Detection of Lung Cancer from X-ray Images using Deep Learning - A Survey

[1]Someswari Perla, [2]G.Balaji
[1, 2]Department of Computer Science and Engineering, GMR Institute of Technology, Rajam, India,
[1]someswari.p@gmrit.edu.in, [2] 20341a0565@gmrit.edu.in

ABSTRACT:

The immense growth of technology has led to great technical advancements in medicine. One of the major focuses of research is detection of cancer in different organs like brain, breast, lung, prostate etc. Lung cancer is a leading cause for cancer deaths, and has higher rate of death when comparison with other cancers. Technology for lung cancer screening may benefit from the advancement of deep learning methods. Utilizing various deep learning techniques, Computer Aided Diagnosis (CAD) system of lung cancer has become one of the increasingly popular research topic. It aims in utilizing the computer system to analyze the possibility of cancer which is beneficial in early stages of recovery and also greatly increases the chance of recovery. This proposes an identification method for distinguishing between normal lung and certain forms of lung cancer. Therefore, development of deep learning algorithms may improve the identification of lung cancer using low-dose Computed Tomography and X-ray images. This study includes a more complex system and finds a better integrated method to improve accuracy and decrease the misdiagnose rate.

Keywords: Deep learning, Computer Aided Diagnosis, Lung cancer screening, Computer vision.

DOI Number: 10.14704/nq.2022.20.9.NQ44622 Neuro Quantology 2022; 20(9):5351-5355

Introduction:

Due to changes in the environment, the changing climate, people's lifestyles, and other factors, the impact of disease on health is fast rising. As a result, health problems are now more likely to occur. In 2016, 3.4 million individuals died from chronic obstructive pulmonary disease (COPD), which is typically brought on by smoking and pollution, while asthma claimed the lives of 400,000 people.

Particularly in emerging nations, where a lot of people struggle with air pollution and poverty, there is a very high risk of lung disorders. The WHO estimates that illnesses linked to home air pollution, such as asthma and pneumonia, cause nearly 4 million preventable deaths each year. As a result, action must be taken to lower carbon emissions and air pollution. The use of effective diagnostic tools that can aid in the early detection of lung disorders is also crucial.

The new corona virus illness has been causing severe damage in lungs and respiratory issues since late December 2019. Additionally, the virus that causes COVID-19 or other viral or bacterial infections may be to blame for pneumonia, a type of lung disease [3]. Early lung illness detection is therefore more crucial than ever. Here, machine learning can extremely useful. Digital technology has recently gained importance on a global scale.

This study can point physicians and other researchers in the right path for using deep learning to identify lung problems. The collection consists of a significant number of lung X-ray images. The system described here can also help with more accurate illness detection, protecting many vulnerable people and lowering the disease rate. Due in part to population increase, the health system has not yet been formed. Numerous studies have looked into a relation between machine learning models and prediction of diagnostic information from X-ray images. This is a critical moment to address this issue given the control of computers and the vast amount of records that are available to the public without restriction. With the expansion of Computer Science for medical and health research initiatives, this approach can result in declining medical expenses. The NIH chest X-ray

picture data is used for the implementation from the Kaggle repository, which is an entirely open source platform. This research introduces a new hybrid algorithm that successfully classifies lung illness using the dataset provided above. Numerous researchers have looked into how humans and machines interact. This study's key contribution is creation of a new combined deep learning method that can accurately diagnose lung illness using X-ray images.

Literature Survey:

1. Mubashir, M., Ahmed, M. R., Ahmad, M., Siddiqui, S. A., & Ahmad, M In order to distinguish between lung cancer and pulse wave signals, they have suggested a revolutionary deep CNN in this research. Proposed DCNN model undergoes supervised training to enable automatic perception of feature extraction and selection. As a result, the three traditional methods that devoured time and money were combined and replaced with a training procedure that was considerably more cohesive and demonstrated accuracy of 99.96%. It conducted a thorough comparison of accuracy, model amplitude, and time usage between the proposed DCNN and other neural networks using SAE, GRU, and LSTM. Findings confirmed that the proposed DCNN enacts the imposing precision and reasonable training cost. The validated accuracy of the submitted 15-layer DCNN model was determined to be 97.67%. Suggested approach would be required to notify lung cancer patients when the disease is still in its early stages, possibly allowing them to receive the recommended treatments. It significantly reduce the proposed DCNN's model magnitude in the future. Believe that a smaller model size will lead to even faster computation and calculation. Additionally, it would execute on embedded devices with greater coherence.

2. Mastouri, R., Khlifa, N., Neji, H., & Hantous-Zannad, S. A CNN model to classify lung nodules on CT scans called B-CNN was suggested in this work. It was composed of two VGG16 architecture-based feature extractors. Those results were pooled together with a bilinear pooling function. Two B-CNN combinations were discussed where the earlier one is depends on transfer learning, and the latter one is on fine-tuning, were put into practice with the goal of enhancing the performance of the suggested model. The methods of transfer learning and fine-tuning were contrasted objectively. In fact, the impact of each process on the model's performance was assessed.

The findings of the suggested method stimulate further research on additional cutting-edge pre-trained architectures in CNN to enhance classification performance. In this method, the best available parameters were assigned in order to minimise classification error rates.

3. Diao, Y., Shi, Y., Hou, Y., Gao, Y., & Yu, H. This research emphasized the importance of early on lung cancer identification using CAD systems as a means of reducing the high mortality rate. Patients may have more time to plan if an illness is caught early, which is beneficial for their rehabilitation. It showed out a complete 3D neural network-based CAD system, and its accuracy was 82.89%. It altered the CAD system's integrated technique, which, with a 90.38% true positive rate, can greatly lower the misdiagnosis rate. Making the appropriate diagnosis is beneficial to doctors. Even though the integrated system has overall satisfactory performance, some CT scans of the lungs are not separated appropriately, indicating that the division methods still have issues. The upcoming endeavor will be to determine a better division technique. In the future, It will be experimented with a more intricate approach in an effort to enhance accuracy and the rate of misdiagnosis is also reduced.

4. Causey, J., Li, K., Chen, X., Dong, W., Walker, K., Qualls In this paper they have discussed Over the past fifty years, a variety of computer-assisted methods for chest image interpretation have been created. Ginneken examines the use of 3 types of techniques — Rule-based image processing, machine and deep learning — in chest imaging and provides examples. The paper also demonstrated how deep learning is quickly taking over as the most effective strategy with highly promising outcomes. The majority of computational methods to date concentrate on locating and examining nodules in lung CT images. The application of these methods for automated screening is, however, constrained since relying on preset objects of interest necessitates detection and segmentation stages that are challenging to automate. In this study, It offer's the method Deep Screener, which uses a volumetric lung CT scan to determine a patient's

cancer status. The algorithm is built using the Spatial Pyramid Pooling model.

5. Anwer, D. N., & Ozbay In this paper, it is natural that all other disciplines, particularly the medical area, have been impacted by technological advancements. It is common knowledge that one of the most challenging diseases to treat is cancer. To boost the likelihood of survival, numerous scientists and researchers are working to develop methods to identify malignant cells early. Machine learning, which is one of the most effective techniques, is one of the methods scientists use to detect cancer early. Deep Network Designer has been used to create a deep learning CNN with seven layers. Transfer learning was used for training and testing, while confusion matrices were used for accuracy, yielding 93.33% accuracy, 92.59% precision, and 100% recall. These data were collated and compared to those from two other comparable works, demonstrating that the present work's overall result is superior.

6. Yu, H., Zhou, Z., & Wang, Q Lung cancer is a risky disease that must be detected at an early stage. Deep learning assisted AHHMM model is presented here to predict cancer in lungs on images of CT. In contrast to the DNN, which will focus on classifying photos within similar images, this study employs the Modified K-means method to classify images to slices in same image. Convolution layer with filtered edges is the next step in a thorough search for lung cancer. The quality of the photos was improved by calculating the weighted mean function that substitutes the pixel using the cumulative distribution and likelihood distribution approach. After the image has been enhanced, the damaged area has been divided by pixel value measurement. Extraction of spectral-related features was done using the similarity computation. The suggested AHHMM method successfully foresees lung cancer in computed tomography scanning pictures. In the future, a hybridized heuristic mathematical model will be used for predicting cancer in lungs at earlier stage.

7. Xie, Y., Xia, Y., Zhang, J., Song, Y., Feng, D., Fulham, M., & Cai, W Developed the MV-KBC model to distinguish benign from malignant lung nodules on CT chest images. It allows the model to be trained end-to-end. The LIDC-IDRI dataset results demonstrate that the model is more accurate than current state-of-the-art methods. To eliminate the necessity for data annotation, that adapts the suggested model to semi supervised learning in subsequent work. This will allow them to employ nodules with varying degrees of malignancy, unlabeled nodules as training sample. To increase the computational efficiency of the model training process, they will study the DCNN structure's ability to be compressed. For a more precise categorization of benign from malignant lung nodules, it will also be important to look into adding additional pathological data to the deep model.

8. Li, Z., Zhang, J., Tan, T., Teng, X., Sun, X., Zhao, H.,& Litjens, G The ACDC Lung HP challenge was outlined in this study. The challenge's current phase concentrated on segmenting lung cancer cases. For this challenge, 200 slides were used, and the best 10 teams strategies were chosen for comparison. In general, multi-model strategies outperformed single model-based techniques. The outcomes demonstrated the viability of deep learning for precise lung cancer diagnosis using WSI. Although deep learning was the foundation of all submitted approaches, the networks used varied greatly. It will concentrate on employing WSI biopsy to categorize the main lung cancer subtypes in the upcoming 2nd stage of this challenge. In 2020, at least 1000 slides gathered from various medical facilities will be made available. Think that the first stage's experiences will substantially aid the digital pathology community in achieving better results for the second stage.

9. Wang, B., Si, S., Cui, E., Zhao, H., Yang, D., Dou, S., & Zhu, J. This study offers a quick and effective method to assess the degree of lung nodule malignancy. Which suggest an unique vascular segmentation strategy to lessen the FPs because a lot of FP nodule regions will impair the efficiency and accuracy of classification in feature in the malignancy evaluations. The method uses a multi-view discriminating scheme to quickly, sensitively, and robustly eliminate vessels. Additionally, the use of the Frangi filter improvement result in this technique to identify tubular-form structure helps in guarantee the completeness of lung nodule region. In order to improve the categorization of lung nodule malignancy, which also optimize the extraction of spatial edge

characteristics and spatial texture data. This approach reduces the amount of redundant feature information, increasing system efficiency. Malignancy level classification and vessel segmentation serve as examples of the method's effectiveness. It contrasts the suggested MSPLBP and joint EOH with various features descriptors that includes single feature and multiple feature technique, for an experiment of malignancy level classification. Presented methodologies enable the CAD system to quickly and accurately classifying the malignancy level of lung nodules.

10. Bhatia, S., Sinha, Y., & Goel, L. In this research, It provide a method for diagnosis of cancer in lungs which makes the use of deep residual networks for feature extraction. It evaluates the effectiveness of classifiers that use trees, like as XGBoost and Random Forest. Also utilized an ensemble method of different classifiers including Random Forest and XGBoost, It achieved an accuracy of 84%. The largest cause of cancer-related fatalities worldwide is lung cancer. Identifying pulmonary nodules in the lungs that could develop into tumors in the near future is one of the crucial phases in the early stage cancer detection process. This study tries to estimate the risk that a particular CT scan of the lungs has malignant tissue.

11. Rahane, W., Dalvi, H., Magar, Y., Kalane, A., & Jondhale, S. A dangerous condition known as lung cancer is characterized by unrestricted cell proliferation in lung tissues. Early lung cancer detection is vital since it can save many lives. Using various machine learning and image processing algorithms, including binarization, grayscale conversion and noise reduction, it describe's lung cancer and its stages in our suggested system. The supplied CT scan image is pre-processed using each of these techniques. The main CT scan image is used to define ROI. The median filter and segmentation provide accurate results for pre-processing phases. These traits are useful for identifying cancer in early stages in lungs. The positive and negative samples of lung cancer imaging samples are classified in this system by a Support Vector Machine classifier for grouping purposes. Future research will use the proposed approach to develop ways to identify cancer in many human organs. The suggested method can be used for a particular malignancy, or for a group of diseases that aid in lowering the growth of aberrant cell or dispersal to other body areas.

Conclusion:

This work is for detecting lung diseases from X-ray images. Early detection allows patients more time to prepare, which aids in recovery. The integrated method of this CAD system, with a 90.38% true positive rate, can significantly reduce the misdiagnosis rate. It aids doctors in making accurate diagnoses. Even if the integrated system has overall satisfactory performance, some CT scans of the lung are not divided correctly, indicating that the dividing methods have flaws. The next step in the process is to figure out a better dividing method. In the future, trials will be more complex in order to find a better integrated method to improve accuracy and reduce the rate of misdiagnosis.

References:

[1] Mubashir, M., Ahmed, M. R., Ahmad, M., Siddiqui, S. A., & Ahmad, M. (2019, April). A novel deep learning approach for lung cancer recognition based on 1-D deep convolutional neural network. In Proceedings of the 2019 4th international conference on mathematics and artificial intelligence (pp. 32-38).

[2] Mastouri, R., Khlifa, N., Neji, H., & Hantous-Zannad, S. (2020, June). Transfer learning vs. fine-tuning in bilinear CNN for lung nodules classification on CT scans. In proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition (pp. 99-103).

[3] Diao, Y., Shi, Y., Hou, Y., Gao, Y., & Yu, H. (2019, December). Lung Cancer Detection with 3D Ensemble Convolution Neural Network. In Proceedings of the 2019 3rd International Conference on Computer Science and Artificial Intelligence (pp. 64-70).

[4] Causey, J., Li, K., Chen, X., Dong, W., Walker, K., Qualls,(2020). Spatial Pyramid Pooling With 3D Convolution Improves Lung Cancer Detection. IEEE/ACM Transactions on Computational Biology and Bioinformatics.

[5] Anwer, D. N., & Ozbay, S. (2020, September). Lung Cancer Classification and Detection Using Convolutional Neural Networks. In Proceedings of the 6th International Conference on Engineering & MIS 2020 (pp. 1-8).

[6] Yu, H., Zhou, Z., & Wang, Q. (2020). Deep learning assisted predict of lung cancer on computed tomography images using the adaptive hierarchical heuristic mathematical model. IEEE Access, 8, 86400-86410.

[7] Xie, Y., Xia, Y., Zhang, J., Song, Y., Feng, D., Fulham, M., & Cai, W. (2018). Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT. IEEE transactions on medical imaging, 38(4), 991-1004.

[8] Li, Z., Zhang, J., Tan, T., Teng, X., Sun, X., Zhao, H., ... & Litjens, G. (2020). Deep learning methods for lung cancer segmentation in whole-slide histopathology images—the acdc@ lunghp challenge 2019. IEEE Journal of Biomedical and Health Informatics, 25(2), 429-44.

[9] Wang, B., Si, S., Cui, E., Zhao, H., Yang, D., Dou, S., & Zhu, J. (2020). A fast and efficient CAD system for improving the performance of malignancy level classification on lung nodules. IEEE Access, 8, 40151-40170.

[10] Bhatia, S., Sinha, Y., & Goel, L. (2019). Lung cancer detection: a deep learning approach. In Soft Computing for Problem Solving (pp. 699-705). Springer, Singapore.

[11] Rahane, W., Dalvi, H., Magar, Y., Kalane, A., & Jondhale, S. (2018, March). Lung cancer detection using image processing and machine learning healthcare. In 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT) (pp. 1-5). IEEE.