International Journal On Engineering Technology and Sciences – IJETS™
ISSN (P):2349-3968, ISSN (O): 2349-3976
Volume VIII, Issue IX, September - 2021

A Survey on Lung Cancer Detection

A.Lavanya Mathiyalagi
Research Assistant,
Francis Xavier EngineeringCollege, Tirunelveli.
Emal:lavanyamathiyalagi@gmail.com

R.MALLIKA@ PANDEESWARI,
Research Scholar
&
RAHUL RAJA. P
Cyber forensics Applied Lab Student
Francis Xavier Engineering College, Tirunelveli.

Dr. R. RAVI
Professor/Dept. of CSE,
Francis Xavier Engineering College,Tirunelveli
directorresearch@francisxavier.ac.in

Abstract— According to statistics, lung cancer is the second most frequent type of cancer. The global and local feature extraction framework for lung cancer detection from CT scan images are proposed in many research papers. Data collection, worldwide training, and local training and testing are the three key elements of this framework. Given the prevalence of lung cancer the limitations of conventional techniques in detecting lung cancer at its early stages. In this study, various machine learning-based techniques for the early detection of lung cancer are compared and analysed. The majority of contemporary approaches for identifying lung cancer rely on CT scan pictures, however some also use x-ray imaging. To identify lung cancer nodules by image recognition, many segmentation approaches are paired with numerous classifier algorithms. This study suggests that CT scan pictures are more likely to yield accurate results. As a result, the bulk of CT scan pictures are used for cancer diagnosis. Marker-controlled watershed segmentation also produces more accurate results as compared to other segmentation methods. Additionally, compared to approaches that were applied using traditional machine learning techniques, the findings from methods based on deep learning techniques had higher accuracy.

Keywords*: CT scan, lung cancer, Machine Learning, local feature extraction.*

I. INTRODUCTION

In recent years, there has been a rapid increase in death rate due to lung cancer. Yesubai Rubavathi Charles and Ravi Ramraj, proposed highest performance with average precision of 88.2% on COIL-100 and 76.3% on Corel, the average recall of 69.9% on COIL and 76.3% on Corel. Thus, the experimental results confirm that the proposed content based image retrieval system architecture attains better solution for image retrieval [1].

Yesubai Rubavathi Charles and Ravi Ramraj proposed that experiments on benchmark databases such as MIT V are Tex database (DB1), Corel 1000 database (DB2). Experimental study shows that the proposed pattern gives average retrieval precision of 99.8 for DB1 and 76.5 for DB2, respectively [2].

It is a challenge for the radiologists to predict the pulmonary nodule growth earlier [4].

This causes major problem in many of the cases, since the CT scan is high in cost most of the rural patients are getting affected. Calculation of malignancy probability is predicting the lung cancer early. Lung cancer is one of the deadliest

diseases among dangerous diseases in the world. Despite significant research efforts and advances in the understanding of tumor biology, there was no reduction of the mortality over the last decades. More researches are going on about early diagnosis of lung cancer recently [5] [6].

Many conservative methods are focusing to select the nodule by enhancement filter only. This article algorithm is constructed neural network classification for reduction of false positive identification including of enhancement filter [7-10].

The motivation behind this paper is to reduce the challenges of the radiologists and to detect the early growth rate of pulmonary nodule by calculating probability of malignancy and segmented of lung nodule.

From the windpipe, the major airway, or the One could develop lung cancer. Unchecked cell growth and the spread of certain cells in the lungs are the causes of it. Lung cancer is more likely to occur in those who have emphysema or other chest diseases. The two primary types of lung cancer are small-cell lung carcinoma (SCLC) and non-small-cell lung carcinoma. The prevalence of SCLC is increasing, and smoking is a significant risk factor. Recent statistics predicted that non-small cell lung cancer, which makes up around 87 percent of all new cases of lung cancer, would rise to about 234,035 cases in 2019.

The growth of lung cancer without symptoms is the main aspect that makes this illness so lethal. A quarter of those surveyed showed no indications of malignancy. Many individuals are aware that lung cancer can also cause X-rays of the lungs. The importance of early diagnosis cannot be overstated because lung cancer spreads swiftly. Technological advances in imaging early lung cancer can now be identified thanks to developments like low-dose computed tomography.

Lung cancer is one type of nodule, a tumour that develops from cells in the respiratory system's airways.

These cells appear as a spherical entity and are always in direct contrast in chest X-rays. The likelihood that a patient will survive can be greatly increased if lung nodules are reliably identified at an early stage. However, evaluating these diagnostic images has become a highly tedious and challenging task because lung nodules utilizing raw chest X-ray imaging cannot be discovered rapidly. The computer-aided diagnostic system would have to be able to identify a small nodule from a huge 3D lung CT scan.

II. Machine Learning

The ability of computers to learn is known as machine learning, in which a machine is created with algorithms from which it may choose its own actions and present the user with the result. It is generally understood that artificial intelligence is involved. For the classification and decision-making of complex data, machine learning is being used. It is closely related to mathematical optimization, a topic that offers methods, theory, and implementation domains for a variety of computational tasks for which explicit algorithms cannot be planned and programmed. Machine learning methods and tasks can be divided into three groups:

Supervised learning:

This type uses the Linear Regression or Random Forest algorithms to tackle regression issues like anticipating weather, population growth, and life experience. Furthermore, supervised learning

employs algorithms from a wide range of disciplines, such as Support Vector Machines, Random Forest, and others, to address a variety of classification techniques. There are few levels in supervised learning. The data sets utilized for training must have known labels. The algorithms try to anticipate the values of the testing data by analyzing the link between the input values and labels.

Unsupervised learning:

This kind of learning focuses on issues like feature elaboration, the discovery of hidden structure, and dimensionality reduction utilized for large-data visualization. Additionally, it is utilized for clustering issues related to targeted marketing, client segmentation, and recommendation framework. This strategy does not use labels, in contrast to supervised learning.

Reinforcement Learning:

According to a set of tuning parameters, algorithms in this category attempt to forecast how an issue will turn out. The output is changed into an input parameter after the best output has been selected, and a new output is then found. This approach was utilised by both depth learning and artificial neural networks. Robot navigation, real-time judgments, and AI gaming are examples of applications that significantly rely on reinforcement learning.

III. Related Works

For diseases like lung cancer, researchers have successfully developed prediction models using statistical and machine learning techniques. Due to the work done by neural network approaches, the field of medical imaging is being used for the diagnosis of lung cancer. Several studies have attempted to classify and identify lung cancer using machine learning and conventional neural network techniques.

Using a chemical sensor array and machine learning techniques, a breath analyzer to detect lung cancer was created. Lung cancer cases and non-tumor controls are followed up in a prospective research between 2017 and 2019, and one alveolar air sample is evaluated utilising arrays of carbon nanotube sensors. The decision was made to use subjects enrolled in 2017 and 2018 for the model's internal validation. In 2019 patients recruited for evaluation used the model, which underwent external evaluation. The pathology records served as the benchmark for measuring the diagnostic accuracy. The model underwent external validation utilising pathological records as the reference standard.

Chest X-ray images are classified using the MAN method and lung CT images is classified using the EFT method. This method can detect 96% lung nodules on JSRT database. To test the neural system using machine learning techniques, the fuzzy neural system method is used. The accuracy of this strategy is 95.82 percent.

IV. Discussion and Analysis

The MAN method can detect 95% lung nodules on JSRT database. This study focuses on machine learning methods for lung cancer detection. Whereas some of the works that have been studied in literature have used X-ray pictures and some of the works have been based on CT scan images. The technique for detecting lung cancer goes through these stages in both situations: -

Pre-processing:

Pre-processing is the initial stage in which the input is either an X-ray or a CT scan image. Some image processing techniques such as Segmentation, which separates the comparable and dissimilar regions from the CT scan image, will come next. It has been proven that using a watershed and the masking method produced better outcomes.

Classification:

The collected features are then passed to a specific classifier at this point, which categorizes them as either normal or malignant. Researchers have employed a variety of classifiers, including different approaches, in the literature. Nearly 98 percent accuracy was the highest result that could be obtained using multi-resolution patch-based CNNs, which was the same for all research that applied Deep Learning methods.

V. Conclusion

Early detection of lung cancer is advantageous since treatment may start right away to stop the disease from having negative effects. This article gives a comprehensive evaluation of different machine learning techniques for classifying lung tumours using CT scan images. Additionally, a variety of classifiers, including Nave Bayes, Neural Network, Gradient Boosted Tree, k-nearest neighbors have been employed by researchers in the literature.

It can be concluded from the extensive survey undertaken in this work that procedures using deep learning techniques gave results that were more accurate than those using other conventional machine learning techniques. Where the best performance when using multi-resolution patch-based CNNs was around 98 percent.

ACKNOWLEDGMENT

Financial support obtained from the All India Council for Technical Education (AICTE) under Research Promotion Scheme (RPS),Sanction order no: F.No 8.9 / RIFD / RPS / Policy-1/2017-18 coordinated by Anna University Recognized Research Centre, Department of Computer Science and Engineering, Francis Xavier Engineering College,Vannarpettai,Tirunelveli-627003,Tamilnadu, India.

REFERENCES

[1] YesubaiRubavathi Charles and Ravi Ramraj, “Opponent Color and Curvelet Transform Based Image Retrieval System Using Genetic Algorithm”, International Journal of Computer and Information Engineering, vol:9, no:9, pp. 2011-2126, 2015.

[2] YesubaiRubavathi Charles and Ravi Ramraj, “A novel local mesh color texture pattern for image retrieval system”, International Journal of Electronics and Communications (AEÜ), vol.70, pp. 225-233, 2016.

[3] S.Shargunam, G. Rajakumar and R. Ravi, “Brain Tumor Identification: A comprehensive survey and perspective on recent works”, International Journal On Engineering Technology and Sciences, vol.7, no.2, pp.1-3, 2020.

[4] Mehdi Aliloub.,VassiliKovalev., Eduard Snezhko.,VahidTaimouri .:A Comprehensive Framework for Automatic Detection of Pulmonary Nodules in Lung CT Images. In: Image Anal Stereol. vol. 33, pp. 13-27. (2014)

[5] Wook-Jin Choi.,Tae-Sun Choi .: Automated Pulmonary Nodule Detection System in Computed Tomography Images: A Hierarchical Block Classification Approach. In: Entropy. vol. 15, pp. 507-523. (2013).

[6] Ezhil E, Nithila.,S.S. Kumar,S, S.: Automatic detection of solitary pulmonary nodules using swarm intelligence optimized neural networks on CT images. In: Engineering Science and Technology. vol. 20, pp. 1192–1202. (2017)

[7] Alex Martins Santos., Antonio Oseasde Carvalho Filho., Aristofanes Correa Silva., Anselmo Cardoso de Paiva ., Rodolfo Acatauassu Nunes., Marcelo Gattass. : Automatic Detection of Small Lung Nodules in 3DCT Data using Gaussian Mixture Models, Tsallis Entropy and SVM. In: Engineering Applications of Artificial Intelligence. vol.36, pp. 27–39. Elsevier (2014)

[8] Pablo G, Cavalcanti., ShahramShirani., Jacob Scharcanski., Crystal Fong., Jane Meng., Jane Castelli., David Koff .: Lung Nodule Segmentation

[9] in Chest Computed Tomography using a Novel Background Estimation Method. In: Quant Imaging Med Surg. Vol.6, pp. 16-24. (2016)

[10] TemesguenMessay., Russell C, Hardie., Timothy R, Tuinstra .: Segmentation Of Pulmonary Nodules In Computed Tomography using a Regression Neural Network Approach and Its Application to the Lung Image Database Consortium and Image Database Resource Initiative Dataset. In: Medical Image Analysis. vol.22, pp. 48–62. Elsevier (2015)

[11] Feifei Zheng., Aaron C, Zecchin., Jeffery P, Newman., HolgerR,Maier., Graeme C, Dandy .: An Adaptive Convergence-Trajectory Controlled Ant Colony Optimization Algorithm with Application to Water Distribution System Design Problems. In: IEEE Transactions on Evolutionary Computation. vol.21, pp.773 - 791.(2017)

[12] S. Ma,W.Wang, B. Xia, S. Zhang, H. Yuan, H. Jiang,W. Meng, X. Zheng, and X. Wang, ``Multiplexed serum biomarkers for the detection of lung cancer," EBioMedicine, vol. 11, pp. 210_218, Sep. 2016.

[13] W.-J. Kuo, R.-F. Chang, D.-R. Chen, and C. C. Lee, ``Data mining with decision trees for diagnosis of breast tumor in medical ultrasonic images," Breast Cancer Res. Treatment, vol. 66, no. 1, pp. 51_57, Mar. 2001.

[14] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S.Moore, S. Phillips, D. Maf_tt, M. Pringle, L. Tarbox, and F. Prior, ``The cancer imaging archive (TCIA): Maintaining and operating a public information repository," J. Digit. Imag., vol. 26, no. 6, pp. 1045_1057, 2013.

[15] A. Baratloo, M. Hosseini, A. Negida, and G. El Ashal, ``Part 1: Simple de_nition and calculation of accuracy, sensitivity and speci_city," Arch. Acad. Emergency Med., vol. 3, no. 2, pp. 48_49, 2015.

[16] P. Monkam, S. Qi, M. Xu, H. Li, F. Han, Y. Teng, andW. Qian, ``Ensemble learning of multiple-view 3D-CNNs model for micro-nodules identi_cation in CT images," IEEE Access, vol. 7, pp. 5564_5576, 2019.

[17] D. Yu, H. Wang, P. Chen, and Z. Wei, ``Mixed pooling for convolutional neural networks," in Proc. Int. Conf. Rough Sets Knowl. Technol., 2014, pp. 364_375.

[18] S. Özekes, ``Rule-based Lung region segmentation and nodule detection via genetic algorithm trained template

matching," Istanbul Commerce Univ. J. Sci., vol. 6, no. 11, pp. 17_30, 2007.

[19] F. Shaei and S. Fekri-Ershad, ``Detection of lung cancer tumor in CT scan images using novel combination of super pixel and active contour algorithms," Traitement du Signal, vol. 37, no. 6, pp. 1029_1035, 2020.

AUTHORS BIOGRAPHY

A.Lavanya Mathiyalagi is currently working as a Research Assistant, Department of Computer Science and Engineering, Francis Xavier Engineering College, Tirunelveli. Her research interest includes Image processing, Machine Learning, Big data.

R.MALLIKA@PANDEESWARI is currently pursuing Phd at Department of Computer Science and Engineering, Francis Xavier Engineering College, Tirunelveli. Her research interest includes Image Processing, Deep Learning and networks.

Dr.R.Ravi is currently working as a Professor &Research Centre Head, Department of Computer Science and Engineering, Francis Xavier Engineering College, Tirunelveli. His research interests include Medical Image Processing, Networks and Deep learning-based algorithm development.