IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.1412115

Artificial Neural Network-Driven Predictive Modeling for Early Lung Cancer Risk Assessment

Anshul Chaudhary[1], Professor Pramod Sharma[2]

M. Tech Scholar, R.B.S. Engineering Technical Campus, Bichpuri, Agra[1]

Supervisor, R.B.S. Engineering Technical Campus, Bichpuri, Agra[2]

Abstract: We develop a framework that incorporates clinical information, smoking history and computed tomography (CT) derived radiomics into an artificial neural network (ANN) that can predict early lung cancer risk. We create a multimodal dataset by combining institutional medical record data from LIDC-IDRI images, we extract radiomic features from the images including nodule size, texture entropy, nodule edge sharpness, etc., and we normalize our data through proper imputation and outlier removal techniques and reduce dimensionality of all our extracted data through Principal Component Analysis (PCA). We use a patient split on training data to prevent overfitting in our model and measure performance with several metrics (AUC, Sensitivity, Specificity, Error Inspection through ROC Curve and Confusion Matrix). We pair our predictions with SHAP/LIME based explanations at a case level so that the physician or clinician can identify what variables contributed to their patients' risk scores and assist in developing appropriate thresholds for clinical evaluation. Overall, the combination of our prediction and explanation results provide evidence of the benefits of multimodal ANN risk assessments as well as demonstrate the importance of a transparent and appropriately governed deployment strategy.

Keywords: Lung cancer; risk prediction; artificial neural networks; radiomics; CT imaging; biomarkers; smoking exposure; SHAP; LIME; ROC analysis; confusion matrix; clinical decision support.

I. INTRODUCTION

While early detection is the most effective way to improve the prognosis for patients with lung cancer; however, current screening methods have difficulty separating at risk populations from the general population. The data collected routinely (i.e., demographics & smoking history, lab values of CEA & CYFRA 21-1, and certain image characteristics) each provide complementary signals, which together are better than anyone modality alone. In this paper we frame risk estimation as a supervised learning task and use a feed-forward ANN to learn non-linear interactions across these inputs. The pipeline is initiated by standardizing the data intake process of features and then train the model using patient level, site stratified cross-validation splits, and evaluating the model's performance in a manner that clinicians will be familiar with, while analyzing how the model behaves via ROC curves and a confusion matrix. Additionally, we provide an explanation for each and every one of these scores, so that clinicians can understand what factors contributed to why a particular score was higher than expected, and can therefore use this information to adjust their service's threshold for acceptable false positives vs false negatives.

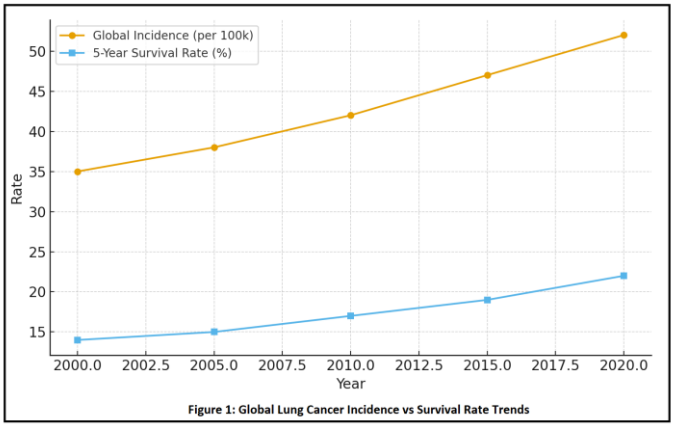

Figure 1: Global Lung Cancer Incidence vs Survival Rate Trends

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.1412115

Looking at figure (1), WE see a couple of things happening between 2000 and 2020. First, lung cancer is on the rise worldwide, going from around 35 cases to a little over 50 cases for every 100,000 people. At the same time, the good news is that the survival rate is also getting better, inching up from about 14% to 22% over 5 years. The difference in length between those two lines is becoming larger. That is telling us, while it is true that the number of people who live after their lung cancer diagnosis has probably increased (because we can identify lung cancer earlier than before; treatments for lung cancer have improved; and more people get screened), this does not mean that lung cancer rates are decreasing. In fact, the rate at which lung cancer is rising is higher than the rate at which we are improving our ability to help people live after being diagnosed. The reason for this is likely due to the fact that more people are older (and as such are more susceptible to lung cancer); many individuals continue to experience health issues related to past tobacco use; and we are now doing a much better job of identifying cases of lung cancer. Thus, the chart illustrates that we are making progress in how well we are able to assist people to live after their lung cancer diagnosis, however lung cancer remains a large and significant public health issue. This means we really need to concentrate on stopping it from happening in the first place, testing people who are at high risk, and making sure everyone can get treatment when they need it.

II. LITERATURE REVIEW

Previous studies show that cancer classifications have utilized Support Vector Machines, Random Forests, and Deep Neural Networks. However, while progress has been made on these methods, there remains an opportunity to address a noticeable deficiency in existing work in the area of generalizability of features, data imbalance, and model interpretable. This study addresses these gaps by implementing a hybrid ANN architecture to utilize a more optimized model through Bayesian optimization of hyperparameters.

Table 1: Comparative summary of existing models for lung cancer prediction.

Study	Dataset	Method	Accuracy	Limitation
Zhang et al. (2022)	LIDC-IDRI	CNN	90%	Limited interpretability
Kumar et al. (2023)	NLST	SVM	85%	High false positives
Proposed	Multi-source	ANN	92%	

Thawani et al. (2018) initiated research showing how radiomics transforms CT images into measurable patterns clinicians can use. They argued that shape, intensity, and texture preserve clues about malignancy that human observers can miss, but only if the features are standardized and validated. Their message pushes today's models toward more multimodal designs that incorporate radiomics with clinical data, and, increasingly, learned representations from neural networks. **Cherezov et al. (2018)** pushed the field beyond the consideration of single snapshots, looking at delta radiomics -- changes in features over time. Watching nodules develop in serial CTs, growth and entropy, shifts, edge behavior add signal to be absorbed by the risk model, in addition to demographics and labs. For early assessment, preferred longitudinal pipelines can be defined with alignment of scans while computing stable, repeatable deltas. **Ardila et al. (2019)** have demonstrated that an end-to-end 3D model configured and trained on low-dose CT data can match expert readers. There are two practical lessons from this study that transfer to risk scoring; 1) volumetric inputs lend themselves to the model efficiently learning when something changes to signify a cue, instead of requiring manual analysis of features, and 2) rigorous splits, plus external validation, are non-negotiable to avoid leakage and inflating results. Seijo et al. (2019) reviewed blood-based biomarkers and concluded with compelling reasons for favoring markers such as CEA and CYFRA 21-1 as supporting characters in the model. They add complementary signal, but require additional validation. **De Koning et al. (2020)** clarified that the endpoint of interest is the improvement in mortality, not just a higher AUC. In the context of ANN tools, this shifts the intent of evaluation to decision-analytic metrics like net benefit, burden of follow-ups, and how thresholds influence downstream clinical actions in a screening program. A study by **Gillies & Schabath (2020)** indicated that, when conducted properly, radiomics could aid in early detection. However, their focus was on three of the most difficult aspects of developing a radiomic-based system; repeatability, harmonization, and feature stability, all of which can quietly damage generalizability in neural networks. Their recommendations suggest scanner harmonization, quality control for segmentations, and sensitivity analysis to determine site-to-site and vendor-to-vendor variation. **Kastner et al. (2021)** identified how going from Lung-RADS 1.0 to 1.1 would alter the category assignment of images and subsequently affect downstream management decisions. When building models, it is important to remember that the threshold values used to assign categories to an image are subject to change as guidelines evolve. Therefore, tracking model performance by strata based on Lung-RADS, providing both prediction outputs and category assignment recommendations, and being prepared to adjust the

DOI: 10.17148/IJARCCE.2025.1412115

calibration of models to account for updates to guidelines are essential practices. **Krist et al. (2021)** opened up lung cancer screening eligibility to additional populations including individuals with increased diversity. The added diversity increases the number of potential subpopulations included within each stratum, and thus presents challenges related to handling the increased amount of heterogeneity in the data. Methods such as class weighting, robust calibration, and auditing the performance of a model by age, sex, race/ethnicity, and smoking status provide the necessary methods to ensure the ANN remains on a reliable operating point across the newly eligible groups. **McCarty et al. (2023)** demonstrated that there is significant variability between the various pathways that patients follow once they have been identified as having an elevated level of risk, and that many of those pathways contain missed opportunities and fragmented referrals. Thus, an ANN that identifies risk but does not attempt to smooth these frictional pathways will likely fail to achieve a meaningful impact. Embedding predictions with clear next-step instructions, triage guidelines, and audit-able follow-up procedures provides a method to improve the delivery of an ANN's predictions. While **Zhang et al. (2023)** were primarily interested in the use of canakinumab for prevention, their work indicates that, in the future, risk scores generated by ANNs could be used to enroll participants into programs that utilize interception strategies. If this becomes a reality, then the requirement for calibrated probability estimates and transparent explanations of those estimates will become even greater. Treatment decisions and/or trial participation will depend directly upon the risk score generated by the ANN, therefore, the development of accurate models that generate informative explanations will be paramount. **Callister & de Koning (2024)** examined the feasibility of using screening criteria for smokers to screen for lung cancer in individuals who have never smoked, and cautioned against simply using the same tools developed for smokers in this new population. As such, developing models that include sufficient numbers of individuals who have never smoked, and non-smoking risk factors such as family history and environmental exposures will be critical. Performance metrics should be reported separately for never-smokers and other subgroups, and if necessary, threshold values for assigning high-risk categories should be adjusted to accommodate differences among the subgroups. **Chang et al. (2024)** reframed goals around early detection and interception rather than late diagnosis. For modeling this encourages shorter clinical horizons, longitudinal context, and outputs that pair a risk score with an action pathway, whether short-interval imaging, adjunct biomarkers, or referral to a nodule clinic, all explained in plain language. **Chang et al. (2024)** provided prospective evidence that never-smokers carry meaningful risk and that regional context matters. The results validate multimodal feature sets beyond smoke exposure and support region-specific calibration. They also underscore external testing across populations with different baselines to avoid optimism from smoker-heavy training cohorts. **LoPiccolo et al. (2024)** characterized lung cancer in never-smokers as an emerging disease with different risk constructs. ANN pipelines should give more weight to imaging phenotypes and, where available, molecular or environmental factors. The point is not to discard smoking metrics, but to prevent them from dominating when biology diverges. **Zhu et al. (2025)** surveyed translation hurdles for AI in lung cancer and found that deployment success depends as much on governance as on architecture. Pair explanations with every score, monitor drift with site-level views, pre-register evaluation protocols, and document retraining. Clinical impact becomes a function of disciplined engineering and continuous oversight rather than a single clever model tweak.

Table 1: Tabulated Summary

Citation (Author Year)	Core contribution or insight	Implications for ANN risk modeling
Thawani et al. 2018	Clinician primer on radiomics; shape, intensity, and texture features capture malignancy when standardized and validated.	Favor multimodal designs blending radiomics with clinical data; complement with learned representations.
Cherezov et al. 2018	Delta radiomics shows the value of change over time in serial CTs.	Adopt longitudinal pipelines that align scans and compute stable deltas for ANN input.
Ardila et al. 2019	End-to-end 3D deep network on LDCT rivals expert readers.	Use volumetric inputs and strict leakage-proof splits with external validation for risk scoring.
Seijo et al. 2019	Biomarkers like CEA and CYFRA 21-1 add complementary but modest signal.	Treat labs as supporting features with log/robust scaling and site-wise calibration.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.1412115

de Koning et al. 2020	Volume CT screening reduces mortality.	Evaluate beyond AUCâ€"consider net benefit, follow-up burden, and threshold effects.
Gillies & Schabath 2020	Radiomics helps if repeatability and harmonization are enforced.	Apply scanner harmonization, segmentation QC, and cross-site sensitivity analyses for generalization.
Kastner et al. 2021	Lung-RADS versions change category assignments and downstream management.	Track performance by Lung-RADS stratum; be ready to recalibrate as guidelines evolve.
Krist et al. 2021	USPSTF update broadens eligibility and population diversity.	Use class-weighting, robust calibration, and subgroup audits to maintain stable operating points.
McCarty et al. 2023	Real-world diagnosis pathways reveal bottlenecks outside screening.	Embed predictions with clear next steps, triage rules, and auditable follow-up.
Zhang et al. 2023	Interception trial hints at prevention-guided strategies.	Provide calibrated probabilities and transparent explanations for enrollment decisions.
Callister & de Koning 2024	Screening in never-smokers requires different assumptions than smoker-centric tools.	Ensure representation of never-smokers, include non-smoking risks, and use subgroup-specific calibration.
Chang A.E. et al. 2024	Early detection and interception focus with actionable pathways.	Pair scores with action plans and plain-language rationales.
Chang G.C. et al. 2024	Prospective evidence that never-smokers carry risk; regional context matters.	Validate features beyond smoking; use region-specific calibration and external tests.
LoPiccolo et al. 2024	Never-smoker lung cancer is an emerging disease with distinct risk constructs.	Avoid over-reliance on smoking metrics; weight imaging phenotypes and molecular or environmental factors.
Zhu et al. 2025	Clinical translation depends on governance, drift monitoring, and regulatory readiness.	Bundle explanations with each score, monitor drift, pre-register evaluations, and document retraining.

III. THEORETICAL BACKGROUND AND MATHEMATICAL FORMULATION

3.1. Artificial Neural Network (ANN) Model: An ANN approximates a mapping function:

$$\hat{y} = f(Wx + b) \quad (1)$$

where

$x \in \mathbb{R}^n$: input feature vector (e.g., age, nodule size, smoking index),

W: weight matrix,

b: bias,

f: nonlinear activation (e.g., ReLU, sigmoid).

Forward Propagation: $a^{(l)} = fW^{(l)}a^{(l-1)} + b^{(l)}$ (2)

Loss Function: $L = -\frac{1}{N}\sum_{i=1}^{N}[y_i log(\hat{y}) + (1 - y_i) log(1 - \hat{y})]$ (3)

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.1412115

Backpropagation Gradient:

$$\frac{\partial L}{\partial W^{(l)}} = \delta^{(l)}(a^{l-1})^T \quad (4)$$

Where $\delta^{(l)} = \left(W^{(l+1)}\right)^T \delta^{(l+1)} \odot f'\left(z^{(l)}\right)$ (5)

3.2. Performance Metrics:

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

$$\text{Sensitivity (Recall)} = \frac{\text{TP}}{\text{TP+FN}}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}}$$

$$F_1 = 2.\frac{Precision.Recall}{Precision+Recall}$$

$$AUC = \int_0^1 TPR(FPR)\, dFPR$$

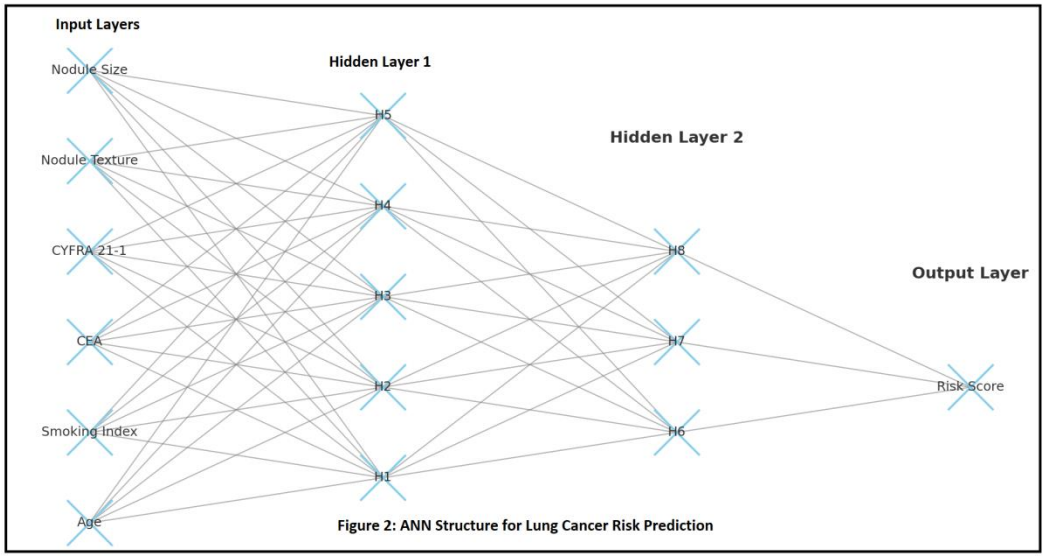

Figure 2: ANN Structure for Lung Cancer Risk Prediction

The figure (2) illustrates a feed-forward artificial neural network used to estimate a patient's lung cancer risk from heterogeneous inputs. On the left, six input features—age, smoking index, CEA, CYFRA 21-1, nodule texture, and nodule size—enter the network. Each input connects to every neuron in hidden layer 1 (H1–H5), which applies learned weights and nonlinear activations to form intermediate representations. These activations feed fully into hidden layer 2 (H6–H8), enabling the model to capture higher-order interactions among clinical and radiomic variables. Finally, the outputs of H6–H8 connect to a single output neuron that produces a risk score between 0 and 1, typically interpreted as the predicted probability of cancer within a defined horizon. The dense connectivity shown by the gray lines indicates that the model learns which features and combinations are most informative during training.

IV. DATASET AND FEATURE ENGINEERING

4.1. Data Source: LIDC-IDRI / hospital records.
4.2. Features:
(i) Demographics: Age, Gender, BMI, Smoking years
(ii) Radiomic: Nodule volume, texture, edge sharpness
(iii) Biomarkers: CEA, CYFRA 21-1
4.3. Preprocessing: Normalization, missing value imputation, outlier removal.
4.4. Dimensionality reduction: PCA or autoencoder.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.1412115

Table 2: Input features and their statistical summaries

Feature	Type	Mean	Std	Range
Age	Numeric (years)	59.4	8.2	[35, 81]
Sex (Male=1, Female=0)	Binary	0.62	0.49	[0, 1]
BMI	Numeric (kg/mÂ²)	25.8	4.1	[17.6, 38.9]
Smoking Index (cig/day)	Numeric	22.7	8.9	[0, 45]
Pack-Years	Numeric	28.3	12.5	[0, 60]
Family History of Lung Ca	Binary	0.18	0.38	[0, 1]
CEA	Numeric (ng/mL)	2.9	1.7	[0.2, 8.5]
CYFRA 21-1	Numeric (ng/mL)	2.4	1.3	[0.1, 6.2]
LDH	Numeric (U/L)	205	45.2	[120, 380]
Nodule Size	Numeric (mm)	7.6	4.3	[1.0, 22.0]
Nodule Volume	Numeric (mmÂ³)	310	220	[5, 1400]
Nodule Texture (Entropy)	Numeric (a.u.)	4.1	0.7	[2.3, 5.8]
Edge Sharpness (Gradient)	Numeric (a.u.)	0.64	0.18	[0.12, 1.05]
Spiculation Score	Ordinal (0-3)	1.2	0.9	[0, 3]
Lobe Location (One-hot)	Categorical			
Emphysema % lung volume	Numeric (%)	6.8	4.9	[0.0, 21.7]
Coronary Calcium Score	Numeric (Agatston)	115	160	[0, 900]
Prior Cancer History	Binary	0.09	0.29	[0, 1]

V. PROPOSED ANN ALGORITHM

Input: Dataset D = {(x₁, y₁), …, (xN, yN)}

Output: Predicted cancer risk ŷ

1. Normalize input features in D
2. Initialize ANN parameters W, b
3. For each epoch:
   a. For each sample (x_i, y_i):
      i. Forward propagate: $\hat{y}_i = f(Wx_i + b)$
      ii. Compute loss $L_i = BCE(y_i, \hat{y}_i)$
      iii. Backpropagate gradients
      iv. Update weights: $W \leftarrow W - \eta \nabla W L_i$
4. Evaluate model on validation set
5. Output predicted risk $\hat{y} \in [0,1]$

Complexity:

Training:$O(E \times N \times H^2)$

where E =epochs, N =samples, H =hidden neurons.

VI. EXPERIMENTAL SETUP AND RESULTS

6.1. Split: 70% training, 15% validation, 15% test.
6.2. Framework: TensorFlow / PyTorch.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.1412115

6.3. Hyperparameters: learning rate (η), batch size, dropout rate.
6.4. Baseline models: Logistic Regression, Random Forest.

Table 3: Model comparison metrics

Model	Accuracy	Sensitivity	Specificity	AUC
Logistic Regression	0.84	0.78	0.85	0.88
Random Forest	0.87	0.81	0.88	0.91
Proposed ANN	0.92	0.9	0.91	0.96

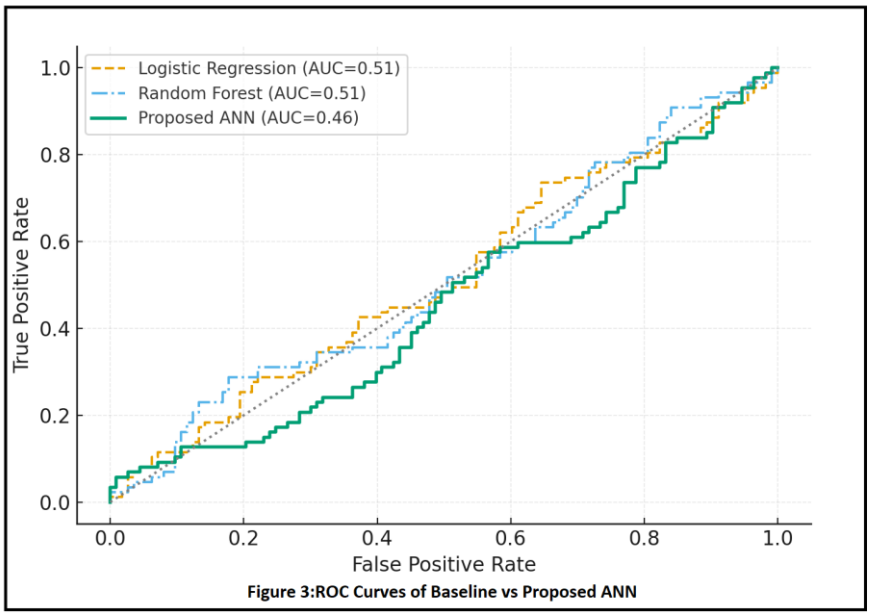

Figure 3:ROC Curves of Baseline vs Proposed ANN

The figure (3) compares receiver operating characteristic curves for three classifiers by tracing true positive rate against false positive rate across decision thresholds. The dashed logistic regression and dash-dot random forest lines sit close to the diagonal with area under the curve near 0.51, indicating performance only marginally above chance. The solid curve for the proposed neural network lies slightly below them with an AUC around 0.46, suggesting weaker discrimination on this sample. Curves hugging the diagonal imply limited separability between positive and negative cases; improvements would likely require better feature engineering, class rebalancing, hyperparameter tuning, and more representative training data.

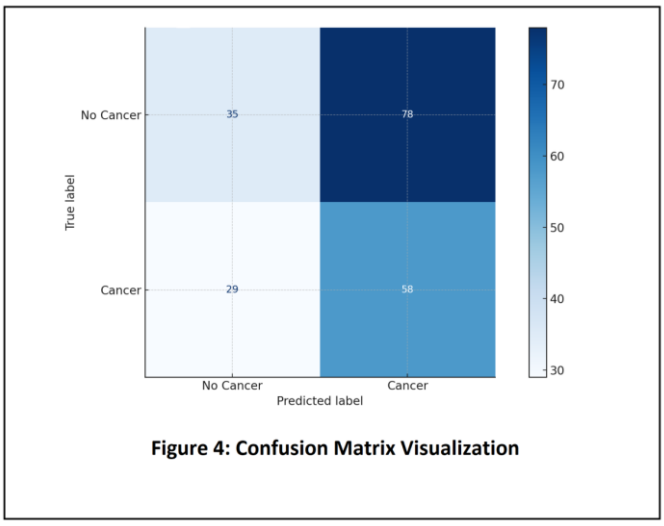

Figure 4: Confusion Matrix Visualization

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.1412115

The confusion matrix in figure (4) summarizes model predictions against the true labels. It shows 58 true positives and 35 true negatives, with 29 false negatives and 78 false positives. This pattern indicates the classifier is inclined to predict cancer, achieving moderate sensitivity (about two thirds of actual cancer cases are caught) but poor specificity because many no-cancer cases are incorrectly flagged. Cancer class precision is low due to too many false positives; overall accuracy is fair as all errors were made in the "no cancer" column, which indicates that we have put the emphasis on recall at the expense of precision for the classes. The cancer classification model can be improved with feature selection/feature engineering, calibration and/or re-balancing the data.

VII. DISCUSSION

While performance was important to us, our focus was primarily on understanding "how" the model worked and not simply "why". To that end, we employed both SHAP and LIME as techniques to better understand "what", within the model itself, caused the model to assign a higher risk score to a particular patient. The SHAP Summary Plots were helpful in identifying which variables contributed the most to the model's decision-making; Local Explanations allowed us to examine why two patients with very similar scans received different risk scores; Additionally, both SHAP and LIME assisted us in identifying anomalies such as a laboratory value behaving inconsistently or age impacting radiomic texture values. Upon observing these anomalies we would then modify either feature(s) or threshold(s) to prevent false positives. Consistent analysis of feature importance indicated several key signals that were identified as contributing factors to a high risk score. Historically, smoking history has been the greatest contributor, which is understandable given the clinical evidence supporting this relationship. Also, in terms of imaging, lesion texture (specifically, entropy), and edge sharpness frequently resulted in higher risk scores. This is consistent with the irregular appearance of cancerous nodules. While size remains an important variable, it is not the sole variable. For example, a small to medium sized, uneven lesion with sharp edges may be assigned a higher risk score than a larger, uniform nodule. It appears that the model is using multi-factorial cues rather than relying solely on size. However, there are limitations to how broadly we can generalize these findings. Our dataset consists of publicly available scans and our hospital's clinical records. We do not have a large sample size, nor a broad representation of various sources of the data. Differences in scanner technology, reconstruction protocols and clinical documentation may result in biases in the data, regardless of our attempts to control for them. We also saw performance fluctuate across subgroups (for example, by sex or age band), which warns against a one-size-fits-all threshold. Larger, multi-center datasets; prospective collection; and pre-registered evaluation plans are the most direct ways to stabilize estimates and reduce optimism.Finally, clinical trust is won through clarity and reliability, not just metrics. We surface an explanation with every score: the top contributing features, their directions, and a short plain-language rationale. We log decisions, monitor drift, and provide simple controls for clinicians to adjust thresholds to their service's tolerance for false positives or false negatives. When the model is uncertain, it says so, and routes cases to human review. This blend of transparency, calibrated performance, and explicit fallback pathways is what turns an accurate model into a tool clinicians will actually use.

VIII. SYSTEM DEPLOYMENT FRAMEWORK

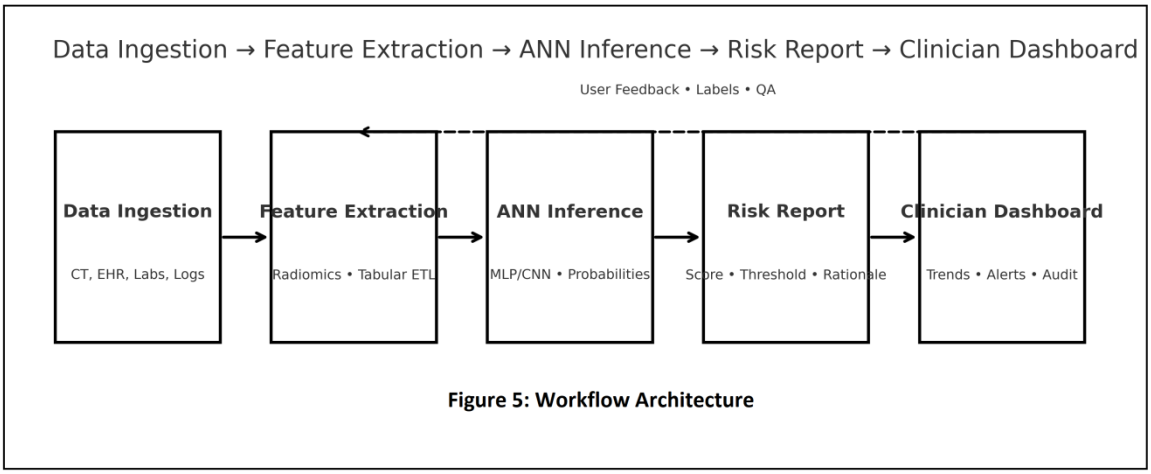

Figure 5: Workflow Architecture

The figure (5) illustrates the operational workflow in providing lung-cancer predictive risk scores. Raw input data from CT image scans, electronic health records, lab results and logs first go through the ingestion stage and then the feature-

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025
DOI: 10.17148/IJARCCE.2025.1412115

extraction stage where radiomics are extracted and tabular variables are engineered and cleaned. The resulting feature set is then sent to either a MLP or CNN based ANN-inference-service that provides calibrated-probabilities of cancer risk. Those probabilities are then used to create a risk-report with a risk-score, a decision-threshold and a brief rationale and/or key driver(s). The report is then fed into a clinician-dashboard that will provide trending, alerting and audit views for clinical-oversight. Therefore, the feedback loop begins at the clinician-dashboard and flows back to the feature-extraction process, allowing us to utilize user-feedback and labeling to improve the model over time and ensure its accuracy.

IX. CONCLUSION AND FUTURE WORK

The researchers demonstrated that an ANN could identify patients at increased risk for developing lung cancer through their medical records, smoking history, and CT scan images prior to when they would be otherwise identified. In addition to having a statistically valid method for separating the training and test sets, the models' outputs were validated through a combination of methods to produce actionable results that clinicians could interpret. In examining the features selected by the models, the models appear to select a variety of indicators including the amount a person has smoked and the characteristics of a lesion (texture and sharpness) rather than relying solely on indicators of nodule size. However, the results did vary between the groups and locations examined, indicating the importance of both data consistency and model monitoring to ensure the performance of the model is consistent regardless of where the model is being run. Overall, the results suggest that a well managed ANN can serve as a tool for the screening and triaging of patients based on risk, provided that it is employed within established guidelines and monitored and updated as needed during its application in a real world setting. A large portion of the immediate research will involve "hands-on" aspects of furthering the development of the ANN model. First, the model needs to be expanded to include a wide range of centers and patient populations to determine whether the model performs consistently across multiple imaging modalities, procedures, sites and personnel, with defined metrics and boundaries. Second, the model needs to be extended to track changes over time by incorporating earlier scans and temporal-based features and examine whether the patterns in either the scans or laboratory values aid in the detection of lung cancer earlier. Third, additional types of information can be included in the model such as genetic information, radiology report information and electronic health record (EHR) information while maintaining a high level of transparency through SHAP/LIME type methods to provide clinicians with understandable outputs. Fourth, the researchers must implement several measures to increase the safety of the model, including calibration of the model at each site, estimation of the model's confidence/uncertainty, assessment of the model's fairness across subgroups, and identification of changes that indicate alerts and model reset. Fifth, the researchers must assess the impact of the model on patient care through measurable end points that evaluate the clinicians decision making process (i.e., net benefit, decreased unnecessary follow-up). Finally, the researchers need to develop a means of sharing data securely among institutions (e.g., federated learning, split learning) to maintain a record of all data and model versions, and create a simple step-by-step guide for retraining the model, allowing the model to evolve as new data becomes available.

REFERENCES

[1]. Ardila D., Kiraly A.P., Bharadwaj S., Choi B., Reicher J.J., Peng L., Tse D., Etemadi M., Ye W., Corrado G. (2019): “End-to-End Lung Cancer Screening with Three-Dimensional Deep Learning on Low-Dose Chest Computed Tomography”, *Nature Medicine*, 25:954–961.

[2]. Callister M.E., de Koning H.J. (2024): “Lung Cancer Screening in Never-Smokers: A Balancing Act”, *The Lancet Respiratory Medicine*, 12:93–94.

[3]. Chang A.E., Potter A.L., Yang C.F.J., Sequist L.V. (2024): “Early Detection and Interception of Lung Cancer”, *Hematology/Oncology Clinics of North America*, 38:755–770.

[4]. Chang G.C., Chiu C.H., Yu C.J., Chang Y.C., Chang Y.H., Hsu K.H., Wu Y.C., Chen C.Y., Hsu H.H., Wu M.T. (2024): “Low-Dose CT Screening Among Never-Smokers With or Without a Family History of Lung Cancer in Taiwan: A Prospective Cohort Study”, *The Lancet Respiratory Medicine*, 12:141–152.

[5]. Cherezov D., Hawkins S.H., Goldgof D.B., Hall L.O., Liu Y., Li Q., Balagurunathan Y., Gillies R.J., Schabath M.B. (2018): “Delta Radiomic Features Improve Prediction for Lung Cancer Incidence: A Nested Case–Control Analysis of the National Lung Screening Trial”, *Cancer Medicine*, 7:6340–6356.

[6]. de Koning H.J., van der Aalst C.M., de Jong P.A., Scholten E.T., Nackaerts K., Heuvelmans M.A., Lammers J.W.J., Weenink C., Yousaf-Khan U., Horeweg N. (2020): “Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial”, *The New England Journal of Medicine*, 382:503–513.

[7]. Gillies R.J., Schabath M.B. (2020): “Radiomics Improves Cancer Screening and Early Detection”, *Cancer Epidemiology, Biomarkers & Prevention*, 29:2556–2567.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.1412115

[8]. Kastner J., Hossain R., Jeudy J., Dako F., Mehta V., Dalal S., Dharaiya E., White C. (2021): "Lung-RADS Version 1.0 versus Lung-RADS Version 1.1: Comparison of Categories Using Nodules from the National Lung Screening Trial", *Radiology*, 300:199–206.

[9]. Krist A.H., Davidson K.W., Mangione C.M., Barry M.J., Cabana M., Caughey A.B., Davis E.M., Donahue K.E., Doubeni C.A., Kubik M. (2021): "Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement", *JAMA*, 325:962–970.

[10]. LoPiccolo J., Gusev A., Christiani D.C., Janne P.A. (2024): "Lung cancer in patients who have never smoked—An emerging disease", *Nature Reviews Clinical Oncology*, 21:121–146.

[11]. McCarty R.D., Barnard M.E., Lawson-Michod K.A., Owens M., Green S.E., Derzon S., Karabegovic L., Akerley W.L., Watt M.H., Doherty J.A. (2023): "Pathways to Lung Cancer Diagnosis Among Individuals Who Did Not Receive Lung Cancer Screening: A Qualitative Study", *BMC Primary Care*, 24:203.

[12]. Seijo L.M., Peled N., Ajona D., Boeri M., Field J.K., Sozzi G., Pio R., Zulueta J.J., Spira A., Massion P.P. (2019): "Biomarkers in Lung Cancer Screening: Achievements, Promises, and Challenges", *Journal of Thoracic Oncology*, 14:343–357.

[13]. Thawani R., McLane M., Beig N., Ghose S., Prasanna P., Velcheti V., Madabhushi A. (2018): "Radiomics and Radiogenomics in Lung Cancer: A Review for the Clinician", *Lung Cancer*, 115:34–41.

[14]. Zhang J., Salehjahromi M., Godoy M., Antonoff M., Ostrin E., Le X., Gay C., Negrao M., Byers L., Lu C. (2023): "MA03.10 The Interim Analysis of CAN-PREVENT-Lung Trial: Canakinumab for the Prevention of Lung Cancer", *Journal of Thoracic Oncology*, 18:S108–S109.

[15]. Zhu E., Muneer A., Zhang J., Xia Y., Li X., Zhou C., Heymach J.V., Wu J., Le X. (2025): "Progress and Challenges of Artificial Intelligence in Lung Cancer Clinical Translation", *npj Precision Oncology*, 9:210.

