, CNN shares its weighs heavily in the convolutional layer, improving performance. To detect candidate nodules in Chest tomography scans, we employ a deep learning methodology. The ultimate goal is to train DCNN to detect lung protuberance in a significant subset of Chest tomography images. In unprocessed Chest tomography images, this can be used to determine the boundary and location of candidate nodules. The Convolutional Neural Network is made up of twenty layers, which are divided into two sections. The first section extracts valuable volumetric information from the input data using a Convolutional Neural Network. The network also includes a number of convolution layers, a Rectified linear unit, and a max pooling layer. The next component of the Convolutional Neural Network is the Classifier. A Deep CNN is made up of many hidden layers, including a fully-connected layer and a thresholding layer. Finally, the softmax layer is employed to determine the outcome. [8].

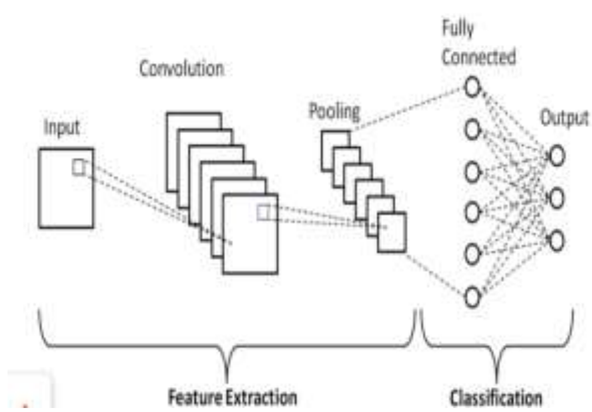

Fig-6: Flow Diagram of Proposed CNN

IV. RESULT COMPARISION

Result Comparison of different methods on lung cancer detection.

Authors	Model	Results
Sharma et al	CNN	Accuracy=0.9733p recision=0.9673A UC = 0.9954
Chon et al	Deep CNN	Sensitivity=0.652 ACC=0.665 AUC=0.663
Yashas-wini et al	CNN & SVM	Wt. average =.90 Precision=0.80 Recall=0.80
W.Zuo et al	CNN with knowledge Transfer	CPM=0.787
Jenuwine et al	3D CNN	AUC=0.722
Lin et al	Taguchi Parametric optimization	Accuracy=94.68
Hongtao et al	2D CNN with faster R-CNN	Sensitivity=86.42

Fig-7: Comparison of the results of the papers.

V. RESULTS

5.1 DEEP CNN MODEL

The CNN model is trained with both Computed tomography (CT) Scans from LIDC-IDRI dataset and Lung X-Ray images from JSRT dataset. The CNN model is trained with 1000 CT-Scans and 300 X-Ray images in 80:20 ratio.

Basic measures derived from the confusion matrix:

Error rate (ERR) is calculated as the number of all incorrect predictions divided by the total number of the dataset. The best error rate is 0.0, whereas the worst is 1.0.

Error Rate=(FP+FN)/(P+N)

Accuracy (ACC) is calculated as the number of all correct predictions divided by the total number of the dataset. The best accuracy is 1.0, whereas the worst is 0. It can also be calculated by 1 – ERR.
Accuracy = (TP+TN)/(P+N)

Precision (PREC) is calculated as the number of correct positive predictions divided by the total number of positive predictions. It is also called positive predictive value (PPV). The best precision is 1.0, whereas the worst is 0.0.

Precision = TP/TP+FP

False positive rate (FPR) is calculated as the number of incorrect positive predictions divided by the total number of negatives. The best false positive rate is 0.0 whereas the worst is 1.0. It can also be calculated as 1 – specificity [9].

FP= FP/TN+FP

The performance analysis results are as follows:

Sl. No	1
Dataset	X-Ray images From JSRT
Data Set Split	Total=300 Training=80% Testing=20%
Error Rate (ERR)= **FP+FN/P+N**	10%
Precision =TP/FP+FN	0.89
Accuracy =(1-ERR)	±90%

Table-1: Performance analysis of X-Ray Images.

Sl. No	2
Dataset	CT-Scans from Kaggle
Data Set Split	Total=1000 Training=80% Testing=20%
Error Rate (ERR)= **FP+FN/P+N**	5%
Precision =TP/FP+FN	0.9
Accuracy =(1-ERR)	±95%

Table-2: Performance analysis of CT-Scans.

Fig-8: Lung cancer detection GUI.

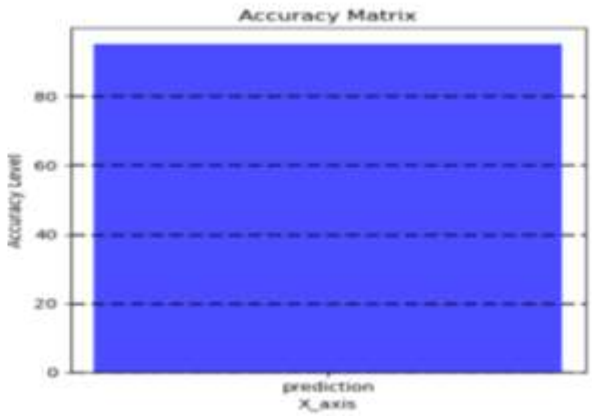

Fig-9: Lung cancer detection Accuracy Prediction.

V. CONCLUSION

It is clear that diagnosing lung cancer in its early stages remains difficult. To address all of these issues, we propose an optimized Deep CNN method [8]. The reason for using CNN is that it can extract spatial information from data using kernels, which other networks cannot. D-CNN is used in the proposed method to detect lung cancer based on CT images. A Deep CNN consists of many hidden layers such as fully-connected layer, thresholding layer. Towards the end, softmax layer is used to determine the result. It is proven that CT scans gives more accurate analysis of the lung nodules and detects lung cancer with almost ±95% accuracy rate compared to ±90% X-Ray images.

VI. REFERENCES

[1] https://www.who.int/newsroom/factsheets/detail/cancer

[2] B. Kumar S and M. Vinoth Kumar, "Detection of Lung Nodules using Convolution Neural Network: A Review," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2020, pp.590-594, doi:10.1109/ICIRCA48905.2020.9183183.

[3] Samhitha, B. K., Mana, S. C., Jose, J., Vignesh, R., & Deepa, D. (2020). Prediction of Lung Cancer Using Convolutional Neural Network (CNN). International Journal, 9(3).

[4] Sharma, Sayana, Mandeep Kaur, and Deepak Saini. "Lung Cancer Detection Using Convolutional Neural Network."

[5] Chon, A., Balachandar, N., & Lu, P. (2017). Deep convolutional neural networks for lung cancer detection. Standford University.

[6] Yashaswini, S. L., & Prasad, K. V. (2019). Lung Cancer Nodules Classification and Detection Using SVM and CNN Classifiers.

[7] W. Zuo, F. Zhou, Z. Li and L. Wang, "Multi-Resolution CNN and Knowledge Transfer for Candidate Classification in Lung Nodule Detection," in IEEE Access, vol. 7, pp.32510-32521,2019, doi:10.1109/ACCESS.2019.2903587

[8] Jenuwine, N. M., Mahesh, S. N., Furst, J. D., & Raicu, D S. (2018, February). "Lung nodule detection rom CT scans using 3D convolutional neural neworks without candidate selection". In Medical Imaging 2018: Computer-Aided Diagnosis (Vol. 10575, p. 1057539). International Society for Optics and Photonics.

[9] Sasikala, S., Bharathi, M., & Sowmiya, B. R. (2018). “Lung Cancer Detection and Classification Using Deep CNN”. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN, 2278-3075.