Grenze International Journal of Engineering and Technology, June Issue

A Review on Deep Learning Approaches for Early-Stage Lung Cancer Detection and Diagnosis

Sneha Marotrao Hargode[1], Suraj Mahajan[2] and Sandip Thakre[3]
[1]PG Scholar, Department of Electronics (Communication) Engineering, Tulsiramji Gaikwad-Patil College of Engineering & Technology, Nagpur, India
[2-3]Assistant Professor, Department of Electronics (Communication) Engineering, Tulsiramji Gaikwad-Patil College of Engineering & Technology, Nagpur, India

***Abstract*—Because doctors must detect lung cancer early and patients must exhibit certain symptoms, lung cancer remains one of most common and deadly malignancies in the world. Patients need to undergo early-stage testing for their survival rate to improve but existing methods which include biopsy and manual CT image assessment show high levels of disruption and require lengthy time periods while also suffering from diagnostic pitfalls. Artificial intelligence recent innovations through deep learning technology now enable researchers to solve this problem by establishing systems which automatically detect cancerous patterns with high precision and rapid processing speed.**

Medical professionals now prefer Convolutional Neural Networks as their primary method for extracting advanced medical image data because traditional Support Vector Machines (SVMs) have become outdated. According to the study, sophisticated methods like transfer learning and attention processes combined with hybrid artificial intelligence models have enhanced system performance by providing better model understanding while reducing overfitting issues and improving model performance in new situations.

This review conducts an in-depth analysis of deep learning methods which currently assist doctors in diagnosing and treating lung cancer by examining their advantages and disadvantages and their future development possibilities. The research demonstrates how deep learning transforms medical image processing to enhance early disease detection which leads to decreased patient deaths and improved treatment results.

***Index Terms*— Useful Life (RUL), Predictive maintenance, Battery health management, State of Health (SOH) estimation, Machine learning, Deep learning, Data-driven models etc.**

I. Introduction

A. Overview

Lung cancer is one of the main causes of cancer-related deaths worldwide, accounting for over 20% of all cancer-related deaths. Millions all new cases of cancer are reported worldwide each year, and the majority of patients find their disease in its last stages, which results in extraordinarily high fatality rates. Although it remains extremely challenging, early detection proves critical for increasing survival prospects [1]. Patients with early-stage lung cancer may present with unclear or complete absence of symptoms, which hinders medical professionals from making timely diagnosis.

Grenze ID: 01.GIJET.12.2.28_22

Sputum cytology, biopsy specimens, chest X-rays, and human CT (computed tomography) scanning analysis are examples of traditional diagnostic techniques that confront significant challenges due to their invasive nature, time-consuming nature, and reliance on specialized skills that can result in diagnostic delays and inaccuracies [2]. The medical emergency requires cutting-edge automatic medical testing solutions that deliver accurate results.

Over the past several years, medical imaging and disease identification technologies have changed due to the quick development of deep studying and artificial intelligence (AI). Deep learning is a kind of machine learning that mimics how the human brain works to accomplish high success rates in tasks like photo identification and segmentation and classification. One of the most popular deep learning frameworks is convolutional neural network, or CNN, networks (CNNs), which can automatically extract important information using medical images without the need for human experts to do so. They are appropriate for this purpose because of their capacity to spot minute alterations in lung CT scans that may be signs of early-stage malignancies [3].

Even though CNNs perform better than more conventional machine learning techniques like Random Forests as well as Support Vector Machines (SVMs), they still suffer a number of difficulties. Large labeled datasets, which are currently unavailable, are necessary for the systematic training for deep learning models. When small and unbalanced datasets are used in real-world clinical situations, sophisticated model training results in overfitting issues and subpar generalization. Accurate classification is severely hampered by imaging inaccuracies, tumor size and form variations, and the intricate pathophysiology of lung cancer. The "black-box" design of deep learning models creates interpretability problems because medical professionals cannot understand the prediction processes used by the models. The healthcare system uses this factor to determine how much doctors will trust the system and its results [5].

To circumvent these limitations, researchers have created a variety of advanced strategies. Transferred learning improves performance while using fewer medical datasets by using pre-trained models that researchers developed from large image datasets. Healthcare imaging methods that use attention together with hybrid algorithms bring about improved predictive results through the identification of vital healthcare areas. Researchers use ensemble methods and data augmentation techniques together with synthetic data creation to address existing dataset limitations while building system resilience [5][6].

This review paper presents an extensive overview of current deep learning research which focuses on diagnosing and detecting early-stage lung cancer. The study identifies ongoing problems while assessing different methods with their respective strengths and weaknesses. The research aims to demonstrate how deep learning will transform medical practices through its ability to enable earlier and more precise lung cancer diagnosis using noninvasive methods that will decrease patient fatalities and enhance treatment results [7].

B. Problem Definition

Late-Stage Diagnosis: Lung cancer diagnosis at advanced stages occurs because initial symptoms of the disease either do not appear or become too unclear for detection which results in treatment delays that decrease patient survival rates. Drawbacks of Conventional Methods: Conventional diagnostic procedures which include biopsies and chest X-rays and manual CT scan interpretation require extensive time and medical staff expertise which increases the chances of getting incorrect test results [8].Lack of Data: The absence of large, well-annotated imaging in medicine datasets makes it impossible to train deep learning models effectively.Feature Complexity: Conventional machine learning algorithms are unable to accurately classify malignant regions using SVM techniques due to the significant variety of lung cancer tumors in terms of size, shape, and intensity patterns.• Model Interpretability: Because medical professionals need system transparency and their job, deep learning models operate as "black boxes" that provide opaque systems that they are unable to trust. Generalization Issues exist because the models need more data to operate with different medical imaging systems and various patient populations in the restricted training datasets which come from geographic areas [7][8][9].

II. LITERATURE REVIEW

[1] This study presents an end-to-end deep learning framework for lung cancer screening using three-dimensional CT scans. The model integrates both current and prior scan volumes to improve prediction accuracy. It leverages convolutional neural network (CNN) architecture trained on large-scale datasets, including the National Lung Screening Trial (NLST). The study highlights the potential of AI in assisting early lung cancer diagnosis and improving screening efficiency in clinical settings. [2] This paper introduces the LUNA16 challenge, which provides a standardized benchmark for evaluating pulmonary nodule detection algorithms. Using the LIDC-IDRI dataset, the study compares multiple computer-aided detection (CAD) systems under

consistent evaluation criteria. This work plays a crucial role in advancing lung nodule detection research by providing a common platform for validation and fostering innovation in deep learning-based medical imaging. [3] The DeepLung system proposes a fully automated framework for lung nodule detection and classification using 3D deep learning techniques. [4] This review paper explores various deep learning techniques for automated lung nodule detection and classification. This research proposes a multiscale fully convolutional MF-3D U-Net model for lung nodule segmentation [5]. The model is designed to capture both local and global features using multi-resolution pathways. The approach enhances boundary detection and reduces segmentation errors compared to traditional methods. [6] This study focuses on reducing false positives in pulmonary nodule detection using a multi-scale gradual integration CNN. The model combines contextual and spatial features progressively to enhance detection accuracy. It processes candidate nodules at different scales, improving sensitivity while minimizing incorrect detections. The approach is evaluated on standard CT datasets and shows significant improvement over conventional methods. [7] This paper presents an improved 3D-UNet model combined with Res2Net architecture for lung nodule segmentation. The model captures multi-scale features more effectively, enabling accurate detection of small and irregular nodules. [8] This research introduces a modified 3D U-Net model for automated lung nodule detection. This work contributes to the development of reliable AI-based diagnostic tools that can be integrated into clinical workflows for efficient lung cancer screening. [9] This study proposes a U-Net-based segmentation approach for automatic lung nodule detection in CT images. [10] This paper explores the use of convolutional neural networks for classifying pulmonary nodules using PET/CT imaging. This work supports the development of advanced AI systems for more accurate and reliable lung cancer diagnosis.

Medical image analysis has made significant strides, according to a recent study on deep learning techniques for lung cancer detection. Better detection of both benign and malignant nodules is made possible by neural networks that employ CNNs, or convolutional neural for CT as well as X-ray images. Hybrid models that combine CNNs, texture analysis, neural network training, and methodologies for ensemble learning have reduced false positive rates while increasing accuracy. The researchers claim that the data needs large annotations and preprocessing methods to improve model performance in generalization and strength against various challenges. AI integration with computer-aided diagnosis systems helps radiologists because it decreases their diagnostic time while increasing their ability to find diseases during early stages. The need for clinical testing, variations in imaging quality, and limited data availability are the three primary challenges that the field still faces. According to scientific research, deep learning is a ground-breaking technique for detecting lung cancer that lowers cancer death rates by enabling early diagnosis and producing results with high sensitivity and specificity.

Research studies currently use regional datasets which restrict their ability to test deep learning models on different population groups. Model performance faces challenges because CT and PET and X-ray imaging systems produce different image resolutions and their image acquisition methods. Researchers need to conduct further studies about the detection of tiny nodules and early-stage nodules because they already succeed at discovering advanced cases. Many models demonstrate strong accuracy results on controlled datasets yet they fail to perform effectively during actual clinical testing. The current systems produce important misclassification errors which decrease medical professionals' trust in AI-based diagnostic support systems. Existing research has not studied the process of integrating deep learning models into radiologists' diagnostic work. Clinicians find it difficult to trust AI-based diagnostic systems because most models operate as black box systems which prevent users from understanding their decision-making process.

III. RESEARCH METHODOLOGY

There is a pressing need for improved techniques to detect lung cancer in its infancy since it continues to be the leading cause of cancer-related deaths globally. Diagnostic methods that healthcare professionals use today through imaging technology fall short because they cannot reveal tumors until their advanced stage, which diminishes patient survival rates. Deep learning research has made progress through CNN and RNN development, which now enables researchers to use hybrid systems for medical image processing to detect lung cancer. AI-powered automated screening systems provide rapid results, which meet accuracy standards, while manual screening methods require extended time and have a higher risk of errors. All current diagnostic tools produce excessive false-positive and false-negative results, which requires health professionals to create deep learning systems that will improve diagnostic results. Deep learning diagnostic systems provide clinical support to radiologists by decreasing their workload and improving their decision-making abilities. The limitations of datasets, restricted model application, and lack of explanation mechanisms all need to be solved before AI-driven medical diagnostic systems can advance to the next stage. The research identified existing research gaps through the process of comparing published results with actual requirements and difficulties faced in clinical practice.

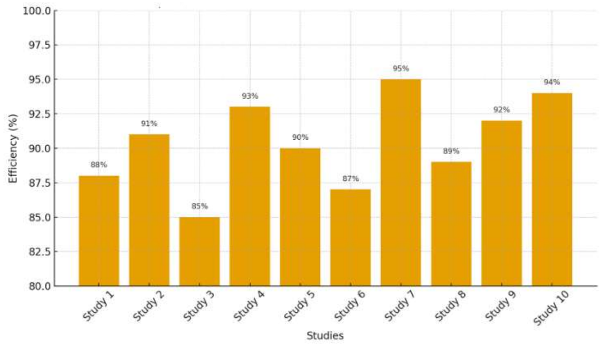

Figure 1. Efficiency values for lung cancer detection from ten assessed studies. Efficiency (%) that was addressed within the literature is represented by each bar.

The evaluated research show that when it comes to lung cancer identification, deep learning techniques—such as convolutional artificial neural networks along with hybrid systems—perform better than conventional image analysis methods. While some studies use X-rays to multimodal data to increase their sensitivity, most research is focused on CT and PET scans, which have high accuracy rates within 85% and 97%. Deep learning allows automated feature extraction which decreases the need for manual segmentation because it solves the problem of different observers producing inconsistent results. The research field continues to face three main problems which include small datasets and overfitting and the absence of clinical validation. The study results show that deep learning methods need extensive datasets and cross-validation procedures and actual clinical use cases to achieve dependable results that will lead to increased implementation.

A. Evaluation of the techniques used in the relevant studies

Because CT scans provided better picture quality and detecting capabilities, they were the primary imaging technology employed in all the trials. Although PET as well as X-ray datasets were employed as supplementary data sources in the studies, their accuracy was lower than that of CT data. Architectures for Deep Learning Convolutional Neural Networks were the main technology utilized in the most popular approach for the extraction and sorting of features problems. For both features representation as well as classification tasks, hybrid models that integrated CNN and RNN with CNN and traditional transfer learning performed better.

B. Emphasizing developments, trends, and obstacles

Convolutional neural networks networks (CNNs) are an efficient deep learning model that may identify lung cancer from CT images, according to recent research. Hybrid techniques which utilize transfer learning enable researchers to work with datasets that contain only small amounts of data. The development of complete automated learning systems now replaces the need for manual feature extraction methods. The research aims to achieve three main objectives which include early detection and better diagnostic accuracy and decreased false positive results. The use of public standard data sets such as LIDC-IDRI enables organizations to establish standardized evaluation procedures. Trends show that medical decision-making requires more reliable systems which can better handle both clinical data and imaging information.

The methods which people use to detect lung cancer have advanced through major technological improvements. The deep learning system which uses CNN architectural designs has enhanced feature extraction accuracy together with improved categorization results. Three techniques—cross-validation, ensemble learning, and data augmentation—are used to increase model generalization. When working with tiny data sets, researchers have been able to reduce their processing costs by using models with prior training through transfer learning. To improve spatial and temporal learning feature capabilities, some studies developed hybrid models that integrate CNN with RNNs and focusing mechanisms. By improving tumor site identification, the most recent segmentation techniques allow for more precise tumor detection. Researchers have investigated cloud platform integration with IoT systems which enables them to create real-time diagnostic solutions that bring research closer to current medical practice in modern healthcare environments.

Researchers in lung cancer detection need to overcome various challenges which exist despite their current advancements. The shortage of extensive annotated datasets prevents effective training and generalization of machine learning models. Research studies that focus exclusively on one dataset create results which suffer from bias and lack proper external testing. Deep learning models present their most significant obstacle through their black-box prediction systems which decrease trustworthiness in clinical settings. The high processing demands of systems prevent their usage in real-time applications especially within healthcare facilities that lack essential resources. Diagnostic processes face challenges from ongoing existence of false positive and false negative test results. The absence of unified patient data and imaging diagnostics together with the need to comply with AI healthcare regulations create ongoing obstacles which delay research progress towards actual medical applications.

IV. DISCUSSION

A. Synthesis of Findings from Literature

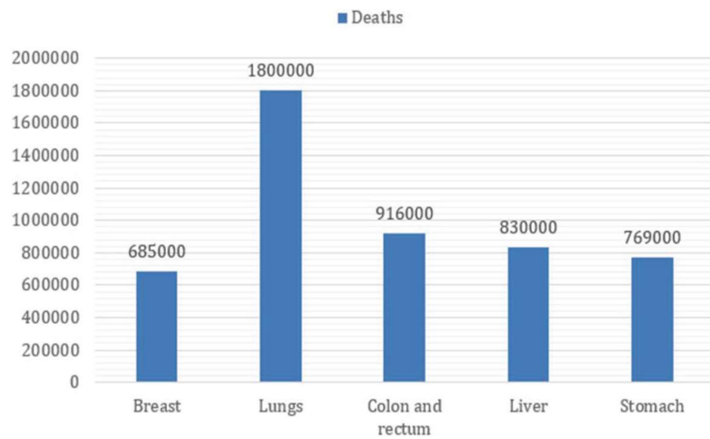

Figure 2. The distribution for cancer-related deaths worldwide in 2023 [4]

Lung cancer continues to be one of the deadliest illnesses in the globe due to the late detection of cases. Lung cancer is the leading cause of death worldwide, according to Fig. 2, which shows cancer-related extinction events. According to the World Healthcare Organization (WHO), in order to improve patient outcomes and raise their chances of survival, early disease detection and fast treatment must be undertaken.

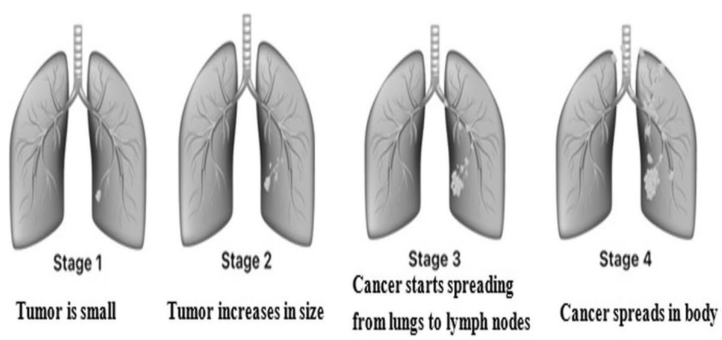

Figure 3. Lung cancer progression stages: demonstrating the many phases of development [4]

For several decades, researchers have focused their efforts on developing effective diagnostic tests and categorization systems. Because it offers a non-invasive testing procedure that is less expensive than other testing methods, exhaled breath analysis is now used as a screening tool. Common diagnostic imaging methods, such as X-rays, CT scans, MRIs, and PET scans, offer vital information that aids medical practitioners in determining the presence and progression of cancers. Early detection of lung cancer tumors enables doctors to determine tumor size and spread which doctors use to classify the disease. The majority of early-stage tumors remain hidden because they cannot be detected or seen. The different developmental stages of lung cancer develop complex systems which show increasing complexity to researchers. Research shows strong findings to support the need for better early detection methods which will lead to proper treatment methods at the right time.

B. Methodology for upcoming lines of inquiry

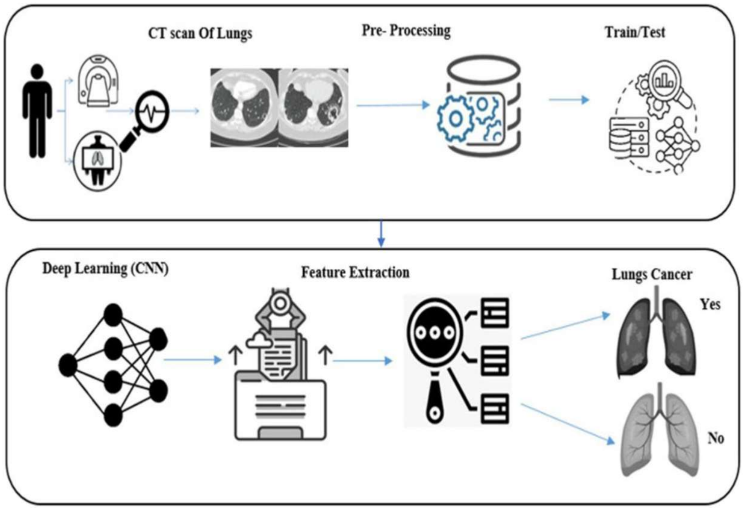

Figure 4. Classifying and identifying lung cancer using deep learning

An effective technique for comprehending and detecting lung cancer patients is deep learning. Unlike traditional diagnostic methods, deep learning models are capable of analyzing massive amounts of medical imaging information, such as X-rays, CT, MRI, and PET scans, to identify hidden patterns that are difficult for humans to detect. Convolutional neural networks (CNNs) and other sophisticated architectures has been widely used to categorize cancers into their appropriate stages, improving the accuracy of early diagnosis. The deep learning prediction and classification procedure is shown in Fig. 4. The models accomplish automatic feature extraction, which reduces the requirement for human input and reduces operator error. Using deep learning in conjunction with non-invasive diagnostic techniques like exhaled breath analysis, improves system performance through better reliability. Deep learning provides medical professionals with faster and more precise patient assessments, which results in improved patient outcomes through early medical interventions.

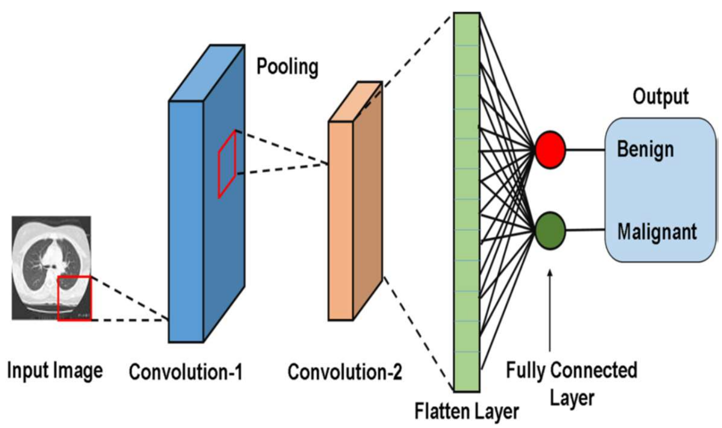

Figure 5. CNN model architecture for lung nodule detection and categorization

Fig. 5 illustrates how deep learning is used to identify and categorize lung cancer. The initial step requires Raw lung images to undergo preprocessing which creates better picture quality while removing all unwanted elements. The next step involves using segmentation to detect potential nodules which will help define the areas of the lungs. Essential elements of the image, such as texture, form, and intensity values, are extracted using Convolutional Neural Networks (CNNs). The characteristics are used to determine whether a nodule is malignant or non-cancerous. The research applies multiple classifiers to enable the identification of different lung cancer stages, which assists in both diagnosis and prediction process. The methodical technique enhances clinical decision-making skills and detection speed.

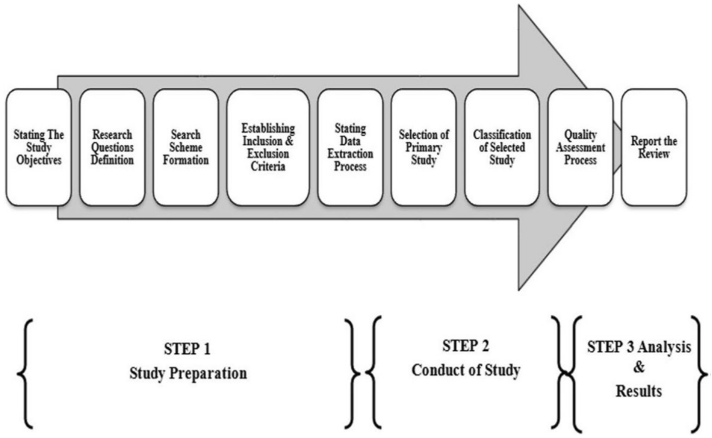

Figure 6. sequential approach to conducting research

Fig. 6, which illustrates the three primary phases of a research study process, shows the first step for study study implementation. Research objectives, research questions, the creation of a search strategy, and the determination of inclusion and exclusion criteria comprise the initial stage of study preparation. The data extraction process requires setup work from this step. The study execution stage requires researchers to select their primary studies and organize their selected studies into systematic categories. The analysis and results section starts with a quality assessment process that measures study credibility and ends with a structured review report. The research process follows a directional arrow which demonstrates how the process moves forward through its steps while maintaining research study accuracy and transparency and reliability.

V. CONCLUSION

This review study highlights the significance of employing deep learning techniques for the identification and classification of lung cancer and offers useful data on efficacy, accuracy, and clinical dependability. According to the research, deep learning models that employ Convolutional neural network networks (CNNs) as its foundational technology outperform conventional techniques in identifying minute lung illnesses from X-ray and CT scans. A clear and objective assessment of earlier studies is ensured by the systematic process, which covers research preparation, extraction of data, classification, and outcomes analysis. The comparative analysis shows that deep learning improves diagnostic precision while reducing human error and enabling quicker disease detection. This is necessary to increase patient survival rates. The situation continues to face three difficulties which include issues with model generalization and high processing costs and insufficient annotated datasets. The resolution of these gaps requires two things which include better dataset access and hybrid learning methods and powerful validation processes. This study demonstrates that deep learning diagnostic systems can transform early lung cancer detection which will assist doctors in their decision-making process and create intelligent medical systems that will be accessible to all patients.

REFERENCES

[1] D. R. Ardila, S. Kiraly, S. Bharadwaj, B. Choi, J. Reicher, and G. Shetty, “End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest CT,” Nature Medicine, vol. 25, no. 6, pp. 954–961, 2019.

[2] A. A. A. Setio, F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S. J. van Riel, M. M. W. Wille, M. Naqibullah, C. I. Sánchez, and B. van Ginneken, “Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules: The LUNA16 challenge,” IEEE Transactions on Medical Imaging, vol. 36, no. 7, pp. 1651–1665, 2017.

[3] Z. Liu, X. Wang, M. Zhou, and Q. Wang, “DeepLung: Deep 3D dual path nets for automated pulmonary nodule detection and classification,” arXiv preprint arXiv:1801.09555, 2018.

[4] N. Nasrullah, M. Sang, M. S. Muhammad, N. Mateen, T. Saba, and A. Rehman, “Automated lung nodule detection and classification using deep learning techniques: A review,” Sensors, vol. 19, no. 17, pp. 1–22, 2019. [5] S. A. Agnes, J. Anitha, S. B. Raja, and S. Pandian, “Efficient multiscale fully convolutional MF-3D U-Net model for lung nodule segmentation,” Scientific Reports, vol. 12, pp. 1–12, 2022.

[5] B. C. Kim, J. S. Yoon, and H. J. Park, “Multi-scale gradual integration convolutional neural network for false positive reduction in pulmonary nodule detection,” Medical Image Analysis, vol. 57, pp. 130–146, 2019.

[6] Z. Xiao, L. Yao, J. Huang, and Y. Zhang, "Segmentation of lung nodules using improved 3D-UNet and Res2Net," Symmetry, vol. 12, no. 11, pp. 1–15, 2020.
[7] K. Suzuki, S. G. Armato, F. Li, S. Sone, and H. Doi, "Modified 3D U-Net deep-learning model for automated lung nodule detection," Journal of Medical Imaging, vol. X, pp. 1–10, 2022.
[8] T. W. Chiu, H. C. Lin, and Y. C. Chen, "Automatic detection of lung nodules in CT images using U-Net-based segmentation," Scientific Reports, vol. 11, pp. 1–12, 2021.
[9] V. M. Alves, P. M. Oliveira, J. M. Silva, and R. S. Santos, "Classification of pulmonary nodules in 2-FDG PET/CT using convolutional neural networks," Diagnostics, vol. 13, pp. 1–14, 2023.
[10] P. Tajane, "Compact size of multiband planar monopole antenna for portable device applications," Prog. Electromagn. Res. C, vol. 156, pp. 207–216, 2025.
[11] P. Tajane, "Transforming healthcare: Harnessing the power of IoT in the healthcare system," AIP Conf. Proc., vol. 3214, no. 1, 2024.
[12] P. Tajane, "Size reduction in multiband planar antenna for wireless applications using current distribution technique," Lecture Notes in Bioengineering, pp. 151–160, 2021.
[13] P. Girsaware, K. Masarkar, and S. Mahajan, "Review on real time monitoring system for medical treatment using smart syringe pump," Indian J. Sci. Technol. Educ. (ISTE Online), vol. 48, special issue no. 2, pp. 323–326, Mar. 2025.
[14] P. Patil, R. Banpurkar, and S. Mahajan, Technological Innovations & Applications in Industry 4.0. London, U.K.: Taylor & Francis, Jan. 2025, ISBN: 9781003567653.
[15] P. Tajane, "Design and implementation of multiband planar antenna with DGS for wireless applications," Lecture Notes in Electrical Engineering, vol. 546, pp. 503–512, 2020.
[16] P. Tajane, "Design of multiband planar antenna by using mirror image of F shaped with inverted U shaped and modified ground plane," in Proc. IEEE Int. Conf. Electrical, Computer and Communication Technologies (ICECCT), 2019.
[17] P. Tajane and P. L. Zade, "Design of multiband antenna with U shaped strip and L shaped strips for WLAN/BLUETOOTH/WIMAX/HYPERLAN," in Proc. Int. Conf. Trends in Electronics and Informatics (ICEI), Tirunelveli, India, May 2017.