engineering proceedings

An Analysis of Machine Learning and Image Processing Techniques for Early Detection of Lung Cancer †

Vishwas V. Patange[1] and Jagadish B. Jadhav[2]

[1] Electronics & Telecommunication Engineering Department, S.T.B. College of Engineering Tuljapur Maharashtra; vishwaspatange@gmail.com
[2] Electronics & Telecommunication Engineering Department, R.C.P.I.T. Shirpur Dhule, Maharashtra; jagadish.rcpit@gmail.com

† Presented at the 4th International Electronic Conference on Applied Sciences, 27 October–10 November 2023; Available online: https://asec2023.sciforum.net/.

Abstract: Lung cancer is a significant global health concern, necessitating accurate and reliable methods for its diagnosis and classification. This survey paper aims to provide a comprehensive overview of the existing research on lung cancer, focusing on the advancements in diagnostic techniques and classification models. Through a systematic literature review, various machine learning algorithms employed for lung cancer classification were examined, highlighting their strengths and limitations. Additionally, the impact of handling Dicom images on accuracy levels was investigated, emphasizing the need for proper image processing techniques. The survey reveals that while several classifiers have demonstrated promising results, achieving close to 100% accuracy remains a challenge. Furthermore, the study emphasizes the effectiveness of ensemble classifiers in outperforming other algorithms. To enhance accuracy levels and gain meaningful insights for tumor diagnosis, the paper suggests the development and application of more sophisticated models. Lastly, it emphasizes the significance of further research in the field of Oncology to enhance the classification of benign and malignant lung tumors. This survey paper serves as a valuable resource for researchers, clinicians, and practitioners working towards improved lung cancer diagnosis and classification.

Citation: Department, T.E.; Maharashtra, S.C.E.T. Vishwas V. Patange[1] and Jagadish B. Jadhav[2]. *Eng. Proc.* **2023**, *5*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s):

Published: date

1. Introduction:

Lung cancer is a formidable global health challenge, accounting for a substantial number of cancer-related deaths worldwide. It is characterized by uncontrolled growth of abnormal cells in the lung tissues, leading to the formation of tumors that can interfere with normal lung function. Lung cancer is a complex and multifaceted disease, with diverse etiological factors and distinct histological subtypes that necessitate comprehensive understanding for effective management and treatment.

The two main types of lung cancer are non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC comprises approximately 85% of all lung cancer cases and is further classified into three subtypes: adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Each subtype has unique characteristics and may exhibit different responses to treatments. Adenocarcinoma, the most common subtype, typically arises in the outer regions of the lung and is often associated with genetic mutations such as EGFR (epidermal growth factor receptor) and ALK (anaplastic lymphoma kinase). Squamous cell carcinoma arises in the lining of the bronchial tubes and is frequently linked to smoking. Large cell carcinoma is a less common subtype that lacks specific features observed in the other subtypes.

SCLC, on the other hand, accounts for about 15% of lung cancer cases and is characterized by its rapid growth, early metastasis, and association with smoking. SCLC cells are small in size, and the cancer tends to spread quickly to other organs. Due to its aggressive nature, SCLC often requires a different treatment approach compared to NSCLC.

Understanding the distinct subtypes of lung cancer is crucial as it guides treatment decisions, including surgery, radiation therapy, chemotherapy, targeted therapies, and immunotherapy. Moreover, advancements in molecular profiling and precision medicine approaches have provided new opportunities for personalized treatment strategies based on the specific genetic alterations exhibited by individual lung cancer patients.

Lung cancer is a complex disease with different subtypes that vary in terms of histology, molecular characteristics, and treatment response. Continued research efforts are necessary to deepen our understanding of the underlying mechanisms, develop novel therapies, and improve patient outcomes. By unraveling the intricacies of lung cancer and its subtypes, we can pave the way for more effective prevention, diagnosis, and treatment strategies to combat this devastating disease.

2. Motivation

According to the stage at which abnormal cells are discovered in the lungs, lung abnormality is the most deadly and pervasive disease in the world. This tells us that early diagnosis of this disease is crucial for preventing serious stages and reducing its global prevalence. The leading cause of mortality from disease worldwide is lung cancer.There is a relatively high fatality rate from lung cancer. There are many different cancerous tumours, including breast and lung cancer. For successful treatment of carcinoma, early stage identification is essential. Computed Tomography (CT) images are used to make the diagnosis.Cancer additionally called tumor should be quick and effectively detected inside the preliminary degree to discover what might be beneficial for its therapy. Despite the fact that modality has different considerations consisting of complicated records, incorrect diagnostics and remedy might be major reasons of deaths [7].A selection of operational, applied and surgical treatments will gain several sufferers if the situation is a tiny, diffuse tumor. Unfortunately, as there are few without any symptoms within the early stages of contamination, a analysis of innovative scientific illness, nodal progression and/or metastatic ailment in seventy five percentage of lung cancers emerge later.

3. Related Work

According to American Cancer Society, it is more probable that lung cancer will respond favorably to treatment if it is discovered early, when it is still tiny and has not spread. Some smokers or former smokers who don't exhibit any symptoms are advised to get screened for lung cancer. In most cases, the absence of symptoms in a lung cancer patient indicates that the illness may be early-detected. Often, lung cancer symptoms do not show up until the illness has progressed significantly. Even when signs of lung cancer do appear, many individuals may mistake them for other conditions, such as an infection or long-term smoking-related consequences. Delays in diagnosis might result from this.

All researchers aimed to develop a system that predicts and detects cancer in its early stages. Efforts were also made to improve the accuracy of the early detection and prediction system through preprocessing, segmentation, feature extraction, and classification techniques of the extracted database.According to G. Sutedja [8], a number of difficulties about this global epidemic have been brought up by the renewed interest in lung cancer screening and the use of innovative approaches for the management of early disease. Conventional chest radiography and sputum cytology tests for screening have not been proved to reduce lung cancer mortality in prior randomized clinical trials. Most lung cancer patients receive an advanced stage diagnosis. If a tumour can be entirely removed during surgery and there are no lymph nodes or distant metastases, surgery offers the best chance of recovery [8]. Accurate staging is the key stage before any treatment. Squamous cell cancer in its "early stage" should be evaluated with relation to the tumour margins and depth of infiltration in the bronchial wall. It's critical to understand that a subset of patients with early-stage SCC have tiny intraluminal and superficial lesions [8]. For the research on the early detection of lung cancer (probably at stage I or II), we are goinh to consider input database as available on internet sites viz. kaggle, LIDC-IDRI (Lung image database consortium and image database resource initiative), Indian Lung CT Image

Database. Lung image datasets play an essential role in evaluating the performance of deep learning-based algorithms for lung nodule classification and detection.

4. An Overview of Some Recent Developments and Approaches in Lung Cancer Early De-Tection Are As Under

i. Low-dose Computed Tomography (LDCT): LDCT screening has emerged as a promising tool for early detection of lung cancer in high-risk individuals. Several large-scale studies, such as the National Lung Screening Trial (NLST), have demonstrated a reduction in lung cancer mortality with LDCT screening compared to conventional methods.
ii. Biomarkers: Researchers are actively investigating biomarkers, such as circulating tumor DNA (ctDNA), microRNAs, and proteins, for their potential use in early detection. These biomarkers can be detected through blood tests or other non-invasive methods and may aid in identifying lung cancer at an early stage.
iii. Artificial Intelligence (AI): AI algorithms are being developed and applied to aid in the early detection of lung cancer. AI can assist in analyzing medical images, such as chest X-rays and CT scans, to identify suspicious lesions or nodules that may indicate lung cancer.
iv. Molecular profiling: Advances in genomic sequencing technologies have allowed researchers to identify specific genetic alterations and molecular signatures associated with lung cancer. Molecular profiling of lung tumors can provide insights into personalized treatment strategies and may aid in early detection through the identification of driver mutations or specific gene expression patterns.
v. Liquid biopsies: Liquid biopsies, which involve the analysis of circulating tumor cells, cell-free DNA, or exosomes in blood samples, are being explored as a non-invasive method for early detection of lung cancer. These tests have the potential to detect genetic mutations or alterations associated with lung cancer.

5. Literature Survey

A detail literature survey is described in Section 5. Section 5.1 deals with the performance evaluation of models.

"Comparative Analysis of learning Algorithms for Lung Cancer Identification" was the topic that Syed et al. [17] offered. The authors of this article focused on the differences between benign and malignant characteristics. The suggested effort attempted to fully automate the procedure of automatically detecting the cancer. The two phases underwent experimentation by the authors: The first phase is selecting the most important features from the CT scan images and mapping them. Machine learning methods are used in the second phase.

Dash et al. [18] suggested the category "Multi-Classifier Structure for Lung Cells." The Authors used High-Resolution Computed Tomography (HRCT) tests to detect lung cancer cells in this study. To achieve the first decision on the input image, the classifier employs the Discrete Wavelet Transform as well as Several Classifier. They extracted the functions from the input photo and fed them into the Semantic Network Classifier, Ignorant Bayes Classifier, and Choice Combination, which resulted in the Accurate Choice.

A study on "Comparison of Lung Cancer Detection Algorithms" was proposed by Günaydin et al. [19]. In this study, authors looked at the prevalence of lung cancer in both men and women. To find anomalies, authors used a variety of classifiers, including Principal Component Analysis, K-Nearest Neighbours, Support Vector Machines, Naive Bayes, Decision Trees, and Artificial Neural Networks. By using machine learning methods, he determined the accuracy, sensitivity, and specificity. The authors used lung section radiography instead of using noise removal techniques on the image.

A study on "Lung Cancer Detection Using CT Scan Images" was offered by Makaju et al. [20]. In this study, the authors claimed that they employed watershed segmentation rather than working with other filters like the Gabor, Median, and Gaussian filters (Table 2, 3).

"A Comparative Study of Lung Cancer Detection Using Machine Learning Algorithms" was the topic that Radhika et al. [21] offered. In this study, the authors developed four classifiers: SVM model, Naive Bayes model, Decision Tree model, and Logistic Tree model. These classifiers were then applied to three datasets from the UCI Machine Learning Repository: BRATS, OASIS, and NBTR. and in Table 1 compared the degrees of accuracy of these methods.

Roy et al. [22] presented a "Comparative study of lung cancer detection using supervised neural network." The authors conducted considerable research on classification methods utilised in the diagnosis and detection of lung cancer in this paper. To classify, he employed bio-medical pictures and machine learning techniques. He took the following technique (Table 1, 3).

Step 1: Converting RGB image to Grayscale

Step 2: Applying Linearization Techniques

Step 3: Detecting and segmentation the Image

Step 4: Extracted the Features

Step 5: Cross-Validation

Song et al. [23] offered their work "Prognosis of Stage I Lung Cancer Patients Using Quantitative Analysis of Centrosomal Features," which shows cancer nodules with markings. In addition, they used additional parameters such as centroid, diameter, and pixel mean intensity to detect the cancer nodule. The following are the strengths of this work: (1) the accuracy of detecting lung cancer nodules has increased.

(2) Determines whether the cancer cell is malignant or benign.

(3) False detection parameters such as salt-pepper noise and speckle noise have been eliminated.

"A Comparative Analysis of Segmentation Techniques for Lung Cancer Detection" was the topic of a study proposed by Tripathi et al. [24]. The authors offered a number of picture segmentation techniques in this paper, including basic thresholding, Otsu thresholding, edge thresholding, DE-based segmentation, and marker controlled watershed segmentation. There have been enough methods for identifying lung cancer established in recent years. They used image processing to integrate several segmentation algorithms with classifier techniques to identify lung cancer nodules. He applied several picture segmentation strategies in Table 1 using the SVM classifier, which was used in this paper.

Lung cancer, according to Mokhled A. Leonardo et. al. [9], is caused by aberrant cells proliferating and developing into tumours. As lymphatic fluid naturally exits the lungs with a flow towards the centre of the chest, lung cancer frequently spreads in that direction.Primary lung cancer is a type of cancer that begins in the lung.

Comparative Study and Performance Metrics

This study extensively surveys papers published between 2014 and 2022. It demonstrates that deep learning-based lung imaging systems have achieved high efficiency and state-of-the-art performance for lung nodule segmentation, detection, and classification using existing medical images. However, deep learning techniques still have many unsolved problems in lung cancer detection. First, clinicians have not fully acknowledged deep learning techniques for everyday clinical exercise due to the lack of standardized medical image acquisition protocols. Second, It is costly and time consuming to collect an enormous annotated image dataset, even performed by experienced radiologists. Third, the clinical application of deep learning requires high interpretability, but current deep learning techniques cannot effectively explain the learned features. Many researchers have applied visualization and parameter analysis methods to explain deep learning models. However, there is still a certain distance from the interpretable imaging markers required by clinical requirements. Finally, some of the current literature has little translation into applicability in clinical practice due to the lack of experience of non-medical investigators in choosing more relevant clinical outcomes.

Table 1. Lung Nodules Detection Techniques with Machine Learning Algorithms Table 1 shows the various models employed by the authors.

Sr. No.	Author	Technique applied
01	Tripathi, Tyagi, and Nath 2019	Image Segmentation Technique
02	Radhika, Nair, and Veena 2019	Decision Trees, Logistic Regression, Support Vector Machines
03	Roy et al. 2019	Random Forest, Support Vector Machines
04	Günaydin, Günay, and Şengel 2019	Artificial Neural Network, Naive Bayes, Decision Tree
05	Makaju et al. 2018)	Support Vector Machines
06	Dash et al. 2014	Neural Network Classifier, Naive Bayes, Decision Fusion
07	Song et al. 2012	Linear Discriminator Analysis (LDA), Support Vector Machines
08	Vijai Anand 2010	Back Propagation Network

Table 2. Comparative study of classification accuracy.

Algorithm	Sensitivity	Specificity	Accuracy
Image Segmentation Technique	65.4%	65.4%	59.12%
Decision Trees	72.67%	72.67%	58.11%
Logistic Regression	84.34%	96.66%	65.4%
Random Forest	64.28%	64.28%	62.5%
Support Vector machines	88.5%	88.5%	62.5%
Artificial Neural Network	65.4%	65.4%	65.4%
Naive Bayes	62.5%	62.5%	62.5%
Ensemble Classifier	72.67%	65.4%	88.5%
Decision Fusion	74.78%	62.5%	72.67%
Linear Discriminator Analysis (LDA)	62.5%	91.66%	74.78%
Back Propagation Network	72.67%	72.67%	72.67%
K-Nearest	94.12%	53.78%	72.24%

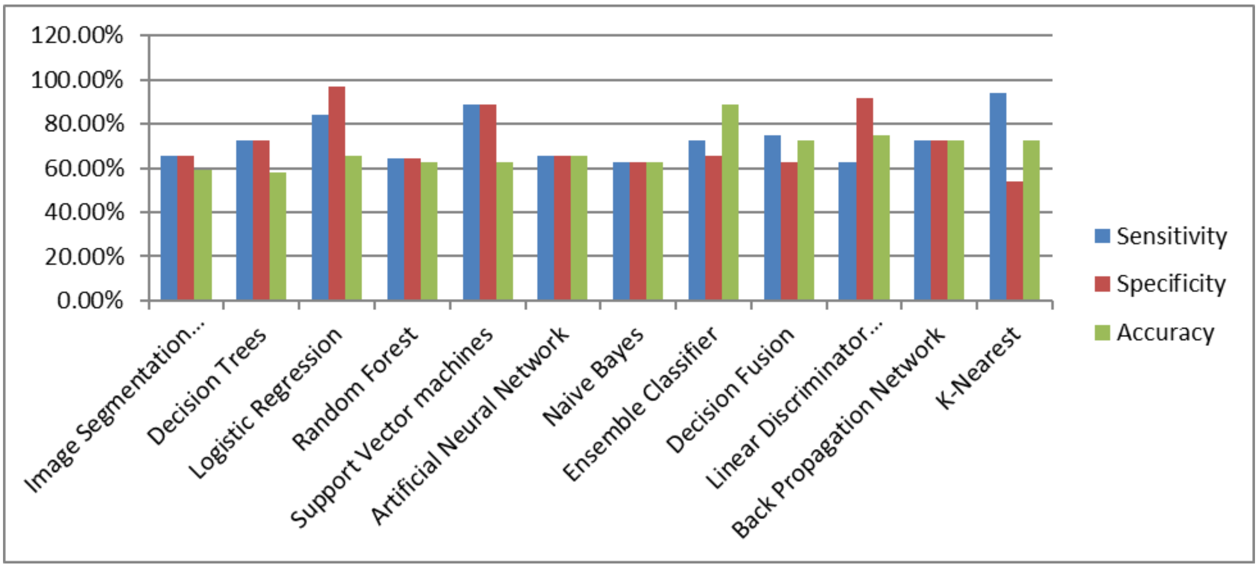

Figure 1. The accuracy levels of various machine learning algorithms for lung cancer classification.

Table 3. Comparison of various machine learning classifiers.

	S. Deviation Train	S. Deviation Test	ROC	AUC	Gmean	Precision	Recall	F1_score
AdaBoost	0.15726	11.74342	0.78319	0.50000	0.65094	0.73469	0.83721	0.78261
Decision Tree	0.15726	9.18525	0.79361	0.50000	0.66322	0.75000	0.83721	0.79121
Extra Forest	0.63311	7.90518	0.80644	0.50000	0.67843	0.74510	0.88372	0.80851
Extra Tree	0.15726	6.78843	0.83648	0.50000	0.71894	0.80435	0.86047	0.83146
Gaussian Naive Bayes	0.63547	5.58525	0.88857	0.50000	0.80128	0.90244	0.86047	0.88095
Gaussian Process	1.05988	6.81512	0.86531	0.50000	0.76870	0.89744	0.81395	0.85366
k-Nearest Neighbors	0.81683	7.24203	0.84448	0.50000	0.73553	0.85366	0.81395	0.83333
LDA	1.06127	8.35666	0.82243	0.50000	0.70560	0.82927	0.79070	0.80952
Linear SVM	0.15726	11.74342	0.78319	0.50000	0.65094	0.73469	0.83721	0.78261
Logistic Regression	1.33854	7.25686	0.85489	0.50000	0.75185	0.87500	0.81395	0.84337
Multilayer Perceptron	0.67524	5.97748	0.88736	0.50000	0.80237	0.92308	0.83721	0.87805

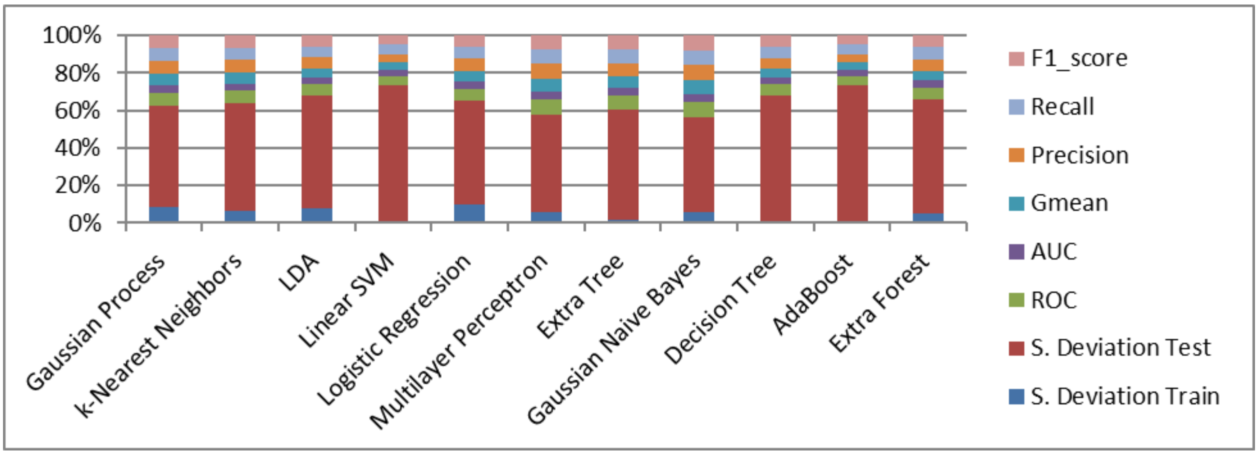

Figure 2. The performance levels of each classifier.

Steps of image processing for lung cancer:

1. The initial step entails gathering both normal and abnormal CT images from the accessible Database.

2. At the second stage, picture segmentation algorithms are used, which is a crucial step in the image processing process.

3. The third stage employs feature extraction methods, which are useful at earlier phases of image processing.

4. The retrieved attributes that indicate whether an image is normal or abnormal are used in the fourth stage to identify lung cancer cells [9].

The various machine learning methods have been studied and we have prepared a research article on this which shows the comparative study analysis of different methods. This work does not presente any new solution right now, but suggests some enhancements and future scope for the problem. We are working on the solution for the early detection of lung cancer so that the most important human lives can be saved or can be extended to a few more years. In my paper, I reviewed recent achievements in deep learning-based approaches for lung nodule segmentation, detection, and classification. CNN is one of the most widely used deep learning techniques for lung disease detection and classification, and CT image datasets are the most frequently used imaging datasets for training networks. The article review was based on recent publications. Radiologists, computer scientists, researchers and engineers need to work more closely to develop more realistic and sensitive models and add more meaning to the research field.

Mokhled et. Al., In his research, Mokhled used three stages: image enhancement, image segmentation, and feature extraction. FFT and the Gabor filter are both employed for image improvement. He employed a marker controlled watershed segmentation approach for image segmentation, and a binarization and masking strategy for feature extraction.

The implementation of the auto enhancement approach heavily relies on subjective observation and mathematical calculations like average and variance. Cancer is caused by a variety of factors, including physical carcinogens like exposure to radiation and ultraviolet rays, behavioural abnormalities like high body mass index and alcohol and tobacco use, as well as specific biological and genetic carcinogens. Ache, exhaustion, vomiting, a chronic cough, respiratory issues, fat loss, muscle discomfort, bleeding, bruises, and numerous other symptoms are frequent cancer signs.

Over diagnosis, according to James H. Finigan and Jeffrey A. Kern et. al. [10], denotes the recognition of malignancies that do not advance or have an impact on mortality. Unnecessary, often invasive examinations and therapies are the result of over diagnosis. Over diagnosis is proven by the Mayo Lung Study. 206 malignancies were discovered in the screening arm after 8.5 years of follow-up versus 160 in the control group. The Prostate, Lung, Colorectal, and Ovarian (PLCO) experiment was the most current trial using chest radiograph screening. The PLCO research presents potential conclusive evidence that annual chest radiographs had no effect on lung cancer mortality or diagnosis when compared to no screening.

Lung cancer is the leading cause of cancer-related death since the majority of patients are identified at a more advanced stage. There is no likelihood of a viable treatment being created in the near future. No matter if a nation is industrialized or developing, lung cancer is constantly rated as one of the worst types of the disease.Lung cancer is becoming more common in developing countries as a result of longer lifespans, urban development, and the acceptance of Western lifestyles. The early detection of cancer, in addition to the survival of cancer patients, is critical to the control of lung cancer [13].

According to the published paper [15] posted by Muthazhagan B et. al. (2021), forecasting and identifying lung cancer at an early stage is a challenging task using current lung cancer prediction technologies. An initial lung nodule prediction could add one to five years to a person's life. They developed a classifier based on Support Vector Machines that given approximately 98% prediction accuracy in a short period of time. However, the images were simply classified as 'abnormal' or 'normal,' with no regard for the different phases [Stage 0 - Stage IV]. Deep residual learning is used by Siddharth Bhatia et al. [14] to detect lung cancer. Using UNet and ResNet models, they present a set of preprocessing techniques for retrieving melanoma lung features. They assess the efficacy of classifiers such as Random forest and Gradient boosting in predicting carcinogenic CT scans. The authors achieve the highest accuracy of 84% when they combine the two classifiers. In this case, the constraint is that the best attainable accuracy could have been higher.

The National Lung Screening Trial (NLST), which was just published, is the first trial to indicate that screening can lower lung cancer mortality. It showed increased lung cancer survival in patients who underwent serial low dose CT scans. This trial demonstrates that lung diagnostic test on a big scale could be effective.

There have been several ancient advancements over the hundred years, with none more significant than Sir Richard Doll and Austin Hill's landmark article, published in the British Medical Journal in 1950, which affirmed apprehensions that cigarette smoking was linked to lung cancer. The second seminal article was the Surgeon General's report in 1964, in which he mentioned that cigarette smoke was a health risk and that it was important to avoid or quit the habit. The most concerning trend is the burgeoning epidemic in developing-world countries. It was estimated in 1985 that there were 921,000 lung cancer deaths worldwide, a 17% increase from 1980 [11].

Preprocessing has been utilised by Sangamithraa and Govindaraju [12] to get rid of undesirable data that is unaffected by median and Wiener filters. This was done to raise the data's overall quality. The segmentation of the CT images is carried out using the K-means approach. Bhatia et al. developed a method for identifying the presence of lung cancer in a CT image using deep residual learning. Forecasts about the probability that a CT scan is malignant are collected using an ensemble of XGBoost and random forest classifiers. The likelihood that a CT scan is malignant is calculated using the combined findings of the predictions made by each classifier. The accuracy of the LIDC-IRDI is 84% greater than that of conventional methods [12].

6. Analysis

Early detection of lung cancer plays a crucial role in improving patient outcomes and reducing mortality rates associated with the disease. Here is an analysis of the significance and impact of early detection:

1. Improved Treatment Options: Detecting lung cancer at an early stage often allows for a wider range of treatment options. Early-stage lung cancer is more likely to be amenable to surgical resection, which offers the highest chance of cure. Additionally, early detection increases the likelihood of successful treatment with less invasive interventions, such as localized radiation therapy or targeted therapies.
2. Increased Survival Rates: Studies have consistently shown that early detection of lung cancer is associated with improved survival rates. When lung cancer is diagnosed at localized stages (Stage 0, I, or II), the five-year survival rate is significantly higher compared to when it is diagnosed at advanced stages (Stage III or IV). Early detection allows for prompt initiation of appropriate treatment, increasing the chances of long-term survival.
3. Screening Programs: The development of lung cancer screening programs, particularly using low-dose computed tomography (LDCT), has demonstrated the potential to detect lung cancer at an early stage in high-risk individuals. Screening programs have shown a reduction in lung cancer mortality rates by detecting cancers when they are smaller and more localized, enabling timely intervention.
4. Cost-Effectiveness: Early detection can lead to cost savings in the long term. By identifying lung cancer at an early stage, treatment options may be less invasive and less costly compared to advanced-stage disease management. Additionally, early detection reduces the need for extensive treatment and palliative care, which can be associated with significant healthcare costs.

By studying the characteristics of early-stage lung cancer and the factors associated with early detection, researchers can identify new biomarkers, develop novel diagnostic tools, and improve screening methodologies. This, in turn, leads to further advancements in early detection strategies and overall management of lung cancer.

To conclusively demonstrate the value of early intervention in the "population at risk," longer follow-up is required. When lung cancer is discovered at the carcinoma in situ stage, the bronchoscopist should take a more proactive approach because early detection, appropriate staging, and intraluminal bronchoscopic treatment may offer the highest opportunity for cure [8].

7. Conclusion

This paper presents an extensive study on machine learning algorithms for the classification of lung cancer malignancies. Through a systematic literature review, three main gaps were identified. Firstly, the accuracy of existing classifiers was found to be unsatisfactory. Secondly, the need for novel techniques, specifically for noise reduction in medical images such as Dicom and MRI, was highlighted. Thirdly, the limitations of earlier machine learning algorithms in capturing hidden patterns from raw images were noted.

Furthermore, the importance of employing proper classifiers with improved algorithms in the field of medical informatics was emphasized to ensure accurate results. The paper underscores the necessity for intensive research in cancer studies to reduce cancer-related mortality. In conclusion, early detection of lung cancer is crucial for improved patient outcomes and increased survival rates. It allows for more treatment options, higher chances of cure, and better quality of life. Screening programs and ongoing research efforts are essential in advancing early detection strategies and combating this devastating disease. Lung cancer survival has only marginally improved over the last several decades, but the availability of low-dose computer tomography screening and early detection, as well as advances in targeted treatments and immunotherapy, will likely reduce mortality rates and improve patient survival outcomes in the near future.

8. Future Enhancements

Ensemble methods with an extensive Deep Neural Network yield more accurate results for tumor prediction than individual classifiers such as SVM, Decision Trees, and Back Propagation Network. Assessing entity criteria and their relation to patient survival, especially in longer-lived individuals, is important for improved and prognostic semi-supervised machine learning algorithms.To improve performance in longer survival times, dividing the problem into subgroups based on survival duration can lead to better models. Focusing on individuals with shorter survival times may uncover important factors for accurate predictions in that subset. Evaluating clinically significant cutoffs and exploring additional variables beyond the ones considered can aid in predicting long-term survival. Further research should investigate different approaches and timeframes.

References

1. SharmilaNageswaran ,G. Arunkumar ,Anil Kumar Bisht ,ShivlalMewada ,J. N. V. R. Swarup Kumar , Malik Jawarneh ,and Evans Asenso "Lung Cancer Classification and Prediction Using Machine Learning and Image Processing" Volume 2022, Article ID 1755460 https://doi.org/10.1155/2022/1755460.
2. Imran NazirIhsanUlHaq , Muhammad Mohsin Khan , Muhammad Bilal Qureshi, Hayat Ullah and Sharjeel Butt "Efficient Preprocessing and Segmentation of Lung cancer detection using fused CT images". Electronics MDPI.https://doi.org/10.3390/electronics11010034.
3. MalayilShanid and Anitha A. "Lung Cancer detection using CT images using Salp Elephant optimization based Deep Learning" Biomedical Engineering: Applications, Basis and Communications, Vol. 32, No. 1 (2020) 2050001 (14 pages) DOI: 10.4015/S1016237220500015.
4. AnumMasooda, Bin Shenga, Ping Lic, XuhongHoud, XiaoerWeid, Jing Qine, Dagan Fengf "Computer Assisted Decision support system in pulmonary cancer detection and stage classification on CT images" Journal of Biomedical Informatics https://doi.org/10.1016/j.jbi.2018.01.005.
5. Atsushi Teramoto · Hiroshi Fujita "Fast Lung nodule detection in chest CT images using cylindrical nodule enhancement filter" DOI 10.1007/s11548-012-0767-5.
6. National Heart Lung and Blood Institute. How lungs work. https://www.nhlbi.nih.gov/.
7. Tanzila Saba, "Recent advancement in cancer detection using machine learning: systematic survey of decades, comparisons and challenges" http:// www.elsevier.com/locate/jiph.
8. G. Sutedja, New Techniques for Early Detection of Lung Cancer European Respiratory Journal ISSN 0904-1850, DOI: 10.1183/09031936.03.00405303.
9. Lung Cancer Detection Using Image Processing Techniques - MokhledAltarawneh Leonardo Electronic Journal of Practices and Technologies · August 2012.
10. Lung Cancer Screening Past, Present and Future By: James H. Finigan, Jeffrey A. Kern.
11. One Hundred Years of Lung CancerStephen G. Spiro and Gerard A. Silvestri*Department of Thoracic Medicine, Middlesex Hospital, London, United Kingdom; and Division of Pulmonary and Critical Care Medicine*.
12. Sangamithraa P. B., Govindaraju S. Lung tumour detection and classification using EK-mean clustering. Proceedings of the 2016 IEEE International Conference on Wireless Communications, Signal Processing and Networking.
13. Halder A., Kumar A. Active learning using Fuzzy-Rough Nearest Neighbor classifier for cancer prediction from microarray gene expression data. *Journal of Biomedical Informatics* . 2020;34(1):p. 2057001. doi: 10.1142/S0218001420570013.
14. Bhatia S., Sinha Y., Goel L. *Soft Computing for Problem Solving* . Singapore: Springer; 2019. Lung cancer detection: a deep learning approach; pp. 699–705.

15. Muthazhagan B, Ravi T, Rajinigirinath D. An enhanced computer-assisted lung cancer detection method using content-based image retrieval and data mining techniques. Journal of Ambient Intelligence and Humanized Computing. 2020 Jun 2:1-9.
16. Cancer Progress and Priorities: Lung Cancer, Matthew B Schabath, Michele L Cote. DOI: 10.1158/1055-9965.EPI-19-0221.
17. Syed, A.A., Fatima, W., Wajahat, R. (2018). Comparative analysis of learning algorithms for lung cancer identification. Indian Journal of Science and Technology, https://doi.org/10.17485/ijst/2018/v11i27/130707.
18. Dash, J.K., Mukhopadhyay, S., Garg, M.K., Prabhakar,N., Khandelwal, N. (2014). Multi-classifier frameworkfor lung tissue classification. 2014 IEEE Students'Technology Symposium, Kharagpur, India, pp. 264-269.https://doi.org/10.1109/TechSym.2014.6808058.
19. Günaydin, Ö., Günay, M., Şengel, Ö. (2019).Comparison of lung cancer detection algorithms. 2019Scientific Meeting on Electrical-Electronics andBiomedical Engineering and Computer Science, Istanbul,Turkey, pp. 1-4.https://doi.org/10.1109/EBBT.2019.8741826.
20. Makaju, S., Prasad, P.W.C., Alsadoon, A., Singh, A.K.,Elchouemi, A. (2018). Lung cancer detection using CTscan images. Procedia Computer Science, 125: 107-114.https://doi.org/10.1016/j.procs.2017.12.016.
21. Radhika, P.R., Nair, R.A.S., Veena, G. (2019). Acomparative study of lung cancer detection usingmachine learning algorithms. 2019 IEEE InternationalConference on Electrical, Computer and CommunicationTechnologies (ICECCT), Coimbatore, India, pp. 1-4.https://doi.org/10.1109/ICECCT.2019.8869001.
22. Roy, K., Chaudhury, S.S., Burman, M., Ganguly, A.,Dutta, C., Banik, R. (2019). A comparative study of lungcancer detection using supervised neural network. 2019International Conference on Opto-Electronics andApplied Optics (Optronix), Kolkata, India, pp. 1-5.https://doi.org/10.1109/OPTRONIX.2019.8862326.
23. Song, D., Zhukov, T.A., Markov, O., Qian, W., Tockman,M.S. (2012). Prognosis of stage I lung cancer patientsthrough quantitative analysis of centrosomal features 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, pp. 1607-1610. https://doi.org/10.1109/ISBI.2012.6235883.
24. Tripathi, P., Tyagi, S., Nath, M. (2019). A comparativeanalysis of segmentation techniques for lung cancerdetection. Pattern Recognition and Image Analysis, 29(1):167-173. https://doi.org/10.1134/S105466181901019X.