International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

www.ijiemr.org

IJIEMR Transactions, online available on 28 Aug 2022. Link

:http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=Issue 08

10.48047/IJIEMR/V12/ISSUE 08/50

Title **Lung cancer detection and classification using deep learning Techniques**

Volume 12, ISSUE 08, Pages: 344-351

Paper Authors **M. Premchander, Dr. M.Venkateshwara Rao,Dr. T.V Rajinikanth,Subba Reddy Borra**

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Lung cancer detection and classification using deep learning Techniques

M. Premchander[1]
Assistant Professor, Department of Information Technology, Mallareddy Engineering College for Women,Hyderabad.

Dr. M.Venkateshwara Rao[2]
Professor In Department of Information Technology, GITAM University, Vishakhapatnam

Dr. T.V Rajinikanth[3]
Professor & Dean in Department of CSE, Srinidhi Institute of Science and Technology,Hyderabad

Subba Reddy Borra[4]
Professor & HOD Department of Information Technology, Mallareddy Engineering College for Women, Hyderabad.

***Abstract*:** The largest cause of cancer-related fatalities worldwide is lung cancer. One type of cancer that starts in the lungs is lung cancer. Your lungs are two pliable organs located in your chest that allow you to breathe in oxygen and exhaust carbon dioxide. Medical imaging technologies are required for both the early detection of lung cancer and the ongoing monitoring of lung cancer throughout treatment. Chest X-rays, magnetic resonance imaging, positron emission tomography, computed tomography, and molecular imaging techniques have all been extensively studied for the detection of lung cancer.These methods have some drawbacks, such as the inability to automatically categorise cancer photos, which makes them unsuitable for patients with other illnesses. The creation of a sensitive and precise method for the early diagnosis of lung cancer is urgently required. With quickly developing applications across, deep learning is one of the medical imaging fields that is expanding the fastest. The detection and classification of lung nodules can be done more rapidly and accurately by physicians.

Keywords: Computed Tomography images, Segmentation, Deep learning, Lung cancer, Convolution neural networks, Classification.

1. Introduction.

In the US, lung cancer has the greatest morbidity and fatality rates and is the most common disease as well. GLOBOCAN predicted that there would be 2.09 million new cases and 1.76 million deaths from lung cancer in 2018. Globally, the number of lung cancer cases and deaths has considerably grown. Small cell lung cancer (SCLC) accounts for 12–15% of lung cancer cases, while non-small cell lung carcinoma (NSCLS) accounts for 85-88% of lung cancer cases. Due to the invasiveness and heterogeneity.Mohanapriya, et al[1].

The detection of lung nodules has been thoroughly studied over the past 20 years using a variety of medical imaging techniques, including chest X-ray, positron emission tomography (PET), magnetic resonance imaging (MRI), Ju, W., Xiang et al[2], computed tomography (CT), low-dose CT (LDCT), and chest radiograph (CRG), Although CT is the gold standard imaging technique for lung nodule diagnosis, it has a high false-positive rate, can only detect apparent lung cancer, and emits dangerous X-ray radiation ,Luan, X. C et al [3]. LDCT has been suggested as a way to identify lung cancer while reducing harmful radiation.

Classifying benign and malignant nodules is difficult, though. The identification, segmentation, and classification of benign and malignant pulmonary nodules are the core topics of research on deep learning-based lung imaging approaches, Henschke.C.I,et al [5]. To enhance the performance of deep learning models, researchers mostly concentrate on creating new network architectures and loss functions. Recent reviews of deep learning methods have been published by several research teams, Haider, W., et al [6].

2. Proposed Method

The automatic identification of lung cancer using deep learning techniques has been proposed as an algorithm. Jiang, J. et al [9].The VGG-16 Architecture was selected by the Base network. Through the use of deep learning algorithms, the classification of lung cancer, including adenocarcinoma, large cell carcinoma, and squamous cell carcinoma, has been distinguished from photographs of healthy lungs in this work. The dataset is applied to two different pre-trained neural networks, GoogleNet and Vgg16, and their performance is assessed in terms of accuracy, sensitivity, and precision. These values are then compared to those of the CNN network, Hussein, S., et al [10]. which has two blocks of layers made up of a convolution layer, a normalisation layer, and a pooling layer. The network is trained and tested using 100 samples of images from each class. among the 70 pictures 30 are used for validation purposes, while the remaining 30 are used for training. It has been noted that only two CNN network blocks were required to meet the accuracy levels of GoogleNet and Vgg16 Network, Henschke, C et al[11]

Due to the fact that GoogleNet and Vgg16 Network are the smallest and heaviest pre-trained neural networks currently available, respectively, they were chosen. GoogleNet, which supports images with a resolution of 224x224 pixels, has a depth of 28 layers, a memory size of 27 MB, and 7.0 million parameters. In contrast, Vgg16 supports images up to 224x224 pixels in size and has 16 layers, 535 MB of memory, and 144 million parameters, A.R., Hemmati, C et al [12]

2.1 Lung Imaging Techniques

Radiologists use medical imaging equipment to identify lung conditions. When compared to the other medical imaging methods, CT offers additional advantages, such as details on lesion size, location, characterization, and growth, which can be utilised to find nodules and lung cancer. More accurate radiation targeting made possible by 4D CT has a substantial impact on lung cancer treatment. Gierada, D et al [13]

The most effective lung imaging method without ionising radiation is MRI, however it is expensive, has time and cost constraints, and delivers insufficient information. Medical imaging techniques are a crucial component of early-stage lung cancer diagnosis strategies that increase survival rates. Roy, K t al [14] However, these methods have some drawbacks, such as large false positive rates and the inability to automatically detect lesions. Lung cancer detection has been made possible by a number of CAD systems. Khdair, H.et al [15].

3. Deep Learning Techniques

The most popular unsupervised learning algorithms in medical images consist of three types of neural networks: convolutional (CNN), deep (DCNN), and recurrent (RNN). CNN architecture is one of the most used supervised deep learning techniques for lesion segmentation and classification since it requires less pre-processing. CNN architectures have been used to classify and segment medical pictures, such as with AlexNet ,VGGNet and Mask R-CNN. Google Deep Mind used the reinforcement learning technique for the first time in 2013. Since then, numerous studies have been conducted to enhance the accuracy, sensitivity, and specificity of lung cancer detection using reinforcement learning techniques. Labelled datasets are used in semi-supervised learning techniques, J Kuruvilla et al [19].

4. Lung Cancer Prediction Using Deep Learning

4.1. Pre-Processing Techniques:

A deep learning algorithm is fed to the pre-processed images, and the algorithm is then trained on the image datasets and tested. The final classifier's precision is impacted by image noise. For pre-processing, a number of noise reduction techniques have been created, Sairamya, N.J., et al [20]. These techniques aim to increase accuracy and generalisation performance. After denoising, rescaling for a normalisation technique, such as min-max normalisation, Bhat, G.M.,et. al[21].

4.2 Performance metrics

Deep learning algorithm performance, Different metric also referred to as performance metrics or evaluation metrics—are used to assess the effectiveness or quality of the model. The category or classes of data are found in a classification problem using training data. The model gains knowledge from the provided dataset before classifying the fresh data into groups or classes in accordance with the training, Sairamya, N.J et al[22]. Accuracy evaluates the capacity for results using the available information features.

The ROC's detection sensitivity can be changed. The area under the receiver operating characteristic curve (AUC) has been used to assess how well the proposed deep learning model works. Nanglia, P A et al [23]. Deep learning-based algorithm performs better when it has higher integrity of AUC, sensitivity, specificity, sensitivity, accuracy Ganeshan, B et al[24].

4.3 Convolution Layers and Convolution Operation

The convolution layer is the first layer of a CNN. A CNN may contain several convolution layers. The input for the first convolution layer is the pictures, which it then starts to process, Dr. T. V. Rajini Kanth et al [25].

Each convolution layer features the convolution procedure. Between an image portion and the filter, the convolution procedure is nothing more than an element wise multiply-sum operation, B. V. Kiranmayee et al [26].Now look at the illustration shown below figure 1.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

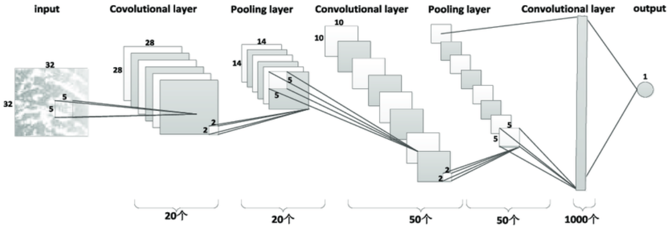

Figure : 1 CNN Model

Filter: This also goes by the names Kernel and Feature Detector. This matrix is quite small. One convolution layer may have several filters. The filters employed in a convolution layer are of the same size. A variety of filters are applied to the image to isolate various features.

Image Section: The image section's size difficult to match the size of the filter(s) we select. On the input image, we can reposition the filter(s) both vertically and horizontally to produce various image parts. The Stride we use determines the amount of image parts, Patela S. et al [28].

Calculation of convolution

An image portion and a single filter are convolutioned in the diagram above. You can multiply elements either row-wise or column-wise, followed by summing. K. Punithavathy, et al [30].

Row-wise

(0*0 + 3*1 + 0*1) + (2*0 + 0*1) + 1* 0) + (0*1 + 1*0 + 3*0) = 3

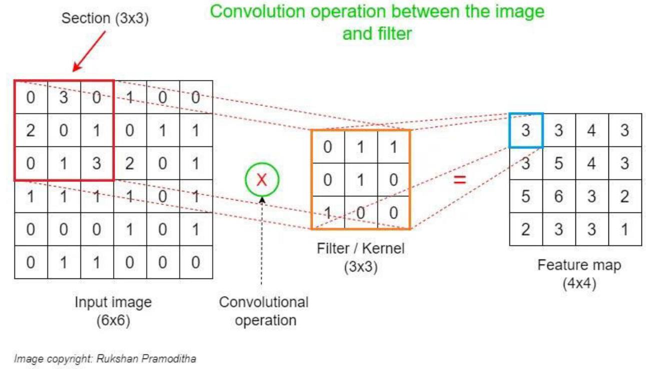

Figure : 2 Convolution operation

5. Results and Discussion

Today, a lot of people have lung disease because of a poor air quality measure. Lung cancer may develop as a result of several mild illnesses. This illness not only causes issues for one particular gender, but also for both genders. Therefore, it is crucial to exercise extra caution before it becomes dangerous. We compared various methods of lung cancer early detection in our review. Some people use CT scan images, while others use x-ray scans, Robert M. et al [34]

Using the CNN model, acquired the highest accuracy 89.78% as compared with other Convolution neural network models, and less accuracy was achived at ADABoost with 68%. CNN is being used to get data than other models like Random forest, Support Vector Machines and Extreme Gradient Boost model, as shown in figure 1.3. Usin g the CNN model, acquired the highest accuracy 89.78% as compared with other Convolution neural network models, and less accuracy was achived at ADABoost with 68%. CNN is being used to get data than other models like Random forest, Support Vector Machines

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

and Extreme Gradient Boost model, as shown in figure 1.1

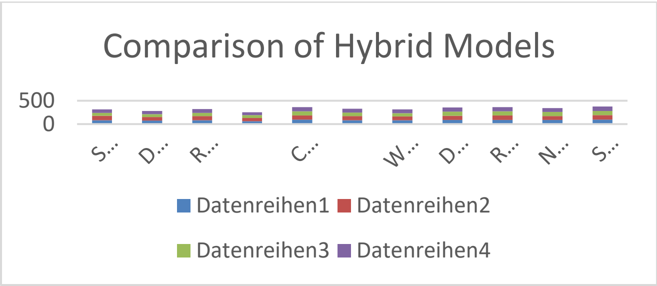

Figure : 4 Comparison with other Models

How well a measurement compares to a known or reference value. The accuracy of the results is important when deciding which algorithm to use going forward. Accuracy improves the research's findings' quality. In Table 1, SVM is shown to perform better than all other classification methods, including Logistic Regression, Gawade Prathamesh et al [36]. As a result, we can state that SVM has the highest accuracy rate of all other classification techniques for these particular datasets as shown in the figure 4.

Table I: Results of different deep learning lung cancer detection models during training

Models	**Training Accuracy**	**Training AUC**	**Training Recall**	**Training Loss**
CNN	99.85%	100%	99.80%	0.003
Inception V3	93.64%	97.50 %	96.50%	1.890
ResNet-50	99.45%	99.97 %	99.50%	0.043
Autoencoder	91.34%	95.23 %	99.48%	0.046
Xception	95.25%	97.67 %	95.89%	1.540

Table II : Validation of deep learning models for finding lung cancer

Model	**Validation Accuracy**	**Validation AUC**	**Validation Recall**	**Validation Loss**
CNN	92.65	97.05	92.50	0.367
Inception V3	83.45	87.36	84.56	0.548
ResNet-50	85.23	95.20	84.26	0.624
Autoencoder	80.34	92.56	82.17	7.248
X-pn	81.24 %	91.20%	83.50 %	8.420

Table III: Test results for various deep learning models for spotting lung cancer.

Model	**Testing Accuracy**	**Testing AUC**	**Testing Recall**	**Testing Loss**
CNN	**93%**	**98.35%**	**92.86 %**	**0.297**
Inception V3	**83.08%**	**89.24%**	**83.07 %**	**14.67%**
ResNet-50	**85.23%**	**96.68%**	**84.56 %**	**0.483**
Autoencoder	**88.54%**	**93.37%**	**84.79 %**	**0.873**
Xception	**83.20%**	**91.06%**	**83.08 %**	**7.830**

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

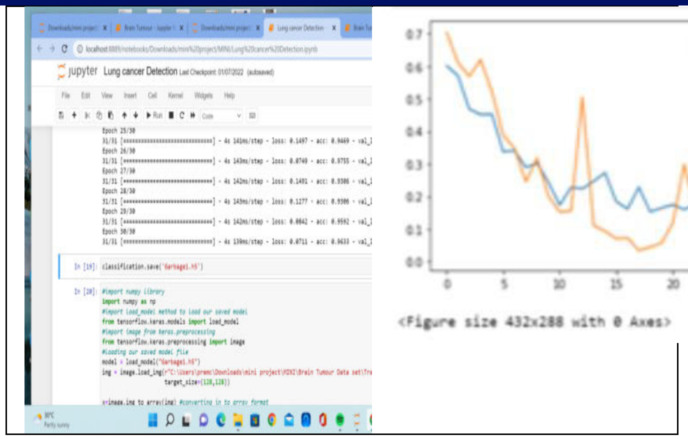

Figure: 5 a) output screen of Accuracy
b) Output screen of Training and Testing

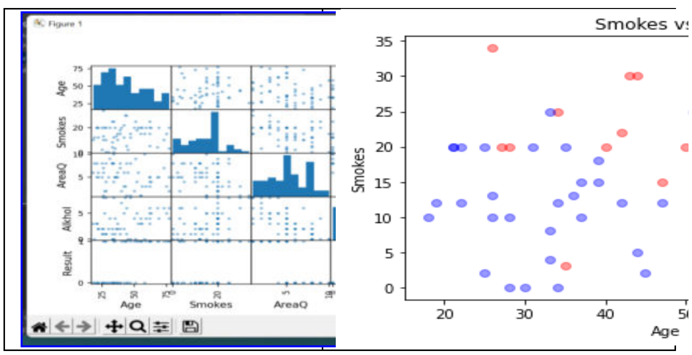

Figure:6a) Training of lung cancer
b) validation of lung cancer

Lung cancer training progress graph for a straightforward CNN method, As the epoch rose and the loss dropped, it can be seen from Figs. 5 and 6 that training and validation of the classifier's accuracy increased quickly, N. Pandey et al[37]. There is a slight discrepancy between the training and validation accuracy for hybrid CNN-SVM, Ada et al [38]. This indicates that the overfitting and underfitting issues are mitigated by the model Dhara et al[39].

F1 Score : 75.0
ACCURACY : 83.33333333333334
Using Decision Tree Algorithm
Classification accuracy on train 95.74468085106
Classification accuracy on test 75.0

6. Conclusion

This research addressed current developments in lung nodule segmentation, detection, and classification methods based on deep learning. CT image datasets are the most often used imaging datasets for training networks, and CNN is one of the most widely used deep learning algorithms for lung disease detection and classification. Results from experiments and clinical trials show that deep learning methods can be more effective than specialised radiologists. Segmentation, identification, and classification of lung nodules are anticipated to be significantly enhanced by deep learning. Radiologists can more accurately analyse pictures with this potent tool. Despite these drawbacks, deep learning is predicted to increase the demand for precise medical diagnosis and treatment given the present trend and the industry's rapid development.

References

[1] Mohanapriya, N., Kalaavathi, B., Kuamr, T.s.: Lung tumor classification and detection from ct scan images using deep convolutional

[2] Ju, W., Xiang, D., Zhang, B., Wang, L., Kopriva, I., Chen, X.: Random walk and graph cut for co-segmentation of lung tumor on pet-

[3] Luan, X. C., Li, C. H., & Liu, R, "Gaussian Noise Removal of Image on the Local Feature", IEEE International Symposium on Intelligent Information Technology Application, Vol. 3, pp. 867-871, 2008.

[4] Gajdhane, A. V., & Deshpande, L. M., "Detection of Lung Cancer Stages on CT scan Images by Using Various Image Processing Tech niques"., Journal of Computer Engineering, Vol. 16, No.5, pp. 28- 35.,2014

[5] Henschke.C.I, Mc Cauley.D.I, Yankelevitz.D.F, Naidich.D.P, McGuinness.G, Miettinen.O.S, Libby.D.M, Pasmantier.M.W, Koizumi.J, Altorki.N.K, Smith.P.J, "Early lung cancer action project : Overall design and findings from baseline screening", The Lancet, Vol. 354, No. 9173, pp.99-105., 1999.

[6] Haider, W., Sharif, M., & Raza, M., "Achieving accuracy in early stage tumor identification systems based on image segmentation and 3D structure analysis", Computer Engineering and Intelligent Systems, Vol. 2, No.6, pp.96-102, 2011.

[7] Zhao, Y., Zhang, Y.: Parallel deep learning algorithms with hybrid attention mechanism for image segmentation of lung tumors. IEEE Transactions on Industrial Informatics 17(4), 2880–2889 (2021) .

[8] Wang, X., Chen, H., Gan, C., Lin, H., Dou, Q., Tsougenis, E., Huang, Q., Cai, M., Heng, P.-A.: Weakly supervised deep learning for whole slide lung cancer image analysis. IEEE Transactions on Cybernetics 50(9), 3950–3962 (2020

[9] Jiang, J., Hu, Y.-C., Liu, C.-J., Halpenny, D., Hellmann, M.D., Deasy, J.O., Mageras, G., Veeraraghavan, H.: Multiple resolution residually connected feature streams for automatic lung tumor segmentation from ct images. IEEE Transac- tions on Medical Imaging 38(1), 134–144 (2019).[

[10] Hussein, S., Kandel, P., Bolan, C.W., Wallace, M.B., Bagci, U.: Lung and pan- creatic tumor characterization in the deep learning era : Novel supervised and unsupervised learning approaches. IEEE Transactions on Medical Imaging 38(8), 1777–1787 (2019) https://doi.org/10.1109/TMI.2019.2894349

[11] Henschke, C. I., Yankelevitz, D. F., Naidich, D. P., McCauley, D. I., McGuinness, G., Libby, D. M., & Miettinen, O. S, "CT screening for lung 125 cancer: Suspiciousness of nodules according to size on baseline scans", Radiology, Vol. 231, No.1, pp.164-168., 2004.

[12] A.R., Hemmati, H.R., Hosseinian, M.Z.: Automatic lung nodules detection in computed tomography images using nodule **filtering and neural net- works. In: 2014 22nd Iranian Conference on Electrical Engineering (ICEE), pp. 1883–1887 (2014).**

[13] Gierada, D. S., Pilgram, T. K., Ford, M., Fagerstrom, R. M., Church, T. R., Nath, H., & Strollo, D. C, "Lung cancer: Inter observer agreement on interpretation of pulmonary findings at low-dose CT screening", Radiology, Vol. 246, No. 1, pp. 265–272, 2008.

[14] Roy, K., Chaudhuri, S., Burman, M., Ganguly, A., Dutta, C., Banik, S., Banik, R.: A comparative study of lung cancer detection using supervised neural network, pp. 1–5 (2019).

[15] Khdair, H., Dasari, N.M.: Exploring machine learning techniques for coronary heart disease prediction. International Journal of Advanced Computer Science and Applications 12(5), 1227–1237 (2021) .

[16] Harikiran, J., Saichandana, B., & Divakar, B, "Impulse noise removal in digital images", International Journal of Computer Applica Harikiran, J., Saichandana, B., & Divakar, B, "Impulse noise removal in digital images", International Journal of Computer Applications , Vol. 10, No. 8, pp. 39-42, 2010.

[17] Hashemi, A., Pilevar, A. H., & Rafeh, R, "Mass detection in lung ct images using region growing segmentation and decision making based on fuzzy inference system and artificial neural network", International Journal of Image, Graphics & Signal Processing, Vol. 5, No.6, pp . 16- 24, 2013

[18] Rathod, S., Ragha, L.: Analysis of CT Scan Lung Cancer Images using Machine Learning Algorithms (2022).

[19] J Kuruvilla and K. Gunavathi, "Lung cancer classification using neural networks for CT images", Computer Methods and Programs in Biomedicine, vol. 113, no. 1, pp. 202–209, 2014.