International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 288-294, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE47
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

An Innovative Hybrid Deep Learning Technique Based on Neural Networks for Early Detection of Lung Cancer

[1]Moksud Alam Mallik, [2]Mohammed Dawood Khan, [3]Muazuddin Syed, [4]Mohd Shoeb Taj, [5]Nurul Fariza Zulkurnain

[1]Dean R&D and Associate Professor, Department of CSE (Data Science), Lords Institute of Engineering and Technology, Hyderabad, India
[2,3,4]UG Student, Department of CSE (Data Science), Lords Institute of Engineering and Technology, Hyderabad, India
[5]Associate Professor, Electrical and Computer Engineering, International Islamic University, Kuala Lumpur, Malaysia
E-mail: [1]malammallik@lords.ac.in, [2]dawoodkr02@gmail.com, [3]muazuddinofficial@gmail.com, [4]shoebtajofficial@gmail.com, [5]nurulfariza@iium.edu.my

***Abstract* - With a high death rate among affected individuals, lung cancer is a deadly illness. Patients can be saved by receiving an early diagnosis and correctly determining the stage of lung cancer. Lung cancer can be detected using a variety of image processing; biomarker based, and machine automation techniques, although early detection and accuracy are difficult for medical professionals to achieve. This work uses the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) to extract the Computed Tomography (CT) scan images. Conventional techniques use manual CT scans to determine if a patient has lung cancer. This study suggests a brand-new technique called Cancer Cell Detection utilizing Hybrid Neural Network (CCDC-HNN) for an early and precise diagnosis. Deep neural networks are used to extract the features from the CT scan images. To save the patient from this deadly illness, early detection of malignant cells depends heavily on feature extraction accuracy. An advanced 3Dconvolution neural network (3D-CNN) is also used in this study to increase diagnosis accuracy. Additionally, the proposed method makes it possible to distinguish between benign and malignant tumors. The outcomes validate the feasibility of the suggested hybrid deep learning (DL) approach for early lung cancer detection when assessed using conventional statistical methods.**

Keywords: Lung Cancer, Computer-Aided Diagnostic, Machine Learning, Deep Learning, Neural network.

I. INTRODUCTION

Lung Cancer is caused by the growth and dissemination of abnormal cells throughout the body. The disease's precise cause is yet uncertain. Bad habits like drinking or smoking can cause it, or it can be genetic and related to the immune system. These days, the rate of technical improvement and the incidence of cancer worldwide are both rising. Our study aims to further understand the high level and discriminative features displayed by cancer cells utilizing CNNs for precise classification of lung cancer subtypes. It is driven by the success of employing CNN algorithms in digital pathology image processing, as previously reported. Doctors can only diagnose lung cancer after closely examining CT scans, which are time-consuming and not always reliable. To create picture that was as precise, useful, and efficient as feasible, contemporary optimization techniques and image processing procedures were required. Physicians will be able to understand more about the interior anatomy and identify lung nodules early with the aid of the proposed technology. The basic kinds of cancer can be treated early. Cancerous cells in the body proliferate more quickly than healthy ones, which leads to lumps becoming into nodules or tumors. An Advanced 3D-Convolution Neural Network (3DCNN) with Recurrent Neural Network (RNN) models based on Deep Learning (DL) have been offered as DL approaches in this research with the primary goal of analyzing and detecting cancer cells [1]. The number of lung nodules in the pulmonary region can be more accurately counted with the use of the recommended technique. Additionally, the proposed method makes it possible to distinguish between benign and malignant tumors. Compared to the traditional 3D-CNN and RNN, the proposed advanced 3D with the RNN model has yielded more aesthetically pleasing outcomes in the early identification and treatment of lung illness. The patient's lung cancer CT scan image is shown in Fig 1.

Fig 1: The lung cancer patient's CT scan image

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 288-294, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE47
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

Through the blood arteries, lung cancer cells may move to other parts of the body. This type of cancer is known as "metastasis" or "additional lung cancer." Studies have shown that lung cancer is among the deadliest types of cancer [2]. Every 30 seconds, someone with lung cancer passes away somewhere on Earth. After being diagnosed with lung cancer, the patient has a 20% chance of surviving [3]. Women who smoke are more likely than men to develop lung cancer, according to several studies [4]. Only women had a significantly elevated risk of lung cancer, despite not smoking. CT scan images may show round or oval lung tissue masses called lung nodules, which have a diameter of less than 30 mm. They represent various forms, densities, locations, and environments. Lung nodes typically have a width of more than 3 mm; nodes with a diameter of less than 3 mm are called micro-nodes. Pulmonary and vascular walls may result in false positives because they are non-nodules that mimic nodules. Lung nodules can be divided into three groups based on their concentration: partially solid nodules, solid lesions, and ground glassware [5,6]. Nodes may have solitary, juxta-pleural, or juxta-vascular terminal locations.

When there is pulmonary oedema, nodules are typically hard to spot. Consequently, it becomes considerably more challenging to accurately identify and diagnose lung nodules. Lung malignancies come in two varieties: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Although NSCLC accounts for 75% of lung cancer occurrences, not all clusters are tumor cells; benign tumors have a diameter of less than 3 mm, while malignant tumors have a diameter of more than 3 mm [5,6].

Medical science uses computer-aided diagnostic (CAD) tools to identify diseases early on [7]. The CAD techniques were once employed by doctors and radiologists to identify a range of illnesses [8]. It is usually difficult to identify several diseases in their early stages, such as prostate, skin, and lung cancer. As a result, early cancer detection is facilitated by the CAD system. Precision, velocity, and automatic level are used to gauge the system's performance [9]. CT scan images are more helpful in identifying malignant lung nodules than x-ray images [9]. X-rays are not used to find cancers, but to pinpoint particular body components. Nodules with bones connected are found when the CT scan images are rotated to encompass a 360-degree area. Radiologists can make more accurate diagnoses by identifying minor lung abnormalities [10]. If you need accurate information from the scanned image, only manual CT identification is feasible.

However, if lung cancer is detected in an early, benign stage, a Machine Learning (ML) method is needed. There are two methods for locating nodules in the lungs. To find small nodules in the chest from the pulmonary vasculature, the lung endothelium is first extracted from the CT scan image [11]. It is crucial to give the health sector more attention and, in particular, to create novel methods for early cancer detection in order to lower the death rate from cancer. Segmentation accuracy is an important consideration when using lung cancer segmentation data. When segmenting pictures manually, radiologists may provide erroneous results due to inter-observer variability and inconsistent procedures. To solve this problem, an automated method for segmenting CT scan images of lung cancer is highly relevant.

In order to detect lung cancer early, this study presents an innovative hybrid deep learning method that combines recurrent neural networks (RNNs) and convolutional neural networks (CNNs). This hybrid model seeks to improve the precision and dependability of lung cancer diagnosis from medical imaging data, such as CT scans and X-ray pictures, by utilizing the advantages of both CNNs for feature extraction and RNNs for sequential pattern recognition. The suggested technique is intended to minimize diagnostic errors, automatically identify cancerous nodules in their early stages, and offer prompt care.

This paper is arranged as: Section 1 providing a short introduction about this paper, Section 2 gives the literature survey of deep learning methods, Section 3 brief about the methodology, in Section 4 elaborates the experimental results and the conclusion and future enhancement of this paper is made in Section 5.

II. LITERATURE REVIEW

Here we gone through various research paper related to early detection of lung cancer.

Title	Authors	Key Findings
Early Stroke Prediction Methods for Prevention of Strokes	M. Kaur, S.R. Sakhare, K. Wanjale, et al. 2022	Reviews early-stage stroke prediction methods, highlighting the role of machine learning in improving risk factor identification and patient outcomes.

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 288-294, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE47
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

Cancer Statistics	R.L. Siegel, et al. 2022	Provides an overview of cancer incidence, mortality, and survival rates in the U.S., emphasizing early detection and prevention strategies.
Lung-RADS: Pushing the Limits	M.D. Martin, J.P. Kanne, L.S. Broderick, et al. 2017	Discusses the Lung-RADS system for lung cancer screening, focusing on risk stratification and clinical decision-making while addressing challenges and future directions.
3-D Segmentation Algorithm of Small Lung Nodules in Spiral CT Images	S. Diciotti, G. Picozzi, M. Falchini, et al. 2008	Proposes a 3-D segmentation algorithm for identifying small lung nodules in spiral CT images, demonstrating improved accuracy over traditional methods.
Towards Automatic Pulmonary Nodule Management in Lung Cancer Screening	R. Yamashita, M. Nishio, R.K.G. Do, et al. 2018	Reviews CNNs and their applications in radiology, highlighting their effectiveness in medical image analysis for diagnosing various conditions.
Using Deep Learning for Classification of Lung Nodules on CT Images	F. Ciompi, et al. 2017	Explores a deep learning framework for detecting and classifying pulmonary nodules, improving efficiency and accuracy in clinical workflows.
A Novel Deep Learning-Based Framework for Breast Cancer Detection	S. Khan, N. Islam, Z.I. Jan, et al. 2019	Proposes a transfer learning-based framework for detecting and classifying breast cancer, emphasizing the effectiveness of pre-trained models in improving diagnostic performance.
Breast Cancer Detection and Classification Using Artificial Neural Networks	Y.A. Hamad, K. Simonov, M.B. Naeem 2018	Presents the use of ANNs for breast cancer detection and classification, showcasing their potential to enhance diagnostic accuracy and support clinical decisions.
Lung Cancer Detection: A Deep Learning Approach	S. Bhatia, N. Mittal, S.K. Sonbhadra, et al. 2019	Introduces a deep learning based method for lung cancer detection, demonstrating improved accuracy and efficiency in medical image analysis.

In previous studies, various automation algorithms and biomarker based techniques have been implemented to detect cancerous cells from lung CT-scan images. However, these methods often fall short in terms of detection accuracy, leading to challenges in reliable diagnosis. Traditional algorithms, including those based on conventional image processing and machine learning approaches, have struggled to effectively identify subtle changes in lung tissues over time. As a result, there has been a pressing need for more advanced methodologies that can enhance detection rates and improve the overall accuracy of cancer diagnosis.

Disadvantages

1. It takes more time and 2. Less accuracy

To address the limitations of existing techniques, the proposed system employs a Hybrid Neural Network that combines Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) algorithms for cancer cell detection. Initially, the UNET algorithm is utilized for accurate segmentation of cancer cells, enabling the identification of tumor boundaries. Following this, the CCDC-HNN (Cancer Cell Detection using Hybrid Neural Network) algorithm classifies the segmented cells based on their features. This approach not only enhances the accuracy of cancer detection but also provides physicians with critical information regarding the size of the detected cancer cells, facilitating a better understanding of the cancer stage. The system is trained using the LIDC-IDRI dataset, which is accessible via a KAGGLE URL, ensuring a robust training process with diverse lung CT scan images.

Advantages

1. It takes less time and 2. More prediction

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 288-294, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE47
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

III. METHODOLOGY

To implement this technique, we have designed following modules:

1) Upload LIDC-IDRI Dataset: This component facilitates the incorporation of the publicly accessible LIDC-IDRI (Lung Image Database Consortium and Image Database Resource Initiative) dataset into the application. It performs checks to ensure the completeness and consistency of metadata, including patient identifiers, image resolution, and scanning parameters. The verification process plays a vital role in ensuring the effective integration of the dataset, thereby laying a robust foundation for the ensuing preprocessing and analytical stages.

Fig 2: The User Interface (UI) of the application

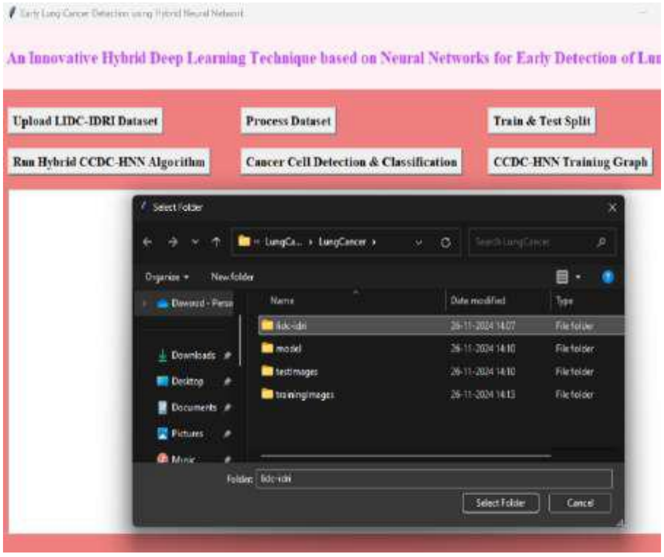

Fig 3: Uploading LIDC-IDRI Dataset

In above screen click on 'Upload LIDC-IDRI Dataset' button to upload dataset and get below page

Fig 4: Dataset uploaded with 19255 images

In above screen total 19255 images loaded from dataset and now click on 'Process Images' to clean and normalize images and then get below output.

2) Process Dataset: This element is tasked with the handling of uploaded images by methodically examining each file, adjusting their sizes to a consistent dimension, and standardizing pixel values to maintain uniformity. These procedures are essential for standardizing the data, thereby rendering it appropriate for deep learning applications. Additional validation steps are implemented to confirm that the processed images are intact and prepared for the following stages.

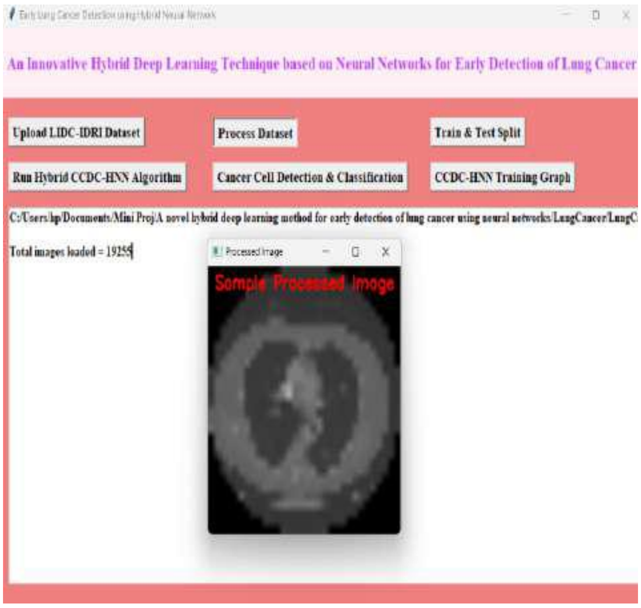

Fig 5: Processing the dataset

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 288-294, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE47
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

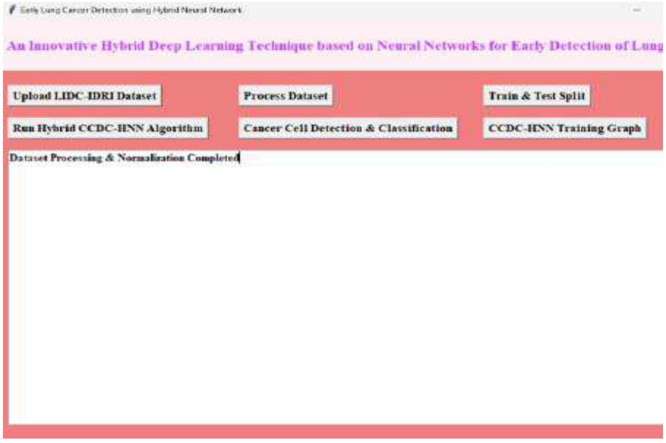

Fig 6: Data processing and Normalization

3) Train & Test Split: This component is tasked with partitioning the dataset into training and testing subsets, designating 80% of the images for training purposes and 20% for testing. This division is critical for enabling the model to learn from a varied dataset while reserving unseen samples for an unbiased evaluation. Appropriate stratification techniques are employed to ensure a balanced distribution of classes within both subsets.

Fig 7: Splitting the dataset into training and testing

In above screen can see train and test size of dataset out of 14599 total images and now click on 'Run Hybrid CCDC-HNN Algorithm' button to train model on training images and tested on testing images.

4) Run Hybrid CCDC-HNN Algorithm: In this section, the Hybrid CCDC-HNN algorithm is trained using 80% of the dataset designated for training purposes. This algorithm analyzes the images to construct a predictive model within its hybrid deep learning architecture. After the training is completed, the model is tested on the remaining 20% of the images to assess its performance and ascertain the accuracy of its predictions.

Fig 8: Running the Hybrid CCDC-HNN Algorithm

5) Cancer Cell Detection & Classification: This module detects and classifies cancer cells in test images. Upon uploading a test image, the UNET model segments the image to identify potential cancerous regions. If a cancer cell is detected, the Hybrid CCDC-HNN algorithm classifies it as either Benign or Malignant, offering actionable insights for early diagnosis.

6) CCDC-HNN Training Graph: This module visualizes the performance of the Hybrid CCDCHNN algorithm during training. It generates graphs showing training accuracy and loss for each epoch, providing insights into the model's learning progress and optimization. These graphs help evaluate the model's stability and detect potential overfitting or underfitting.

Fig 9: Hybrid CCDC-HNN Graph

7) In above screen propose algorithm got 91% accuracy and can see other metrics like precision, recall and etc. in confusion matrix graph x-axis represents Predicted Labels and y-axis represents True labels and then yellow and light blue

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 288-294, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE47
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

color boxes in diagnol represents correct prediction count and all blue boxes in diagnol represents incorrect prediction count which are very few. In ROC graph x-axis represents False Positive Rate and y-axis represents True Positive Rate and if blue lines comes below orange line then all predictions are incorrect or false and if goes above orange line then all predictions are correct or true. Now close above graph and then click on 'Cancer Cell Detection & Classification' button to upload test image and get below output

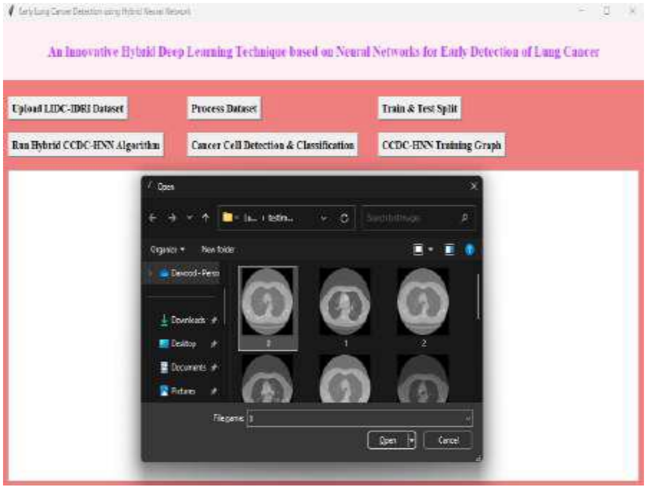

Fig 10: Testing the model on sample image '0.png'

8) In above screen selecting and uploading '0.png' image and then click on 'Open' button to get below output

The above figure shows the model construction, where the training tuples is stated as input to the classification algorithm, which builds the classifier.

IV. RESULTS & DISCUSSION

Results demonstrate the potential for hybrid deep learning techniques to support healthcare professionals in making timely and accurate diagnoses, ultimately contributing to better patient outcomes.

Fig 11: Identifying the malignant cancer cell

In above screen first image is the original image and second image is UNET segmented image of cancer cell detection and then propose 'CCDC-HNN' algorithm classify cell as Malignant. Similarly, you can upload and test other images.

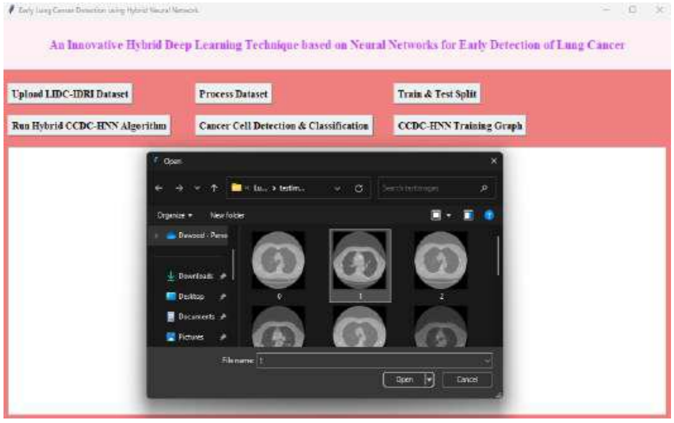

Fig 12: Testing the model on sample image '1.png'

In above screen selecting and uploading another image and below is the output.

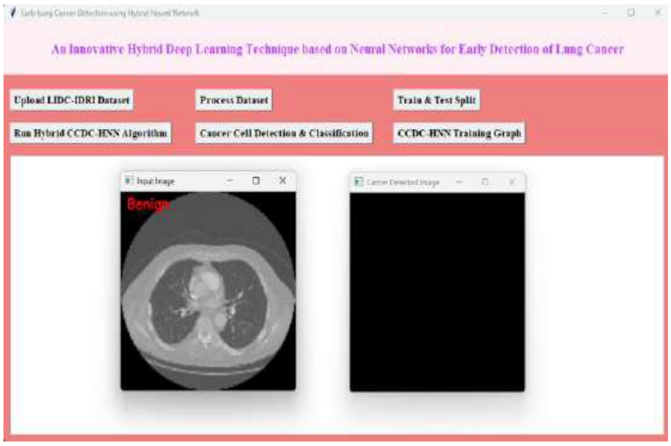

Fig 13: Classifying the image as 'Benign'

In above screen image classify as Benign as there is no cancer cell detected.

V. CONCLUSION

Lung cancer is a very fatal disease which needs an attention from the medical practitioners and researchers to find innovative solutions for early diagnosis of this disease to save the lives of the patients. Early detection of lung nodules aids in diagnosing illness and averts disease-related mortality. With the assistance of medical professionals, computer diagnosing methods have been devised to shorten diagnostic times, increase effectiveness and accuracy, and decrease disease-related mortality. The approach in this research suggests a sophisticated 3DCNN with an RNN algorithm for classifying

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 288-294, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE47
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

cancerous lung nodules. The system makes use of the LUNA 16 database. The proposed novel technique in this article is able to process the CT scan images with high accuracy by devising an innovative method for feature extraction from the CT scan images and then by employing the hybrid DL method for classification of images. The empirical results obtained from the proposed mechanism demonstrate that the proposed improved model provides single 3D-CNN and RNN classifications with 90% selectivity (SP), 87% sensitivities (SE), and 95% accuracy.

REFERENCES

[1] M. Kaur, S.R. Sakhare, K. Wanjale, et al., Early stroke prediction methods for prevention of strokes, Behav. Neuro. (2022) 7725597, 1–9.

[2] R.L. Siegel, et al., Cancer statistics, Cancer J. Clin. 72 (2022) 7–33.

[3] M.D. Martin, J.P. Kanne, L.S. Broderick, et al., Lung-RADS: Pushing the limits, RadioGraphics 37 (7) (2017) 1975–1993.

[4] S. Diciotti, G. Picozzi, M. Falchini, et al., 3-D segmentation algorithm of small lung nodules in spiral CT images, IEEE Trans. Inf. Technol. Biomed. 12 (1) (2008) 7–19.

[5] R. Yamashita, M. Nishio, R.K.G. Do, et al., Convolutional neural networks: An overview and application in radiology, Insights Imaging 9 (2018) 611–629.

[6] F. Ciompi, et al., Towards automatic pulmonary nodule management in lung cancer screening with deep learning, Sci. Rep. 7 (1) (2017) 1–11.

[7] Q.Z. Song, L. Zhao, X.K. Luo, et al., Using deep learning for classification of lung nodules on computed tomography images, J. Healthc. Eng. (2017) 8314740, 2017.

[8] S. Khan, N. Islam, Z.I. Jan, et al., A novel deep learning based framework for detecting and classifying breast cancer using transfer learning, Pattern Recognit. Lett. 125 (2019) 1–6.

[9] Y.A. Hamad, K. Simonov, M.B. Naeem, Breast cancer detection and classification using artificial neural networks, in: Proc. 1st Annual Int. Conf. on Info. and Sc. (AiCIS), 2018, pp. 51–57.

[10] S. Bhatia, N. Mittal, S.K. Sonbhadra, et al., Lung cancer detection: a deep learning approach, in: Soft Comp. for Prob. Solving, 2019, pp. 699–705.

[11] D. Palani, K. Venkatalakshmi, An IoT based predictive modelling for predicting lung cancer using fuzzy cluster based segmentation and classification, J. Med. Syst. 43 (2) (2019) 21.

[12] A.Masood, et al., Automated decision support system for lung cancer detection and classification via enhanced RFCN with multilayer fusion RPN, IEEE Trans. Ind. Inform. 16 (12) (2020) 7791–7801.

[13] V. Fredriksen, et al., Teacher-student approach for lung tumor segmentation from mixed-supervised datasets, PLoS One 17 (1) (2022) e0261917.

[14] T. Saba, et al., Cloud-based decision support system for detecting and classifying malignant cells in breast cancer using breast cytology images, Micro. Res. Tech. 82 (6) (2019) 775–785.

[15] J. Talukdar, et al., A survey on lung cancer detection in CT scans images using image processing techniques, Int. J. Curr. Trends Sci. Technol. 8 (3) (2018) 20136–20140.

[16] R. Krithiga, P. Geetha, Deep learning-based breast cancer detection and classification using fuzzy merging techniques, Mach. Vis. Appl. 31 (7) (2020) 1–18.

[17] N. Shrivastava, J. Bharti, et al., Breast tumour detection and classification based on density, Multimedia Tools Appl. 79 (35) (2020) 26467–26487.

[18] R. Suresh, A.N. Rao, B.E. Reddy, et al., Detection and classification of normal and abnormal patterns in mammograms using a deep neural network, Concurr. Comput.: Pract. Exper. 31 (14) (2019) e5293.

[19] AT. Saba, et al., Lung nodule detection based on ensemble of hand crafted and deep features, J. Med. Syst. 43 (12) (2019) 332.

Citation of this Article:

Moksud Alam Mallik, Mohammed Dawood Khan, Muazuddin Syed, Mohd Shoeb Taj, & Nurul Fariza Zulkurnain. (2025). An Innovative Hybrid Deep Learning Technique Based on Neural Networks for Early Detection of Lung Cancer. In proceeding of International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25), published by *IRJIET*, Volume 9, Special Issue of INSPIRE'25, pp 288-294. Article DOI https://doi.org/10.47001/IRJIET/2025.INSPIRE47

****