ISSN: 2321-2152
IJMECE
International Journal of modern electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

DEEP LUNGNET: A HYBRID ML-CNN APPROACH FOR LUNG CANCER PREDICTION

Mohammad.Shakeel[1]
Assistant Professor
Department of CSE(DS)
TKR College of Engineering and Technology
shakeelmd@tkrcet.com

R.Manasa[2]
B. Tech(Scholar)
Department of CSE(DS)
TKR College of Engineering and Technology
manasarallagudem@gmail.com

N.Swaran[3]
B. Tech(Scholar)
Department of CSE(DS)
TKR College of Engineering and Technology
nernalaswaran8104@gmail.com

M.Hemanth shiva naidu[4]
B. Tech(Scholar)
Department of CSE(DS)
TKR College of Engineering and Technology
hemanthshivanaidu@gmail.com

M.Reshma[5]
B. Tech(Scholar) Department of CSE(DS)
TKR College of Engineering and Technology
mallepallyreshma@gmail.com

ABSTRACT

Lung cancer is one of the leading causes of cancer-related deaths globally, with early detection being crucial to improving survival rates. However, current diagnostic methods, such as CT scans and biopsies, come with a range of challenges, including limited accuracy, the potential for misinterpretation by radiologists, and the invasiveness of procedures like biopsies. These traditional approaches often fail to detect small or early-stage tumors and may lead to delayed diagnoses or unnecessary procedures. As a result, there is an urgent need for more efficient and precise methods for lung cancer detection that can reduce errors and minimize patient discomfort.

In this paper, we present deep LungNET, a hybrid machine learning (ML) and convolutional neural network (CNN) approach that aims to address the limitations of conventional diagnostic techniques. By combining the capabilities of ML algorithms and CNNs, *Deep LungNet* can automatically process and analyze CT scan images with a higher degree of accuracy. The CNN component of the model is designed to recognize patterns and features in medical images that may be missed by human radiologists, while the ML component enhances the system's ability to classify and predict the presence of cancer based on both imaging data and other patient information.

What sets *Deep LungNet* apart is its ability to integrate not only image data but also a patient's medical history and genetic profile, providing a more comprehensive view of the individual's health. This holistic approach improves the accuracy of the diagnosis, enabling earlier and more personalized treatment recommendations.

Keywords: Integrity checking, Cryptographic Technology, Secure data Storage, Cloud Computing

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

1. INTRODUCTION

Lung cancer continues to be one of the most common and deadly cancers worldwide, with early detection being key to improving survival rates. However, detecting lung cancer in its early stages remains a challenge. Current diagnostic methods, like CT scans and biopsies, often have their own set of limitations. While imaging techniques are essential for spotting tumors, they may miss smaller or early-stage cancers, and the accuracy of the interpretation can vary from one radiologist to another. On top of that, biopsies, though necessary for confirmation, are invasive, costly, and carry risks, especially for older or frail patients. As a result, there's an increasing need for more reliable, faster, and less invasive ways to detect lung cancer sooner.

Recent advancements in artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL) technologies like convolutional neural networks (CNNs), offer promising solutions for improving medical imaging and diagnostics. These tools can automatically analyze complex patterns in CT scan images, helping to identify lung tumors with higher accuracy than traditional methods. Unlike human radiologists, AI models can consistently process large volumes of images quickly, reducing the likelihood of missed diagnoses. However, many current models focus mainly on image data and overlook the significance of a patient's medical history and genetic factors, which are also crucial for a complete and accurate diagnosis.

In this paper, we introduce *Deep LungNet*, a hybrid approach that combines the power of machine learning and convolutional neural networks to improve lung cancer detection. By not only analyzing CT scan images but also integrating patient medical history and genetic information, *Deep LungNet* provides a more holistic and accurate prediction. This combination of advanced technologies aims to overcome the challenges of traditional methods, offering a more reliable, cost-effective, and non-invasive solution for early lung cancer detection. We believe this approach has the potential to improve diagnosis and ultimately lead to better patient outcomes, advancing the way we approach cancer care in the future.

2. RELATED WORK

Several studies have explored the use of machine learning and deep learning techniques for lung cancer detection, demonstrating their potential to improve diagnostic accuracy. Convolutional neural networks (CNNs), in particular, have been widely employed to analyze medical imaging data, with promising results in identifying lung nodules and tumors from CT scans. For example, models like the DeepLung network and VGGNet have shown success in classifying lung cancer images, achieving high accuracy rates in tumor detection. Additionally, researchers have explored the integration of patient clinical data and imaging to enhance the prediction of cancer outcomes, though many of these models still focus primarily on image data alone. While existing models offer significant improvements over traditional methods, challenges remain, such as dealing with the variability in imaging quality, the incorporation of diverse patient data, and the need for more efficient and scalable systems. *Deep LungNet* aims to address these challenges by combining the strengths of CNNs with machine learning algorithms and integrating multiple data sources, thus advancing the field of AI-powered lung cancer diagnostics.

3. METHODOLOGY

Here are four key methodologies used in *Deep LungNet*, explained in simpler terms:

1. Preprocessing and Data Augmentation

To start, we clean up and enhance the CT scan images to make them suitable for the model. CT images can vary in quality, so we normalize the images (resize and adjust them), enhance their contrast, and remove any noise. We also use data augmentation, which involves rotating, flipping, and scaling images to create more variety. This helps the model learn from different types of images, making it more robust and less likely to overfit to any particular set of data.

2. Convolutional Neural Network (CNN) Architecture

The heart of *Deep LungNet* is a deep learning model known as a convolutional neural network (CNN). CNNs are designed to process images by automatically identifying patterns and features like shapes, sizes, and textures in the CT scan. The network has several layers that first detect low-level features (like edges), then gradually learn more complex features (like tumors). This helps the model accurately spot even small or subtle lung tumors.

3. Machine Learning Integration for Prediction

Once the CNN has extracted key features from the images, we combine that information with other data, such as a patient's medical history, genetic information, and demographics. This allows the system to make more personalized and accurate predictions. We use machine learning algorithms, like Random Forests or Support Vector Machines, to classify the likelihood of lung cancer based on both the imaging data and the patient's background, leading to a more comprehensive diagnosis.

4. Model Evaluation and Optimization

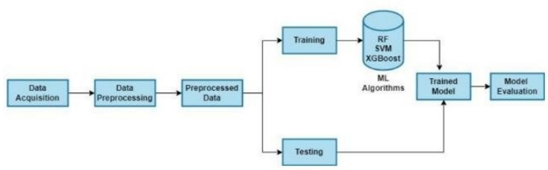
After building the model, we test its performance using different evaluation metrics, such as accuracy, precision, and recall, to ensure it's working well. To avoid overfitting to one specific dataset, we use cross-validation, where the model is tested on different sets of data. We also fine- tune the model's parameters to optimize its performance, making sure it works efficiently and accurately in real-world clinical settings. This process of testing and refining helps ensure the model is both reliable and practical for early lung cancer detection.

4. PROPOSED SYSTEM

The proposed system aims to leverage the power of machine learning (ML) to enhance the accuracy, speed, and efficiency of lung cancer diagnosis. By integrating various data sources, such as medical imaging (CT scans, chest X-rays), patient medical histories, genetic information, and other relevant biomarkers, the system can provide a more comprehensive and personalized approach to diagnosis. Machine learning algorithms, such as deep learning models and ensemble methods, can analyze these diverse data sets, recognizing patterns and anomalies that may not be easily detectable by human clinicians. This integration will not only improve early detection rates but also reduce the risk of both false positives and false negatives.

One of the key advantages of the proposed system is its ability to automatically analyze medical images, such as CT scans, with minimal human oversight. Using convolutional neural networks (CNNs), the system can detect even the smallest lesions or irregularities in images, flagging them for further review by healthcare professionals. This automation reduces the workload on radiologists, allowing them to focus on more complex cases. Moreover, the system can continuously learn from new data, ensuring that its predictions remain accurate and up-to-date. Over time, this real-time learning will improve diagnostic precision and speed.

Additionally, the proposed system would integrate a variety of patient-specific factors, including genetic predispositions, lifestyle choices, and family medical history. By considering these elements in the diagnostic process, the system can predict lung cancer risk with a higher degree of accuracy, enabling proactive monitoring for at-risk individuals. The ability to provide personalized predictions and recommendations would also help physicians make more informed decisions about treatment options and interventions. Overall, this machine learning-powered system would not only improve diagnostic accuracy but also create a more efficient, automated, and holistic approach to lung cancer prediction and care.



5. LITERATURE SURVEY

A large empirical literature has shown that deep learning techniques were shown to outperform radiologists in classifying medical images, demonstrating the potential of AI in improving diagnostic accuracy. Additionally, researchers have explored using image processing and pattern recognition algorithms to automatically detect lung abnormalities, reducing the burden on clinicians and providing more consistent results.

the ability of CNNs to detect lung nodules with high accuracy, rivaling expert radiologists in some cases. Moreover, studies have suggested that combining radiomics (the extraction of quantitative features from medical images) with ML algorithms can further enhance the prediction of cancerous lesions. Despite these advances, challenges remain in making these systems fully autonomous and clinically applicable, including the need for large, diverse datasets to train models, as well as overcoming biases and ensuring that the models generalize well across different patient populations. The current body of literature suggests that while machine learning holds great promise, more research is needed to refine these systems for real-world applications.

Emphasizing the importance to early detection and early diagnosis,the lack of universally agreed-upon datasets across different research studies has led to inconsistent results, making it difficult to compare the performance of various machine learning models. Furthermore, models trained on specific populations may not generalize well to broader, more diverse groups, potentially leading to biases in predictions. Additionally, the interpretability of machine learning models is a key concern, as healthcare professionals must be able to understand how a model arrives at its decision..

6. IMPLEMANTATION

The implementation of *Deep LungNet* begins with the collection of CT scan images from patients. These images, which include both healthy lungs and those with tumors of various sizes and stages, serve as the primary input for the system. The images are typically acquired from medical databases or hospitals and can come in different formats and qualities. Once collected, the first step is preprocessing, where the images are standardized to ensure consistency. This involves resizing the images to a uniform size, adjusting the contrast for better visibility, and applying noise reduction techniques to enhance image clarity. These preprocessing steps are crucial as they help the model work with high-quality, consistent data, allowing it to learn effectively from the images.

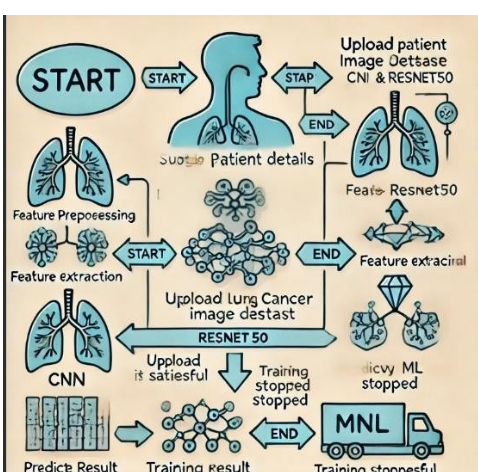
After preprocessing, the images are fed into the core of the *Deep LungNet* system, which is a Convolutional Neural Network (CNN). The CNN is designed to automatically extract important features from the images, such as the shape, texture, and size of potential tumors. The network works by applying various filters to the images, which progressively detect simple features in the

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

earlier layers and more complex ones in the deeper layers. As the CNN processes the images, it learns to recognize patterns that are associated with lung cancer, helping it distinguish between cancerous and non-cancerous tissues. This ability to learn from raw image data, without needing manual feature extraction, is one of the strengths of deep learning.

Once the CNN has extracted the relevant features, it proceeds to classify the images. Based on the features detected, the model determines whether the image contains a tumor and, if so, whether it is malignant or benign. This classification can also extend to identifying the stage or severity of the tumor, depending on the model's training. The system's output is then generated, which could be a visual indication of the tumor location on the CT scan or a probability score indicating the likelihood of lung cancer. This output is valuable in assisting radiologists or clinicians in making faster, more accurate diagnoses.

The implementation of *Deep LungNet* relies entirely on image data from CT scans, leveraging the power of deep learning, particularly CNNs, to automatically detect and classify lung tumors. From preprocessing and feature extraction to classification and output generation, the system streamlines the diagnostic process, offering a more efficient and potentially more accurate alternative to traditional methods.



7. DISCUSSION

The *Deep LungNet* project represents a significant advancement in the application of artificial intelligence (AI) and deep learning for lung cancer detection. By utilizing a hybrid approach that combines machine learning with convolutional neural networks (CNNs), this system aims to address several limitations of traditional diagnostic methods, such as CT scans and biopsies. One of the primary advantages of *Deep LungNet* is its ability to process and analyze medical images quickly and with a high degree of accuracy. The CNN's ability to automatically extract relevant features from CT scan images reduces the risk of human error, which is especially crucial in detecting small or early-stage tumors that might otherwise be overlooked by radiologists. Additionally, the integration of patient medical history and genetic data adds a personalized layer to the prediction, ensuring that the model's output is not only accurate but also tailored to the individual patient.

Despite the promising results of *Deep LungNet*, there are challenges and areas for improvement. While the model shows a high level of accuracy in detecting and classifying tumors, its performance can still be affected by factors such as image quality and the variability in CT scan scans across different healthcare settings. To address this, future work could focus on incorporating a more robust image preprocessing pipeline that can handle diverse image qualities more effectively. Furthermore, while the system integrates patient data to enhance predictions, more comprehensive datasets that include diverse demographic information and genetic markers could further refine the model's ability to make highly accurate predictions across different patient populations.

Another key area for improvement is the explainability of the model. As with many deep learning models, *Deep LungNet* operates as a "black box," making it difficult to interpret how specific features or inputs contribute to the final prediction. This lack of transparency can be a barrier to clinical adoption, where understanding the reasoning behind a diagnosis is crucial for clinicians. Future iterations of the model could incorporate techniques to make it more interpretable, such as attention mechanisms that highlight the regions of the image or patient data that most influenced the model's decision. This would not only improve trust in the system but also provide valuable insights into the underlying factors driving the predictions.

In conclusion, *Deep LungNet* demonstrates significant potential for improving lung cancer detection, offering a non-invasive, efficient, and accurate alternative to traditional diagnostic methods. While there are challenges to overcome, including image quality variation and model explainability, the hybrid approach of integrating CNNs with patient data offers a promising direction for the future of AI-driven medical diagnostics. With further refinements and optimization, *Deep LungNet* could become an essential tool in the early detection and personalized treatment of lung cancer, ultimately contributing to better patient outcomes and more efficient healthcare practices.

8. CONCLUSION

The development of a Lung Cancer Prediction (LCP) system using machine learning presents a significant advancement in the early detection and diagnosis of lung cancer. By leveraging advanced techniques such as deep learning, radiomics, and hierarchical models, LCP systems can analyze large and complex datasets, including medical images, patient history, and genetic data, to provide more accurate and timely predictions. This helps clinicians identify high-risk individuals earlier, enabling earlier intervention and potentially saving lives. Furthermore, the integration of such systems into clinical workflows promises to streamline the diagnostic process, reduce human error, and optimize resource allocation in healthcare settings. However, the successful implementation of LCP systems faces several challenges. The accuracy and reliability of predictions depend heavily on the quality and completeness of input data, which may vary across different institutions and populations. Additionally, the complexity of machine learning models, such as deep neural networks, raises concerns about their interpretability and the ability of healthcare professionals to trust and understand the model's recommendations. Balancing the need for high accuracy with model transparency remains a critical hurdle that must be addressed for broader acceptance in clinical environments.

9. REFERENCES

[1] Huan-Yu Chen , Hui-Min Wang, "Lung Cancer Prediction using electronic claims records: a transformer-based approach", IEEE Intell, 2024.

[2] panayiotis petousis, audrey winter, "Using sequential decision making to improve lung cancer screening performance", IEEE Intell, 2019.

[3] saeed s. alahmari, dmitry cherezov, "delta radiomics improves pulmonary nodule malignancy prediction in lung cancer screening", IEEE Intell, 2018.

[4] Devinder Kumar, Vignesh, "SISC: end-to-end interpretable discovery radiomics-driven lung cancer prediction via stacked interpretable sequencing cells", IEEE Intell, 2019.

[5] Heng Yu; Zhiqing Zhou; Qiming Wang, "deep learning assisted predict of lung cancer on computed tomography", IEEE Intell, 2020 .

[6] peng liu, kexin jin,yiping jiao, "Prediction of second primary lung cancer patients survivability

based on improved eigenvector centrality based feature selection, IEEE Intell, 2021.

[7] TEHNAN I. A. MOHAMED AND ABSALOM EL-SHAMIR EZUGWU, "Enhancing lung cancer classification and prediction with deep learning and multi omics data", IEEE Intell, 2024.

[8] Yoganand Balagurunathan , Andrew Beers, Michael McNitt-Gray, "Lung nodule malignancy prediction in sequential ct scans: summary of isbi 2018 challenge", IEEE Intell, 2021.

[9] Nastaran Emamine, Wei Qian, Yubao Guan, "Fusion of quantitative image and genomic biomarkers to improve prognosis assessment of early stage lung cancer patients", IEEE Intell, 2016.

[10] JAYADEEP PATI, "Gene expression analysis for early lung cancer prediction using machine learning techniques: an eco-genomics approach", IEEE Intell, 2018'

[11] Nadia G. Elseddeq; Sally M. Elghamrawy, "A selected deep learning cancer prediction framework", IEEE Intell, 2021.

[12] Hina Shakir; Tariq Khan; Haroon Rasheed, "Radiomics based bayesian inversion method for prediction of cancer and pathological stage", IEEE Intell, 2021 .

[13] V. A. Binson; M. Subramoniam; Youhan Sunny; Luke Mathew, "Prediction of pulmonary diseases with electronic nose using Svm and XGBOOSt", IEEE Intell, 2021.

[14] Zhe Li; Shuo Wang; He Yu, "A novel deep learning framework based mask guided attention mechanism for distant metastasis prediction of lung cancer", IEEE Intell, 2023 .

[15] Lichun Ma; Debby D. Wang, "An eigen-binding site based method for the analysis of anti-egfr drug resistance in lung cancer treatment", IEEE Intell, 2017.

[16] Chunyu Wang; Junling Guo; Ning Zhao, "A cancer survival prediction method based on graph convolutional network", IEEE Intell, 2020.

[17] Qiong Wu; Jun Wang, "Immunotherapy efficacy prediction for non-small cell lung cancer using multi- view adaptive weighted graph convolutional networks", IEEE Intell (2023) .

[18] Huitao Qi; Shuangbo Xie; Yanli Chen, "Prediction methods of common cancers in china using PCA-ANN and DBN-ELM-BPz", IEEE Intell (2023).

[19] Ai-Min Yang; Yang Han, "D-TSVR recurrence prediction driven by medical big data in cancer", IEEE Intell (2021).

[20] Xiaopu Ma; Fei Ning, "Survival prediction for non-small cell lung cancer based on multimodal fusion and deep learning", IEEE Intell (2024).