https://doi.org/10.46335/IJIES.2024.9.6.21 e-ISSN: 2456-3463
Vol. 9, No. 6, 2024, PP. 97-106

International Journal of Innovations in Engineering and Science, www.ijies.net

A Review of Intelligent Systems for Lung Cancer Diagnosis Covering Imaging Modalities, Deep Learning Models and Clinical Applications

Jaya Shrivastav[1], Dr. Ranu Pandey[2]

[1]Research Scholar, Department of Computer Science, Shri Rawatpura Sarkar University, Raipur Chhattisgarh, India
[1]shrijay15@gmail.com
[2]Assistant Professor, Department of Computer Science, Shri Rawatpura Sarkar University, Raipur Chhattisgarh, India
[2]ranu_pandey8@hotmail.com

***Received on**: 5 May,2024* ***Revised on**: 29 June,2024* ***Published on**: 02 July ,2024*

***Abstract –** Lung cancer continues to be one of the most lethal malignancies worldwide, primarily due to its late-stage detection and the complexity of accurate diagnosis using traditional methods. Early diagnosis significantly enhances treatment outcomes, yet current clinical practices, which often rely on the manifestation of symptoms and manual evaluation of imaging, are time-consuming and prone to human error. With the proliferation of advanced imaging modalities such as Computed Tomography (CT), Low-Dose CT (LDCT), Positron Emission Tomography (PET), and Magnetic Resonance Imaging (MRI), automated systems have emerged to assist in the identification and classification of lung cancer. This review provides a comprehensive examination of the latest developments in automated lung cancer detection, with a focus on machine learning and deep learning approaches, including Convolutional Neural Networks (CNNs), 3D CNNs, Capsule Networks, and hybrid models. It discusses essential pipeline components such as image preprocessing, region segmentation, feature extraction and selection, and classification. Key datasets utilized for training and validation are also outlined, highlighting their role in benchmarking diagnostic models. Moreover, the review identifies critical research gaps, including inadequate handling of low-resolution images, high computational complexity, lack of interpretability, and insufficient integration of diagnostic stages. Recent efforts toward improving segmentation accuracy, reducing false positives, and enhancing model generalizability are also explored. The review concludes by outlining potential future research directions, emphasizing the need for standardized imaging protocols, interpretability through explainable AI, resource-efficient architectures, and privacy-preserving methods such as federated learning. Collectively, these insights aim to bridge the gap between experimental diagnostic systems and their practical deployment in clinical environments, thereby facilitating more reliable and early detection of lung cancer.*

***Keywords-** Artificial Intelligence, Computed Tomography, Convolutional Neural Networks, Deep Learning, Low-Dose CT, Lung Cancer, Machine Learning, Magnetic Resonance Imaging, Positron Emission Tomography, Segmentation.*

INTRODUCTION

Early diagnosis of lung cancer significantly improves treatment outcomes and patient survival rates [1]. Traditionally, lung cancer is identified through the observation of clinical symptoms such as coughing up blood, chest pain, shortness of breath, fatigue, unexplained weight loss, memory loss, bone fractures, joint pain, headaches, neurological issues, bleeding, facial swelling, voice changes, and discoloration of sputum [2].

Upon the manifestation of such symptoms, patients typically undergo a range of diagnostic screenings including genetic testing, bronchoscopy, reflex testing, fluid biopsy, standard biopsy, and blood tests [3]. These methods are widely supported by national healthcare guidelines which aim to standardize and improve the accuracy of lung cancer staging and diagnosis.

Among these techniques, Computerized Tomography (CT) has emerged as one of the most reliable screening tools. CT scans utilize X-rays to examine internal organs and tissue abnormalities over a 30-minute scan period, offering more detailed insights compared to PET and MRI scans [4].

With the growing availability of CT images, researchers have developed automatic lung cancer detection systems to assist in identifying disease more efficiently [5]. These systems generally follow a multi-step pipeline: image preprocessing (noise removal), region segmentation, cancer feature extraction, feature selection, and classification [6].

Among these steps, region segmentation is particularly crucial as it isolates the abnormal areas for further analysis. Accurate segmentation enables the extraction of meaningful features, thereby reducing system complexity and enhancing diagnostic precision. Subsequently, feature selection techniques are employed to filter out redundant data, which helps lower computation time and prevents overfitting [7].

A variety of segmentation algorithms are utilized to process X-ray and CT images, including k-means clustering, distributed clustering, Canny and Sobel edge detection, fuzzy c-means, fuzzy k-means, self-organizing maps, and Hopfield neural networks [8]. These methods help in identifying critical regions indicative of malignancy.

Following segmentation, feature selection is applied using optimization algorithms such as wrapper methods, ant colony optimization, particle swarm optimization, genetic algorithms, firefly algorithms, and bacterial foraging optimization [9]. The selected features are then classified using machine learning models such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and other intelligent classifiers. While these traditional systems show promising results in lung cancer prediction, they still face notable challenges. These include lower recognition accuracy when handling large datasets, high processing time, and poor performance with low-quality CT images, which can lead to false feature detection and increased misclassification rates [10-11].

LITERATURE REVIEW

As a result, various researchers have proposed enhancements and new frameworks to address these limitations, contributing valuable insights for the development of more intelligent and robust lung cancer prediction systems.

Here is a tabular literature review summarizing various methods, outcomes, and limitations in the field of automatic lung cancer detection:

Table 1- Comparative Summary of Recent Approaches in Automated Lung Cancer Detection

Study / Method	Approach / Technique	Outcomes	Limitations
Handcrafted Methods [12]	Domain-specific, manual feature extraction	Simple and interpretable features	High false negatives, poor generalization, missing true labels
Discriminative Models (General) [13]	Deep learning, automatic feature learning	Models relationship between features and labels effectively	Require large labeled datasets, high computational cost
Faster R-CNN [14]	Two-stage 2D object detection	High accuracy, reliable bounding box proposals	Slower inference, limited 3D contextual understanding
YOLO [15], SSD [16]	Single-stage 2D object detection	Faster detection, real-time performance	Slightly less accurate than two-stage methods
Partially Supervised Segmentation [17]	Weight transfer + limited mask annotations	Trains on large datasets with few detailed labels	Performance limited by sparse mask availability
Group Normalization [18], Cascade [19]	Advanced normalization and cascaded training	Improved performance in small-batch training	Added architectural complexity
Fast CapsNet [20]	Capsule networks on CT data	3x faster than CNNs, better screening accuracy on small samples	Requires specialized architecture, generalization not fully validated
3D CNN Models [21] [22] [23]	3D convolutional neural networks on CT scans	Exploit full volumetric context, encouraging detection results	Very high memory and computational requirements
Multi-Path CNN [24]	Parallel paths in CNN for lung CT analysis	Robust cancer detection performance	High training cost, not optimized for real-time or daily clinical deployment

2.1 Problem Definition

Lung cancer remains one of the leading causes of cancer-related deaths worldwide, largely due to its late-stage diagnosis and complex presentation in medical imaging. Traditional diagnostic methods depend heavily on symptom observation, clinical judgment, and manual interpretation of imaging data, which are time-consuming, error-prone, and inconsistent across practitioners. Although automated systems using CT images and machine learning techniques have been introduced, they still face significant limitations. These include high computational requirements, reduced accuracy when handling low-quality images, misclassification due to poor feature selection, and inefficient segmentation of cancer-affected regions. Therefore, there is a pressing need to develop more efficient, accurate, and robust automated systems for early and reliable detection of lung cancer.

2.2 Motivation

Early and accurate detection of lung cancer significantly increases the chances of successful treatment and survival. Automation in diagnosis can reduce the workload of radiologists, lower healthcare costs, and minimize diagnostic delays. With the growing volume of CT imaging data and advancements in artificial intelligence, there is strong potential to build intelligent systems that can outperform traditional methods. However, to realize this potential, the system must be optimized for feature extraction, segmentation, classification, and must be capable of handling real-world challenges such as low-quality scans and imbalanced datasets. These motivations drive the effort to improve the existing diagnostic frameworks and develop a more intelligent, resource-efficient solution.

2.3 Research Gap

Despite significant advancements, current research on automated lung cancer detection reveals several critical gaps:

- Inadequate accuracy in detecting cancer from low-resolution or noisy CT images.
- Inefficient segmentation methods, leading to poor localization of affected regions and inaccurate feature extraction.
- High computational complexity and memory usage in most deep learning-based methods, making them unsuitable for real-time clinical use.
- Limited integration of denoising, segmentation, feature selection, and classification into a cohesive and optimized pipeline.
- Underutilization of hybrid or ensemble approaches that can potentially combine strengths of multiple algorithms to improve performance.

Addressing these gaps is essential for advancing lung cancer diagnostic systems from experimental setups to reliable tools used in routine clinical practice.

LUNG IMAGING TECHNIQUES

Medical imaging plays a vital role in assisting radiologists with the diagnosis and management of lung diseases, particularly lung cancer. Among the various imaging modalities, Computed Tomography (CT) has proven highly effective due to its ability to provide detailed insights into the size, location, characteristics, and growth of lung lesions and nodules. 4D CT, in particular, enhances targeting precision during radiation therapy, thereby improving lung cancer treatment outcomes [25].

3.1 CT and LDCT

Lakshmanaprabu et al. [26] developed an automatic lung cancer detection system using Linear Discriminant Analysis (LDA) for feature reduction and an Optimal Deep Neural Network (ODNN) optimized by a Modified Gravitational Search Algorithm. This approach achieved higher classification accuracy by reducing the dimensionality of features while preserving diagnostic information.

Low-Dose CT (LDCT) offers higher sensitivity in detecting early-stage lung nodules and cancers compared to conventional CT, with significantly lower radiation exposure. However, studies have shown that LDCT does not significantly reduce lung cancer mortality. Current guidelines recommend annual LDCT screening for high-risk individuals aged 55 to 74 who have a history of heavy smoking [27].

3.2 PET Imaging

Positron Emission Tomography (PET) provides superior sensitivity and specificity in detecting lung nodules, especially in cases involving granulomatous or reactive nodal diseases [28]. PET imaging, particularly with 18F-

https://doi.org/10.46335/IJIES.2024.9.6.21 e-ISSN: 2456-3463
Vol. 9, No. 6, 2024, PP. 97-106

FDG, has been effective in diagnosing solitary pulmonary nodules [29], guiding radical radiotherapy planning in advanced NSCLC patients [30], and managing around 32% of stage IIIA lung cancer cases [31]. Additionally, 18F-FDG PET is valuable in assessing treatment response during induction chemotherapy.

3.3 MRI

Magnetic Resonance Imaging (MRI) offers a radiation-free alternative for lung imaging. However, it has limitations, including high cost, long scanning times, and difficulty in detecting small nodules—missing up to 10% of nodules in the 4–8 mm range [32]. The introduction of Ultra-Short Echo Time (UTE) MRI techniques has improved image quality by enhancing signal intensity and reducing lung susceptibility artifacts. MRI with UTE has shown promising sensitivity for small nodule detection [33].

Different MRI pulse sequences, including T1-weighted and T2-weighted imaging, have been explored to improve nodule detection [34-35]. 1.5T MRI has shown superior capability in identifying ground-glass opacities (GGOs) compared to 3T systems [36]. In particular, SSFP sequences with 1.5T MRI detected GGOs in 75% of lung fibrosis patients [37], and T2-weighted fast spin echo MRI performed comparably or better than CT in immunocompromised patients [38].

3.4 MIT (Magnetic Induction Tomography)

Emerging imaging modalities such as Magnetic Induction Tomography (MIT) have also been investigated for lung disease detection [39-40]. While MIT shows potential due to its non-invasive nature, the technology is still in its infancy. Major limitations include the lack of established measurement systems, high computational cost due to complex electromagnetic models, low image resolution, and overall system instability. These challenges currently prevent MIT from being a viable commercial option.

DEEP LEARNING-BASED IMAGING TECHNIQUES

Deep learning-based computer-aided diagnosis (CAD) systems have emerged as powerful tools for the automatic detection and classification of lung diseases, particularly lung cancer, from medical imaging data. These systems have demonstrated notable improvements in diagnostic accuracy and are increasingly being integrated into clinical workflows [38-40].

Deep learning models, particularly Convolutional Neural Networks (CNNs), are capable of learning hierarchical representations of image data through multiple layers of abstraction. These models automatically extract and refine features such as edges, textures, shapes, and patterns from raw input images without requiring manual intervention or handcrafted features. This capability makes them highly effective in analyzing complex medical images like CT and PET scans.

Some of the commonly used deep learning architectures for lung disease detection include:

- CNNs (e.g., ResNet, VGG, DenseNet) for classification and segmentation.
- Autoencoders for unsupervised feature learning and anomaly detection.
- U-Net and its variants for precise segmentation of lung regions and lesions.
- 3D CNNs for volumetric analysis of CT scans.
- Capsule Networks (CapsNet) for preserving spatial hierarchies in lung nodule detection.

Advantages:

- *High Accuracy:* Deep learning models often outperform traditional machine learning and rule-based systems.
- *Automation:* Minimizes the need for domain-specific feature engineering.
- *Scalability:* Can handle large-scale imaging datasets.
- *Generalization:* When trained properly, they adapt to varied image sources and patient demographics.

Challenges and Limitations: Despite their advantages, deep learning-based CAD systems face several challenges:

- *Data Dependency:* Require large, well-annotated datasets for effective training.
- *Computational Cost:* Training deep networks demands substantial processing power and memory.
- *Interpretability:* Most models operate as "black boxes," making clinical validation and trust more difficult.

https://doi.org/10.46335/IJIES.2024.9.6.21 e-ISSN: 2456-3463
Vol. 9, No. 6, 2024, PP. 97-106

- *Overfitting:* Especially in small or imbalanced datasets, models may perform well on training data but fail in real-world applications.
- *Robustness:* Performance can degrade with poor image quality or noise, which is common in real clinical settings.

Recent research is focused on addressing these challenges through transfer learning, attention mechanisms, data augmentation, ensemble models, and explainable AI (XAI) approaches to improve both performance and transparency.

LUNG IMAGING DATASETS

Lung imaging datasets are fundamental resources for the development, training, and evaluation of deep learning-based algorithms in lung nodule classification and detection. These datasets provide annotated CT images and clinical metadata, enabling researchers to benchmark the performance of CAD systems and validate their effectiveness in real-world diagnostic scenarios. Over the years, several publicly available datasets have been introduced, each contributing uniquely to advancements in lung cancer research.

- One of the most widely used datasets is the Lung Image Database Consortium (LIDC) [41], which consists of 399 CT images annotated by multiple radiologists. This dataset laid the groundwork for consistent labeling practices and was later expanded through collaboration with the Image Database Resource Initiative, resulting in the LIDC-IDRI dataset [42]. The LIDC-IDRI includes 1,018 CT scans from 1,010 patients and features detailed annotations for lung nodules, including malignancy ratings and segmentation masks.
- To encourage standardized evaluation, the Lung Nodule Analysis Challenge 2016 (LUNA16) [43] was introduced. It builds upon the LIDC-IDRI dataset by selecting 888 CT scans that meet specific quality and consistency criteria, making it a benchmark dataset for nodule detection algorithms.
- The Early Lung Cancer Action Program (ELCAP) [44] provides a smaller but significant collection of 50 low-dose CT (LDCT) scans and 379 unique lung nodule images. This dataset focuses on early-stage lung cancer, offering insight into subtle pathological features that are challenging to detect.
- Another notable dataset is the Lung Nodule Database (LNDb) [45], which includes 294 CT scans collected from the Centro Hospitalar e Universitário de São João in Portugal. It contains annotations performed by multiple radiologists, providing a robust set of clinical-grade images for algorithm testing.
- The Indian Lung CT Image Database (ILCID) [46] adds diversity to the available datasets by including CT images from 400 patients in the Indian population, addressing the need for demographic variety in model training and evaluation.
- The Japanese Society of Radiological Technology (JSRT) dataset [47] includes both nodular (154 cases) and non-nodular (93 cases) chest X-ray images, each with ground truth labels. Though it focuses on radiography rather than CT, it remains a valuable resource for lung disease classification studies.
- A large-scale dataset comes from the Nederland-Leuvens Longkanker Screenings Onderzoek (NELSON) study [48], which contains CT scans from 15,523 human subjects. This dataset supports longitudinal analysis and screening-based studies, providing a foundation for real-world model deployment in population-wide screening programs.
- Lastly, the Automatic Nodule Detection 2009 (ANODE09) dataset [49] offers a smaller benchmark with 5 annotated examples and 50 test images, typically used for validating nodule detection systems on limited but well-labeled data.

These datasets collectively support a wide range of research efforts, from basic lung nodule detection to complex classification tasks, and continue to play a vital role in advancing the field of AI-driven lung cancer diagnostics.

LUNG IMAGE SEGMENTATION

Image segmentation is a foundational process in medical imaging that involves identifying and delineating specific structures or regions within an image—typically at the voxel or pixel level. In the context of clinical diagnosis and treatment planning, segmentation enables the precise extraction of anatomical features, such as organs or lesions, which is critical for quantifying disease extent, navigating surgical procedures, and guiding radiotherapy.

In lung imaging, segmentation is particularly vital for accurately isolating the lungs from surrounding thoracic structures. This process usually involves thoracic region extraction, which helps eliminate irrelevant artifacts, followed by lung extraction to distinguish between the left and right lungs. Effective segmentation allows for detailed lesion analysis and supports the automation of downstream tasks such as detection and classification of lung abnormalities.

Traditionally, various thresholding methods have been employed for lung segmentation, including basic thresholding [50], iterative thresholding [51], Otsu's method [52], and adaptive thresholding techniques [53-54]. While these methods are computationally efficient, they often struggle with variations in image intensity and complex anatomical structures.

To improve segmentation accuracy, researchers have also explored region-based approaches, including 3D region growing methods [55-56], which expand a segmented area based on predefined similarity criteria. Active contour models, initially introduced by Kass et al. [57], have also been utilized for lung segmentation tasks. Lan et al. [58], for instance, applied active contour techniques to capture the irregular boundaries of lung regions. However, these traditional approaches are largely semi-automatic or manual, making them time-consuming, susceptible to human error, and dependent on high-quality ground truth annotations. Furthermore, they often suffer from issues such as class imbalance and low reproducibility.

More recently, deep learning-based techniques have shown great promise in lung segmentation. For example, Wang et al. [59] proposed a Multi-View Convolutional Neural Network (MV-CNN) designed specifically for lung nodule segmentation. Evaluated on the LIDC-IDRI dataset, their model achieved an average Dice Similarity Coefficient (DSC) of 77.67% and an Average Surface Distance (ASD) of 0.24, indicating a solid performance in capturing the structural boundaries of lung nodules.

Deep learning models offer several advantages over traditional methods, including higher accuracy, better generalization across diverse datasets, and the ability to learn complex features automatically. As the field progresses, segmentation models continue to evolve, incorporating advanced techniques such as attention mechanisms, 3D convolutions, and hybrid architectures to further enhance the precision and robustness of lung image analysis.

Table 2- Performance Comparison of Lung Nodule Classification and Segmentation Models

Reference	Method	Outcome	Limitations
SIFT + SVM [61-61]	SIFT feature extraction + SVM classifier	Sensitivity: 86-91.38%, Specificity: 89.56-97%	Feature-based, less robust to complex variations
Multi-scale CNN [62]	Multi-scale CNN architecture	Accuracy: 90.63%, Sensitivity: 92.30%, Specificity: 89.47%	High accuracy but computationally expensive
Multi-crop CNN [63]	Multi-crop CNN model	Accuracy: 87%, Sensitivity: 77%, Specificity: 93%	Lower sensitivity, potential under-detection
Deep Semantic Net [64]	Deep-level semantic network	Accuracy: 84.2%	Moderate accuracy, needs deeper training
Multi-scale CNN [65]	Multi-scale CNN variant	Accuracy: 86.84%	Less exploration of generalizability
CAD by Cheng et al. [66]	CAD system using CNN	Accuracy: 95.6%, Sensitivity: 92.4%, Specificity: 98.9%	Excellent results, but possibly dataset-specific
CNN vs DBN [67]	Comparison: CNN and DBN	CNN: 73.4%/73.3%, DBN: 82.2%/78.7%	Lower performance of CNN compared to DBN
CNN vs ResNet [68]	Comparison: CNN and ResNet	CNN: 76.64%/89.5%, ResNet: 81.97%/89.38%	Still room for improvement in CNN
CNN + RNN [69]	Combined CNN and RNN model	Accuracy: 94.78%, Sensitivity: 94.66%, Specificity: 95.14%	Complex model, high training cost
Ensemble CNN [70]	Ensemble of multiple deep CNNs	Accuracy: 84% on LIDC-IDRI	Better than others but not best-in-class
Multi-section CNN [71]	Lightweight multi-section CNN	Accuracy: 93.18% on LIDC-IDRI	High accuracy, limited to LIDC-IDRI
Transferable Texture CNN [72]	Transferable texture CNN (9 layers)	Accuracy: 96.69%, Recall: 97.19%	Small standard deviation but fixed architecture
Multi-task CNN [73]	Multi-task CNN	AUC: 0.783	Moderate AUC, needs improvement

CNN on JSRT [74]	CNN-based classifier on JSRT	Accuracy: 86.67% on JSRT	Dataset-specific, no cross-validation
ML on LNDb [75]	Machine learning classifier on LNDb	Accuracy: 94%, F1-score: 92%	Good performance, lacks generalizability

CHALLENGES AND FUTURE RESEARCH DIRECTIONS

Despite remarkable progress in deep learning-based computer-aided diagnosis (CAD) systems for lung cancer, several key challenges continue to hinder their full integration into clinical practice. These challenges affect both the development and deployment of models and must be addressed for CAD systems to achieve widespread adoption and clinical reliability.

One of the foremost challenges is the lack of standardized imaging protocols across medical institutions. Variations in image resolution, slice thickness, scanning techniques, and contrast use introduce inconsistencies that hinder model generalizability and reproducibility. Standardizing these imaging parameters would significantly enhance the comparability of datasets and improve the robustness of deep learning models. Another major barrier is the heavy reliance on annotated data. Supervised deep learning models require vast quantities of high-quality, labeled medical images to achieve reliable performance. However, producing these annotations is a labor-intensive and costly process, often requiring input from experienced radiologists. This data scarcity limits model development and hinders progress, especially in rare or complex cases where examples are few.

While 3D convolutional neural networks (3D CNNs) offer greater capacity for capturing volumetric and spatial information from CT scans, they remain underutilized. The primary reason is their computational complexity and the high hardware requirements associated with training and deploying such models. Consequently, many studies continue to use 2D CNNs, which may not fully capture the 3D context critical for accurate diagnosis. Additionally, there is limited clinical acceptance of AI-based diagnostic tools. Many deep learning systems function as "black boxes," providing predictions without transparent explanations. This lack of interpretability makes it difficult for clinicians to trust the outputs, especially when making critical treatment decisions. Moreover, the legal and ethical implications of relying on opaque algorithms further contribute to resistance in clinical adoption.

Another technical issue is the presence of imbalanced and noisy datasets. Lung imaging datasets often have a disproportionate number of benign versus malignant cases, which skews model training and can lead to increased false positive or false negative rates. Furthermore, low-quality or artifact-laden scans introduce noise that compromises model accuracy.

Looking ahead, several future research directions can help overcome these challenges. First, establishing standardized imaging guidelines across healthcare facilities would ensure data consistency and enhance model transferability across institutions. Second, there is a strong need to explore semi-supervised, unsupervised, and self-supervised learning methods. These approaches can leverage large volumes of unlabeled data, thereby reducing reliance on manually annotated datasets while maintaining performance.In parallel, advancements in 3D deep learning architectures—supported by more efficient training algorithms and hardware—could enable the broader use of 3D CNNs and hybrid models that better utilize volumetric data. Such models could significantly improve diagnostic accuracy by capturing detailed spatial patterns. Integrating explainable AI (XAI) techniques is also crucial. Methods such as saliency maps, attention mechanisms, and feature visualizations can offer interpretability and transparency, helping clinicians understand and trust model predictions. Moreover, techniques like data augmentation and the use of generative models (e.g., GANs) for synthetic data generation can help mitigate the limitations of small and imbalanced datasets, especially for underrepresented classes.

Another promising avenue is federated learning, which allows models to be trained collaboratively across multiple institutions without sharing patient data. This privacy-preserving approach is particularly relevant in the context of stringent healthcare data regulations. Finally, there is a need to develop lightweight, real-time, and resource-efficient deep learning models that can operate in environments with limited computational resources. Such models would be especially valuable in low-resource clinical settings or for deployment in mobile diagnostic tools. By addressing these challenges and pursuing the suggested research directions, the field can move closer to building robust, interpretable, and clinically viable AI solutions for lung cancer diagnosis.

CONCLUSION

The advancement of automated lung cancer detection systems, particularly those leveraging deep learning technologies, marks a transformative shift in diagnostic radiology. These systems have demonstrated considerable potential in improving detection accuracy, reducing diagnostic delays, and assisting clinicians in managing the increasing volume of imaging data. However, the current landscape still presents significant limitations that hinder their widespread clinical adoption. Among these are the dependency on large annotated datasets, computational inefficiency of high-dimensional models, and reduced performance on noisy or low-quality images. Additionally, the lack of transparency in deep learning decision-making processes poses a barrier to trust and regulatory acceptance.

This review has synthesized a wide array of imaging techniques, segmentation strategies, classification models, and benchmark datasets, providing a structured overview of the state-of-the-art in lung cancer diagnostics. The analysis reveals that while individual components—such as 3D CNNs for volumetric analysis or Capsule Networks for spatial feature preservation—have achieved notable successes, the absence of an integrated, optimized diagnostic pipeline remains a key challenge. Moreover, the underutilization of hybrid and ensemble approaches leaves untapped potential for enhancing model robustness and performance.

To address these challenges, future research must focus on developing lightweight and interpretable models that are both computationally efficient and clinically reliable. The implementation of explainable AI, data augmentation, and self-supervised learning can mitigate current data-related bottlenecks, while federated learning can support model training without compromising patient privacy. Ultimately, realizing the full potential of AI-driven diagnostics will require multidisciplinary collaboration and standardized clinical frameworks to ensure that these technologies move from academic prototypes to trusted clinical tools.

Future research can explore the integration of AI-based diagnostic tools into cloud-enabled healthcare platforms for scalable deployment. There is also scope for developing real-time, mobile-compatible systems to assist early diagnosis in remote or under-resourced clinical settings.

REFERENCES

[1] *R. Manser, A. Lethaby, L. B. Irving, C. Stone, G. Byrnes, M. J. Abramson, and D. Campbell, "Screening for lung cancer," Cochrane Database Syst. Rev., vol. 6, no. 6, p. CD001991, 2013.*

[2] *M. V. Brock et al., "DNA methylation markers and early recurrence in stage I lung cancer," N. Engl. J. Med., vol. 358, pp. 1118–1128, 2008.*

[3] *C. C. Wang et al., "HOXA5 inhibits metastasis via regulating cytoskeletal remodelling and associates with prolonged survival in non-small-cell lung carcinoma," PLoS ONE, vol. 10, p. e0124191, 2015.*

[4] *A. Hulbert, I. Jusue-Torres, A. Stark, C. Chen, K. Rodgers, B. Lee, and S. A. Belinsky, "Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum," Clin. Cancer Res., vol. 23, no. 8, pp. 1998–2005, 2017.*

[5] *H. Y. Lee, S. S. Han, H. Rhee, J. H. Park, J. S. Lee, Y. M. Oh, S. S. Choi, S. H. Shin, and W. J. Kim, "Differential expression of microRNAs and their target genes in non-small-cell lung cancer," Mol. Med. Rep., vol. 11, no. 3, pp. 2034–2040, 2015.*

[6] *G. Manogaran, P. M. Shakeel, A. S. Hassanein, M. K. Priyan, and C. Gokulnath, "Machine-learning approach based gamma distribution for brain abnormalities detection and data sample imbalance analysis," IEEE Access, 2018.*

[7] *P. M. Shakeel, A. Tolba, Z. Al-Makhadmeh, and M. M. Jaber, "Automatic detection of lung cancer from biomedical data set using discrete AdaBoost optimized ensemble learning generalized neural networks," Neural Comput. Appl., 2019.*

[8] *O. Gambino, V. Conti, S. Galdino, C. F. Valenti, and W. P. dos Santos, "Image segmentation techniques for healthcare systems," J. Healthc. Eng., vol. 2019, Article ID 2723419, 2019.*

[9] *A. I. Pratiwi and Adiwijaya, "On the feature selection and classification based on information gain for document sentiment analysis," Appl. Comput. Intell. Soft Comput., vol. 2018, Article ID 1407817, 2018.*

[10] *K. P. Sridhar, S. Baskar, P. M. Shakeel, and V. S. Dhulipala, "Developing brain abnormality recognize system using multi-objective pattern producing neural network," J. Ambient Intell. Humaniz. Comput., 2018.*

[11] *P. M. Shakeel, T. E. E. Tobely, H. Al-Feel, G. Manogaran, and S. Baskar, "Neural network based brain tumor detection using wireless infrared imaging sensor," IEEE Access, vol. 7, pp. 5577–5588, 2019.*

[12] *H. C. Roberts, D. Patsios, M. Kucharczyk, N. Paul, and T. P. Roberts, "The utility of computer-aided detection (CAD) for lung cancer screening using low-dose CT," Int. Congr. Ser., vol. 1281, pp. 1137–1142, 2005.*

[13] *L. Abdul et al., "Deep-lumen assay-human lung epithelial spheroid classification from brightfield images using deep learning," Lab Chip, vol. 21, pp. 447–448, 2021.*

[14] *S. G. I. Armato, "Deep learning demonstrates potential for lung cancer detection in chest radiography," Radiology, vol. 297, pp. 697–698, 2020.*

[15] *S. Ali et al., "State-of-the-art challenges and perspectives in multi-organ cancer diagnosis via deep learning-based methods," Cancers, vol. 13, p. 5546, 2021.*

[16] *D. Riquelme and M. A. Akhloufi, "Deep learning for lung cancer nodules detection and classification in CT scans," AI, vol. 1, pp. 28–67, 2020.*

[17] *T. A. Zhukov et al., "Discovery of distinct protein profiles specific for lung tumors and pre-malignant lung lesions by SELDI mass spectrometry," Lung Cancer, vol. 40, pp. 267–279, 2003.*

[18] *F. A. Zeiser et al., "Deepbatch: A hybrid deep learning model for interpretable diagnosis of breast cancer in whole-slide images," Expert Syst. Appl., vol. 185, p. 115586, 2021.*

[19] *M. Mandal and S. K. Vipparthi, "An empirical review of deep learning frameworks for change detection: Model design,*

https://doi.org/10.46335/IJIES.2024.9.6.21 **e-ISSN: 2456-3463**
Vol. 9, No. 6, 2024, PP. 97-106

experimental frameworks, challenges and research needs," IEEE Trans. Intell. Transp. Syst., vol. 23, pp. 6101–6122, 2022.

[20] *H. Alireza, M. Cheikh, A. Annika, and V. Jari, "Deep learning for forest inventory and planning: A critical review on the remote sensing approaches so far and prospects for further applications," Forestry, vol. 95, pp. 451–465, 2022.*

[21] *G. Stamatis, W. Eberhard, and C. Pottgen, "Surgery after multimodality treatment for non-small-cell lung cancer," Lung Cancer, vol. 45, pp. S107–S112, 2004.*

[22] *N. Journy et al., "Are the studies on cancer risk from CT scans biased by indication? Elements of answer from a large-scale cohort study in France," Br. J. Cancer, vol. 112, pp. 1841–1842, 2015.*

[23] *National Lung Screening Trial Research Team, "Reduced lung-cancer mortality with low-dose computed tomographic screening," N. Engl. J. Med., vol. 365, pp. 395–409, 2011.*

[24] *F. Highamcatherine and J. Highamdesmond, "Deep learning," SIAM Rev., vol. 32, pp. 860–891, 2019.*

[25] *K. Latifi et al., "Impact of dose on lung ventilation change calculated from 4D-CT using deformable image registration in lung cancer patients treated with SBRT," J. Radiat. Oncol., vol. 4, pp. 265–270, 2015.*

[26] *S. K. Lakshmanaprabu, S. N. Mohanty, K. Shankar, N. Arunkumar, and G. Ramirez, "Optimal deep learning model for classification of lung cancer on CT images," Future Gener. Comput. Syst., vol. 92, pp. 374–382, 2019.*

[27] *S. S. Shim et al., "Non-small cell lung cancer: Prospective comparison of integrated FDG PET/CT and CT alone for preoperative staging," Radiology, vol. 236, pp. 1011–1019, 2005.*

[28] *G. D. C. Ab et al., "PET-CT in presurgical lymph node staging in non-small cell lung cancer: The importance of false-negative and false-positive findings," Radiologia, vol. 59, pp. 147–158, 2017.*

[29] *S. Yaturu and R. A. Patel, "Metastases to the thyroid presenting as a metabolically inactive incidental thyroid nodule with stable size in 15 months," Case Rep. Endocrinol., vol. 2014, p. 643986, 2014.*

[30] *S. M. Eschmann et al., "18F-FDG PET for assessment of therapy response and preoperative re-evaluation after neoadjuvant radio-chemotherapy in stage III non-small cell lung cancer," Eur. J. Nucl. Med. Mol. Imaging, vol. 34, pp. 463–471, 2007.*

[31] *W. K. Lee et al., "Modern diagnostic and therapeutic interventional radiology in lung cancer," J. Thorac. Dis., vol. 5, pp. 511–523, 2013.*

[32] *M. Zurek, A. Bessaad, K. Cieslar, and Y. Crémillieux, "Validation of simple and robust protocols for high-resolution lung proton MRI in mice," Magn. Reson. Med., vol. 64, pp. 401–407, 2010.*

[33] *N. S. Burris et al., "Detection of small pulmonary nodules with ultrashort echo time sequences in oncology patients by using a PET/MR system," Radiology, vol. 278, pp. 239–246, 2016.*

[34] *C. Fink et al., "Lung MRI at 1.5 and 3 tesla: Observer preference study and lesion contrast using five different pulse sequences," Investig. Radiol., vol. 42, pp. 377–383, 2007.*

[35] *A. Cieszanowski et al., "Characterization of focal liver lesions using quantitative techniques: Comparison of apparent diffusion coefficient values and T2 relaxation times," Eur. Radiol., vol. 22, pp. 2514–2524, 2012.*

[36] *D. Hughes, H. Tiddens, and J. M. Wild, "Lung imaging in cystic fibrosis," Imaging Decis. MRI, vol. 13, pp. 28–37, 2009.*

[37] *M. Groth et al., "Intraindividual comparison of contrast-enhanced MRI and unenhanced SSFP sequences of stenotic and non-stenotic pulmonary artery diameters," Rofo, vol. 183, pp. 47–53, 2011.*

[38] *A. L. Chong et al., "Proton density MRI increases detection of cervical spinal cord multiple sclerosis lesions compared with T2-weighted fast spin-echo," Am. J. Neuroradiol., vol. 37, pp. 180–184, 2016.*

[39] *S. Alzeibak and N. H. Saunders, "A feasibility study of in vivo electromagnetic imaging," Phys. Med. Biol., vol. 38, pp. 151–160, 1993.*

[40] *R. Merwa, K. Hollaus, P. Brunner, and H. Scharfetter, "Solution of the inverse problem of magnetic induction tomography (MIT)," Physiol. Meas., vol. 26, pp. 241–250, 2006.*

[41] *Lung Image Database Consortium (LIDC), "Lung Image Database Consortium (LIDC)," accessed Oct. 5, 2022.*

[42] *S. G. Armato et al., "Data from LIDC-IDRI," 2015, accessed Oct. 5, 2022.*

[43] *A. A. A. Setio et al., "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 Challenge," Med. Image Anal., vol. 42, pp. 1–13, 2017.*

[44] *ELCAP Public Lung Image Database, 2014, accessed Oct. 5, 2022.*

[45] *J. Pedrosa et al., "LNDb: A lung nodule database on computed tomography," arXiv preprint, arXiv:1911.08434, 2019.*

[46] *D. Prasad, B. Ujjwal, and T. Sanjay, "LNCDS: A 2D-3D cascaded CNN approach for lung nodule classification, detection and segmentation," Biomed. Signal Process. Control., vol. 67, p. 102527, 2021.*

[47] *J. Shiraishi et al., "Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," Am. J. Roentgenol., vol. 174, pp. 71–74, 2000.*

[48] *D. D. Costa et al., "Biomarker identification for early detection of lung cancer by proteomic techniques in the NELSON lung cancer screening trial," Cancer Res., vol. 68, p. 3961, 2008.*

[49] *B. van Ginneken et al., "Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study," Med. Image Anal., vol. 14, pp. 707–722, 2010.*

[50] *S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Accurate lung segmentation for accurate quantization of volumetric X-ray CT images," IEEE Trans. Med. Imaging, vol. 20, pp. 490–498, 2001.*

[51] *A. Dawoud, "Lung segmentation in chest radiographs by fusing shape information in iterative thresholding," Comput. Vis. IET, vol. 5, pp. 185–190, 2011.*

[52] *N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Syst. Man Cybern., vol. 9, pp. 62–66, 1979.*

[53] *T. Peng, C. Wang, Y. Zhang, and J. Wang, "H-SegNet: Hybrid segmentation network for lung segmentation in chest radiographs using mask region-based convolutional neural network and adaptive closed polyline searching method," Phys. Med. Biol., vol. 67, p. 075006, 2022.*

[54] *L. Y. Tseng and L. C. Huang, "An adaptive thresholding method for automatic lung segmentation in CT images," in Proc. IEEE Africon, Nairobi, Kenya, 23–25 Sept. 2009, pp. 1–5.*

[55] *J. Dehmeshki, H. Amin, M. Valdivieso, and X. Ye, "Segmentation of pulmonary nodules in thoracic CT scans: A region growing approach," IEEE Trans. Med. Imaging, vol. 27, pp. 467–480, 2008.*

[56] *A. Fabijacska, "The influence of pre-processing of CT images on airway tree segmentation using 3D region growing," in Proc. 5th Int. Conf. Perspective Technol. Methods MEMS Design, Lviv, Ukraine, 22–24 Apr. 2009.*

[57] *M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," Int. J. Comput. Vis., vol. 1, pp. 321–331, 1988.*

[58] *Y. Lan, N. Xu, X. Ma, and X. Jia, "Segmentation of pulmonary nodules in lung CT images based on active contour model," in Proc. 14th IEEE Int. Conf. Intell. Human-Machine Syst. Cybern. (IHMSC), Hangzhou, China, 20–21 Aug. 2022.*

[59] *S. Wang et al., "A multi-view deep convolutional neural networks for lung nodule segmentation," in Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jeju, Korea, 11–15 July 2017, pp. 1752–1755.*

[60] *A. Farag, A. Ali, J. Graham, A. Farag, S. Elshazly, and R. Falk, "Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose CT scans of the chest," in Proc. IEEE Int. Symp. Biomed. Imaging: From Nano to Macro, Chicago, IL, USA, Mar. 30–Apr. 2, 2011, pp. 169–172.*

[61] *S. S. Parveen and C. Kavitha, "Classification of lung cancer nodules using SVM Kernels," Int. J. Comput. Appl., vol. 95, 2014, pp. 975–8887.*

[62] *E. Dandıl, M. Çakıroğlu, Z. Ekşi, M. Özkan, Ö. K. Kurt, and A. Canan, "Artificial neural network-based classification system for lung nodules on computed tomography scans," in Proc. 6th Int.*

Conf. Soft Comput. Pattern Recognit. (SoCPar), Tunis, Tunisia, Aug. 11–14, 2014, pp. 382–386.

[63] *K. L. Hua, C. H. Hsu, S. C. Hidayati, W. H. Cheng, and Y. J. Chen, "Computer-aided classification of lung nodules on computed tomography images via deep learning technique," OncoTargets Ther., vol. 8, 2015, pp. 2015–2022.*

[64] *S. Shen, S. X. Han, D. R. Aberle, A. A. Bui, and W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification," Expert Syst. Appl., vol. 128, 2019, pp. 84–95.*

[65] *W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," Inf. Process. Med. Imaging, vol. 24, 2015, pp. 588–599.*

[66] *J. Z. Cheng, D. Ni, Y. H. Chou, J. Qin, C. M. Tiu, Y. C. Chang, C. S. Huang, D. Shen, and C. M. Chen, "Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in US images and pulmonary nodules in CT scans," Sci. Rep., vol. 6, 2016, Art. no. 24454.*

[67] *T. L. Kwajiri and T. Tezuka, "Classification of Lung Nodules Using Deep Learning," Trans. Jpn. Soc. Med. Biol. Eng., vol. 55, 2017, pp. 516–517.*

[68] *W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, J. Tian, and Y. Zang, "Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification," Pattern Recognit., vol. 61, 2017, pp. 663–673.*

[69] *Q. Abbas, "Lung-deep: A computerized tool for detection of lung nodule patterns using deep learning algorithms," Int. J. Adv. Comput. Sci. Appl., vol. 8, 2017, pp. 112–116.*

[70] *B. Zhang, S. Qi, P. Monkam, C. Li, and W. Qian, "Ensemble learners of multiple deep CNNs for pulmonary nodules classification using CT images," IEEE Access, vol. 7, 2019, pp. 110358–110371.*

[71] *P. Sahu, D. Yu, M. Dasari, F. Hou, and H. Qin, "A lightweight multi-section CNN for lung nodule classification and malignancy estimation," IEEE J. Biomed. Health Inform., vol. 23, 2019, pp. 960–968.*

[72] *I. Ali, M. Muzammil, D. I. Ulhaq, A. A. Khaliq, and S. Malik, "Efficient lung nodule classification using transferable texture convolutional neural network," IEEE Access, vol. 8, 2020, pp. 175859–175870.*

[73] *S. Marques, F. Schiavo, C. A. Ferreira, J. Pedrosa, A. Cunha, and A. Campilho, "A multi-task CNN approach for lung nodule malignancy classification and characterization," Expert Syst. Appl., vol. 184, 2021, Art. no. 115469.*

[74] *V. Thamilarasi and R. Roselin, "Automatic classification and accuracy by deep learning using CNN methods in lung chest x-ray images," IOP Conf. Ser. Mater. Sci. Eng., vol. 1055, 2021, Art. no. 012099.*

[75] *I. D. Kawathekar and A. S. Areeckal, "Performance analysis of texture characterization techniques for lung nodule classification," J. Phys. Conf. Ser., vol. 2161, 2022, Art. no. 012045.*