A SURVEY ON LUNG CANCER DETECTION USING CONVOLUTION NEURAL NETWORK IN CT IMAGES

[1]Shankara C, [2]Dr.Hariprasad S A, [3] Karumuru Krishna Pavan, [4] Govardhan, [5] Kullay Basha, [6]Jakeer Hussain

[1]Research Scholar, [2] Director & Professor, [3]Student, [3]Student, [4]Student, [5]Student, [6]Student
[1]ECE department
[1]Jain University, Bangalore, India

Abstract: Lung Cancer in today's world is one among the major outspread disease which is the subject of maximum deaths every year. Early and accurate lung cancer detection could boost the endurance rate. Medical image processing play a significant role in an identification of lung tumor using Computer Tomography (CT) scans images. At present, many approaches are present for definite disclosure of lung nodules. Especially, Convolution Neural Network on the basis of deep learning boosted the success amount of recognition and analysis of lung tumor. This paper reviewed the different approaches for lung cancer detection with Convolution Neural Network (CNN). The different objections in CNN for further increasing the accurate detection of lung tumor are addressed.

IndexTerms **-Convolution Neural Network, LIDC, Lung Cancer Detection, 3D CNN.**

I. INTRODUCTION

In today's world lung cancer is the leading widespread diseases, because of which the rate of mortality is increasing rapidly. It is estimated that by 2030 the mortality rate due to lung cancer may reach up to 2.45 million[1] (which is 39% more than that of the cases reported in 2018). By 2040 the number of people suffering from lung cancer could increase up to 32, 99,640. More than 1.3 million people across the globe die due to lung cancer and approximately around 1.8 million new cases are added every year. In the year 2018, the number of total lung cancer cases reported in the entire world was approximately 2.09 million (out of which, 1.36 million were men and 0.73 million were women). It was also reported that the mortality rate due to lung cancer (in the year 2018) was 1.76 million[2]. It is estimated that 1 out of every 68 male and 1 out of 201 female are affected with lung cancer in India. Also, it is stated that 1 person in out of every 101 is affected with lung cancer[3]. In 2020, there were 98,278 new instances of lung cancer cases (out of which 71,788 are men and 26,490 are women). In the year 2018 the total number of cases reported in India was 67,795 (out of which 48,698 were men and 19,097 were women). Around 45,363 deaths were recorded (which included 27251 men and 18,112 women) .The Survival rate of lung cancer patients (including all stages) is roughly about 14% and the time span of survival is for about 5-6 years only[4].

American cancer society states about 80 % deaths from lung cancer are due to smoking and passive smoking. The increased rate of pollution in the environment is also the major cause for lung carcinoma[5]. Lung cancer is of two types, Non- Small-Cell Lung Cancer (NSCLC) and Small-Cell-Lung- Cancer (SCLC)[6]. Almost 85% cases of lung cancer is mainly caused by the NSCLC and SCLC considered for the remaining 15% and is usually caused by smoking or chewing tobacco [7].Because of low cost, short imaging time, less distortion and widespread availability Computer Tomography (CT) scan is used as the efficient medical screening test for easy diagnosis and recognition of lung cancer over other imaging modalities like chest radiograph (X-Ray), Sputum Cytology, Magnetic Resonance and Imaging (MRI) scans and Position Emission Tomography[8]. A Computed Tomography (CT) scan can be used to get the detailed information of cancerous cells in lungs, which can be used to detect lung cancer and it also decreases the chance of death due to lung cancer.

In traditional method of nodule detection, there are many different stages which may increase the error rate and computational time. But in CNN as it is an end to end process where pre-processing, the segmentation and the feature extraction do not exist due to which computational time will be less. In the traditional method only image processing method is used for the detection of lung nodule, whereas in convolution neural network method the neural network is used for training and testing of the classification nodules[9].There is necessity to reveal the existence of early stage cancer in lung CT scan images by applying image processing operation and CNN method.

In this paper the remaining part is described as follows: The Section 2 explain about the convolution neural network. The Section 3 explain the literature survey on convolution neural network method to detect the tumor in the lung .The Section 4 explains various results obtained by different researchers and Summary of techniques used to detect lung cancer. The Section 5 explains the conclusion of this survey work.

II. CONVOLUTION NEURAL NETWORK

Convolution Neural Network which is a concept of deep learning method. CNN learns automatically using various stack layers of convolution. It considers all the useful content from the given data without focus on the pre-processing and feature extraction procedures. Convolution Neural Network consists of 3 main components such as the convolution layer, the pooling layer as well as layer with fully –connected [11] as shown in figure 1.

In traditional method of nodule detection, there are many different stages which may increase the error rate and computational time. But in CNN as it is an end to end process where pre-processing, the segmentation and the feature extraction do not exist due to which computational time will be less. In the traditional method only image processing method is used for the detection of lung nodule, whereas in convolution neural network method the neural Network is used for training and testing of the classification nodules.

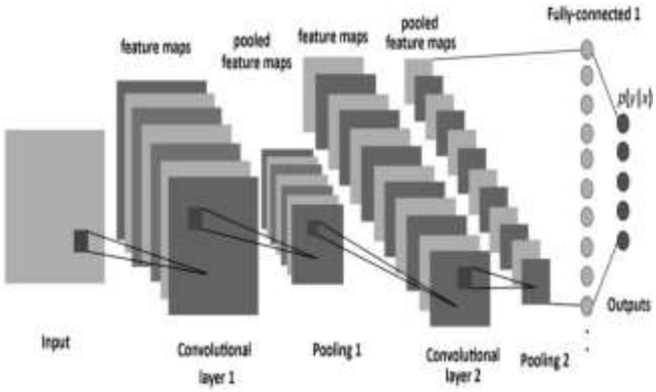

Figure 1: Convolution Neural Network with pooling layer and Fully Connected Layer

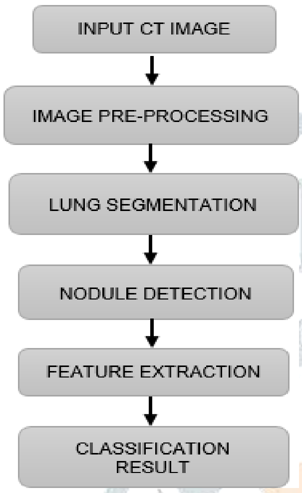

Figure 2: Traditional method

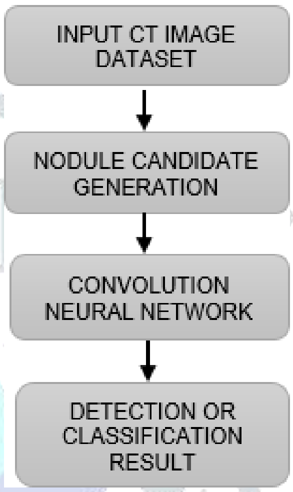

Figure 3: Convolution Neural Network method

III. LITERATURE SURVEY

N.Mohanapriya et al.[12] proposed a four stage system for the detection of lung cancer using Deep Convolutional Neural Networks. It consists of tumor detector stage, malignancy detector stage, tumor classifier and patient classifier. Initially the images collected from the LIDC dataset are preprocessed to remove the noise then applied to three different DCNN architectures. All the three architectures are experimented with different patch size. The system performance is evaluated with parameters such as accuracy, recall, precision, dice score coefficient. The three DCNN architecture with fully connected layer along with Softmax classifier performs well than other architecture with 82.55% accuracy, 85.79 % of precision, 85.24% of recall and 83.14% of dice score coefficient. With the proposed system the tumor is detected either benign or malignant.

Rohit Y. Bhalerao et al. [13] proposed a novel method for detection of cancer in the lung by applying image processing and CNN method . Initially the images collected from LIDC database are preprocessed using gray scale conversion, binarization, and median filter. The image is segmented using watershed segmentation and then dilation is performed. Boundaries in the image are generated using boundary plotting. The tumor is detected as normal, benign or malignant using convolution neural network with input layer, convolution layer, the rectified linear unit, layer of max-pooling, layer with fully connected, the Softmax layer as well as Classification (output) Layer. The system achieved an accuracy of 94.34%.

Prajwal Rao et al.[14] has proposed a system which contains Can Net layers along with CNN to analyze tumor with domain knowledge of medicine as well as neural networks. The Can Net consists of input layer, two convolution layers, the pooling layer, the dropout layer and finally layer with fully connected. The formation of Can Net is performed by author using framework of Caffe deep learning which is developed by the Berkeley Vision and Learning Center. In the proposed method the system performs better than traditional neural network and existing CNNs. For the implementation used LIDC and achieved 76%.

M. Bikromjit Khumancha et.al.[15] has proposed a 3D CNN for lung cancer detection. The proposed method used LUNA 16 grand Challenge dataset which also included LIDC/IDRI dataset. A total of 888 CT scans were used in which sample slices size were less than or equal to 2.5 mm. Pre-processing is done by converting each pixel values into Hounsfield units (HU). Nodule detection is done by segmenting small 3D cubes around lung coordinates, later data augmentation is done to handle low amount of nodules. Cancer detection is done by using segmentation of CT scan images. Segmentation includes binarization, clearing borders, labeling images, morphological operations such as erosion and closing and finally filling holes. Later masking is applied for the segmented lung. A 3D CNN is trained using these segmented images, with ReLu activation function, sigmoid activation function and dropouts of 0.2. The proposed system obtained precision of 89.24%, recall of 82.17% and an accuracy of 90.78%.

Adarsh Pradhan et.al.[16] has proposed CNN model using SPIE-AAPM lung CT challenge dataset for training. Pre-processing is done by thresholding the CT images with HU value of lung and the air pixels are removed then holes are filled and nodule mask is applied. CNN is trained which consists of ReLu function, Max Pooling and Softmax optimizer. 30 epochs were used in the training model with a batch size of 2. Training, validation and testing were in 60% of images, 20% of images and 20% of images. The model has obtained an accuracy of 83.33% in training and obtained an accuracy of 100 % in testing. The model has also obtained precision value of 1, kappa value of 1, recall value of 1and F- score value of 1.

Azmira Krishna et al.[17] proposed CNN based model for classification of CT images. Images are collected from LIDC database and are used as input for the proposed system. Gaussian Smoothing filter is used for pre-processing and then watershed segmentation is used .To decrease the number of assets required to predict, feature extraction is done. CNN classifier is used to give accurate results and the proposed system gives an accuracy rate of 94.5%.

Thiago Jose Barbose Lima et al. [18] proposed an automated algorithm to classify lung nodules into benign and malignant using LIDC dataset. To extract the characteristics and for classification of nodule, three 3D CNN architecture with distinct input size as well as different number of convolution layer as well as pooling layers are used. The performance of the system is evaluated using metrics such as accuracy; kappa, sensitivity, specificity, and area under the ROC. The features of the nodules are extracted using 3D CNN architecture and are classified as benign and malignant. The segmented images are collected from github with input size 32×32×32 are given to architecture consists of two number of convolutional blocks as well as two number of pooling layers. The proposed system achieved an efficiency of 90.77%.

3D CNN based system proposed by Pouria Moradi et al. [19] in detection of lung cancer and to decrease false positive rate using CT images which helps to get high sensitivity. For false positive reduction, used multi-level contextual CNNs. Due to the different size of nodules designed four different size 3D CNN so that the fusion of these 4 networks can provide a good coverage error of classifiers. Input images are collected from LUNA 16 and used for training. In all layers except for the last layer, the rectified linear units (ReLu) function was used. In addition, in the last layer, the soft max activation function is used to generate the final probability. The proposed system achieved an accuracy of 91.23%.

IV. RESULTS AND DISCUSSION

N.Mohanapriya et al.[12] collected images from LIDC dataset and achieved an accuracy of 82.55%. Rohit Y. Bhalerao et al. [13] collected images from LIDC database and obtained an accuracy of 94.34%. Prajwal Rao et al.[14] collected images from LIDC database and achieved an accuracy of 76 %. M. Bikromjit Khumancha et.al.[15] collected images from LUNA16 database and obtained an accuracy of 90.78%. Adarsh Pradhan et.al.[16] collected images from SPIE-AAPM lung CT challenge dataset and achieved an accuracy of 100%. Azmira Krishna et al.[17] collected images from LIDC database and obtained an accuracy of 95%. Thiago Jose Barbose Lima et al. [18] collected images from LIDC dataset and achieved an accuracy of 90.77%. Pouria Moradi et al.[19] collected images from LUNA16 dataset and achieved an accuracy of 91.23 %.

Table 1: Summary of techniques used by researchers to detect lung cancer.

Author	**Techniques used**	**Performance Measure**	**Limitations**
N.Mohanapriya et al.[7]	Three DCNN architecture with different patch size.	Accuracy - 82.55%	The proposed system not validated with the real time database.
Rohit Y. Bhalerao et al.[8]	Gray scale conversion, binarization, median filter, watershed segmentation, boundary plotting, CNN	Accuracy - 94.34%.	Static database is used. Dynamic real time database is not used. Number of convolution and batch size are not sufficient.
Prajwal Rao et al.[9]	Can Net layer with CNN classifier, Caffe deep learning framework	Accuracy- 76%	Experimentation on different CNN hyper parameters cannot be done to rise the accuracy
M.Bikromjit Khumancha et.al.[15]	Binarization, clearing borders, labeling images, morphological operations such as erosion and closing, 3D CNN	Accuracy- 90.78%	CNN architectures like VCG, Google Net and ResNet are not applied. Different input sizes to increase the accuracy are not implemented.
Adarsh Pradhan et.al.[16]	Conversion to Hounsfield Unit, Thresolding, CNN	Accuracy- 100%	Limited dataset used and not validated with dynamic database
Azmira Krishna et al.[12]	Erosion, binary image slicing, Gaussian smoothing filtering Watershed segmentation, CNN classifier.	Accuracy- 95%	Limited dataset is used for detection of lung cancer.
Thiago Jose Barbose Lima et al. [18] José Barbosa Lima	3D CNN	Accuracy- 90.77%	Data augmentation technique is not applied. Auto encoders and the generative adversarial network techniques are not used
Pouria Moradi et al. [19]	Multi-level contextual CNN, 3D CNN	Accuracy- 91.23%	Limited dataset is used and better DCNN architecture can be applied to increase the accuracy.

V.CONCLUSION

The work which is performed in this paper mainly focuses on the various methods used by researcher to predict lung cancer using convolution neural network. Various techniques used in convolution neural network, data set used by researchers are reviewed along with limitations**.** From the survey it is observed that lung CT images are detected as normal, benign or malignant using convolution neural network and it also help the researchers to overcome the limitations exist in the current CNN method.

VI.ACKNOWLEDGMENT

The researcher would like to express respect and gratitude towards the respected guide and mentor Dr. Hariprasad.S.A, Research Director, Director of Faculty of Engineering and Technology, Jain (Deemed-to-be University) Bangalore for his continuous motivation, technical advice, guidance, support and valuable suggestions to carry out the research work.

REFERENCES

[1] B. Abdillah, A. Bustamam, and D. Sarwinda, "Image processing based detection of lung cancer on CT scan images," in *Journal of Physics: Conference Series*, 2017, vol. 893, no. 1, pp. 1–7, doi: 10.1088/1742-6596/893/1/012063.

[2] M. H. Jony, F. Tuj Johora, P. Khatun, and H. K. Rana, "Detection of Lung Cancer from CT Scan Images using GLCM and SVM," in *1st International Conference on Advances in Science, Engineering and Robotics Technology 2019, ICASERT 2019*, 2019, pp. 978-1-7281-3445–1/19, doi: 10.1109/ICASERT.2019.8934454.

[3] A. E. Kaucha, Deep Prakash, P.W.C. Prasad, Abeer Alsadoon, "Early Detection of Lung Cancer using SVM Classifier in Biomedical Image Processing," in *IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017)*, 2017, pp. 3143–3148, doi: 978-1-5386-0814-2/17.

[4] S. J. S.Kalaivani, Pramit Chatterjee, "Lung Cancer Detection Using Digital Image Processing and Artificial Neural Networks," in *International Conference on Electronics, Communication and Aerospace Technology ICECA 2017*, 2017, pp. 100–103, doi: 978-1-5090-5686-6/17.

[5] R. S. Manika Jha, Richa Gupta, "A Review on Non-Invasive Biosensors for Early Detection of Lung Cancer," in *IEEE*, 2020, pp. 162–166, doi: 978-1-7281-5493-0/20/$31.00 ©2020 IEEE.

[6] J. Madhura, Y. Amith, A. Gurjar, K. Ayappa, M. Bharath, and M. L. Goutham, "Methods of impulsive noise reduction using image processing," *Proc. Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2017*, vol. 2017-Janua, pp. 296–302, 2017, doi: 10.1109/ICECA.2017.8203691.

[7] M. Islam, A. H. Mahamud, and R. Rab, "Analysis of CT Scan Images to Predict Lung Cancer Stages Using Image Processing Techniques," in *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2019*, 2019, pp. 961–967, doi: 10.1109/IEMCON.2019.8936175.

[8] N. S. Nadkarni and S. Borkar, "Detection of lung cancer in CT images using image processing," *Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019*, vol. 2019-April, no. Icoei, pp. 863–866, 2019, doi: 10.1109/icoei.2019.8862577.

[9] N. K. Chauhan and K. Singh, "A review on conventional machine learning vs deep learning," *2018 Int. Conf. Comput. Power Commun. Technol. GUCON 2018*, pp. 347–352, 2019, doi: 10.1109/GUCON.2018.8675097.

[10] N. D. T. Kanitkar, Sayali Satish, "Detection of Lung Cancer Using Marker-Controlled Watershed Transform," in *2015 International Conference on Pervasive Computing (ICPC)*, 2015, pp. 1–5, doi: 978-1-4799-6272-3/15/.

[11] B. S. Kumar and M. Vinoth Kumar, "Detection of Lung Nodules using Convolution Neural Network: A Review," *Proc. 2nd Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2020*, pp. 590–594, 2020, doi: 10.1109/ICIRCA48905.2020.9183183.

[12] N. Mohanapriya, B. Kalaavathi, and T. S. Kuamr, "Lung Tumor Classification and Detection from CT Scan Images using Deep Convolutional Neural Networks (DCNN)," *Proc. 2019 Int. Conf. Comput. Intell. Knowl. Econ. ICCIKE 2019*, pp. 800–805, 2019, doi: 10.1109/ICCIKE47802.2019.9004247.

[13] R. Y. Bhalerao, H. P. Jani, R. K. Gaitonde, and V. Raut, "A novel approach for detection of Lung Cancer using Digital Image Processing and Convolution Neural Networks," *2019 5th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2019*, pp. 577–583, 2019, doi: 10.1109/ICACCS.2019.8728348.

[14] P. Rao, N. A. Fereira, and R. Srinivasan, "Convolutional neural networks for lung cancer screening in computed tomography (CT) scans," *Proc. 2016 2nd Int. Conf. Contemp. Comput. Informatics, IC3I 2016*, pp. 489–493, 2016, doi: 10.1109/IC3I.2016.7918014.

[15] M. B. Khumancha, A. Barai, and C. B. R. Rao, "Lung Cancer Detection from Computed Tomography (CT) Scans using Convolutional Neural Network," *2019 10th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2019*, pp. 1–7, 2019, doi: 10.1109/ICCCNT45670.2019.8944824.

[16] A. Pradhan, B. Sarma, and B. K. Dey, "Lung Cancer Detection using 3D Convolutional Neural Networks," *2020 Int. Conf. Comput. Perform. Eval. ComPE 2020*, pp. 765–770, 2020, doi: 10.1109/ComPE49325.2020.9200176.

[17] A. Krishna *et al.*, "Lung Tumor Classification and Detection from CT Scan Images using Deep Convolutional Neural Networks (DCNN)," *Proc. 2nd Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2020*, pp. 18–21, 2020, doi: 10.1109/ICCIKE47802.2019.9004247.

[18] D. Ushizima, A. Oseas, and D. C. Filho, "LUNG CT SCREENING WITH 3D CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES Thiago José Barbosa Lima ⋆ Electrical Engineering , Federal University of Piauí (UFPI), Teresina , PI , Brazil Information Systems , Federal University of Piauí (UFPI), Picos , PI , ," 2020.

[19] M. J. Pouria Moradi, "Detecting Lung Cancer Lesions in CT Images using 3D Convolutional Neural Networks," in *4th International Conference on Pattern Recognition and Image Analysis (IPRIA), 6 and 7 March, Tehran, Iran*, 2019, pp. 114–118, doi: 978-1-7281-1621-1/19/$31.00 ©2019 IEEE.