http://www.jctjournals.com

ISSN(ONLINE):2278 - 3814

Volume-12, Issue-09, September 2023

JOURNAL OF COMPUTING TECHNOLOGIES (JCT)

International Journal

Page Number: 01-05

A Literature Survey of Lung Cancer Detection Using Different Machine Learning Techniques : A Review

[1]*Seema Bhalavi*, [2]*Prof. Shatendra Dubey*

[1]*M.Tech Scholar*, [2]*Assistant.professor*

[1,2]*Department of Information Technology Engineering*

[1,2]*NRI Institute of Information Science & Technology, Bhopal, India,*

[1]kathak.seema@gmail.com, [2]shatendradubey@gmail.com

***Abstract*— In this paper, we discuss on lung cancer is so common and dangerous, early diagnosis is essential for improving survival rates. The use of CT scans for lung cancer detection has been well validated. In order to help in the early detection of lung cancer nodules, this research offers a complete Lung Cancer Detection System that employs cutting-edge image processing algorithms on lung CT images. We have also addressed what has been covered before in this study, and we have showed that the size, cell structure or histology, and its molecular profile are all used to classify lung cancer. lung cancer of every kind persists and spreads in unpredictable ways.**

Keywords*— *Lung Cancer Detection, Computed Tomography (CT) and Image Processing etc.

I. INTRODUCTION

Image processing is an imperative branch in bio-medical applications, as this process is required in elaboration and distinguishing of pixels originating from numerous image sources, for example, Computed Tomography (CT) scans, X-rays, Magnetic resonance (MR) imaging etc. Exceptions are information which can be viewed as a typical because of a few causes (e.g. abnormal process conditions). Hence, the image processing helps to identify and detect such anomaly things like cancer that limit the effect of special cases in the last model to make, or as a preliminary pre-taking care of stage. As per literature, the image processing in medical field determines how cancer is to be cured, what are all the methods that must adapt to detect cancer in easy way.

On the other hand in recent surveys, cancer is determined in various parts of human body. The detection mechanism which is available results in serious limitations like compatibility is missing between each cancer analyses. In terms of improving such aspects and detection of cancer with high accuracy is must. A past methodology presents weaknesses, hence in the current years numerous commitments have been proposed to beat them and enhance the nature of the detection. Established strategies are regularly not 2 reasonable to treat some specific databases, along these lines late examinations have been directed on exception location for these sorts of dataset.

There are many researchers to take care of the lung cancer detection.

The methodologies that are all the more generally pertinent are unsupervised calculations as they don't require specific preparing information meeting the necessities of reasonable frameworks. The algorithms were implemented efficiently in order to ease the comparison of the approaches for researchers as well as making it possible for non-experts, investigators and specialists to spontaneously utilize the calculations and process complex datasets.

A. Lung Cancer

The term “lung disease” is utilized for tumour‟s emerging from the respiratory epithelium (bronchi, bronchioles, and alveoli). In lung Cancer, the cells of the lung develop in an abnormally uncontrolled way. It is one of the major dangerous neoplasms found throughout the world. It represents more malignant growth passing‟s consistently than breast, colon and prostate diseases. Most lung malignant growths are brought about via cancer-causing agents and tumour ingested by means of cigarette smoking.

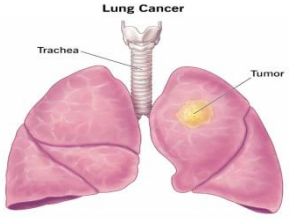

Figure 1 Lung cancer

In 2011, the National Lung Screening Trial Research Team reduced the lung-cancer transience with the help of computed topographic screening. Similarly, the Chronic Obstructive pulmonary disease is one of the smoking-related issues that further build the danger of generating lung malignancy. Other than cigarette smoking, there are different variables which are involved in the lung malignant growth. These incorporate air contamination, radioactivity, etc. Major effects of smoking have more vulnerability among ladies than men. There are lot of ways in our food culture to be changed to avoid cancer. Apart from that, the air pollution is one of the major issues for lung cancer.

II. LITERATURE REVIEW

Overview

This Section discuss about the exploration work done by different analysts on lung cancer analysis. Moreover, this area reviews vision concerning the utilization of detection mechanisms, therapeutic information, disease investigation, image segmentation and classification concepts are utilized by different analysts in their exploration work. The restorative information discussed here incorporates information identified with lung cancer or tumor in human body, which are utilized in different investigations. Also the order and clustering calculations are considered for research utilized in different inquires about diagnosing therapeutic issues are examined.

Karthick Prasad Gunasekaran et.al. (2023) - Lung cancer poses a significant global public health challenge, emphasizing the importance of early detection for improved patient outcomes. Recent advancements in deep learning algorithms have shown promising results in medical image analysis. This study aims to explore the application of object detection particularly YOLOv5, an advanced object identification system, in medical imaging for lung cancer identification. To train and evaluate the algorithm, a dataset comprising chest X-rays and corresponding annotations was obtained from Kaggle. The YOLOv5 model was employed to train an algorithm capable of detecting cancerous lung lesions. The training process involved optimizing hyperparameters and utilizing augmentation techniques to enhance the model's performance. The trained YOLOv5 model exhibited exceptional proficiency in identifying lung cancer lesions, displaying high accuracy and recall rates. It successfully pinpointed malignant areas in chest radiographs, as validated by a separate test set where it outperformed previous techniques. Additionally, the YOLOv5 model demonstrated computational efficiency, enabling real-time detection and making it suitable for integration into clinical procedures. This proposed approach holds promise in assisting radiologists in the early discovery and diagnosis of lung cancer, ultimately leading to prompt treatment and improved patient outcomes [01].

Yahia Said et.al. (2023)- Lung cancer presents one of the leading causes of mortalities for people around the world. Lung image analysis and segmentation are one of the primary steps used for early diagnosis of cancer. Handcrafted medical imaging segmentation presents a very time-consuming task for radiation oncologists. To address this problem, we propose in this work to develop a full and entire system used for early diagnosis of lung cancer in CT scan imaging. The proposed lung cancer diagnosis system is composed of two main parts: the first part is used for segmentation developed on top of the UNETR network, and the second part is a classification part used to classify the output segmentation part, either benign or malignant, developed on top of the self-supervised network. The proposed system presents a powerful tool for early diagnosing and combatting lung cancer using 3D-input CT scan data. Extensive experiments have been performed to contribute to better segmentation and classification results. Training and testing experiments have been performed using the Decathlon dataset. Experimental results have been conducted to new state-of-the-art performances: segmentation accuracy of 97.83%, and 98.77% as classification accuracy. The proposed system presents a new powerful tool to use for early diagnosing and combatting lung cancer using 3D-input CT scan data [02].

Iftikhar Naseer et.al. (2023) - Lung cancer is the most dangerous and death-causing disease indicated by the presence of pulmonary nodules in the lung. It is mostly caused by the instinctive growth of cells in the lung. Lung nodule detection has a significant role in detecting and screening lung cancer in Computed tomography (CT) scan images. Early detection plays an important role in the survival rate and treatment of lung cancer patients. Moreover, pulmonary nodule classification techniques based on the convolutional neural network can be used for the accurate and efficient detection of lung cancer. This work proposed an automatic nodule detection method in CT images based on modified AlexNet architecture and Support vector machine (SVM) algorithm namely LungNet-SVM. The proposed model consists of seven convolutional layers, three pooling layers, and two fully connected layers used to extract features. Support vector machine classifier is applied for the binary classification of nodules into benign and malignant. The experimental analysis is performed by using the publicly available benchmark dataset Lung nodule analysis 2016 (LUNA16). The proposed model has achieved 97.64% of accuracy, 96.37% of sensitivity, and 99.08% of specificity. A comparative analysis has been carried out between the proposed LungNet-SVM model and existing stateof-the-art approaches for the classification of lung cancer. The experimental results indicate that the proposed LungNet-SVM model achieved remarkable performance on a LUNA16 dataset in terms of accuracy [03].

Ashwini Pawar et.al. (2023) - Lung cancer is quite possibly the most risky illness in the world. Various determinations and recognition of cancer in the lungs have

been made utilizing different arrangement methods and information investigation. In this paper, a lung cancer detection framework utilizing a support vector machine and image processing are utilized to group the presence of lung cancer in CT scan images and blood tests. The work introduced in this paper is centered on the plan and improvement of a framework for primer conclusion and recognition of lung cancer from CT scan images. Arrangement of various tumor types is in this way imperative to guarantee higher endurance rates. The cycle of arrangement of lung cancer is continually testing. The framework recognizes the various phases of lung cancer which will help the specialists to identify lung cancer precisely and quickly from a lot of information [04].

Dr. P. Nancy et.al. (2022) **-** The primary contributor to lung cancer is an abnormal proliferation of lung cells. Tobacco usage and smoking cigarettes are the primary contributors to the development of lung cancer. The most common forms of lung cancer fall into two distinct types. Non-small-cell lung cancers and small-cell lung cancers are the two primary subtypes of lung cancer. A computed tomography, or CT, scan is an essential diagnostic technique that may determine the kind of cancer a patient has, its stage, the location of any metastases, and the degree to which it has spread to other organs. Other diagnostic tools include biopsies and pathology tests. The creation of algorithms that allow computers to gain information and abilities by seeing and interacting with the world around them is the core emphasis of the field of machine learning. This article demonstrates how to detect lung cancer via the use of machine learning by using improved feature selection and image processing. Image quality may be improved with the help of the CLAHE algorithm. The K Means technique is used in order to segment a picture into its component components. In order to determine which traits are beneficial, the PSO algorithm is utilised. The photos are then categorised using the SVM, ANN, and KNN algorithms respectively. It uses images obtained from a CT scan. When it comes to detecting lung cancer, PSO SVM provides more accurate results [05].

Imran Shaf et.al. (2022) **-** The diagnosis of early-stage lung cancer is challenging due to its asymptomatic nature, especially given the repeated radiation exposure and high cost of computed tomography(CT). Examining the lung CT images to detect pulmonary nodules, especially the cell lung cancer lesions, is also tedious and prone to errors even by a specialist. This study proposes a cancer diagnostic model based on a deep learning-enabled support vector machine (SVM). The proposed computer-aided design (CAD) model identifies the physiological and pathological changes in the soft tissues of the cross-section in lung cancer lesions. The model is first trained to recognize lung cancer by measuring and comparing the selected profile values in CT images obtained from patients and control patients at their diagnosis. Then, the model is tested and validated using the CT scans of both patients and control patients that are not shown in the training phase. The study investigates 888 annotated CT scans from the publicly available LIDC/IDRI database. The proposed deep learning-assisted SVM-based model yields 94% accuracy for pulmonary nodule detection representing early-stage lung cancer. It is found superior to other existing methods including complex deep learning, simple machine learning, and the hybrid techniques used on lung CT images for nodule detection. Experimental results demonstrate that the proposed approach can greatly assist radiologists in detecting early lung cancer and facilitating the timely management of patients [06].

Gopi Kasinathan et.al (2022) **-** Artificial intelligence (AI), Internet of Things (IoT), and the cloud computing have recently become widely used in the healthcare sector, which aid in better decision-making for a radiologist. PET imaging or positron emission tomography is one of the most reliable approaches for a radiologist to diagnosing many cancers, including lung tumor. In this work, we proposed stage classification of lung tumor which is a more challenging task in computer-aided diagnosis. As a result, a modified computer-aided diagnosis is being considered as a way to reduce the heavy workloads and second opinion to radiologists. In this paper, we present a strategy for classifying and validating different stages of lung tumor progression, as well as a deep neural model and data collection using cloud system for categorizing phases of pulmonary illness. The proposed system presents a Cloud-based Lung Tumor Detector and Stage Classifier (Cloud-LTDSC) as a hybrid technique for PET/CT images. The proposed Cloud-LTDSC initially developed the active contour model as lung tumor segmentation, and multilayer convolutional neural network (M-CNN) for classifying different stages of lung cancer has been modelled and validated with standard benchmark images. The performance of the presented technique is evaluated using a benchmark image LIDC-IDRI dataset of 50 low doses and also utilized the lung CT DICOM images. Compared with existing techniques in the literature, our proposed method achieved good result for the performance metrics accuracy, recall, and precision evaluated. Under numerous aspects, our proposed approach produces superior outcomes on all of the applied dataset images. Furthermore, the experimental result achieves an average lung tumor stage classification accuracy of 97%-99.1% and an average of 98.6% which is significantly higher than the other existing techniques [07].

Sneha S. Nair et.al. (2022) **-** One of the most serious and deadly diseases in the world is lung cancer. On the other hand, prompt diagnosis, as well as care, could save lives. Probably the most capable imaging method in the medical world, computed tomography (CT) scans are challenging for clinicians to analyze as well as detect cancer. In recent years, there has been an increase in the use of image analysis techniques for the detection of CT scan images matching cancer tissues. Using a Computer aided detection (CAD) system employing CT scans to aid inside the early lung cancer diagnosis as well as to differentiate among

benign/malignant tumors is thus interesting to address. The primary objective of this study would be to assess several computer aided approaches, analyze the right methodology already in use, and afterward propose a new approach that integrates enhancements to the best system currently in use. This research improves the performance of the existing retrieval system by combining various image feature extraction processes and modifying the internal layer section of the classifier. The segmentation method proposed here to identify cancer is Improved Random Walker segmentation along with Random Forest (RF) classifier and K-Nearest Neighbors (KNN) classifier. Here, the research is accomplished on the Lung Image database consortium (LIDC) datasets which is a collection of CT images and is utilized as the input images to verify the effectiveness of the suggested strategy. The accuracy of the proposed method for the detection of lung cancer with the aid of the RF classifier is 99.6 % as well as the KNN classifier is 96.4% accordingly [08].

Smita Raut et.al. (2021) - The cancer detection is doing with the aid of the skilled expert docs and earlier tiers it may be helpful. The opportunity of human error must be there. It produces the probability of error in the lung cancer detection which necessitate an automatic manner. Afterwards, the paperaims at early detection of cancer through an automatic procedure to decrease human error and making the system greater accurate and error free. In this system we use digital image processing and machine learning algorithm to discover the tumor in the images. Specially there are steps detection manner is performed one is digital image processing and other is machine learning algorithm. In digital image processing image acquisition, grey scale conversion, noise reduction, binarization of picture, segmentation, characteristic extraction, machine studying and the remaining step is most cancers mobile identification. In second step machine learning set of rules this is C 4.5 is used [09].

Lalitha, S. et.al. (2021) - Cancer has been one of the most serious health challenges to the human kind for a long period of time. Lung cancer is the most prevalent type of cancer which shows higher death rates. However, lung cancer mortality rates can be tracked by periodic screening. With the advanced medical science, the society has reaped numerous benefits with respect to screening equipments. Computed Tomography (CT) is one of the popular imaging techniques and this work utilizes the CT images for lung cancer detection. An early detection of lung cancer could prolong the lifetime of the patient and is made effortless by the latest screening technology. Additionally, the accuracy of disease detection can be enhanced with the help of the automated systems, which could support the healthcare experts in effective diagnosis. This article presents an automated lung cancer detection system equipped with machine learning algorithm, which can differentiate between the benign, malignant and normal classes of lung cancer. The accuracy of the proposed lung cancer detection method is around 98.7%, which is superior to the compared approaches [10].

III. TYPES OF LUNG CANCER

Generally, lung cancer is determined with the help of size, cell structure or histology and its molecular profile. Since, each kind of lung malignant growth carries on and is dispensed within an unexpected way. It has much data as could be expected about an individual lung disease. To get that data, a specialist should give tests of the tumor from a biopsy to a pathologist. Having a pathologist who is knowledgeable about taking a look at lung disease will give the most exact data.

The cancer is classified into two types based on its size namely, Small cell lung cancers (SCLC) and non-small cell lung cancers (NSCLC). Khuder (2001) described about the lung cancer and its causes. As shown in the Figure 1.2, the lung cancer is separated into various types of cell lung cancer with its possibilities. Here, the non-small lung cancer is determined as adenocarcinomas, squalors and large cell carcinomas.

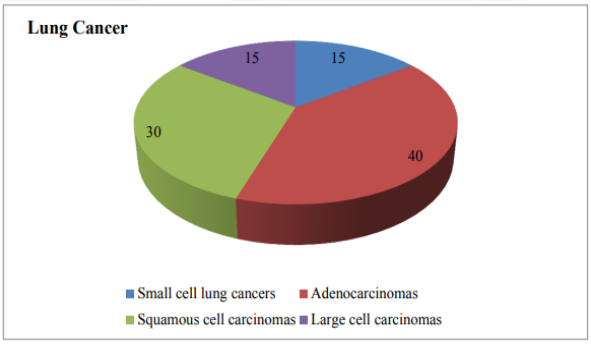

Figure 2 Types of Lung cancer and its possibilities range

There are different stages in lung cancer. Based on the cell regions, the process is varied. Initially, the cancer is present only in the lungs and hasn't spread to any of the lymph nodes. Later, the nearby regions will get affected. Cancer is found in the lung and in lymph which amidst the chest, additionally portrayed as privately propelled infection. It is divided into two subtypes: If the malignant growth has spread just to lymph hubs on a similar side of the chest where the disease began, it is called arrange IIIA. In the event that the disease has spread to the lymph hubs on the contrary side of the chest or over the bone present in neckline, it is called arrange IIIB. Finally, the most exceptional phase of lung malignancy, and is likewise portrayed as cutting edge infection. This is the point at which the malignancy has spread to the two lungs and to fluid in the territory around the lungs, or to another piece of the body, for example, the liver or different organs.

IV. CONCLUSION

The development of a lung cancer detection system using lung CT image processing holds significant promise in improving the early detection and diagnosis of

lung cancer, one of the most prevalent and deadliest forms of cancer worldwide. Through the utilization of advanced image processing techniques, this system aims to assist radiologists and medical professionals in accurately identifying potential cancerous lesions within the lungs. It also discusses about different types of cancer. The cancer is classified into two types based on its size namely, Small cell lung cancers (SCLC) and non-small cell lung cancers (NSCLC). Khuder (2001) described about the lung cancer and its causes

.

References

[1] Karthick Prasad Gunasekaran "Leveraging object detection for the identification of lung cancer"2023.

[2] Yahia Said ,Ahmed A. Alsheikhy, Tawfeeq Shawly and Husam Lahza "Medical Images Segmentation for Lung Cancer Diagnosis Based on Deep Learning Architectures" Volume 13, Issue 3, 5 January 2023.

[3] Iftikhar Naseer1, Tehreem Masood1, Sheeraz Akram1 , Arfan Jaffar, Muhammad Rashid and Muhammad Amjad Iqbal "Lung Cancer Detection Using Modified AlexNet Architecture and Support Vector Machine" vol.74, no.1 2023.

[4] Ashwini Pawar, Sankirti Shiravale, Swati Shekapure, Geetha Chillarge "Lung Cancer Detection Using Support Vector Machine (Svm) Classifier And Image Processing" ISSN: 1004-9037, Vol. 38 (2) 2023.

[5] Dr. P. Nancy, S Ravi Kishan, Kantilal Pitambar Rane, Dr. Karthikeyan Kaliyaperumal, Dr. Meenakshi and I Kadek Suartama "Optimized Feature Selection and Image Processing Based Machine Learning Technique for Lung Cancer Detection" Volume 10, Issue 4, Pages 888-894 2022.

[6] Imran Shafi , Sadia Din 2, Asim Khan , Isabel De La Torre Díez , Ramón del Jesús Palí Casanova , Kilian Tutusaus Pifarre and Imran Ashraf "An Effective Method for Lung Cancer Diagnosis from CT Scan Using Deep Learning-Based Support Vector Network" Volume 14 , Issue 21, 6 November 2022.

[7] Gopi Kasinathan and Selvakumar Jayakumar "Cloud-Based Lung Tumor Detection and Stage Classification Using Deep Learning Techniques" Volume 2022.

[8] Sneha S. Nair, Dr. V. N. Meena Devi and Dr. Saju Bhasi "Prediction and Classification of CT images for Early Detection of Lung Cancer Using Various Segmentation Models" Volume 10, Issue 4,2022.

[9] Shatendra Kumar Dubey "Effective Heart Disease Prediction ClassificationUsing Machine Learning" October 2022, Volume 20, Issue 09 Page 01-06

[10] Lalitha, S. "An automated lung cancer detection system based on machine learning algorithm" vol. 40, no. 4, pp. 6355-6364, 12 April 2021.

[11] Hamdalla F.Kareem , Muayed S. AL-Husieny , Furat Y. Mohsen , Enam A. Khalil , Zainab S. Hassan "Evaluation of SVM performance in the detection of lung cancer in marked CT scan dataset" Vol. 21, No. 3, March 2021, pp. 1731~1738.

[12] Kavitha B, Dr. Naveen K, Dr. Naveen "Investigation on Image Processing Techniques for Lung Cancer Detection" Issue: 6 | Pages: 8147- 8156, Year 2021.

[13] Imran Nazir ,Ihsan U Haq ,Muhammad Mohsin Khan ,Muhammad Bilal Qureshi, Hayat Ullah andSharjeel Butt "Efficient Pre-Processing and Segmentation for Lung Cancer Detection Using Fused CT Images" Volume 11, Issue 1, 23 December 2021.

[14] Mohammad A. Alzubaidi , Mwaffaq Otoom And Hamza Jaradat "Comprehensive and Comparative Global and Local Feature Extraction Framework for Lung Cancer Detection Using CT Scan Images" November 19, 2021.

[15] Meraj Begum and Shaikh Ismail "Lung Cancer Detection and Classification using Machine Learning Algorithm" Vol.12 No.13 (2021), 7048-7054.

