Journal of Computer Science

Review Article

Machine Learning Advancements for Lung Cancer Detection: An In-Depth Review and Future Prospects

Aanchal Vij[1], Kuldeep Singh Kaswan[1] and Anand Nayyar[2]

[1]*School of Computing Science and Engineering, Galgotias University, Greater Noida, India*

[2]*Graduate School, Du Tan University, Dà Nang 550000, Vietnam*

Article history
Received: 03-04-2025
Revised: 19-08-2025
Accepted: 15-09-2025

Corresponding Author:
Aanchal Vij
School of Computing Science and Engineering,
Galgotias University,
Greater Noida, India
Email:
aanchal.vij04@gmail.com

Abstract: Lung cancer is a major cause of cancer-related mortality globally, underscoring the necessity for efficient diagnostic instruments. The review paper summarizes the role of ML and DL in detection, staging, and prognostication of lung cancer. We evaluate the relative efficacy of several models, such as CNNs, SVMs, and ensemble approaches, by analyzing publically accessible imaging and molecular information. We emphasize difficulties including class imbalance, model interpretability, and generalizability across clinical environments. We also talk about new trends that could improve clinical translation, like vision transformers, explainable AI, and federated learning. This interdisciplinary approach highlights the revolutionary potential of AI-driven techniques in lung cancer therapy and delineates critical future research directions to enhance clinical integration.

Keywords: Lung Cancer, Medical Image Processing, CNN, Explainable AI, Artificial Intelligence

Introduction

Lung cancer is the most deadly cancer among all cancer types, characterized by its prevalence and severe effects on morbidity and death rates, hence highlighting the urgent necessity for effective diagnosis and intervention measures. Histologically, Lung Cancer (LC) is divided into two groups: Non-Small Cell Lung Cancer (NSCLC) and Small Cell Lung Cancer (SCLC). Each of these groups has several subgroups and shows a lot of differences in how they present, how their genes change, and how they respond to treatment (Vij and Kaswan, 2023). The complex etiology, shaped by genetic predisposition, environmental exposure, and lifestyle factors, complicates diagnosis and management. The emergence of precision medicine underscores the significance of early detection, illustrating that tailored therapies founded on a tumor's unique genetic profile can significantly improve patient outcomes (Tan et al., 2025).

This review aims to investigate the intricate field of lung cancer diagnosis, with a particular emphasis on the integration of machine learning techniques. It is important to understand lung cancer and know how to treat it well. The disease includes complicated subgroups such as Small Cell Lung Cancer (SCLC) and Non-Small Cell Lung Cancer (NSCLC). It has a complex landscape with different genetic profiles, clinical presentations, and responses to treatment (Wang et al., 2025b). Early detection and intervention are crucial, as they are directly associated with enhanced patient outcomes and elevated survival rates. It is necessary to confirm the clinical significance and dependability of these computational methods in real-world settings. Researchers, clinicians, and data scientists must continue to work together in order to improve computational oncology and bring these improvements into clinical practice. The use of artificial intelligence and deep learning in lung cancer management could change the way patients are treated by allowing for earlier detection, more accurate prognosis prediction, and better treatment outcomes. Thus, it is essential to dedicate resources to research and development in this domain to fully exploit the capabilities of computational methods in lung cancer treatment, thereby improving patient quality of life and survival rates.

Lung Cancer: A Clinical Perspective

Lung cancer is a major global health concern that is characterized by an uncontrolled growth of cancerous cells in lung tissues (Sujitha and Seenivasagam, 2021). It is a lethal disease, constituting the primary cause of cancer-related fatalities globally. The fatality rate remains excessively high due to various variables, including late-stage diagnosis, restricted treatment options, and the disease's aggressive aggressiveness. The overall prognosis for lung cancer patients continues to be unfavorable, particularly for those identified at advanced stages. Moreover, elevated rates of recurrence and metastasis exacerbate the lethality of lung cancer. Table 1 presents the projected lung cancer mortality rates for 2023 as calculated by the American Cancer Society, which indicates that lung cancer ranks among the most lethal cancers for both genders.

Table 1: Cancer-specific mortality rates till 2023

Cancer Type	Deaths (till 2023)
Lung and Bronchus	127,070
Colorectum	52,550
Pancreas	50,550
Breast	43,700
Prostate	34,700
Liver and Intrahepatic bile duct	29,380
Leukaemia	23,710
Non-Hodgkin lymphoma	20,180
Brain and other nervous system	18,990
Urinary Bladder	16,710
Oesophagus	16,120
Kidney and renal pelvis	14,890
Ovary	13,270
Uterine corpus	13,030
Myeloma	12,590
Oral cavity and pharynx	11,580
Stomach	11,130
Melanoma of the skin	7,990
Soft tissue (including heart)	5,140
Gallbladder and other biliary	4,510
Cervix	4,310
Larynx	3,820
Bones and Joints	2,140
Thyroid	2,120
Small Intestine	2,070
Anus, anal canal and anorectum	1,870
Vagina and other female genital	1,740
Vulva	1,670
Ureter and other urinary organs	990
Hodgkin lymphoma	900
Penis and other male genital	470
Testis	470
Eye and orbit	430

At its core, lung cancer exhibits multifaceted manifestations encompassing distinct types with varying characteristics and implications for treatment. Understanding these classifications is pivotal for navigating effective interventions and prognostic strategies (Ansari et al., 2025). Moreover, exploring the different stages of lung cancer aids in elucidating its progression and in devising tailored treatment regimens. Existing diagnostic methods, while foundational to the identification and evaluation of lung cancer, confront hurdles in enabling timely detection. The intricacies surrounding early diagnosis and precise detection constitute pivotal challenges that affect the treatment efficacy and patient outcomes.

Types and Stages of Lung Cancer

Lung cancer comprises different histological categories that significantly impact treatment strategies and patient outcomes. The main types are Non-Small Cell Lung (NSCLC) and Small Cell Lung Cancer (SCLC).

Figure 1 provides a comprehensive visual representation of various lung cancer types and their associated subtypes, elucidating the inherent variability within this disease spectrum.

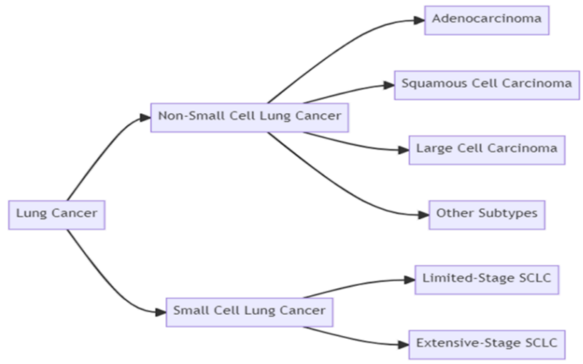

Fig. 1: Types of Lung Cancer (Cords et al., 2024)

NSCLC includes about 85% of all cases and includes a range of histological subtypes. Each subtype exhibits unique clinical characteristics and distinct genetic profiles, influencing variations in disease progression and treatment response (Alduais et al., 2023). In the context of NSCLC, adenocarcinoma, squamous cell carcinoma, and large cell carcinoma have become the primary subtypes, each distinguished by unique molecular signatures and clinical characteristics. Figure 2 displays the principal cause severity of its subtypes, alongside additional subtypes categorized under NSLC (Giaccone and He, 2023).

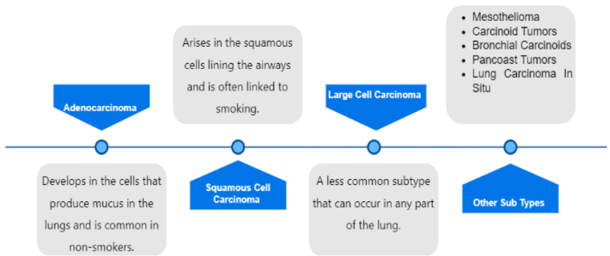

Fig. 2: Types of Non Small Lung Cancer (Liu et al., 2024)

Adenocarcinoma: A Diverse Subtype

Adenocarcinoma displays various histological characteristics and genetic modifications in NSCLC. These mutations are crucial for understanding the initiation and progression of malignancies, as well as for tailoring treatments to patients according to their distinct genetic profiles (Chen et al., 2020). ALK rearrangements, EGFR mutations, ROS1 fusions, and KRAS mutations are some of the most important genetic changes that happen in adenocarcinoma. These changes have made a big difference in the treatment options for people with NSCLC. Targeted medications, including EGFR Tyrosine Kinase Inhibitors (TKIs) like gefitinib and osimertinib, have proven highly effective in treating individuals with EGFR mutations (Xu et al., 2010). Also, discoveries like the fact that crizotinib works well in patients with ALK-positive NSCLC (Solomon et al., 2023) have made it possible to create personalized treatments that focus on these genetic changes. Finding and targeting these mutations has changed how adenocarcinoma is treated. Now, patients have more personalized and effective treatment options that can improve their quality of life and outcomes.

Squamous Cell Carcinoma

Squamous cell carcinoma, frequently linked to a history of smoking, demonstrates unique genetic changes. Alterations in the TP53, CDKN2A, and FGFR genes are prevalent in squamous cell carcinoma, despite the infrequency of driver mutations compared to adenocarcinoma (Ermakov et al., 2023). It is extremely challenging to effectively target this type of cancer in the absence of distinctly identifiable actionable alterations. Immunotherapeutic methods have emerged as viable approaches to address this issue. Immune checkpoint inhibitors that target PD-1 and PD-L1 have been shown to enhance the immune response against squamous cell carcinoma. The KEYNOTE-407 trial, which was very important, showed that giving pembrolizumab and chemotherapy together helped people with squamous NSCLC. This gives hope for better results from treatment for this type of cancer (Chen et al., 2023).

Large Cell Carcinoma: Uncommon and Highly Aggressive

LCC is a rare and aggressive form of NSCLC characterized by its high level of aggressiveness and inability to differentiate between cells. We don't know as much about genetic changes that are unique to this subtype as we do about other NSCLC subtypes, even though they are clinically important (Lantuejoul et al., 2020). The way it is classified has changed as molecular profiling tools have gotten better. This has helped to show what makes it special. Recent studies indicate that immune checkpoint inhibitors may be effective in managing this challenging subtype (Birnboim-Perach and Benhar, 2024). The study indicates that immunotherapy may be a promising approach for treating this aggressive form of lung cancer.

Machine Learning in Healthcare

The utilization of Machine Learning (ML) in healthcare has transformed medical paradigms by utilizing data-driven insights to improve the overall health of patients, the efficacy of treatment, and the quality of diagnosis. Various machine learning techniques have shown significant potential in the identification of complex patterns in medical data. These advancements have increased the power of predictive modeling, enhanced the detail of image analysis, and made customized medicine possible (Xia et al., 2022). In order to identify lung cancer, ML algorithms are used to examine complex imaging data. Figure 3 demonstrates the different types of medical images that are used to detect disease. This ushers in a new era of medical diagnostics, enabling medical professionals to detect, understand, and treat potential cancers with previously unheard-of speed and precision.

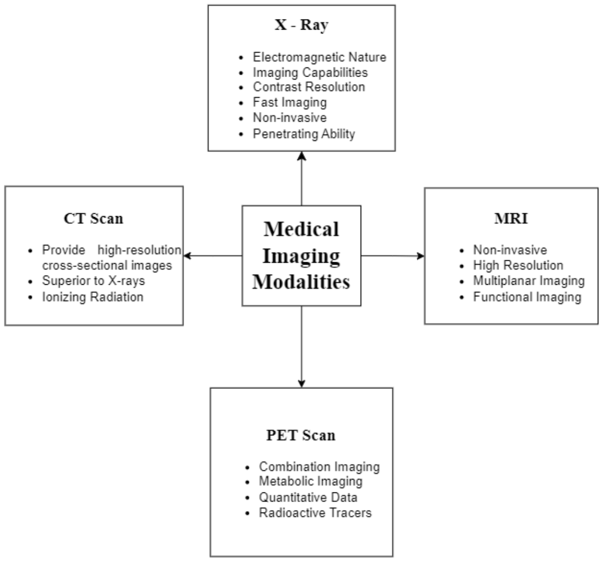

Fig. 3: Medical Image Types (Saw et al., 2025)

Convolutional Neural Networks (CNNs) have demonstrated an unparalleled level of sensitivity and specificity in detecting lung nodules, as demonstrated by studies such as Vij and Kaswan (2023). This is an excellent illustration of how this ever-changing environment has evolved. Notably, the development of transfer learning has accelerated model building by applying knowledge gathered from multiple datasets and optimizing pre-trained models on medical images. This approach is described in the studies by Huang et al. (2023); Ito et al. (2022); Yaqoob et al. (2023), it has resulted in cross-specialty applications that extend beyond disciplines such as radiology and pathology.

Machine Learning for Lung Cancer Detection

Machine Learning (ML) techniques have transformed lung cancer detection by offering powerful tools for analyzing medical data and extracting critical insights. Several ML algorithms, each with unique strengths, contribute to more accurate and early detection of lung cancer (Wang et al., 2023). Supervised methods such as SVMs create optimal class separation. Decision trees and random forests offer hierarchical classification and higher accuracy (Nafea et al., 2024). Unsupervised techniques, such as k-means clustering, identify data groups, whereas PCA and t-SNE simplify complex structures (Almahdi et al., 2025; Vikas et al., 2024). CNNs excel in image analysis for lung tumor detection. RNNs can handle sequential data. Ensemble methods, such as gradient boosting and bagging, improve accuracy (Sonthalia et al., 2025). Semi-supervised learning merges the labeled and unlabeled data. Hybrid approaches combine deep learning with traditional ML to improve lung cancer detection. Figure 4 shows the different machine learning algorithms that are used to detect LC, which have been compared in the literature review section of the paper. techniques for detecting lung cancer.

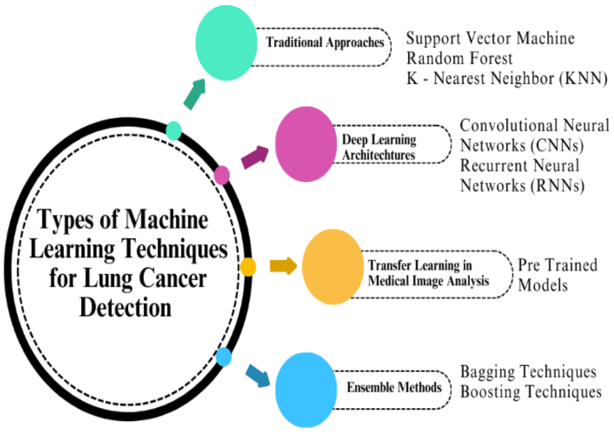

Fig. 4: Types of Machine Learning Techniques (Dubey and Sikarwar, 2025)

Evaluation Metrics for Machine Learning Models

The evaluation of machine learning models involves the use of task-specific evaluation metrics. In the realm of medical diagnostics, various terminologies elucidate the outcomes of diagnostic testing:

True Positive (TP): This designation pertains to instances where the diagnostic test accurately identifies individuals afflicted with the condition under scrutiny. In essence, the test yields a positive result, indicative of the presence of the condition, and corresponds with the actual clinical status of the individual (Liu et al., 2023).

True Negative (TN): This designation characterizes scenarios in which the diagnostic test accurately identifies individuals devoid of the condition under investigation. The test yields a negative result, denoting the absence of the condition, aligning with the veritable clinical condition of the individual.

False Positive (FP): This label comes from a diagnostic test that incorrectly identifies people who don't have the condition as having it. Consequently, the test yields a positive result, erroneously suggesting the presence of the condition, despite its actual absence.

False Negative (FN): This designation specifies cases in which the diagnostic test erroneously classifies individuals with the condition as negative. As a result, the test gives a false negative, which means it says the disease isn't there when it really is.

Classification models usually use more than one metric (Hyunwoo et al., 2024). These metrics work together to see how well the model works and help you choose an algorithm for tasks that involve finding lung cancer.

Accuracy is a common metric, but it might not always give a full picture, especially when the datasets are uneven or when the importance of each class changes. So, for a complete evaluation of a model, it's important to look at metrics like precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC).

Accuracy

Accuracy measures the proportion of correctly predicted instances to the overall number of instances in a dataset. It is given as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision

Precision checks how accurate positive predictions are. It tells you how many of the predicted positives are actually true positives. This is especially important when it's important to keep false positives to a minimum:

$$\text{Precision} = \frac{TP}{TP+FP}$$

Recall

Recall, which is also called sensitivity, shows how well a model can find all the real positive cases. It tells you how many of the actual positives were correctly predicted, which is important when you need to cut down on false negatives:

$$\text{Recall} = \frac{TP}{TP+FN}$$

F1-Score

The F1-score is a harmonic mean of precision and recall. This makes it very useful when classes aren't evenly distributed:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Area Under the Curve (AUC)

The AUC evaluates a classifier's proficiency in binary classification by quantifying its ability to differentiate between positive and negative classes at different thresholds. It can be anywhere from 0.5 (a random guess) to 1 (a perfect classification).

Receiver Operating Characteristic (ROC) Curve

The ROC curve shows how well a model works by plotting sensitivity (recall) against $1 - \text{specificity}$ at different levels. The model's ability to tell the difference between things gets stronger the closer the curve gets to the top-left corner (Singh and Kaswan, 2016).

Current Diagnostic Modalities

Usually, diagnosing lung cancer requires the use of many different tests. Imaging modalities, including chest

X-rays and CT scans, facilitate direct examination and tissue sampling for validation, whereas invasive procedures, such as bronchoscopy and biopsy, offer comprehensive insights into lung tissue (Sherin et al., 2023). While these methodologies are beneficial, they possess limitations regarding their sensitivity, specificity, and ability to detect tumors in their initial stages (Hamza et al., 2024). Table 2 summarizes the different approaches and their differences to extract the best approach to be followed for detecting lung cancer.

Data Sources and Pre-Processing

Medical imaging modalities such as X-rays, CT scans, and MRIs are indispensable tools for detecting lung cancer, providing detailed images of lung anatomy to identify anomalies and lesions essential for diagnosis. Figure 5 shows the procedure of lung cancer data set classification.

Public Databases for Lung Cancer Research

Publicly available databases are a major force behind lung cancer research. The Lung Image Database Consortium (LIDC) and the National Lung Screening Trial (NLST) are two examples of resources that provide annotated medical images that help create and test ML algorithms (Jacobs et al., 2021). Researchers can use these datasets to compare their models to clinical standards and work together on projects with other researchers (Table 3).

Literature Review

For more than twenty years, researchers have been working hard to find ways to find lung cancer early. Conventional diagnostic methods like CT, X-ray, bronchoscopy, and biopsy have persisted as the clinical standard; however, their restricted sensitivity, specificity, and invasiveness prompted the adoption of computational techniques. Machine Learning (ML) and Deep Learning (DL) have made a lot of progress in automated cancer analysis.

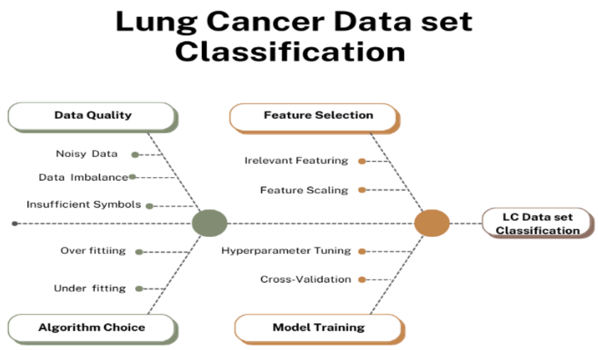

Fig. 5: Lung Cancer Data Set Classification (HaghighiKian et al., 2025)

Table 2: Comparison of diagnostic modalities for lung cancer detection

Aspect	CT Scans	X-rays	Bronchoscopy	Biopsy
Imaging Principle	Cross-sectional imaging using X-rays and computer processing (Wason and Nagarajan, 2019)	Projectional radiography using X-rays	Visual examination of airways using a thin, flexible tube (bronchoscope)	Removal of tissue samples for microscopic examination
Image Quality	High-resolution, detailed images of soft tissues and bones	Provides images of bones and dense tissues	Real-time visualization of airways and lung tissue (Chen et al., 2021)	Provides tissue samples for accurate diagnosis
Diagnostic Range	Detects abnormalities in various body regions, including lungs (Wang et al., 2025a)	Useful for skeletal and lung pathology detection (Bradley et al., 2019)	Direct visualization of airway anomalies	Diagnoses cellular and molecular abnormalities (Di Capua et al., 2021)
Invasive Nature	Non-invasive procedure (Widodo et al., 2020)	Non-invasive procedure (Tiwari et al., 2020)	Invasive procedure, requires sedation	Invasive procedure, tissue extraction
Clinical Applications	Commonly used for lung cancer detection and staging (Widodo et al., 2020)	Used for identifying fractures, lung infections	Diagnoses lung diseases, sources of bleeding	Determines histopathological characteristics

Table 3: Publicly available datasets for lung cancer research and their characteristics

Release Date	Summary	Count	Type	Size	Dataset / Ref
2011	CT scans with annotated lesions for lung cancer; benchmark dataset.	1018 cases	DICOM	125 GB	LIDC-IDRI (Jacobs et al., 2021)
2012	Radiomics + genomics for NSCLC patients; links imaging and mutations.	1,019 pts	DICOM	6.6 GB	NSCLC-Radiomics-II (Lin et al., 2023)
2016	CT scans curated for ML competitions (LUNA).	888 scans	PNG/DICOM	47.4 GB	Kaggle LUNA (Alsaadi and Rahman, 2023)
2002-2017	CheXpert: 224k chest radiographs with radiologist labels.	224,316 imgs	PNG (X-ray)	439 GB	CheXpert (Bojer and Meldgaard, 2021)
2018	NSCLC Radiogenomics dataset (TCIA); imaging + genomics.	286,754	TCIA	97.6 GB	NSCLC-Radiogenomics-I (Bove et al., 2023)

Evolution of Imaging-Based Approaches

Earlier studies (2007-2012) primarily relied on handcrafted radiomic descriptors obtained from CT and PET scans. These methods provided initial insight into tumor characterization but were constrained by reproducibility and observer variability. The release of annotated repositories such as LIDC-IDRI and LUNA16 enabled reproducible benchmarking and systematic evaluation of computer-aided detection systems.

Rise of Machine Learning Models

Between 2013 and 2018, classical ML methods such as support vector machines, random forests, and ensemble classifiers were widely used for nodule detection and malignancy prediction, often achieving accuracies above 85%. Radiogenomics also gained attention. For example, Shiri et al. (2020) combined PET and CT features with ML algorithms to predict EGFR and KRAS mutation status in NSCLC patients, reporting AUC values above 0.80. This work demonstrated the potential of imaging-driven biomarkers as non-invasive alternatives to molecular profiling.

Deep Learning and Multi-Modal Integration

Since 2018, DL architectures have dominated research, with CNNs and 3D networks delivering superior accuracy in classification and staging. Multi-modal fusion approaches that integrate imaging, genomics, and clinical variables have further improved performance. Jiang et al. (2025) compared 2D and 3D CNNs on the NLST dataset, with AUROC values of 0.79 and 0.86, respectively. Similarly, Lu et al. (2025) developed a CT-based DL radiomics model for predicting PD-L1 status in NSCLC patients, achieving AUC values approaching 0.91 when clinical features were included.

Prognosis and Survival Prediction

Recent studies emphasize prognosis alongside diagnosis. Yang et al. (2025) proposed a CT-based radiomics nomogram for small cell lung cancer, achieving a C-index of 0.74 in progression-free survival prediction. Wang et al. (2025c) developed a radiomics-dosiomics nomogram for forecasting radiation-induced pneumonia in NSCLC, yielding external validation AUCs above 0.80. In addition, Salmanpour et al. (2025) demonstrated that semi-supervised learning enhances robustness in survival prediction across multiple datasets, reaching accuracies near 88%.

Multi-Center and Harmonization Efforts

A major challenge remains generalization across diverse clinical settings. Mali et al. (2025) addressed this issue by harmonizing multi-regional radiomic features and foundation-model representations across five centers. Their consensus model achieved time-dependent AUCs above 0.90, underscoring the importance of harmonization and reproducibility for clinical adoption.

Table 4 displays various papers from 2018 to 2025 that discuss the enhancement of lung cancer detection through the implementation of multiple algorithms and the potential for further improvement to achieve superior results.

Table 4: Selected machine learning and radiomics studies for lung cancer

Ref	Open Access	Results Obtained	Data Set	Country
Peng et al. (2025)	TRUE	Prognostic CT radiomics model for NSCLC; strong performance integrating clinical features.	NSCLC cohort	China
Jiang et al. (2025)	TRUE	Compared 2D vs 3D DL models; best AUROC 0.86 (3D), 0.79 (2D).	NLST (National Lung Screening Trial)	USA
Yang et al. (2025)	TRUE	CT radiomics + clinical nomogram for PFS in SCLC; C-index ~0.744.	Small Cell Lung Cancer patients	China
Wang et al. (2025b)	TRUE	High AUC and C-index for lung cancer classification and survival prediction using 3D CNN + Cox model.	NLST (National Lung Screening Trial)	USA
Lu et al. (2025)	TRUE	CT-based DL model predicted PD-L1 status in NSCLC; AUC ~0.85, combined model with clinical data ~0.91.	352 NSCLC patients	China
Wang et al. (2025c)	TRUE	DL radiomics + dosiomics nomogram predicted radiation pneumonia Grade ≥2; AUCs ~0.891 (train), ~0.801 (external).	245 NSCLC patients (multi-center)	China
Wang et al. (2025a)	TRUE	Prognostic model predicted postoperative survival based on imaging + lung + muscle features.	NSCLC postoperative cohort	China
Ouraou et al. (2025)	TRUE	Progression-free survival prediction; AUC ~0.72 using multiple CV folds.	NSCLC cohort	France/International
Salmanpour et al. (2025)	TRUE	Semi-supervised radiomics improved survival prediction; external accuracy ~0.88.	977 patients across 12 datasets	Multi-center
Mali et al. (2025)	TRUE	Harmonized multi-region radiomics + foundation models; C-index ~0.76, t-AUC ~0.92 with consensus model.	876 NSCLC patients from 5 centers	Multi-center
Prasad et al. (2024)	TRUE	95% accuracy; 90% specificity; 87% sensitivity.	LUNA 16	Europe
D'Ambrosi et al. (2023)	TRUE	Area under the curve (AUC) of 0.88 and 0.81 for mRNA and circRNA.	mRNA and circRNA	Switzerland

Table 4: Continued

Ref	Open Access	Results Obtained	Data Set	Country
Moitra and Mandal (2020)	FALSE	Average ROC-AUC is 0.94.	The Cancer Imaging Archive (TCIA)	United Kingdom
Zheng et al. (2023)	FALSE	Six CT radiomics features correlated significantly with progression-free survival (HR 4.531, 95% CI 3.524-5.825, p<0.001). Integration into the nomogram improved PFS estimation (C-index 0.799), yielding robust AUCs (0.885 at 6 months, 0.846 at 12 months). Enhanced classification accuracy was demonstrated by net reclassification improvement (33.7%, p<0.05) and integrated discrimination improvement (22.7%, p<0.05).	The study utilizes a local dataset comprising 558 patients sourced from three distinct medical centers.	Italy
Anderson et al. (2023)	FALSE	The model achieved a gamma pass rate exceeding 90% for both 3%/3mm and 5%/3mm criteria, with a rapid 20-millisecond prediction speed.	193 IMRT fields/images from 118 patients	Netherlands
Liu et al. (2023)	TRUE	AUC values exceeding 0.914, with high sensitivity and specificity. The validation group showed AUC values above 0.732, maintaining good sensitivity and specificity.	A cohort of 416 patients with pathologically confirmed lesions was taken in the form of CT scan images.	Hong Kong
Challab and Mardukhi (2023)	TRUE	99.9487% accuracy; 99.9485% specificity; 99.9485% sensitivity; and 99.8787% F1-score.	4575 CT scan images (1525 COVID-19)	Iran
Donga et al. (2022)	TRUE	Precision = 0.957; Recall = 0.91; F1 = 0.941; Validation accuracy = 95.67%.	Lung Image Database Consortium (LIDC-IDRI)	India
Oshita et al. (2022)	TRUE	Accuracy - 80%.	CT images from 173 lung cancer patients were utilized, focusing on the slice with the largest tumor size.	Japan
Saygili et al. (2025)	FALSE	Study findings highlight baseline disease burden as the strongest predictor of survival (c-index = 0.68). Follow-up burden (c-index = 0.65) and lung lesion count (c-index = 0.62) also predict survival. TRAQinform Profile effectively distinguishes treatment responders (c-index = 0.76).	69 patients; median age 50; 67% female, 33% male; 80% metastatic, 59% lung, 16% localized, 4% NED.	United Kingdom
Liu et al. (2021)	TRUE	Six CT radiomics features correlated significantly with progression-free survival (HR 4.531, 95% CI 3.524-5.825, p<0.001). Integration into the nomogram improved PFS estimation (C-index 0.799), yielding robust AUCs (0.885 at 6 months, 0.846 at 12 months). Enhanced classification accuracy was demonstrated by net reclassification improvement (33.7%, p<0.05) and integrated discrimination improvement (22.7%, p<0.05).	Forty-six patients with previously treated NSCLC were taken.	Switzerland
Nakajo et al. (2022)	TRUE	AUC = 0.872; accuracy = 0.780; F1 score = 0.781; precision = 0.788; recall = 0.780.	50 patients; four clinical and 41 18F-FDG-PET-based radiomic features were ranked.	United States
Tsou et al. (2021)	FALSE	The model demonstrated superior performance with an accuracy, sensitivity, specificity, and area under the curve (AUC) of 0.89, 0.82, 0.94, and 0.95, respectively.	148 lung cancer patients and 168 healthy volunteers	Switzerland
Tang et al. (2021)	TRUE	Accuracy - 66%; Accuracy: 0.794; AUROC Achieved: 0.863; Algorithm Implemented: Linear Discriminant Analysis (LDA).	Contrast-enhanced and unenhanced CT images from 201 and 273 patients, respectively, were analyzed.	Switzerland
Han et al. (2021)	FALSE	Accuracy: 0.792; AUROC Achieved: 0.863; Algorithm Implemented: Support Vector Machine.	A total of 867 ADC and 552 SCC patients were analyzed.	Germany
Carvalho et al. (2025)	FALSE	Texture features exhibit superior association with lung function and COPD diagnosis compared to densitometric measures (p < 0.001). Texture predicts rapid lung function decline uniquely (AUC = 0.538, p < 0.05).	1915 low-dose lung CT scans were analyzed.	United States
Zhang et al. (2021)	FALSE	The model, evaluated on the LUNGx Challenge database, attains high sensitivity (0.887) and specificity (0.924), with an AUC of 0.948.	73 lung nodules (37 benign and 36 malignant)	Netherlands
Zhao et al. (2019)	TRUE	The 3D CNN system, 3D DenseNets achieved AUCs of 75.8% and 75.0% respectively.	The retrospective analysis involved a dataset comprising 579 nodules with EGFR mutation status labels, categorized as either mutant (Mut) or wild-type (WT).	United Kingdom

Table 4: Continued

Ref	Open Access	Results Obtained	Data Set	Country
Sattar and Majid (2019)	TRUE	0.8187 AUC; 0.8261 Sn; 0.8115 Sp; 0.8191 G-Mean; 0.8188 F-Score.	865 Images	Saudi Arabia
Shiri et al. (2019)	TRUE	PET + CT radiomics predicted EGFR/KRAS mutation status; AUC up to ~0.82-0.83	211 NSCLC patients	USA (mixed)

Discussion

A review of the recent literature in Tables 2, 3, and 4 shows consistent trends in using Convolutional Neural Networks (CNNs) and ensemble learning techniques as the most effective methods for detecting lung cancer. This is especially true for CT imaging datasets like LUNA16, LIDC-IDRI, and NSCLC-Radiogenomics. CNN-based models often achieved AUC values above 0.90, showing better performance in image classification tasks. Notably, hybrid deep learning frameworks and radiomics-based nomograms have produced promising results by combining clinical and imaging data for better prognostic prediction. Studies that used larger, open-access datasets usually reported more reproducible and clinically relevant outcomes. In contrast, models trained on limited or proprietary datasets showed more performance variation. Despite these advancements, the absence of standardized testing and inconsistent reporting of evaluation metrics across studies make direct comparisons between models difficult. This highlights the need for unified validation methods and increased transparency in developing algorithms and documenting datasets.

Result and Analysis

The research findings analyzed that lung cancer is one of the deadliest diseases, which leads to numerous harms on mankind. According to the WHO, Asia is one of the prominent areas for lung cancer for all the ages and both genders. Figure 6 presents the estimated global number of cancer cases (in millions) from 2020 to 2040, based on projections by the International Agency for Research on Cancer (IARC) for the World Health Organization (WHO).

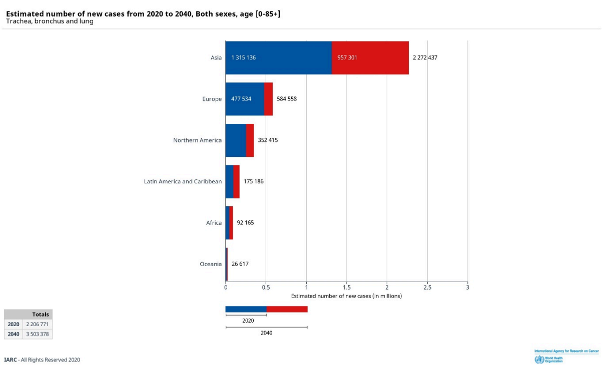

Fig. 6: Estimated number of new cases of Lung Cancer in millions worldwide

The integration of machine learning (ML) into medical image processing represents a pivotal strategy for enabling early lung cancer detection, with the ultimate goal of reducing disease incidence and mortality. This review synthesizes key findings that underscore the transformative potential of Artificial Intelligence (AI) in oncology.

Beyond detection, radiomics and deep learning show significant promise for improving prognosis and personalizing treatment. Zheng et al. (2023) identified six CT radiomics features with a strong correlation to progression-free survival (HR 4.531, 95% CI 3.524-5.825, *p* < 0.001). Integrating these into a nomogram improved prognostic estimation (C-index 0.799) and yielded high AUCs for predicting 6- and 12-month outcomes. Similarly, Anderson et al. (2023) developed a deep-learning model for Intensity-Modulated Radiation Therapy (IMRT) that predicted transit images with >90% gamma pass rate and a 20-millisecond latency, highlighting its potential for real-time, adaptive radiotherapy.

In conclusion, the convergence of machine learning algorithms, deep learning architectures, and radiomic feature analysis is markedly advancing the landscape of lung cancer care. These technologies enhance not only early detection but also prognostic stratification and treatment planning. Future work must focus on the multi-center validation of these models, standardization of feature extraction, and rigorous clinical trials to translate these promising computational tools into routine practice.

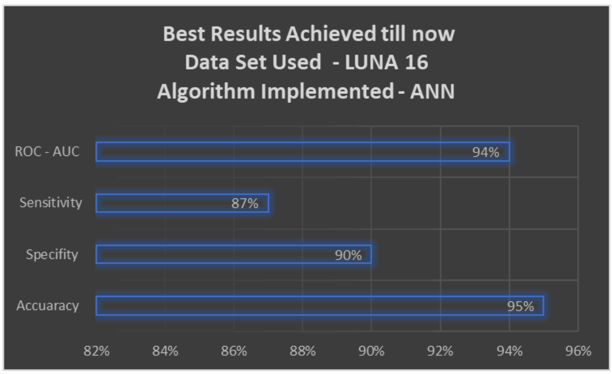

Fig. 7: Result Analysis - Algorithm, Data Set and Results

The results of the literature survey are summarized graphically in Figure 7. A key finding is that an Artificial Neural Network (ANN) applied to the LUNA16 dataset yielded the highest accuracy (95%) among the reviewed studies. This performance benchmark underscores the considerable potential of advanced machine learning models to augment diagnostic precision, which could translate into better patient management and clinical results.

Future Prospects

The next significant advancement in AI-driven lung cancer detection lies at the convergence of emerging technologies: Vision Transformers (ViTs), Explainable AI (XAI), and federated learning. Vision Transformers, which utilize attention mechanisms to model long-range dependencies within medical images, have demonstrated superior performance to conventional Convolutional Neural Networks (CNNs) in recent imaging benchmarks. Their application to lung cancer screening, particularly across heterogeneous and multimodal datasets, holds strong potential to improve diagnostic accuracy and subtle lesion detection.

The adoption of Explainable AI (XAI) is critical for fostering clinical trust and satisfying regulatory requirements. Techniques such as saliency maps, Grad-CAM, and SHAP are increasingly employed to interpret model predictions and highlight decisive image regions. This transparency does more than demystify the "black box"; it actively supports clinical decision-making by correlating model outputs with recognizable radiological features.

Furthermore, federated learning presents a transformative framework for multi-institutional collaboration. By enabling model training on decentralized datasets without sharing sensitive

Conclusion

Deep learning and machine learning have the potential to greatly improve patient outcomes and clinical practice by finding and predicting lung cancer. Our research underscores the capacity of computational methodologies to enhance the diagnostic accuracy, prognostic assessment, and individualized treatment approaches for lung cancer patients. By using a wide range of datasets and advanced algorithms, clinicians can learn a lot about how diseases progress and tailor treatments to fit the needs of each patient. Nonetheless, further investigation is necessary to ascertain the clinical significance and dependability of these computational techniques in real-world settings. Researchers, clinicians, and data scientists must continue to work together in order to move computational oncology forward and use these new ideas in clinical practice. The use of artificial intelligence and deep learning in lung cancer management could change the way patients are treated by allowing for earlier detection, more accurate prognosis prediction, and better treatment outcomes. Thus, it is essential to dedicate resources to research and development in this domain to fully exploit the capabilities of computational methods in lung cancer treatment, thereby improving patient quality of life and survival rates.

Funding Information

No specific grant was awarded to this research by any funding agency in the public, commercial, or not-for-profit sectors.

Author's Contributions

Aanchal Vij: Conceived and designed the study, conducted the review, and prepared the figures and tables.

Kuldeep Singh Kaswan: Provided critical insight into the selection and categorization of machine learning techniques and guided the overall structure of the review.

Anand Nayyar: Oversaw the research direction and critically reviewed the manuscript for intellectual content.

Ethics

The authors have not conducted any novel studies with human participants or animals, and this article is a review of previously published literature.

References

Alduais, Y., Zhang, H., Fan, F., Chen, J., & Chen, B. (2023). Non-small cell lung cancer (NSCLC): A review of risk factors, diagnosis, and treatment. *Medicine*, *102*(8), e32899. https://doi.org/10.1097/md.0000000000032899

Almahdi, Y. A., Schwark, E. R., Mclaughlin, A. J., & Xhabija, B. (2025). Dissecting Morphological and Functional Dynamics of Non-Tumorigenic and Triple-Negative Breast Cancer Cell Lines Using PCA and t-SNE Analysis. *Cancer Reports*, *8*(7). https://doi.org/10.1002/cnr2.70257

Alsaadi, E. M. T. A., & Rahman, Z. H. A. A. (2023). Retraction: Automatic lung cancer recognition in chest CT-scan images using SVM classifier. *AIP Conference Proceedings*, 040033. https://doi.org/10.1063/5.0182244

Anderson, B. M., Moore, L., Moore, K. L., & Bojechko, C. (2023). EPIDEEP: Using a Deep Learning Model to Predict In Vivo Electronic Portal Imaging Device (EPID) Transit Images. *International Journal of Radiation Oncology*Biology*Physics*, *117*(2), e645. https://doi.org/10.1016/j.ijrobp.2023.06.2060

Ansari, M. M., Kumar, S., Heyat, M. B. B., Ullah, H., Bin Hayat, M. A., Sumbul, Parveen, S., Ali, A., & Zhang, T. (2025). SVMVGGNet-16: A Novel Machine and Deep Learning Based Approaches for Lung Cancer Detection using Combined SVM and VGGNet-16. *Current Medical Imaging Formerly Current Medical Imaging Reviews*, *21*, 1-19. https://doi.org/10.2174/0115734056348824241224100809

Birnboim-Perach, R., & Benhar, I. (2024). Using Combination therapy to overcome diverse challenges of Immune Checkpoint Inhibitors treatment. *International Journal of Biological Sciences*, *20*(10), 3911-3922. https://doi.org/10.7150/ijbs.93697

Bojer, C. S., & Meldgaard, J. P. (2021). Kaggle forecasting competitions: An overlooked learning opportunity. *International Journal of Forecasting*, *37*(2), 587-603. https://doi.org/10.1016/j.ijforecast.2020.07.007

Bove, S., Fanizzi, A., Fadda, F., Comes, M. C., Catino, A., Cirillo, A., Cristofaro, C., Montrone, M., Nardone, A., Pizzutilo, P., Tufaro, A., Galetta, D., & Massafra, R. (2023). A CT-based transfer learning approach to predict NSCLC recurrence: The added-value of peritumoral region. *PLOS ONE*, *18*(5), e0285188. https://doi.org/10.1371/journal.pone.0285188

Bradley, S. H., Abraham, S., Callister, M. E., Grice, A., Hamilton, W. T., Lopez, R. R., Shinkins, B., & Neal, R. D. (2019). Sensitivity of chest X-ray for detecting lung cancer in people presenting with symptoms: a systematic review. *British Journal of General Practice*, *69*(689), e827-e835. https://doi.org/10.3399/bjgp19x706853

Carvalho, A. R. S., Guimarães, A., Basilio, R., Conrado da Silva, M. A., Colli, S., Galhós de Aguiar, C., Pereira, R. C., Lisboa, L. G., Hochhegger, B., & Rodrigues, R. S. (2025). Automatic Quantification of Abnormal Lung Parenchymal Attenuation on Chest Computed Tomography Images Using Densitometry and Texture-based Analysis. *Journal of Thoracic Imaging*, *40*(2), e0804. https://doi.org/10.1097/rti.0000000000000804

Challab, J. M., & Mardukhi, F. (2023). Ant Colony Optimization-Rain Optimization Algorithm Based on Hybrid Deep Learning for Diagnosis of Lung Involvement in Coronavirus Patients. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, *47*(3), 887-902. https://doi.org/10.1007/s40998-023-00611-y

Chen, C., Bai, C., Lee, K., Chou, Y., Pan, S., & Wang, Y. (2021). Evaluation of the diagnostic accuracy of bronchial brushing cytology in lung cancer: A meta-analysis. *Cancer Cytopathology*, *129*(9), 739-749. https://doi.org/10.1002/cncy.22436

Chen, C., Tian, X., Zhao, X., & Ren, L. (2020). Clinical study of serum IGFBP7 in predicting lymphatic metastasis in patients with lung adenocarcinoma. *Current Problems in Cancer*, *44*(6), 100584. https://doi.org/10.1016/j.currproblcancer.2020.100584

Chen, H., Yumoto, K., Kashizaki, F., Koizumi, H., Ikeda, I., Horita, N., Takahashi, K., & Kaneko, T. (2023). KEYNOTE-407: an effective and safe first-line treatment option for metastatic squamous non-small cell lung cancer. *Translational Lung Cancer Research*, *12*(8), 1830-1833. https://doi.org/10.21037/tlcr-23-271

Cords, L., Engler, S., Haberecker, M., Rüschoff, J. H., Moch, H., de Souza, N., & Bodenmiller, B. (2024). Cancer-associated fibroblast phenotypes are associated with patient outcome in non-small cell lung cancer. *Cancer Cell*, *42*(3), 396-412. https://doi.org/10.1016/j.ccell.2023.12.021

D'Ambrosi, S., Giannoukakos, S., Antunes-Ferreira, M., Pedraz-Valdunciel, C., Bracht, J. W. P., Potie, N., Gimenez-Capitan, A., Hackenberg, M., Fernandez Hilario, A., Molina-Vila, M. A., Rosell, R., Würdinger, T., & Koppers-Lalic, D. (2023). Combinatorial Blood Platelets-Derived circRNA and mRNA Signature for Early-Stage Lung Cancer Detection. *International Journal of Molecular Sciences*, *24*(5), 4881. https://doi.org/10.3390/ijms24054881

Di Capua, D., Bracken-Clarke, D., Ronan, K., Baird, A.-M., & Finn, S. (2021). The Liquid Biopsy for Lung Cancer: State of the Art, Limitations and Future Developments. *Cancers*, *13*(16), 3923. https://doi.org/10.3390/cancers13163923

Donga, H. V., Karlapati, J. S. A. N., Desineedi, H. S. S., Periasamy, P., & TR, S. (2022). Effective Framework for Pulmonary Nodule Classification from CT Images Using the Modified Gradient Boosting Method. *Applied Sciences*, *12*(16), 8264. https://doi.org/10.3390/app12168264

Dubey, S., & Sikarwar, S. S. (2025). Applications of AI in Cancer Detection - A Review of the Specific Ways in which AI Is Being Used to Detect and Diagnose Various Types of Cancer. *AI in Disease Detection: Advancements and Applications*. https://doi.org/10.1002/9781394278695.ch7

Ermakov, M. S., Kashofer, K., & Regauer, S. (2023). Different Mutational Landscapes in Human Papillomavirus-Induced and Human Papillomavirus-Independent Invasive Penile Squamous Cell Cancers. *Modern Pathology*, *36*(10), 100250. https://doi.org/10.1016/j.modpat.2023.100250

Giaccone, G., & He, Y. (2023). Current knowledge of small cell lung cancer transformation from non-small cell lung cancer. *Seminars in Cancer Biology*, *94*, 1-10. https://doi.org/10.1016/j.semcancer.2023.05.006

HaghighiKian, S. M., Shirinzadeh-Dastgiri, A., Vakili-Ojarood, M., Naseri, A., Barahman, M., Saberi, A., Rahmani, A., Shiri, A., Masoudi, A., Aghasipour, M., Shahbazi, A., Ghelmani, Y., Aghili, K., & Neamatzadeh, H. (2025). A Holistic Approach to Implementing Artificial Intelligence in Lung Cancer. *Indian Journal of Surgical Oncology*, *16*(1), 257-278. https://doi.org/10.1007/s13193-024-02079-6

Hamza, M. N., Tariqul Islam, M., Lavadiya, S., ud Din, I., Cavalcante de Souza Sanches, B., Koziel, S., Iffat Naqvi, S., Farmani, A., & Islam, Md. S. (2024). Development of a Terahertz Metamaterial Micro-Biosensor for Ultrasensitive Multispectral Detection of Early Stage Cervical Cancer. *IEEE Sensors Journal*, *24*(20), 32065-32079. https://doi.org/10.1109/jsen.2024.3447728

Han, Y., Ma, Y., Wu, Z., Zhang, F., Zheng, D., Liu, X., Tao, L., Liang, Z., Yang, Z., Li, X., Huang, J., & Guo, X. (2021). Histologic subtype classification of non-small cell lung cancer using PET/CT images. *European Journal of Nuclear Medicine and Molecular Imaging*, *48*(2), 350-360. https://doi.org/10.1007/s00259-020-04771-5

Huang, S., Arpaci, I., Al-Emran, M., Kılıçarslan, S., & Al-Sharafi, M. A. (2023). A comparative analysis of classical machine learning and deep learning techniques for predicting lung cancer survivability. *Multimedia Tools and Applications*, *82*(22), 34183-34198. https://doi.org/10.1007/s11042-023-16349-y

Ito, Y., Nakajima, T., Inage, T., Otsuka, T., Sata, Y., Tanaka, K., Sakairi, Y., Suzuki, H., & Yoshino, I. (2022). Prediction of Nodal Metastasis in Lung Cancer Using Deep Learning of Endobronchial Ultrasound Images. *Cancers*, *14*(14), 3334. https://doi.org/10.3390/cancers14143334

Jacobs, C., Setio, A. A. A., Scholten, E. T., Gerke, P. K., Bhattacharya, H., M. Hoesein, F. A., Brink, M., Ranschaert, E., de Jong, P. A., Silva, M., Geurts, B., Chung, K., Schalekamp, S., Meersschaert, J., Devaraj, A., Pinsky, P. F., Lam, S. C., van Ginneken, B., & Farahani, K. (2021). Deep Learning for Lung Cancer Detection on Screening CT Scans: Results of a Large-Scale Public Competition and an Observer Study with 11 Radiologists. *Radiology: Artificial Intelligence*, *3*(6). https://doi.org/10.1148/ryai.2021210027

Jiang, Y., Ebrahimpour, L., Després, P., & Manem, V. SK. (2025). A benchmark of deep learning approaches to predict lung cancer risk using national lung screening trial cohort. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-024-84193-7

Lantuejoul, S., Fernandez-Cuesta, L., Damiola, F., Girard, N., & McLeer, A. (2020). New molecular classification of large cell neuroendocrine carcinoma and small cell lung carcinoma with potential therapeutic impacts. *Translational Lung Cancer Research*, *9*(5), 2233-2244. https://doi.org/10.21037/tlcr-20-269

Lin, P., Lin, Y., Gao, R., Wan, W., He, Y., & Yang, H. (2023). Integrative radiomics and transcriptomics analyses reveal subtype characterization of non-small cell lung cancer. *European Radiology*, *33*(9), 6414-6425. https://doi.org/10.1007/s00330-023-09503-5

Liu, C., Gong, J., Yu, H., Liu, Q., Wang, S., & Wang, J. (2021). A CT-Based Radiomics Approach to Predict Nivolumab Response in Advanced Non-Small-Cell Lung Cancer. *Frontiers in Oncology*, *11*. https://doi.org/10.3389/fonc.2021.544339

Liu, R.-S., Ye, J., Yu, Y., Yang, Z.-Y., Lin, J.-L., Li, X.-D., Qin, T.-S., Tao, D.-P., Song, W., Wang, G., & Peng, J. (2023). The predictive accuracy of CT radiomics combined with machine learning in predicting the invasiveness of small nodular lung adenocarcinoma. *Translational Lung Cancer Research*, *12*(3), 530-546. https://doi.org/10.21037/tlcr-23-82

Liu, X., Xi, X., Xu, S., Chu, H., Hu, P., Li, D., Zhang, B., Liu, H., Jiang, T., & Lu, Z. (2024). Targeting T cell exhaustion: emerging strategies in non-small cell lung cancer. *Frontiers in Immunology*, *15*. https://doi.org/10.3389/fimmu.2024.1507501

Lu, J., Liu, X., Ji, X., Jiang, Y., Zuo, A., Guo, Z., Yang, S., Peng, H., Sun, F., & Lu, D. (2025). Predicting PD-L1 status in NSCLC patients using deep learning radiomics based on CT images. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-91575-y

Mali, S. A., Salahuddin, Z., Khan, D., Zhang, Y., Woodruff, H. C., Ibor-Crespo, E., & Lambin, P. (2026). Consensus in the parliament of AI: Harmonized multi-region CT-radiomics and foundation-model signatures for multicentre NSCLC risk stratification. *ArXiv*. https://doi.org/10.48550/arXiv.2505.17893

Moitra, D., & Kr. Mandal, R. (2020). Classification of non-small cell lung cancer using one-dimensional convolutional neural network. *Expert Systems with Applications*, *159*, 113564. https://doi.org/10.1016/j.eswa.2020.113564

Nafea, A. A., Alameri, S. A., Majeed, R. R., Khalaf, M. A., & AL-Ani, M. M. (2024). A Short Review on Supervised Machine Learning and Deep Learning Techniques in Computer Vision. *Babylonian Journal of Machine Learning*, *2024*, 48-55. https://doi.org/10.58496/bjml/2024/004

Nakajo, M., Jinguji, M., Tani, A., Yano, E., Hoo, C. K., Hirahara, D., Togami, S., Kobayashi, H., & Yoshiura, T. (2022). Machine learning based evaluation of clinical and pretreatment 18F-FDG-PET/CT radiomic features to predict prognosis of cervical cancer patients. *Abdominal Radiology*, *47*(2), 838-847. https://doi.org/10.1007/s00261-021-03350-y

Oshita, Y., Takeuchi, N., Teramoto, A., Kondo, M., Imaizumi, K., Saito, K., & Fujita, H. (2022). Prognosis Prediction of Lung Cancer Patients Using CT Images: Feature Extraction by Convolutional Neural Network and Prediction by Machine Learning. *Japanese Journal of Radiological Technology*, *78*(8), 829-837. https://doi.org/10.6009/jjrt.2022-1224

Ouraou, É., Tonneau, M., Le, W., Filion, É., Campeau, M.-P., Vu, T., Doucet, R., Bahig , H., & Kadoury, S. (2025). Predicting early stage lung cancer recurrence and survival. *Wiley Online Library*, *52*(10), e17586. https://doi.org/10.1002/mp.17586

Peng, Z., Wang, Y., Qi, Y., Hu, H., Fu, Y., Li, J., Li, W., Li, Z., Guo, W., Shen, C., Jiang, J., & Yang, B. (2025). Application of prediction model based on ct radiomics in prognosis of patients with non-small cell lung cancer. *Frontiers in Oncology*, *15*(12329999), 1273. https://doi.org/10.1186/s12885-025-14544-8

Prasad, U., Chakravarty, S., & Mahto, G. (2024). Lung cancer detection and classification using deep neural network based on hybrid metaheuristic algorithm. *Soft Computing*, *28*(15-16), 8579-8602. https://doi.org/10.1007/s00500-023-08845-y

Salmanpour, A. (2025). *Robust semi-supervised ct radiomics for lung cancer prognosis*. https://doi.org/10.48550/arXiv.2507.08189

Sattar, M., & Majid, A. (2019). Lung Cancer Classification Models Using Discriminant Information of Mutated Genes in Protein Amino Acids Sequences. *Arabian Journal for Science and Engineering*, *44*(4), 3197-3211. https://doi.org/10.1007/s13369-018-3468-8

Saw, S. N., Yan, Y. Y., & Ng, K. H. (2025). Current status and future directions of explainable artificial intelligence in medical imaging. *European Journal of Radiology*, *183*, 111884. https://doi.org/10.1016/j.ejrad.2024.111884

Saygili, E. S., Elhassan, Y. S., Prete, A., Lippert, J., Altieri, B., & Ronchi, C. L. (2025). Machine learning-based survival prediction tool for adrenocortical carcinoma. *The Journal of Clinical Endocrinology & Metabolism*, 096. https://doi.org/10.1210/clinem/dgaf096

Sherin, K., Jaspin, K., S, H., Sivanantham, S., Senthilvelan, S., & Theerthagiri, S. (2023). Optimising Lung Cancer Screening Through Advanced Techniques for Image Analysis: An Efficient Deep Learning Approach. *2023 4th International Conference on Smart Electronics and Communication (ICOSEC)*, 1528-1535. https://doi.org/10.1109/icosec58147.2023.10275884

Shiri, I., Maleki, H., Hajianfar, G., Abdollahi, H., Ashrafinia, S., Hatt, M., Zaidi, H., Oveisi, M., & Rahmim, A. (2020). Next-Generation Radiogenomics Sequencing for Prediction of EGFR and KRAS Mutation Status in NSCLC Patients Using Multimodal Imaging and Machine Learning Algorithms. *Molecular Imaging and Biology*, *22*(4), 1132-1148. https://doi.org/10.1007/s11307-020-01487-8

Singh, H., & Kaswan, K. S. (2016). Clinical decision support systems for heart disease using data mining approach. *International Journal of Computer Science and Software Engineering*, *5*(2), 19-23.

Solomon, B. J., Bauer, T. M., Mok, T. S. K., Liu, G., Mazieres, J., de Marinis, F., Goto, Y., Kim, D.-W., Wu, Y.-L., Jassem, J., López, F. L., Soo, R. A., Shaw, A. T., Polli, A., Messina, R., Iadeluca, L., Toffalorio, F., & Felip, E. (2023). Efficacy and safety of first-line lorlatinib versus crizotinib in patients with advanced, ALK-positive non-small-cell lung cancer: updated analysis of data from the phase 3, randomised, open-label CROWN study. *The Lancet Respiratory Medicine*, *11*(4), 354-366. https://doi.org/10.1016/s2213-2600(22)00437-4

Sonthalia, A., Josephin JS, F., Aloui, F., & Varuvel, E. G. (2025). Evaluating the Effectiveness of Boosting and Bagging Ensemble Techniques in Forecasting Lithium-Ion Battery Useful Life. *Energy Storage*, *7*(1). https://doi.org/10.1002/est2.70118

Sujitha, R., & Seenivasagam, V. (2021). RETRACTED ARTICLE: Classification of lung cancer stages with machine learning over big data healthcare framework. *Journal of Ambient Intelligence and Humanized Computing*, *12*(5), 5639-5649. https://doi.org/10.1007/s12652-020-02071-2

Tan, S. L., Selvachandran, G., Paramesran, R., & Ding, W. (2025). Lung Cancer Detection Systems Applied to Medical Images: A State-of-the-Art Survey. *Archives of Computational Methods in Engineering*, *32*(1), 343-380. https://doi.org/10.1007/s11831-024-10141-3

Tang, X., Li, Y., Yan, W., Qian, W., Pang, T., Gong, Y., & Yang, Z. (2021). Machine Learning-Based CT Radiomics Analysis for Prognostic Prediction in Metastatic Non-Small Cell Lung Cancer Patients With EGFR-T790M Mutation Receiving Third-Generation EGFR-TKI Osimertinib Treatment. *Frontiers in Oncology*, *11*. https://doi.org/10.3389/fonc.2021.719919

Tiwari, A., Srivastava, S., & Pant, M. (2020). Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019. *Pattern Recognition Letters*, *131*, 244-260. https://doi.org/10.1016/j.patrec.2019.11.020

Tsou, P.-H., Lin, Z.-L., Pan, Y.-C., Yang, H.-C., Chang, C.-J., Liang, S.-K., Wen, Y.-F., Chang, C.-H., Chang, L.-Y., Yu, K.-L., Liu, C.-J., Keng, L.-T., Lee, M.-R., Ko, J.-C., Huang, G.-H., & Li, Y.-K. (2021). Exploring Volatile Organic Compounds in Breath for High-Accuracy Prediction of Lung Cancer. *Cancers*, *13*(6), 1431. https://doi.org/10.3390/cancers13061431

Vij, A., & Kaswan, K. S. (2023). Prediction of Lung Cancer using Convolution Neural Networks. *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*, 737-741. https://doi.org/10.1109/aisc56616.2023.10085058

Vikas, B., Makkapati, S. S., Bogireddy, S. R., Balamurugan, K. S., & Deepa, M. (2024). Advancements in Lung Cancer Diagnosis: A Comprehensive Study on the Role of PCA, LDA, and t-SNE in Deep Learning Frameworks. *2024 Asian Conference on Communication and Networks (ASIANComNet)*, 1-7. https://doi.org/10.1109/asiancomnet63184.2024.10811058

Wang, H., Hong, Y., Zhang, Z., Cheng, K., Chen, B., & Zhang, R. (2025a). Study on postoperative survival prediction model for non-small cell lung cancer. *Journal of Thoracic Disease*, *17*(2), 456-468.

Wang, X., Qiao, Z., Aramini, B., Lin, D., Li, X., & Fan, J. (2023). Potential biomarkers for immunotherapy in non-small-cell lung cancer. *Cancer and Metastasis Reviews*, *42*(3), 661-675. https://doi.org/10.1007/s10555-022-10074-y

Wang, X., Sharpnack, J., & Lee, T. C. M. (2025b). Improving lung cancer diagnosis and survival prediction with deep learning and CT imaging. *PLOS One*, *20*(6), e0323174. https://doi.org/10.1371/journal.pone.0323174

Wang, X., Zhang, A., Yang, H., Zhang, G., Ma, J., Ye, S., & Ge, S. (2025c). Multicenter development of a deep learning radiomics and dosiomics nomogram to predict radiation pneumonia risk in non-small cell lung cancer. *Scientific Reports*, *15*(1), 17106. https://doi.org/10.1038/s41598-025-02045-4

Wason, J. V., & Nagarajan, A. (2019). Image processing techniques for analyzing CT scan images towards the early detection of lung cancer. *Bioinformation*, *15*(8), 596-599. https://doi.org/10.6026/97320630015596

Widodo, S., Rosyid, I., Faizuddin, M., & Bin, R. (2020). Texture feature extraction to improve accuracy of malignant and benign cancer detection on ct scan images. *International Journal of Psychosocial Rehabilitation*, *24*(9), 3540-3553.

Xia, M., Xu, T., & Jiang, H. (2022). Progress and Perspective of Artificial Intelligence and Machine Learning of Prediction in Anesthesiology. *Journal of Shanghai Jiaotong University (Science)*, *27*(1), 112-120. https://doi.org/10.1007/s12204-021-2331-3

Xu, Y., Liu, H., Chen, J., & Zhou, Q. (2010). Acquired resistance of lung adenocarcinoma to EGFR-tyrosine kinase inhibitors gefitinib and erlotinib. *Cancer Biology & Therapy*, *9*(8), 572-582. https://doi.org/10.4161/cbt.9.8.11881

Yang, N., Ma, Z. X., Wang, X., Xiao, L., Jin, L., & Li, M. (2025). Development and validation of a CT-based radiomics nomogram for predicting progression-free survival in patients with small cell lung cancer. *BMC Medical Imaging*, *25*(1), 154. https://doi.org/10.1186/s12880-025-01691-4

Yaqoob, A., Musheer Aziz, R., & verma, N. K. (2023). Applications and Techniques of Machine Learning in Cancer Classification: A Systematic Review. *Human-Centric Intelligent Systems*, *3*(4), 588-615. https://doi.org/10.1007/s44230-023-00041-3

Zhang, C., Li, J., Huang, J., & Wu, S. (2021). Computed Tomography Image under Convolutional Neural Network Deep Learning Algorithm in Pulmonary Nodule Detection and Lung Function Examination. *Journal of Healthcare Engineering*, *2021*, 1-9. https://doi.org/10.1155/2021/3417285

Zhao, W., Yang, J., Ni, B., Bi, D., Sun, Y., Xu, M., Zhu, X., Li, C., Jin, L., Gao, P., Wang, P., Hua, Y., & Li, M. (2019). Toward automatic prediction of *EGFR* mutation status in pulmonary adenocarcinoma with 3D deep learning. *Cancer Medicine*, *8*(7), 3532-3543. https://doi.org/10.1002/cam4.2233

Zheng, X., Liu, K., Li, C., Zhu, C., Gao, Y., Li, J., & Wu, X. (2023). A CT-based radiomics nomogram for predicting the progression-free survival in small cell lung cancer: a multicenter cohort study. *La Radiologia Medica*, *128*(11), 1386-1397. https://doi.org/10.1007/s11547-023-01702-w