International Journal of Computer Trends and Technology
ISSN: 2231 – 2803 / https://doi.org/10.14445/22312803/IJCTT-V71I4P111

Volume 71 Issue 4, 91-96, April 2023

Original Article

Lung Cancer Detection using CT Scan Images in CNN Algorithm

Velagapudi Sreenivas[1], Anupoojitha Gadila[2], Triveni Attunuru[2], Snowja Chodavarapu[2], Sathvika kesamaneni[2]

[1]*Department of Computer Science and Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, A.P, India*

[2] *B. Tech Students, Department of Computer Science and Engineering, Dhanekula, Institute ofEngineering and Technology, Vijayawada, A.P, India*

Received: 03 March 2023 Revised: 06 April 2023 Accepted: 16 April 2023 Published: 30 April 2023

Abstract *- A terrible illness that claims lives quickly all around the world is lung cancer.Because deaths from lung cancer are happening more frequently, the second most frequentform of cancer is lung cancer, according to studies. An automated method is required to predictthe illness and save a person's life. The accuracy and quality of the images are the main factorsin this study. The enhancement step is when low-level—employing pre-processing methods that use the Gabor filter inside the Gaussian rules to assess and enhance the image quality. Thefeatures of normal and abnormal photos are extracted using the segmentation and enhancement technique. A comparison of normalcy is performed based on shared traits. In this study, the key criteria to check for in order to accurately compare photographs are pixelspercentage and mask labelling. Because We chose to use CNN algorithms to analyse CT scanpictures in order because the lung tissue is where this specific type of cancer frequently first appears to identify lung cancer.*

Keywords *- CT scan, lung cancer, SVM, K-Means, KNN, CNN Algorithms.*

1. Introduction

Due to the unrestrained growth of breathing problems in the inhale and exhale regions in the chest, lung cancer frequently affects males and females. The main causes of lung cancer, according to the World Health Organization, are cigarette smoking and passive smoking. Deaths from lung cancer are rising daily, particularly among young people. In this manner because lymphatic fluid typically leaves lymph nodes in the lung on its way to the middle of the chest. The phenomenon known as Metastasis occurs when a cancer cell leaves its original location and moves via the bloodstream to a bodily part, such as a lymph node. The first step entails collecting various CT scans the Home database for the IMBA (International Master of Business Administration). To attain the highest level of clarity, the second step makes use of multiple image enhancement algorithms. The enhanced segmented images' general features are extracted in the fourth stage, which also offers a clue as to whether the images are normal or pathological. Algorithms for picture segmentation, which are helpful in earlier phases of image processing, are really used in the third stage. According to the stage at which cancer cells are discovered in the lungs, the world's worst and most prevalent kind of cancer is lung cancer. Because of this, early disease detection is crucial for stopping the disease's primary phases of progression and restricting its ability to spread throughout the middle of the chest. Infections are usually disseminated by lung cancer.

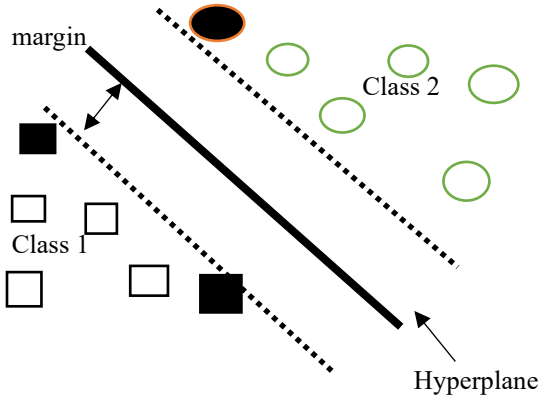

Fig. 1 Architecture

2. Literature Survey

This is their review of related literature for the suggested work. To identify lung cancer, a lot of researchers have developed and implemented various image processing and machine learning algorithms. A model to discriminate between nodulesand the normal anatomy of the lung was put out by Aggarwal, Furqan, and Kalra [1]. The approach takes cues from grey scales, statistics, and geometry. The classifier with the most effective segmentation and thresholding is LDA. The method has an accuracy of 84%, a specificity of 53.33%, and a sensitivity of 97.14%. Although the technique does locate the cancer nodule, its precision is still insufficient. Simple segmentationtechniques were utilised instead of machine learning approaches for classification. Because of this, increasing the likelihood of an enhancement by incorporating any of its phases in our newmodel is insignificant.

By using a convolution neural network as a classifier in his CAD system, Jin, Zhang, and Jin [2] were able to recognise lung cancer. The approach has an accuracy, sensitivity, and specificity of 84.6%, 82.5%, and 86.7%, respectively. This model has the benefit of applying a circular filter to isolate the ROI, reducing the cost of education and reward. Although implementation costs have decreased, accuracy has not improved. Academics have investigated deep learning, convolutional neural networks (Conv Net/CNN), and other machine learning techniques in-depth for diagnosing lung cancer. Srinivas Aru Konda [3] states that highly developed deep neural networks. Due to the disease's high survival rate, the system can detect lung cancer in its early stages.

Currently, Ignatius and Joseph [4] represent the best option. The image is enhanced using the Gabor filter during image pre-processing, and the cancer nodule is located using the marker-controlled watershed approach. This technique also allows extracting the cancer nodules' area, perimeter, and eccentricity. It emphasises the suggested model's 90.1% accuracy rate and contrasts it with other models that have been offered.

In order to isolate the juxta pleural lung nodules, CNN was used by Huo, Y. M., Liang, Z. R., Li, L. H., and Tan, J. X. [5]. They have provided a comparative study about the results of reducing false positives in systems for detecting lung cancer based on deep learning. Creating a technique to identify lung cancer using lung CT scans Jennifer D. Cruz et al. (2015) [6] utilising neural networks and genetic algorithms.

Initially, Pre-processing was applied to the image to improve the picture's quality. The augmented image was then subjected to a feature extraction and selection phase utilising a genetic algorithm. The text image was then classified as malignant or not using the Back Propagation Neural Networks approach.

A method to categorise lung cancers as benign or malignant was presented by S Shashikala et al. in 2018 [7]. The CT scan picture used as input is pre-processed using median filter techniques. Then, the Network was trained using the backpropagation technique to recognise lung cancers in CT scans.

To train the model, lung pictures with varying forms and sizes of cancerous tissues were fed, and the CNN-based method was able to accurately and 96% detect the presence or absence of malignant cells.

As a method for lung cancer feature extraction and diagnosis, deep residual networks [8] were put out by G. Huang, Z. Liu, and L. Van Der Matten. U-Net and Res-Net models are used to extract features, which are then input into various classifiers. Also, the individual forecasts, along with XG-Boost and Random Forest, a CT scan's likelihood of malignancy can be predicted. For LIDC-IRDI datasets, the study work's accuracy was 84%. WA "denoising first" two-path CNN was created by J. Sori et al. [9] and evaluated using a Kaggle dataset 2017 Data Science Bowl competition. This method combines the processes of denoising and detection. A comprehensive strategy that improved lung cancer detection.

A deep CNN was recommended by Golan et al. [10] for the detection of lung nodules. A sensitivity (71.2%) was purchased, and a true positive rate. The most common type of cancer, according to Sunyi Zheng Guo[11] specialists in lung cancer. Lung cancer cases increased by 57,795 in 2012, and by 2020, it is anticipated that 67,000 new cases will be diagnosed annually.

3. Implementation

An effective text summarisation technique is one that captures the most important points and conveys them succinctly and accurately. Techniques like page rank algorithms and other natural language processing algorithms are used to summarise the text.

3.1. Images for Database

The 40 digital pictures (samples) used in the research presented here are Twenty images of small-cell lung cancer, for each of which twenty pictures of non-small-cell lung cancer are displayed. The image is 200 by 200 pixels. The images were obtained from private organisations and located online in the public database of the Japanese Society of Radiation Technologies (JSRT). JPEG images have a resolution of 8 bits per plane and are preserved digitally. The original 200 X 200 X 256JPEG format is used to save each image.

3.2. Image Enhancement

Image improvement is the process of making an image sharper or smoother. It enhances image quality and gets rid of the noise. For handling the processing of digital pictures, it offers the optimum input. Picture enhancement is one of the techniques used in image pre-processing. The purpose of image enhancement is to alter the image to make it better suited for processing or analysis (for instance, by sharpening it or enhancing the contrast). Two categories of image-improving techniques exist:

3.3. Gabor Filter

Dennis Gabor is the one who invented the Gabor filter, which we employed for CT imaging. A linear filter called a Gabor filter—named after Dennis Gabor—is employed for texture analysis when processing images, which essentially analyses whether there are any particular frequencies. Certain directions in a small area surround the point or portion of the images.

In a 2D Gabor filter, a sinusoidal plane wave modulates a Gaussian kernel. The extraordinary placement property of the Gabor function in the partial and frequency domains makes it a vital tool in computer visibility and image processing, especially for texture study.

3.4. Thresholding Strategy

Thresholding is one of the best methods for segmenting images. An augmented image derived from the threshold has advantages over a grayscale image, which typically has 256 levels, such as requiring less storage space, processing data quickly, and facilitating easier manipulation. Black and white, the only two hues of grey that signifylevels 0 and 1, are combined to combine this pair of grayscale hues into a binary image or a black-and-white image.

This binary image is then used as input for a thresholding approach. Due to the limitednumber of levels, the threshold value will vary between 0 and 1. When attained, the image will be divided into segments based on it.

Features are very important in the realm of image processing. Many image processing techniques are applied to the sampled image, including binarisation, thresholding, normalisation, masking approach, etc., before the features are obtained. Finally, features that can be used to categorise and identify photos are collected using feature extraction techniques.

The sampled image is put through a numberof processes before being given the attributes, including. Ultimately, featuresthat can be used to classify and identify photographs are obtained using feature extraction techniques.

4. Algorithms

In this, we used various algorithms to compare them one after another to identify which algorithm has better accuracy.

4.1. Feature Extraction

The importance of features in the realm of image processing. Many images processing methods, including masking, thresholding, binarisation, and normalisation

4.2. K-Means Clustering

K-Means the un-labelled dataset is grouped into various clusters by the process of clustering, which is an Unsupervised Learning technique. The amount K, which must be formed during the operation, is a number of pre-defined clusters. It offers an easy technique to identify the groupings of the data in the un-labelled clusters, allowing us to partition the data into distinct groups without the need for any training, using the dataset alone. This centroid-based approach seeks to minimise the total separation between the data points and the clusters they are a part of.

4.3. K-Nearest Neighbor

The k-nearest neighbor method is a pattern recognition technique that may be used for both classification and regression. The abbreviation k-NN is widely used to refer to the positive integer that makes up the first part of the phrase "k-nearest neighbor.". k closest training examples within a space will be the input for classification or regression.

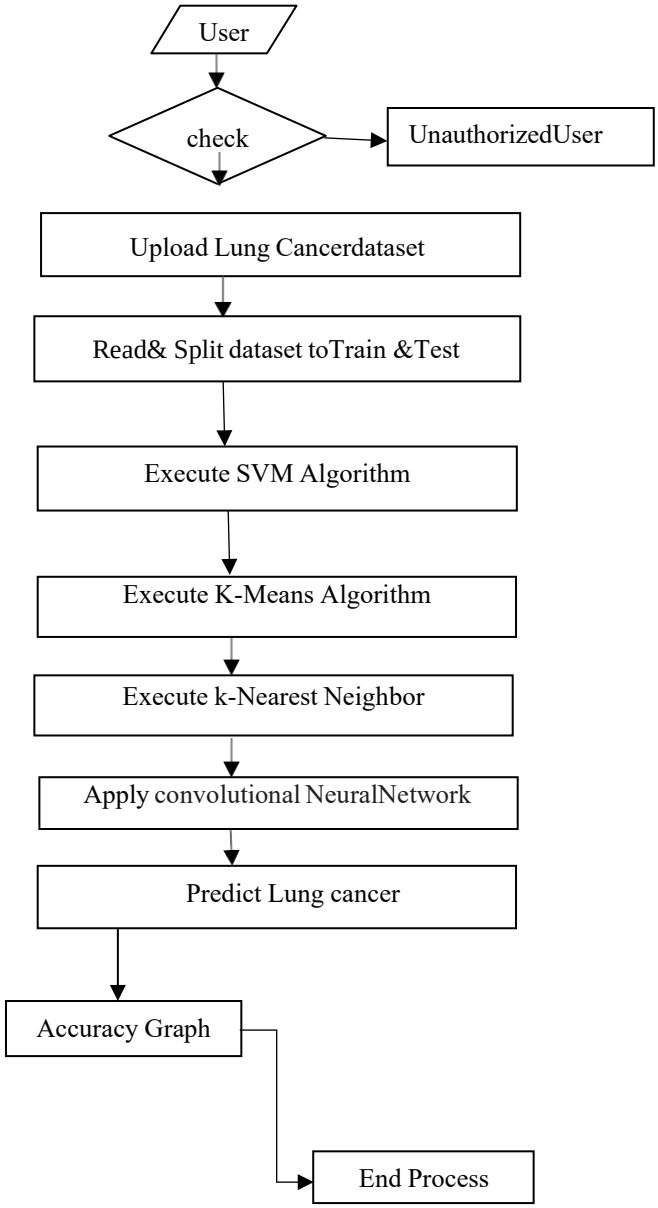

Fig. 2 Flow chart

4.4. CNN Algorithm

A CNN is a specific kind of network design for learning algorithms used in tasks like image identification and pixel data processing. The best network architecture for recognising and identifying objects is one that uses CNNs., despite the fact that many different types of neural networks are employed in learning. CNN is a type of learning model that can comprehend data having a grid pattern, such as images.

5. Proposed System

We proposed implementing a method in convolutional neural networks to identify lung cancer (CNN). By contrasting the SVM with k-means, k-nn., and CNN algorithms using the CT scan pictures, it is possible to identify algorithms with higher diagnostic accuracy. Compared to MRI and X-ray, CT pictures have less noise disruption and are therefore used. Median filtering is applied to these CT scans to enhance the image quality.

In order to train the external images, is used to detect the disease by comparing the images to those from a CT scan. So, we have used different lung segmentation and nodules segmentation methods. Using image processing, feature extraction, and classification techniques, we have been able to determine whether a patient has cancer or not.

6. Experimental Results

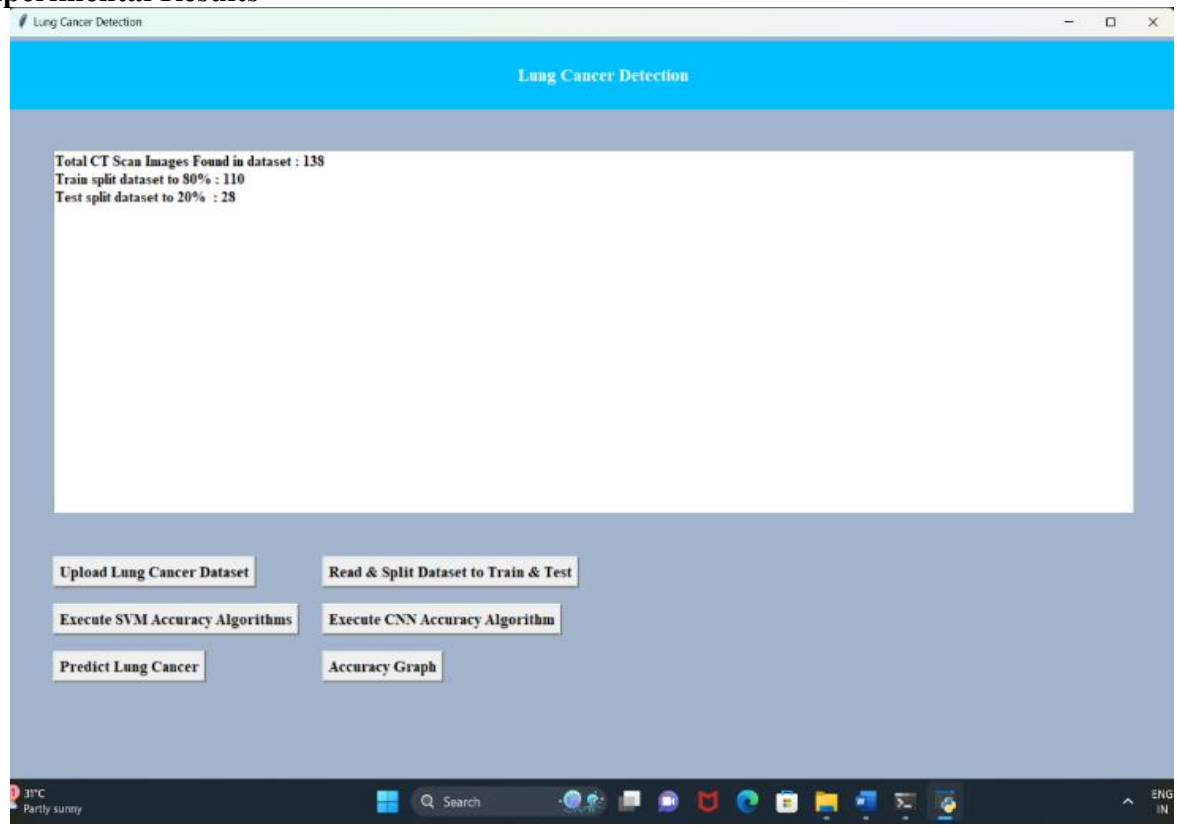

Fig. 3 Read and Split Dataset

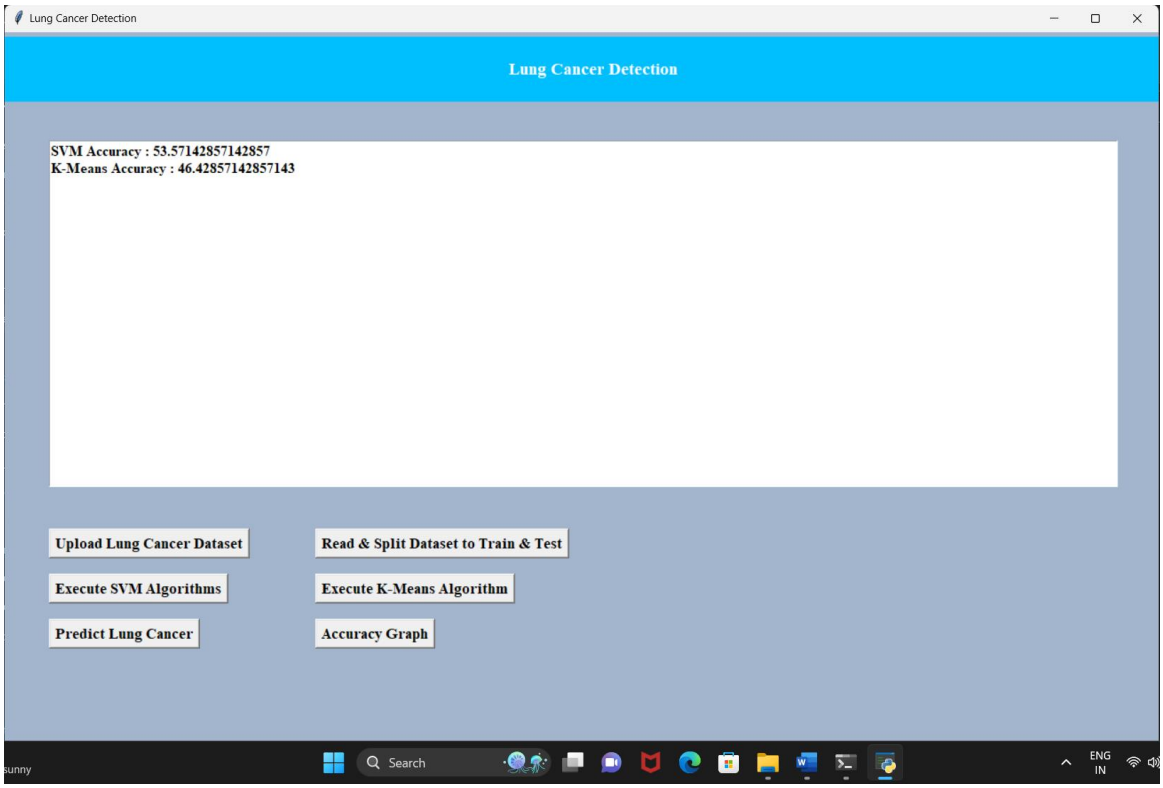

Fig. 4 SVM with k-Means Algorithm Accuracy

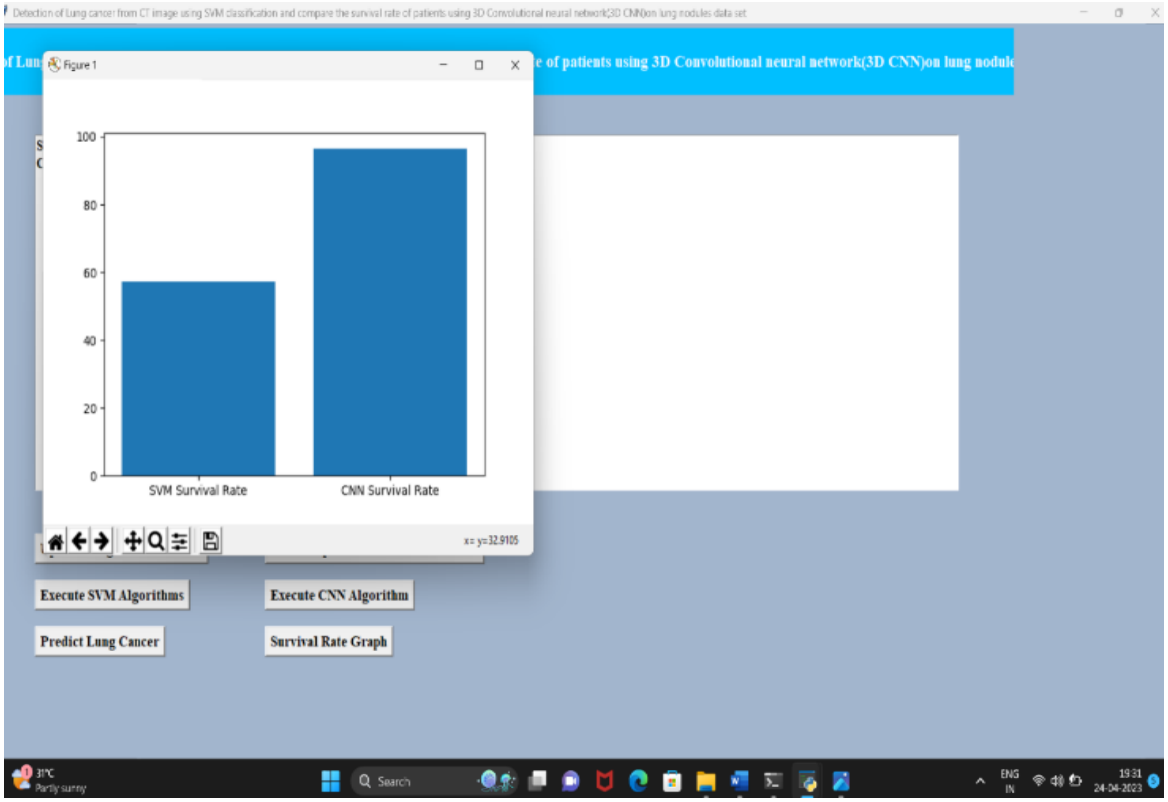

Fig. 6 Accuracy Graph

7. Conclusion

These findings demonstrate that the suggested method outperforms other methods described. In terms of precision, the literature well as image quality, whichis a key aspect of this study. There were low-preprocessing methods based on the Gabor filter accepted for both image quality and image enhancement. In order for patients to receive therapy at an early stage, With image processingtechnologies, diseases like cancer can be found early on. Finding aberrant tissue in target x-ray pictures requires a lot of time.One of the key components of this research is the accuracy and image quality.This method is effective for segmenting the picture, acquiring determining the area of interest for feature extraction, contrasting the normality and abnormalitythe image on the basis of general characteristics. Pixel percentage and masklabelling are the primary features for reliable image comparison detection. Thistells us that early detection of this disease is crucial to preventing its development into a serious stage and lowering its global prevalence.

To secure more precise outcomes in the three primary stages of image enhancement, image segmentation, and feature extraction

References

[1] Taruna Aggarwal, Asna Furqan, and Kunal Kalra, "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)*, 2015. [CrossRef] [Google Scholar] [Publisher link]

[2] Xin-Yu Jin, Yu-Chen Zhang, and Qi-Liang Jin, "Pulmonary Nodule Identification Based on CT Images Using Convolution Neural Network," *2016 Ninth International Conference on Computational Design and Intelligence (ISCID)*, 2016. [CrossRef] [Google Scholar] [Publisher link]

[3] Srinivas AruKonda, and S. Soundararajan, "Investigation of Lung Cancer Detection Using 3D Convolutional Deep Neural Network," *2020 2nd International Conference on Advances in Computing, Communication Control and Networking*, 2020. [CrossRef] [Google Scholar] [Publisher link]

[4] Sruthi Ignatius, and Robin Joseph, "Computer Aided Lung Cancer Detection System," *2015 Global Conference on Communication Technologies (GCCT)*, 2015. [CrossRef] [Google Scholar] [Publisher link]

[5] Y.M. Huo et al., "CNN vs.DNN Comparison of the Effect of False Positive Reduction in Deep Learning Based Detection for Juxta Pleural Lung Nodules," *Conference Proceedings on Modeling and Simulationin Medicine (MSM)*, 2017.

[6] J.D'Cruz et al., "Detection of Lung Cancer using Genetic Algorithm and Back Propagation Neural Networks," *Computer Technology Application*, vol. 6, no. 5, pp. 823-827, 2016.

[7] M. Bharathi, B.R. Sowmya, and S. Shashikala, "Lung Cancer Detection and Classification Using Deep CNN," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 8, no. 2, 2018. [Google Scholar]

[8] Gao Huang et al., "Densely Connected Convolutional Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017. [CrossRef] [Publisher link]

[9] Worku J. Sori et al., "DFD-Net: Lung Cancer Detection from Denoised CT Scan Picture Using Deep Learning," *Frontiers of Computer Science*, vol. 15, p. 152701, 2021. [CrossRef] [Publisher link]

[10] Christian Jacob, Jorg Denzinger, and Rotem Golan, "Lung Nodule Identification Using Deep Convolutional Neural Network," *International Joint Conference on Neural Networks (IJCNN)*, 2016. [CrossRef] [Publisher link]

[11] Sunyi Zheng et al., "Automatic Pulmonary Nodule Detection in CT Scans using Convolutional Neural Networks Based on Maximum Intensity Projection," *IEEE Transactions on Medical Imaging*, vol. 39, no. 3, pp. 797-805, 2020. [CrossRef] [Google Scholar] [Publisher link]

[12] Elmar Rendon-Gonzalez, and Volodymyr Ponomaryov, "Automatic Lung Nodule Segmentation and Classification in CT Images Based on SVM," *9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves*, 2016. [CrossRef] [Google Scholar] [Publisher link]

[13] P.B. Sangamithraa, and S. Govindaraju, "Lung Tumor Detection and Classification using EK-Mean Clustering," International Conference on Wireless Communications, Signal Processing, and Networking (WISPNET), 2016. [CrossRef] [Google Scholar] [Publisher link]

[14] Marjolein A. Heuvelmans et al., "Lung Cancer Prediction by Deep Learning to Identify Benign Lung Nodules," *Lung Cancer*, vol. 154, pp. 1-4, 2021. [CrossRef] [Google Scholar] [Publisher link]

[15] PrajwalRao, Nishal Ancelotti Pereira, and Raghuram Srinivasa, "Convolutional Neural Networks for Lung Cancer Screening in Computed Tomography (CT)Scans," *2nd International Conference on Contemporary Computing and Informatics*, 2017. [CrossRef] [Google Scholar] [Publisher link]

[16] Arvind T, K. Radhika, "XGBoost Machine Learning Model-Based DDoS Attack Detection and Mitigation in an SDN Environment," *International Journal of Engineering Trends and Technology*, vol. 71, no. 2, pp. 349-361, 2023. [CrossRef] [Publisher link]

[17] Anita Chaudhary, and Sonit Sukhraj Singh, "Multi Resolution Analysis Technique for Lung Cancer Detection in Computed Tomographic Images," *International Journal of Researching Engineering and Applied Sciences, IJREAS*, vol. 2, no. 2, 2012.

[18] A Manikandarajan, and S. Sasikala, "Detection and Segmentation of Lymph Nodes for Lung Cancer Diagnosis," *National Conference on System Design and Information Processing*, 2013. [Google Scholar]

[19] Albert Chon, Peter Lu, and Niranjan Balachandar, "Deep Convolutional Neural Networks for Lung Cancer Detection," Wavelet Recurrent Neural Networkfor Lung Cancer Classification": 3rd ICST Computer, 2017.

[20] S. Farjana Farvin, and S. Krishna Mohan, "A Comparative Study on Lung Cancer Detection using Deep Learning Algorithms," *SSRG International Journal of Computer Science and Engineering*, vol. 9, no. 5, pp. 1-4, 2022. [CrossRef] [Publisher link]

[21] Matko Šarić et al., "CNN-based Method for Lung Cancer Detection in Whole Slide Histopathology Images," *2019 4th International Conference on Smart and Sustainable Technologies (SpliTech)*, 2019. [CrossRef] [Google Scholar] [Publisher link]

[22] Wasudeo Rahane et al., "Lung Cancer Detection Using Image Processing and Machine Learning Health Care," *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*, 2018. [CrossRef] [Google Scholar] [Publisher link]

[23] Worawate Ausawalaithong et al., "Automatic Lung Cancer Prediction from Chest X-ray Images Using the Deep Learning Approach," *2018 11th Biomedical Engineering International Conference (BMEiCON)*, 2018. [CrossRef] [Google Scholar] [Publisher link]

[24] Febr Mokhled S. Al-Tarawneh, "Lung Cancer Detection Using Image Processing Techniques," *Leonardo Electronic Journal of Practices and Technologies*, 2012.

[25] Siddharth Bhatia, Yash Sinha, and Lavika Goel, "Lung Cancer Detection: A Deep Learning Approach," *Advances in Intelligent Systems and Computing*, vol. 817, pp. 699–705, 2019. [CrossRef] [Google Scholar] [Publisher link]