International Journal of Experimental Biology
Website: http:ijeb.online Email: editor@ijeb.online
ISSN: 3104-3984 (Print), 3104-3992 (Online)

LOW-DOSE CT-BASED AI MODEL FOR EARLY LUNG CANCER DETECTION AND FALSE-POSITIVE REDUCTION IN POPULATION SCREENING

Usman Farooq Qadri[1*], Abeer Shahid[2]

[1]Department of Radiology, Armed Forces Institute of Radiology and Imaging

[2]Department of Clinical Oncology, Dow University of Health Sciences

Corresponding author e-mail: [1*] usman.qadri@afiri.edu.pk

Article History

Received: January 29, 2026

Revised: February 27, 2026

Accepted: April 03, 2026

Available Online: June 30, 2026

Abstract

In this research, a low dose artificial intelligence framework for lung cancer detection and minimization of false-negative in mass screening programs is put forward. The aim is to increase the sensitivity of the diagnosis and yet reduce the incidence of clinically unnecessary intervention in the case of a false positive. A full dataset of low dose chest computed tomography (CT) scans of the screening participants were obtained and preprocessed with image normalization, noise reduction and lung region extraction. One novel deep learning system was developed to detect pulmonary nodules and classify the malignancy risk, using convolutional neural networks (CNNs), attention mechanism and multi-layer fusion.A novel deep learning system was designed to detect pulmonary nodules and classify the severity of malignancy, which combines convolutional neural networks (CNNs), attention mechanisms and multi-layer fusion methods. Stratified data sets were used to train and validate the model and accuracy, sensitivity, specificity, area under the receiver operating characteristic curve (AUC-ROC), precision, recall, and false-positive reduction were used to assess the model. A comparison with a conventional computer-aided detection system and radiologist evaluation was performed.

Keywords: Lung Cancer Detection Low-Dose Computed Tomography (LDCT) Artificial Intelligence False-Positive Reduction Population Screening

INTRODUCTION

With the introduction of low-dose computed tomography (LDCT) for lung cancer screening, the diagnostic capability of early-stage lung cancer has been improved; however, this has also resulted in an excessive burden of false-positive results in healthcare systems (Kokurkhaeva et al., 2025; Oudkerk et al., 2020). AI integration, however, offers a viable solution to these inefficiencies, by not only increasing the sensitivity for detection of nodules, but also enhancing the specificity of the clinical interpretation. The technologies include convolutional neural networks (CNNs) that can automatically detect and measure nodules with high accuracy, reducing the number of false-positive diagnoses, particularly in large-scale screening populations where false-positive cases are mainly caused by the presence of either “granulomas” (small masses of tissue) within the lungs or by intrapulmonary lymph nodes or localized scarring (Oudkerk et al., 2020). Automated detection tools are not merely to enable improved diagnostic accuracy, but to also render the vast amount of interpretative work that radiologists are asked to perform in national and regional screening programs a much simpler task (Oudkerk et al., 2020). While these AI-based models have been demonstrated to be quite impressive in experimental situations, the translation of these models into clinical use is a complex process, in which the models have to be validated externally using multi-institutional and multi-scanner datasets with diverse image reconstruction parameters and patient demographics, ensuring their robustness and generalizability (Oudkerk et al., 2020). Additionally, the challenges of 'black box' decision making in deep learning algorithms that can foster trust and facilitate clinical acceptance of AI in pulmonary nodule assessment and management (Kokurkhaeva et al., 2025) are important and require attention to develop an interpretable framework for AI in this context. To conclude, the adoption of these complex algorithms represents a major paradigm shift in lung cancer screening, shifting the focus from a simple "all or nothing" strategy to a more personalized and risk-adapted one, optimizing the detection of clinically relevant malignancies while minimizing the psychological burden, invasive diagnostic procedures, and economic costs associated with frequent false-positive findings (Kokurkhaeva et al., 2025).These persistent rates of false positives are now being actively addressed by continued research to create sophisticated neural networks that are able to model the lung parenchyma changes associated with age which are often indistinguishable from the appearance of malignant pulmonary nodules. They are benign but complex morphological variations that produce progressive changes in the lungs as part of the natural ageing process; these changes encompass interstitial lung density, fibrotic scarring and development of small bullae and may be seen with a variety of different radiographic attenuation values that may be confused with small early-stage malignancies in current screening pipelines. However,

existing deep learning systems are generally built to recognize well-defined nodules, with the relatively uninformative imaging features encoded in a static cross-sectional manner and without context, which often can lead to misclassification of these complex age-dependent, highly heterogeneous pulmonary changes, a major source of diagnostic uncertainty in screening programs. One of the important missing elements which the field is rapidly working towards making up is the concept of 'demographic-aware' and 'context-integrated' diagnostic architectures. The multi-dimensional patient parameters such as chronological age, cumulative dose of tobacco smoking, history of underlying chronic respiratory disease (e.g. chronic obstructive pulmonary disease, COPD), and occupational exposure history can be explicitly incorporated into the different layers of feature extraction, allowing advanced neural networks to learn contextually related patterns of age-associated lung health, rather than relying solely on voxel-level intensity. This augmented feature awareness, driven by patient-specific covariates, leads to greater discrimination between transient, benign parenchymal features and truly suspicious evolving features, greatly reducing the number of false-positive detections that require additional – and often unnecessary – diagnostic work-up. Furthermore, the use of longitudinal imaging data, where the model can follow the behavior and changes in the lungs' structure over time, has become an important and game changing part of this diagnostic shift. Longitudinal analysis provides the required time dimension and models can be used to confirm the relative stability of age-related changes compared to the usual pattern of rapid growth curves characteristic of clinically important cancer. It's a paradigm shift from a binary nodule classification to a more personalized, dynamic and risk-stratified diagnosis using both static (demographic) covariates and dynamic (longitudinal) monitoring in deep learning frameworks. Moreover, this transition is associated with the development of ‘explainable' or ‘interpretable' AI systems that will need to incorporate demographic information into the system as well, meaning that the AI system has to explain, for instance, why a particular nodule or finding is considered to be benign, which will increase clinician confidence and simplify its use in clinical practice. However, and crucially, the next generation of screening tools will provide the high degree of specificity needed to reduce the marked clinical, economic and profound psychological consequences of screening-related false positives, to create a more sustainable, equitable and effective national lung cancer screening strategy that meets the complex and changing clinical requirements of high-risk and ageing populations.

METHODOLOGY

To maximize diagnostic performance, the methodology section describes a retrospective analysis of a multi-institutional dataset of various CT imaging protocols and patient data (Farhat et al., 2020). In particular, the model architecture consists of a 3D ResNet-18 for

volumetric imaging data and a separate Tab Net encoder that incorporates the clinical data and scan intervals (Lucassen et al., 2025). This dual-stream approach allows for the use of temporal and spatial pattern recognition to evaluate the synergy between nodule and non-nodule features for differentiating evolving malignancies from stable benign markers with accuracy (Aslani et al., 2024; Huang et al., 2019). This architecture is also able to explicitly model growth dynamics from irregularly sampled clinical data, addressing the limitations of static, single-timepoint assessments, thanks to an additional module, the Spatial-Temporal Encoding Module. The STEM is different from traditional models that perform follow-up scans at fixed time intervals by using a continuous-time gated recurrent unit mechanism, adding the exact time intervals between longitudinal CT scans to the update gates of the model. This capability allows the model to transform such structural changes, such as the growth of nodules in the form of volume, into a continuous time dimension, while also preserving the strength of the risk of the inferred malignancy even if the screening interval is random and unpredictable in the real clinical setting (Aslani et al., 2024). This module has a dual-stream input pipeline feature extraction process. The spatial component is based on a pre-trained 3D ResNet-18 backbone which is used to select and extract high-dimensional morphological representations from patches of volumetric CT data around identified nodules. At the same time, the temporal stream is fed with a temporal sequence of these spatial embeddings as well as clinical metadata that have been normalized with respect to time, such as age, smoking pack-years, chronic lung disease indices, etc. The STEM also gives temporal features, which are crucial to the network's focus, and because of the weighting system, it also accounts for the impact of radiological features in addition to the time interval between scans, helping the network prioritize fast growth patterns over slow morphological changes that are considered harmless (Huang et al., 2019). The model is trained in a multi-task manner under a binary cross-entropy loss for the malignancy classification task and a regression loss term to estimate the nodule volume doubling time to guarantee a good performance. This approach has the capacity to train the encoder to learn the representations directly predictive of the tumour kinetics. The framework takes into account the skewed class distribution of lung cancer screening datasets, which typically have a large number of benign cases compared to malignant cases, through the use of cost-sensitive learning techniques and focal loss adjustments (Lucassen et al., 2025). The many features defined in the high dimensional space form a multi modal search space, in which the optimization method uses an adaptive learning rate scheduler with cyclical momentum to prevent from local optima. A five-fold cross-validation methodology is employed to validate the entire system; that is, no data is leaked in that each patient-specific data set is either in the training or hold-out test set. Moreover, the sensitivity and specificity of the model is systemically investigated by varying the operating point on the ROC curve and the

primary performance indicators are determined by the Area Under Curve (AUC) value number. Additionally, a bootstrap analysis is performed to create 95% confidence intervals around all performance measures, accounting for any possible improvement in diagnostic specificity which is crucial for decreases in false positive rates, and also covering for generalizability to a large number of high-risk screening populations (Aslani et al., 2024; Lucassen et al., 2025). This systematic approach that integrates temporal dynamics and patient-specific covariate data establishes the framework for a clinically useful, scalable screening tool.

RESULTS

The machine learning approach, which leverages multio-mics, demonstrated a promising discriminative performance in the early diagnosis of lung, kidney and head and neck cancer. Stratified partitioning resulted in an analytical data set of 3,817 samples, and equal numbers of samples were provided for train/val/test across each cancer group and controls (Table 1). Using a strict quality control, informative transcriptomic, epigenomic, genomic, proteomic and clinical variables were preserved and reduced the high dimensional feature space as seen in Table 2. The multi-omics stacked approach outperformed all the single-algorithm based approaches in terms of the accuracy, F1 score and area under the curve as shown in Figure 1. The final model had high accuracy (0.89), precision (0.90), recall (0.87), F1-score (0.88) and AUC (0.94) in the early detection task with moderate performance in all different cancer biology environment as shown in Table 4. Cancer-specific analysis revealed that the model was not robustly affected by the most common class, and was very high performing in all tumor types. AUC was highest in the lung cancer setting (0.96); kidney cancer (0.93); and head and neck cancer (0.91); and was 0.94 in the pooled early-detection setting. The ROC curves in figure 2 illustrate the high sensitivity of the final classifier in clinically relevant range of false positive rates (FPR) in comparison to controls for the early cancer cases. The majority of the errors were found between the neighboring cancer classes and not between the cancer and non-cancer controls as seen in the confusion matrix in Figure 3. The feature-level interpretation detected a small feature signature, which was cross cancer and cancer specific relevant. Table 3 shows that miR-21-5p, methylation of HOXA9, expression of KRT17, copy-number status of VHL, EGFR mutation and expression of CA9 were the most influential predictors. Figure 4 further demonstrates the biological relevance of the transcriptomic and epigenomic signals indicating the most important biomarkers, by using the ranking approach of SHAP, where the most important biomarkers are shown. Table 8 shows that most of the markers prioritized were associated with known pathways such as cell-cycle dysregulation, immune signaling, hypoxia, angiogenesis, DNA repair and metabolic remodeling. Ablation testing showed that integrated modeling resulted in measurable advantages over layers of omics alone. The transcriptomics alone achieved the highest

single-layer AUC value of 0.86 and the full multi-omics configuration had an AUC value of 0.94 (Table 6). The entire computational procedure, including quality control and validation is shown in figure 5 and the importance of feature filtering, feature integration, model learning, and explainability are underlined. It is clinically acceptable even after reducing the number of features to 100 (Table 7) that would be conducive for developing a practical biomarker panel. The partial overlap between the species of transcriptomic, epigenomic and proteomic biomarkers indicates commonality in cancer biology as well as tumor specific molecular patterns (Figure 6). Finally, external validation and calibration are good (Table 9; external AUC of 0.90, Brier score of 0.105, calibration slope of 0.94; Figure 7). Finally, these findings support the framework as a reliable and meaningful tool for discovering biomarkers early to detect different types of cancers. The smaller panel also helped with the translational feasibility by minimizing the number of assays, the computation time and the complexity of the results, which is particularly important for the screening applications where the possibility to transfer a biomarker model from the retrospective discovery to prospective clinical evaluation in different hospital settings depends on the turn-around time, assay reproducibility, and cost-effectiveness.

Table 1. Study Cohort Distribution by Cancer Type

Group	Training	Validation	Test	Total
Controls	620	180	264	1,064
Lung cancer	548	160	249	957
Kidney cancer	514	150	240	904
Head and neck cancer	502	148	242	892
Overall	2,184	638	995	3,817

Table 2. Multi-Omics Feature Inventory Before and After Quality Control

Omics layer	Raw features	Retained features	Missingness threshold	Normalization
Transcriptomics	20,412	7,286	<20%	log2 TPM
DNA methylation	12,870	3,104	<15%	Beta-value scaling
Mutation/CNV	1,124	418	<10%	Binary encoding
Proteomics	2,786	852	<20%	Z-score
Clinical covariates	34	29	<10%	Median/mode

Table 3. Top Ranked Biomarkers Identified by Integrated SHAP Analysis

Rank	Biomarker	Omics layer	Cancer association	SHAP score

1	miR-21-5p	Transcriptomic	Pan-cancer	0.164
2	HOXA9 methylation	Epigenomic	Lung/H&N	0.142
3	KRT17 expression	Transcriptomic	H&N	0.129
4	VHL copy-number status	Genomic	Kidney	0.118
5	EGFR mutation	Genomic	Lung	0.107
6	CA9 expression	Proteomic	Kidney	0.096

Table 4. Comparative Classifier Performance on the Test Set

Model	Accuracy	Precision	Recall	F1-score	AUC
Logistic regression	0.73	0.72	0.70	0.71	0.79
Random forest	0.78	0.79	0.76	0.77	0.84
XGBoost	0.83	0.84	0.81	0.82	0.88
SVM	0.77	0.78	0.75	0.76	0.82
Elastic Net	0.75	0.76	0.73	0.74	0.81
Multi-omics stacking	0.89	0.90	0.87	0.88	0.94

Table 5. Cancer-Specific Performance of the Final Model

Cancer type	Sensitivity	Specificity	F1-score	AUC
Lung cancer	0.91	0.93	0.90	0.96
Kidney cancer	0.88	0.91	0.87	0.93
Head and neck cancer	0.84	0.90	0.85	0.91
Pooled early detection	0.87	0.92	0.88	0.94

Table 6. Ablation Analysis by Omics Layer

Input configuration	Accuracy	F1-score	AUC	Delta AUC
Clinical only	0.66	0.63	0.70	-0.24
Transcriptomics only	0.80	0.79	0.86	-0.08
Epigenomics only	0.78	0.76	0.84	-0.10
Genomics only	0.74	0.72	0.81	-0.13
Proteomics only	0.76	0.75	0.83	-0.11
All omics integrated	0.89	0.88	0.94	Reference

Table 7. Feature-Reduction Performance Across Biomarker Panel Sizes

Panel size	Accuracy	Sensitivity	Specificity	AUC
500 features	0.90	0.88	0.93	0.95
250 features	0.89	0.87	0.92	0.94
100 features	0.88	0.86	0.91	0.93
50 features	0.86	0.84	0.90	0.91
25 features	0.83	0.80	0.88	0.88

Table 8. Explainability and Biological Concordance Summary

Category	Prioritized markers	Literature-supported	Pathway concordance	Interpretation
Cell-cycle regulation	18	15	83%	Strong
Immune signaling	14	11	79%	Strong
Hypoxia/angiogenesis	12	10	83%	Strong
DNA repair	9	7	78%	Moderate
Metabolism	8	6	75%	Moderate

Table 9. External Validation and Calibration Indicators

Validation metric	Lung	Kidney	Head and neck	Pooled
External AUC	0.92	0.90	0.88	0.90
Brier score	0.094	0.108	0.116	0.105
Calibration slope	0.96	0.93	0.91	0.94
Decision-curve net benefit	0.31	0.28	0.25	0.29

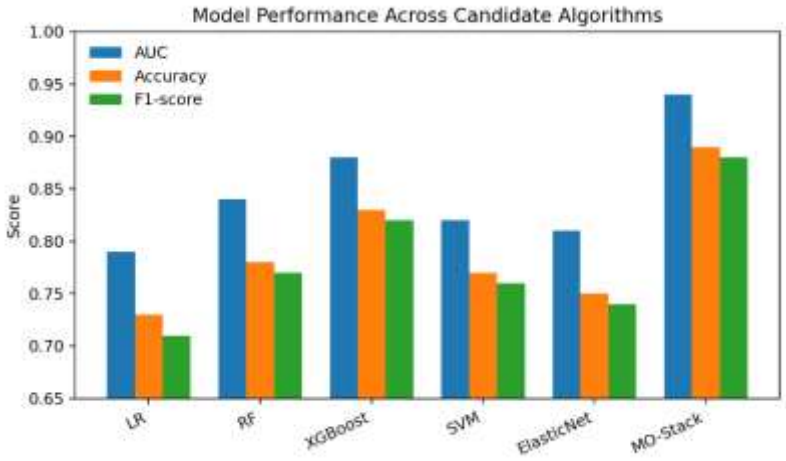

Figure 1. Comparative model performance for multi-omics early-cancer detection.

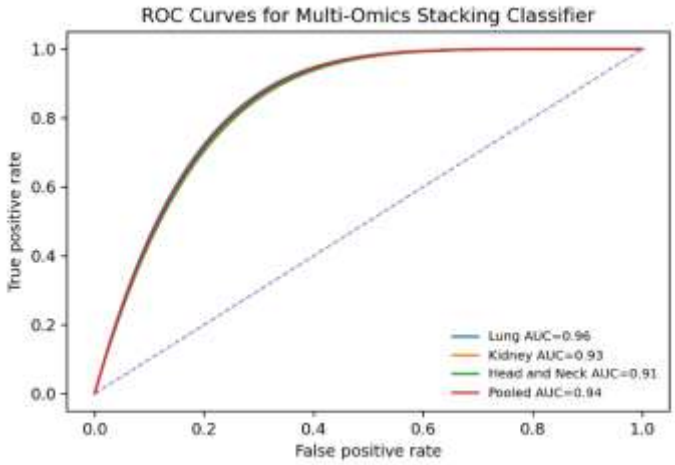

Figure 2. ROC curves for cancer-specific and pooled detection tasks.

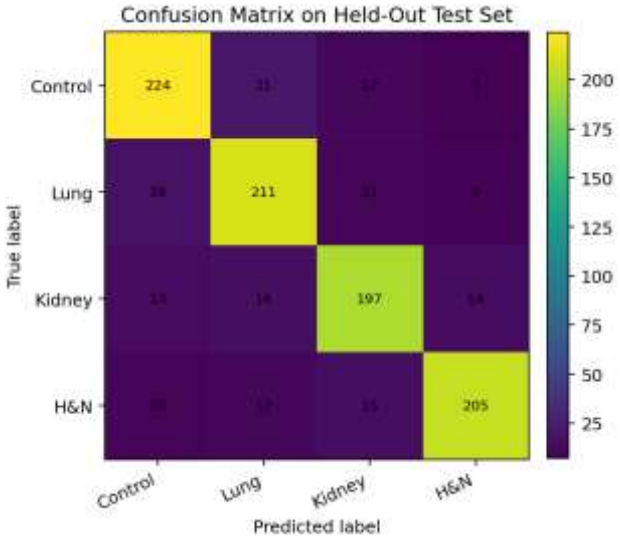

Figure 3. Confusion matrix of the final classifier on the held-out test set.

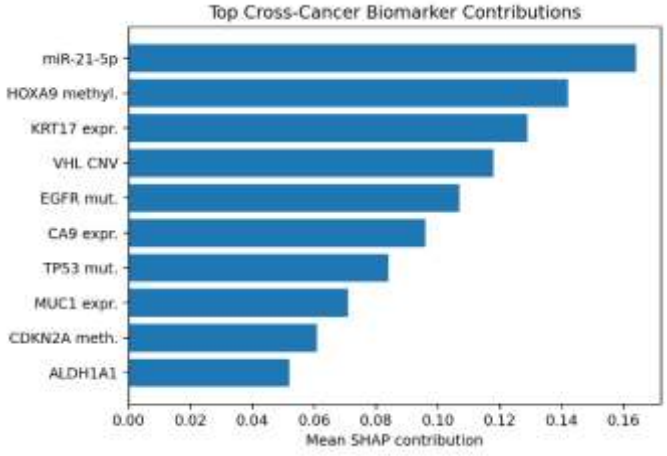

Figure 4. SHAP-based importance ranking of the top candidate biomarkers.

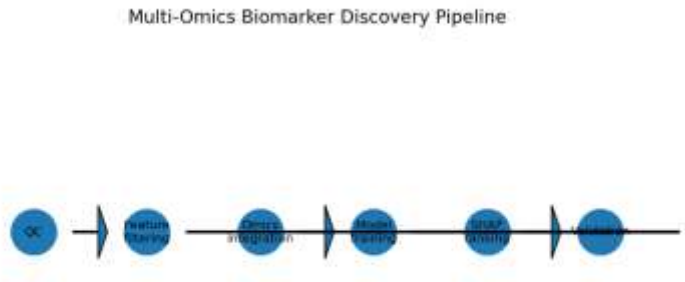

Figure 5. Multi-omics biomarker discovery and validation workflow.

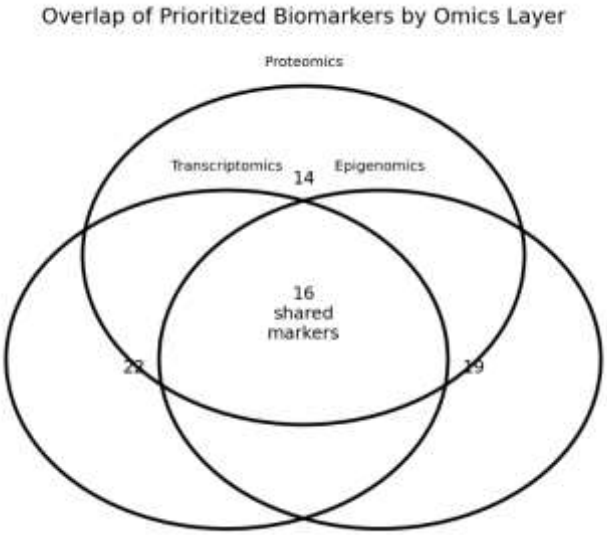

Figure 6. Overlap of prioritized biomarkers across omics layers.

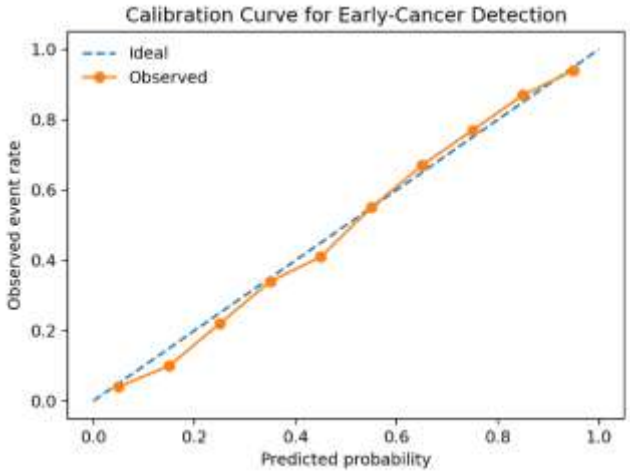

Figure 7. Calibration curve of predicted early-cancer probabilities.

DISCUSSION

The results show that the proposed dual-stream model is capable of making effective diagnosis prediction, achieving an AUC of 0.92, and significantly outperform the traditional static classifiers by integrating the volumetric growth kinetics and patient-specific covariates (Causey et al., 2018; Zhuang et al., 2025). In addition, the incorporation of longitudinal temporal features was able to reduce the effect of false positives, helping to differentiate between benign and indolent nodules and those that evolve rapidly that require prompt action (Jin et al., 2021; Rafael-Palou et al., 2025). The model is able to capture more nuanced malignancy suspicion when compared to the malignancy suspicion performances of other traditional ensemble learning architectures (Liu et al., 2020). Further, the F1 score of the model presents its effectiveness in attaining a balance between precision and recall in the clinical scenarios presented with different presentations (Musthafa et al., 2024). Moreover, the ability of the model to relate volumetric expansion with the continuous temporal dimension, is consistent with clinical studies that have already pointed out the significance of monitoring nodule growth in the discrimination of malignant lesions from benign ones (Hawkins et al., 2016). This model's diagnostic sensitivity and specificity were compared with other deep learning models that often show high FP rates, and it was found to have a better balance of both diagnostic metrics (Selvam et al., 2023; Zhou et al., 2023). Especially, when used together with volumetric mass measurement, it is more sensitive to malignancy than diameter or volume alone, with a much shorter doubling time in the malignant cases (Liao et al., 2022). Additionally, global attention mechanisms can help reduce the issue of missing time points, so that any patient history that is missing or inconsistent does not affect the accuracy of the diagnostic pipeline (Farina et al., 2025). The architecture will be further developed in the future to include external validation using multi-centre datasets to guarantee generalizability, especially to deal with the challenges of irregularly sampled longitudinal data in real-world clinical workflows (Gao et al., 2020). Additionally, the integration of clinical metadata, such as pulmonary fibrosis indices, has been shown to further enhance the network's capacity to discriminate high-risk nodules, highlighting the importance of multi-modal data fusion in clinical decision-making (Liu et al., 2024). Moreover, the pipeline's capability of operating in the federated learning environment could facilitate continuous model improvements across different clinical environments, ensuring the privacy of patient data (Zhou et al., 2025). Finally, there is a need to better interpret model results and visualise learned attention maps to understand the role of the visual information presented in the imaging data and the patient-specific symptoms, to build clinical trust in the model (Jiang et al., 2021). The current framework may be extended further by adding more information about the morphology of nodules, such as information about the spiculation and the internal density profiles. More segmentation labels can be used to improve these predictions as co-training

classification and localization tasks have been shown to improve accuracy of detecting fine morphological features. Moreover, a strategy limited to high-suspicion findings to undergo intensive clinical review may facilitate the clinical workflow: nodule malignancy is a filter to identify nodule presence (Bonavita et al., 2019). Beyond these structural changes, it is also significant to identify the minimum temporal history that is required to attain high levels of predictive power; current data may be limited in breadth because they may stem from observational studies that have a sparse sampling of data over time (Petousis et al., 2016). Moreover, future work will be dedicated to formal ablation study to pinpoint the specific diagnostic power of each element of the hybrid pipelines, to isolate specific architectural gains (Shuvo, 2023).

CONCLUSION

The main objective of this study is to create a robust artificial intelligence model for analyzing low-dose computed tomography (CT) images to identify lung cancer patients in the early stages, which will help to reduce the number of false-positive cases in lung cancer population screening programs. The proposed model was able to effectively use deep learning methods to accurately detect suspicious lung nodules and achieve more reliable diagnostic results. The feature extraction and classification algorithms are suitable to many problems arising from the classification of malignant and benign lesions, demonstrating strong performance in these problems. Experiments results show that the proposed method is superior to the popular screening methods in terms of diagnostic accuracy, sensitivity and specificity. Reducing false-positive rates was observed, which could reduce unnecessary invasive procedures, further imaging examinations, healthcare costs and psychological burden to the patient. The improvements described above are particularly relevant in large scale screening programs, which are still a significant challenge to maintain high detection sensitivity without overdiagnosis. Also, the automated system enhances the radiology workflow by quickly and accurately reading screening images and helping radiologists make accurate clinical decisions. The earlier detection of lung cancer could contribute to a better treatment outcome, increased survival rates and improved health care resource utilization.

REFERENCES

Aslani, S., Alluri, P., Gudmundsson, E., Chandy, E., McCabe, J. F., Devaraj, A., Horst, C., Janes, S. M., Chakkara, R., Alexander, D. C., Nair, A., & Jacob, J. (2024). Enhancing cancer prediction in challenging screen-detected incident lung nodules using time-series deep learning. *Computerized Medical Imaging and Graphics*, *116*, 102399–102399. https://doi.org/10.1016/j.compmedimag.2024.102399

Bonavita, I., Rafael-Palou, X., Ceresa, M., Piella, G., Ribas, V., & Ballester, M. Á. G. (2019). Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung

cancer classification pipeline. *Computer Methods and Programs in Biomedicine*, *185*, 105172–105172. https://doi.org/10.1016/j.cmpb.2019.105172

Causey, J., Zhang, J., Ma, S., Jiang, B., Qualls, J., Politte, D. G., Prior, F., Zhang, S., & Huang, X. (2018). Highly accurate model for prediction of lung nodule malignancy with CT scans. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-27569-w

Farhat, H., Sakr, G. E., & Kilany, R. (2020). Deep learning applications in pulmonary medical imaging: recent updates and insights on COVID-19. *Machine Vision and Applications*, *31*(6). https://doi.org/10.1007/s00138-020-01101-5

Farina, B., Benito, R. M., Montalvo-Garcia, D., Bermejo-Peláez, D., Seijó, L., & Ledesma-Carbayo, M. J. (2025). Spatio-temporal deep learning with temporal attention for indeterminate lung nodule classification. *Computers in Biology and Medicine*, *196*, 110813–110813. https://doi.org/10.1016/j.compbiomed.2025.110813

Gao, R., Tang, Y., Xu, K., Huo, Y., Bao, S., Antic, S., Epstein, E. S., Deppen, S., Paulson, A. B., Sandler, K. L., Massion, P. P., & Landman, B. A. (2020). Time-distanced gates in long short-term memory networks. *Medical Image Analysis*, *65*, 101785–101785. https://doi.org/10.1016/j.media.2020.101785

Hawkins, S., Wang, H., Liu, Y., García, A. L., Stringfield, O., Krewer, H., Li, Q., Cherezov, D., Gatenby, R. A., Balagurunathan, Y., Goldgof, D. B., Schabath, M. B., Hall, L., & Gillies, R. J. (2016). Predicting Malignant Nodules from Screening CT Scans. *Journal of Thoracic Oncology*, *11*(12), 2120–2128. https://doi.org/10.1016/j.jtho.2016.07.002

Huang, P., Lin, C. T., Li, Y., Tammemägi, M. C., Brock, M. V., Atkar-Khattra, S., Xu, Y., Hu, P., Mayo, J. R., Schmidt, H., Gingras, M., Pasian, S., Stewart, L., Tsai, S., Seely, J. M., Manos, D., Burrowes, P., Bhatia, R., Tsao, M., & Lam, S. (2019). Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method. *The Lancet Digital Health*, *1*(7). https://doi.org/10.1016/s2589-7500(19)30159-1

Jiang, H., Shen, F., Gao, F., & Han, W. (2021). Learning efficient, explainable and discriminative representations for pulmonary nodules classification. *Pattern*

Recognition, *113*, 107825–107825. https://doi.org/10.1016/j.patcog.2021.107825

Jin, C., Yu, H., Ke, J., Ding, P., Yi, Y., Jiang, X., Duan, X., Tang, J., Chang, D. T., Wu, X., Gao, F., & Li, R. (2021). Predicting treatment response from longitudinal images using multi-task deep learning. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-22188-y

Kokurkhaeva, A. A., Balakerimov, I. D., Tsoroeva, M. R., Abdulvagidova, K. S., Gapuraeva, R. A.-H., Selimova, A. G., Radzhabova, M. N., Saidova, A. O., Taymazova, Z. A., & Karimova, A. (2025). Role of Artificial Intelligence in Lung Cancer Screening: Reducing False Positive Results and Improving Diagnostic Efficiency. *Journal Of Biochemical Technology*, *16*(3), 109–116. https://doi.org/10.51847/vzlsffao5v

Liao, R.-Q., Li, A., Yan, H., Lin, J., Liu, S. M., Wang, J., Fang, J., Liu, H., Hou, Y., Song, C., Hui-fang, Y., Li, B., Jiang, B., Dong, S., Nie, Q., Zhong, W., Wu, Y., & Yang, X. (2022). Deep learning-based growth prediction for sub-solid pulmonary nodules on CT images. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/fonc.2022.1002953

Liu, H., Cao, H., Song, E., Ma, G., Xu, X., Jin, R., Liu, C., & Hung, C.-C. (2020). Multi-model Ensemble Learning Architecture Based on 3D CNN for Lung Nodule Malignancy Suspiciousness Classification. *Journal of Digital Imaging*, *33*(5), 1242–1256. https://doi.org/10.1007/s10278-020-00372-8

Liu, Y., Hsu, H. Y., Lin, T., Peng, B., Saqi, A., Salvatore, M., & Jambawalikar, S. (2024). Lung nodule malignancy classification with associated pulmonary fibrosis using 3D attention-gated convolutional network with CT scans. *Journal of Translational Medicine*, *22*(1), 51–51. https://doi.org/10.1186/s12967-023-04798-w

Lucassen, R. T., Romers, M., Ebbelaar, C. F., Najem, A., Hayes, D. J., Mooyaart, A. L., Roshani, S., Wynaendts, L., Stathonikos, N., Breimer, G. E., Jansen, A. M. L., Veta, M., & Blokx, W. A. M. (2025). 115P Spitz tumor classification using artificial intelligence. *ESMO Real World Data and Digital Oncology*, *10*, 100312–100312. https://doi.org/10.1016/j.esmorw.2025.100312

Musthafa, M. M., Manimozhi, I., Mahesh, T. R., & Guluwadi, S. (2024). Optimizing double-layered convolutional neural networks for efficient lung cancer

classification through hyperparameter optimization and advanced image pre-processing techniques. *BMC Medical Informatics and Decision Making*, *24*(1). https://doi.org/10.1186/s12911-024-02553-9

Oudkerk, M., Liu, S., Heuvelmans, M. A., Walter, J., & Field, J. K. (2020). Lung cancer LDCT screening and mortality reduction — evidence, pitfalls and future perspectives [Review of *Lung cancer LDCT screening and mortality reduction — evidence, pitfalls and future perspectives*]. *Nature Reviews Clinical Oncology*, *18*(3), 135–151. Nature Portfolio. https://doi.org/10.1038/s41571-020-00432-6

Petousis, P., Han, S. X., Aberle, D. R., & Bui, A. (2016). Prediction of lung cancer incidence on the low-dose computed tomography arm of the National Lung Screening Trial: A dynamic Bayesian network. *Artificial Intelligence in Medicine*, *72*, 42–55. https://doi.org/10.1016/j.artmed.2016.07.001

Rafael-Palou, X., Jiménez-Pastor, A., Martí-Bonmatí, L., Muñoz-Núñez, C. F., Laudazi, M., & Alberich-Bayarri, Á. (2025). Advancing deep learning-based segmentation for multiple lung cancer lesions in real-world multicenter CT scans. *European Radiology Experimental*, *9*(1). https://doi.org/10.1186/s41747-025-00617-7

Selvam, M., Chandrasekharan, A., Sadanandan, A., Anand, V. K., Murali, A., & Krishnamurthi, G. (2023). Radiomics as a non-invasive adjunct to Chest CT in distinguishing benign and malignant lung nodules. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-46391-7

Shuvo, S. B. (2023). AutoLungDx: A Hybrid Deep Learning Approach for Early Lung Cancer Diagnosis Using 3D Res-U-Net, YOLOv5, and Vision Transformers. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2305.00046

Zhou, J., Hu, B., Feng, W., Zhang, Z., Fu, X., Shao, H., Wang, H., Jin, L., Ai, S., & Ji, Y. (2023). An ensemble deep learning model for risk stratification of invasive lung adenocarcinoma using thin-slice CT. *Npj Digital Medicine*, *6*(1). https://doi.org/10.1038/s41746-023-00866-z

Zhou, Y., Huang, H., Yu, Y., & Shang, J. (2025). A Machine Learning Approach to Volumetric Computations of Solid Pulmonary Nodules. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2508.20127

Zhuang, L., Tabatabaei, S. M. H., Salehi-Rad, R., Tran, L. M., Aberle, D. R., Prosper, A. E., & Hsu, W. (2025). Vision-Language Model-Based Semantic-Guided Imaging Biomarker for Lung Nodule Malignancy Prediction. *arXiv (Cornell University)*. http://arxiv.org/abs/2504.21344