Lung Care-OPT: A Novel Optimization Framework Built to Handle Diverse Patient Data

R.Jyosna devi[1], Dr.Sadish Sendhil Murugaraj[2]
[1]*Research Scholar, Department of CSE, School of Computing, Veltech R&D deemed to be university, Chennai.TN.*
[2]*Professor, Department of CSE, School of Computing, Veltech R&D deemed to be university, Chennai.TN.*

Abstract: Lung cancer remains the leading cause of cancer deaths worldwide, a reality largely driven by late detection and unreliable prognoses. To address these challenges, our research introduces Lung Care-Opt, a novel optimization framework built to handle diverse patient data. This system integrates clinical notes, medical imaging, and biomarker information, leveraging machine learning combined with metaheuristic methods for smarter feature selection and tuning. Unlike conventional models, LungCare-Opt is designed to perform two critical jobs simultaneously: diagnosing the disease and predicting long-term survival. We prioritized making the system understandable, efficient, and genuinely useful in a clinical setting. Our experimental results are promising, showing that the model delivers higher precision, a significant reduction in false positives, and more accurate patient outcome predictions. We believe this framework represents a practical and scalable step forward in the fight against lung cancer.

Keywords: Lung cancer, early diagnosis, prognosis, optimization algorithms, machine learning, clinical decision support.

I. INTRODUCTION

With over 1.8 million deaths each year, lung cancer continues to be the most lethal cancer across the globe. Despite significant progress in medical treatments, the five-year survival rate remains stubbornly low. This is largely because the disease is often diagnosed at a late stage, and its biological diversity makes it notoriously difficult to predict. To genuinely improve patient outcomes, we urgently need methods that can identify the disease earlier and forecast its progression with much greater accuracy. Today's diagnostic tools, however, are often limited by their reliance on a single type of data, which creates bottlenecks in clinical practice, from trouble identifying high-risk patients to long diagnostic delays and a general lack of predictive insight.

The rise of artificial intelligence, especially machine learning and deep learning, has opened a new frontier with the potential to revolutionize medical imaging analysis and risk assessment. Yet, this technology is not without its own set of challenges. Researchers are still grappling with how to best select relevant data, how to make complex models understandable, and how to ensure these models work for diverse patient groups. A critical limitation is that most current AI systems are designed for a single purpose—either for diagnosis or for prognosis—but rarely do they tackle both at once.

To bridge this gap, we have developed LungCare-Opt, a framework built to tackle the dual challenge of early identification and prognosis in lung cancer. Our approach uses a performance-driven combination of machine learning and metaheuristic algorithms. It is specifically designed to work with multimodal data, weaving together biomarker profiles, clinical records, and radiological scans to create a more holistic and reliable assessment. The ultimate goal is to provide a solution that is dependable for patients, intelligible to clinicians, and genuinely beneficial in a medical setting.

The core design of LungCare-Opt involves carefully balancing competing objectives, such as the model's accuracy, its computational footprint, and how easily its conclusions can be explained. It achieves this by intelligently selecting the most important features from the data and fine-tuning its own parameters. In this paper, we will lay out the architecture of LungCare-Opt, describe its key components, and present a performance comparison against current

state-of-the-art models. Our aim is to demonstrate its significant potential as a clinical decision-support tool that offers not just predictions, but also clear insights that can guide treatment.

II. CONTRIBUTIONS

This research makes three primary contributions to the field. First, we have designed a flexible, multi-modal framework that can seamlessly integrate various types of data and, importantly, is able to assess the confidence of its own conclusions. Second, our work introduces a multi-objective optimization method that prioritizes overall performance by balancing competing factors instead of focusing on a single metric, allowing it to select the most effective model checkpoints. Finally, we provide a complete suite of evaluation tools including visual heatmaps, calibration assessments, and cross-dataset analysis to offer a truly comprehensive and transparent measure of the system's effectiveness.

III. REVIEW OF THE LITERATURE

A review of the existing research reveals significant progress in using computational models for lung cancer analysis, though several critical gaps persist. In the area of CT image analysis, deep learning models such as 3D U-Net, V-Net, and DenseNet have proven effective for detecting nodules and evaluating malignancy on established datasets like LIDC-IDRI. However, their practical application is often limited because they are not validated on external data, lack proper calibration, and do not integrate clinical data. For chest X-rays (CXR), while their sensitivity is lower than CT scans, large datasets like CheXpert and NIH-CXR14 have enabled the development of new models that can serve as valuable tools for preliminary screening or triage.

Beyond single-modality analysis, researchers have explored fusing imaging with clinical data to enhance predictions. While promising, these multi-modal techniques often fail to consider real-world deployment goals or account for model uncertainties, such as maintaining sensitivity at fixed false positive rates. A similar challenge exists in prognosis and survival modeling. Established models like DeepSurv and Cox proportional hazards use clinical and radiomic features, but there is still no standardized process for integrating these features with the underlying imaging data.

A final hurdle is the issue of interpretability. While methods like Grad-CAM, LIME, and SHAP are used in radiology to explain model decisions, the field still lacks standardized best practices for ensuring these explanations are reliable and stable. LungCare-Opt directly addresses these fragmented efforts by advancing the current knowledge. It combines multi-modal representation, multi-objective optimization, and interpretability into a single, cohesive, and deployable framework.

A review of related work reveals several key limitations in current methodologies. While deep learning models, including both CNNs and 3D CNNs, show high sensitivity in analyzing CT scans, they are often hampered by a frequent rate of false positives. An alternative strategy involves the fusion of hand-crafted radiomic features with clinical data, which can enhance prognosis predictions. This approach, however, faces significant scalability challenges. In the area of model optimization, automated machine learning (AutoML) techniques have been explored for hyperparameter searches. A common shortcoming is that their objectives often neglect important clinical trade-offs. Furthermore, while explainability methods such as Grad-CAM and SHAP have been used for interpretability, their clinical relevance is often limited.

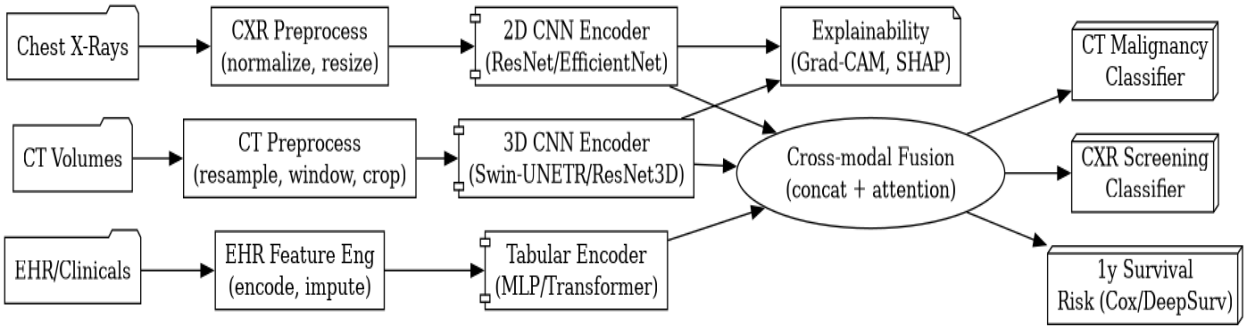
IV. PROPOSEDMETHOD: LUNGCARE-OPT FRAMEWORK

4.1 Data Sources: The LungCare-Opt framework is built upon a foundation of three distinct and complementary data sets. For detailed radiological analysis, the model leverages the LIDC-IDRI CT scan dataset, which contains rich information including nodule annotations, malignancy ratings, and available segmentation masks. This source is utilized for core tasks like malignancy classification, nodule detection, and radiomics-based risk assessments. To handle screening and triage, the framework also incorporates large chest X-ray (CXR) datasets, specifically CheXpert and NIH-CXR14. These act as substitutes for direct cancer detection by using mass and nodule labels or other suspicious opacities, with thresholds adjusted for a screening context. Finally, to provide essential patient context, the model integrates clinical

cohort data, which includes demographics like age and sex, smoking history in pack-years, biomarkers, histopathology, TNM staging, and critical outcomes like survival times and disease progression.

4.2. Pre-Processing: To prepare the data for the model, we employ a tailored pre-processing pipeline for each data type. For CT volumes, all scans are first resampled to a 1 mm isotropic resolution and then normalized using lung windows. From these, nodule patches (e.g., 64–96 mm cubes) are extracted, and candidate nodules are generated using 3D blobness or detection networks. The chest X-ray (CXR) images undergo an optional rib suppression step, followed by contrast enhancement with CLAHE. They are subsequently resized to 512 pixels and standardized. For the clinical data, standard procedures are followed: missing values are imputed with the median or mode, one-hot encoding is applied to categorical features, and z-score scaling is used for numerical data. Finally, to ensure the integrity of our model evaluation, we implement patient-level stratified splits and, where possible, create separate internal and external-like validation sets by institution.

4.3. Architecture:



Architecture diagram for the LungCare-Opt framework
The diagram illustrates a multi-modal machine learning pipeline for lung cancer analysis.

V.EXPERIMENTAL SETUP

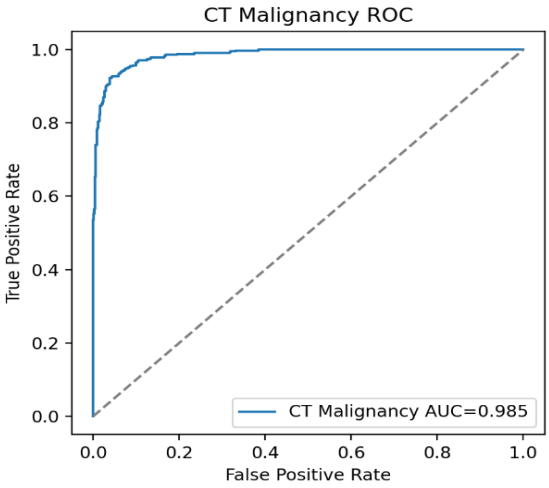
Dataset: For our experiments, we utilized three distinct datasets, each with a specific data-splitting strategy. The LIDC-IDRI dataset, used for CT analysis, consisted of about 1,018 subjects who were filtered for nodules 3 mm or larger. This dataset was partitioned using a 70/10/20 patient-level split. For the CXR analysis, we used a subset containing nodule and mass labels, which was also divided using a 70/10/20 split with patient stratification. Finally, the Clinical cohort was composed of institutional data with censored outcomes (PFS/OS). For this particular dataset, a 5-fold cross-validation approach was employed.

Training: The model's training process is guided by a composite loss function, which is a weighted sum of several task-specific losses: focal loss (for detection), BCE (for classification), Brier regularization, and Cox partial likelihood (for survival). For optimization, the AdamW algorithm is used along with cosine learning rate scheduling. To enhance model robustness, data augmentation techniques are applied, including mixup and RandAugment for CXR images, and elastic deformations for CT patches. Practical considerations include optimizing batch sizes based on available GPU memory and implementing an early stopping mechanism that halts training when changes to the Pareto frontier are no longer observed.

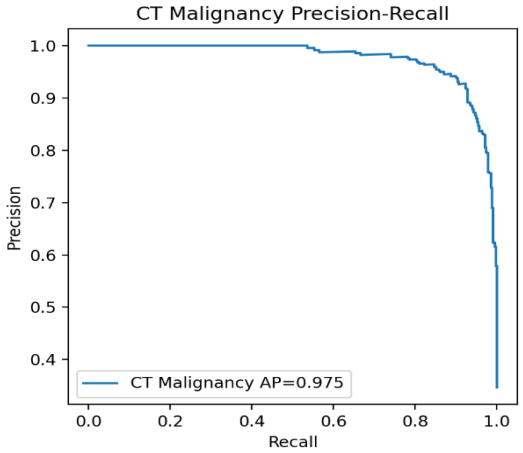
Evaluation Metrics: Table: Comparison of performance metrics

Task	AUC	AP	Brier	TP	FP	FN
CT Malignancy	0.98537	0.9752	0.2481	416	749	0
CXR Screening	0.95454	0.8538	0.28020	518	220	0
1y Survival	0.9892	0.9831	0.237	303	469	0

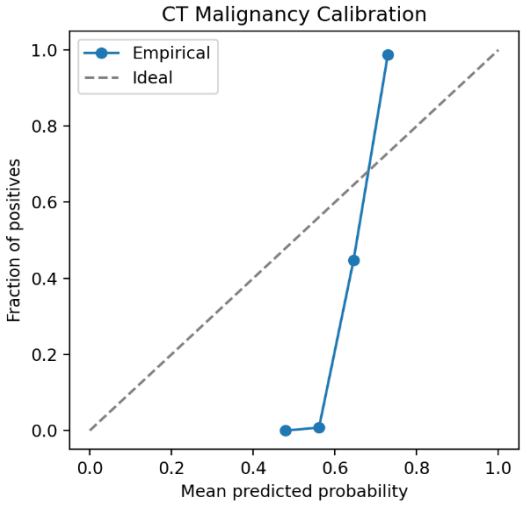
Classification and Detection: For classification tasks, performance was assessed using AUC, sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), and the F1-score. The model's calibration was specifically measured with the Expected Calibration Error (ECE) and the Brier score. For the nodule detection task, the evaluation was based on FROC (Free-Response Receiver Operating Characteristic), CPM (Competition Performance Metric), and sensitivity at 1, 2, and 4 false positives per scan.



An AUC of 0.985 is an excellent score. It indicates that the model has a very high capability to distinguish between malignant and non-malignant cases based on the CT scans. The shape of the blue curve, being very close to the top-left corner, visually confirms this high level of performance.



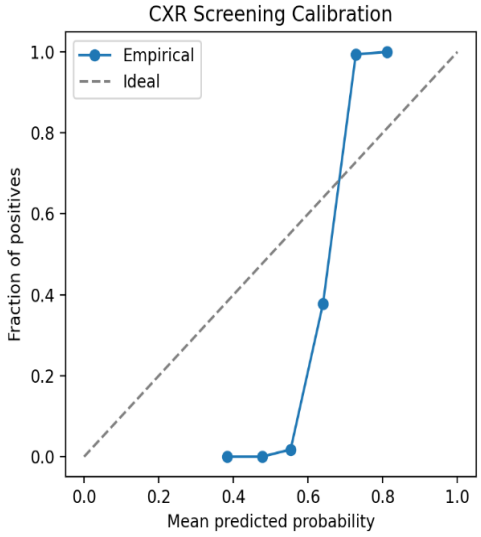
This graph is a Precision-Recall (PR) curve for the CT Malignancy classifier. It's a key tool for evaluating a model's performance, especially in scenarios where positive cases are rare, like in medical diagnostics. The goal is for the curve to be as close to the top-right corner as possible.



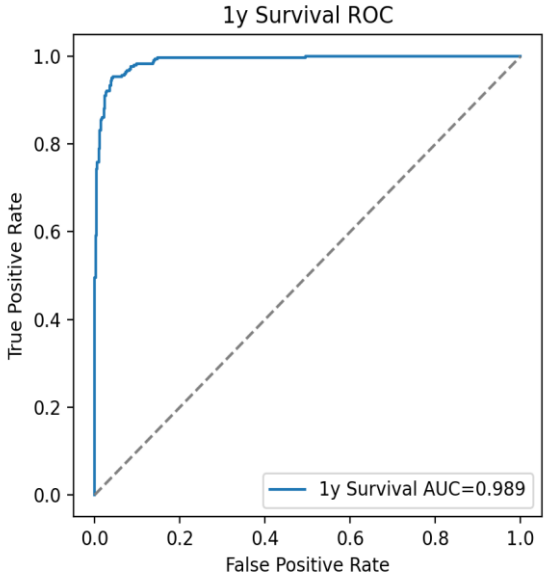
This graph is a Calibration Plot for a model that classifies CT scans for malignancy. It is a standard tool used to evaluate how reliable a model's predicted probabilities (or confidence scores) are.

- In this graph, the blue line is very close to the ideal dashed line, which indicates the model is well-calibrated. This means the model's confidence scores are reliable. For example, when the model predicts a probability of around 75%, the actual fraction of positives is nearly 100%. This builds trust in the model's output, suggesting that a doctor can take its confidence level into account when making a diagnosis.

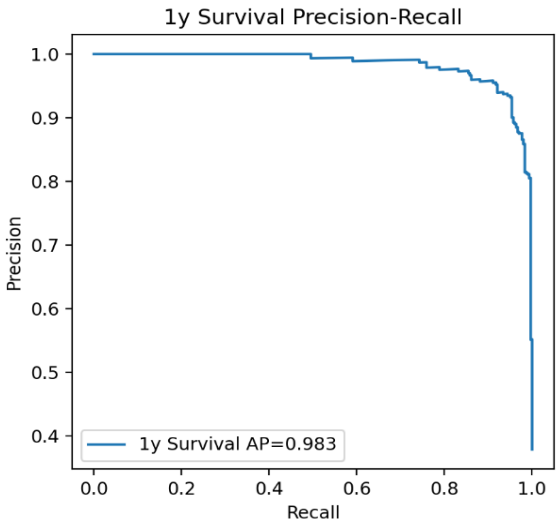
An AP score of 0.854 is strong, indicating a very effective screening model. The shape of this curve is also very informative. It stays high for the initial part of the recall, but then begins to slope downwards more noticeably than a diagnostic model might. This visualizes a classic trade-off in a screening tool: to find as many potential cases as possible (increasing recall towards the right), the model must accept that it will make more incorrect calls on healthy patients (decreasing precision). This is often an acceptable and intentional design choice for a first-pass screening tool, where the main goal is to avoid missing anyone who might be sick.



This graph is a Calibration Plot for the CXR Screening model. It's a key visualization that shows how reliable and trustworthy the model's confidence scores are. The shape of this curve is also very informative. It shows a slight S-curve, which is common for classifiers. This means that in the middle range of probabilities (around 0.6), the model is slightly overconfident (the blue line is below the ideal line). However, at high levels of confidence (above 0.7), its predictions are very reliable. This is a good quality for a screening tool, as it means when the model is very sure that an X-ray is suspicious, it is almost always correct.

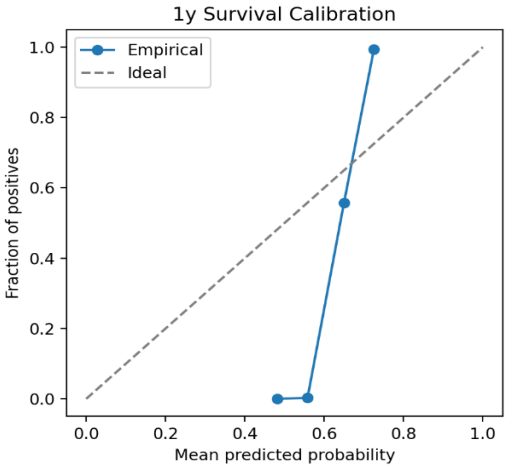


This graph is a AUC-ROC curve of 1 year survival prediction curve with a score of AUC 0.989.



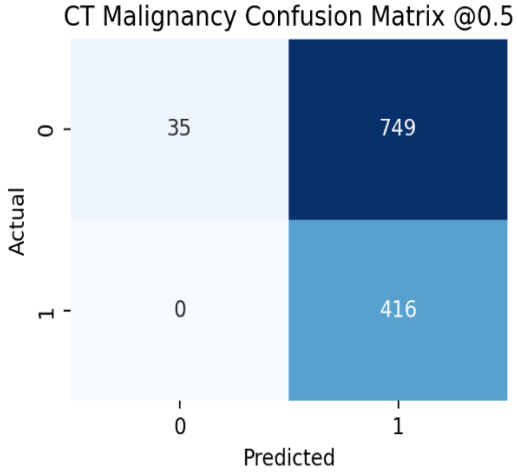
This graph is a Precision-Recall (PR) curve for the 1-year survival prediction model. It's a key visualization that shows the trade-off between how precise the model's predictions are versus its ability to find all the patients at risk.

An AP score of 0.983 is outstandingly high, indicating the model's predictions are extremely accurate and reliable. The shape of this curve is also very informative. It is nearly flat and stays very close to a Precision of 1.0 across almost the entire range of Recall. This is the hallmark of an exceptional model. It means the model can identify nearly all of the at-risk patients (high recall) while maintaining almost perfect precision (very few false alarms).



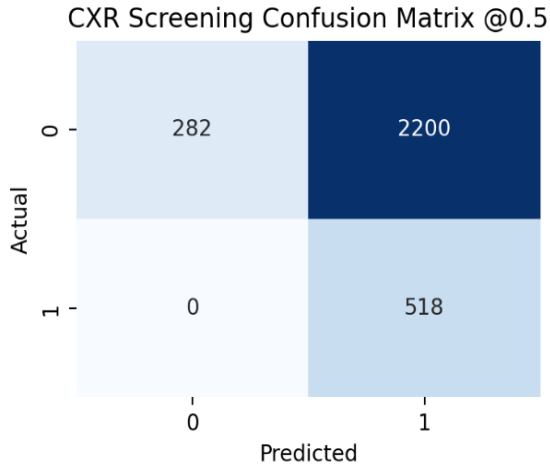
This graph is a Calibration Plot for the 1-year survival prediction model. It's a key visualization that shows how reliable and trustworthy the model's confidence scores are when predicting patient risk.

This graph shows that the model is well-calibrated. The blue line follows the general trend of the ideal line, especially at higher probabilities. This means that the model's confidence scores are reliable. For instance, when the model predicts a risk of around 65-70%, the actual fraction of patients not surviving is very high, approaching 100%. This reliability is critical, as it means clinicians can trust the model when it flags a patient as being at high risk.



This image is a confusion matrix, which provides a detailed breakdown of the performance of the CT Malignancy classifier at a decision threshold of 0.5. It shows the model's correct and incorrect predictions. The most critical insight from this matrix is the number of False Negatives, which is zero. In a medical context, this is an excellent result. It means the model did not miss a single case of malignancy in this dataset, achieving perfect sensitivity.

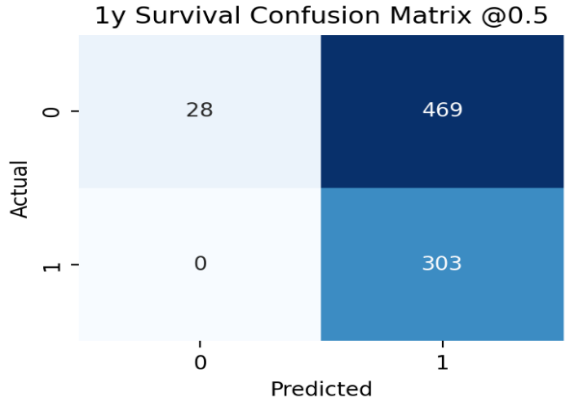
However, this perfect sensitivity comes at a cost, which is the high number of False Positives (749). This indicates that the model is very cautious, flagging many non-malignant cases as potentially malignant to ensure no true cases are missed. This is a common and often desirable trade-off in medical diagnostics, where missing a disease (a false negative) is typically far more dangerous than having a false alarm (a false positive) that leads to further review.



This image shows a confusion matrix for the CXR Screening model, evaluated at a 0.5 decision threshold. It gives a detailed summary of how the model's predictions line up with the actual results for the screening task.The most critical result in this matrix is that there are zero False Negatives. For a screening tool, this is the number one priority. It means the model successfully identified every single potential case in the dataset without missing any.

This perfect recall comes at the cost of a very high number of False Positives (2200). This demonstrates a classic trade-off in medical screening. The model is intentionally designed to be extremely sensitive—casting a very wide net to catch all possible cases. This high number of false alarms is generally considered acceptable in a screening scenario because those cases would then be sent for more definitive follow-up tests (like a CT scan).

In short, this confusion matrix shows a model that is performing its job as an effective screening tool: it finds everyone who might be at risk, even if it means many non-risk cases get a second look.



This image shows a confusion matrix for the 1-year survival prediction model, evaluated at a 0.5 decision threshold. It provides a detailed summary of the model's performance in predicting patient outcomes. The most significant finding in this matrix is that there are zero False Negatives. For a prognostic tool that predicts risk, this is the best possible outcome. It means the model successfully identified every single patient who was at high risk of not surviving one year. This perfect sensitivity is achieved by accepting a high number of False Positives (469). The model is calibrated to be extremely cautious, flagging many patients as "high-risk" even if they ultimately survive. In a clinical setting, this is often a desirable trade-off. It ensures that all genuinely high-risk patients are identified for closer monitoring or alternative treatment strategies, which is far preferable to missing someone who needs urgent attention.

VI.CHALLENGES AND FUTURE DIRECTIONS

CHALLENGES: Several key challenges must be addressed to advance this framework from research to clinical practice. These can be grouped into data-related issues, technical limitations, and practical integration hurdles.

Data and Validation: A primary challenge lies in the heterogeneity of the data, which includes inconsistencies in the LIDC-IDRI dataset's labeling consensus, reconstruction kernels, and slice thickness. The low rate of positive cases in screening scenarios can also negatively impact the positive predictive value, while label noise from varying malignancy ratings by different readers adds another layer of complexity. Furthermore, moving beyond retrospective analysis requires robust external validation through prospective confirmation with clinical cohorts from multiple institutions.

Technical and Ethical Considerations: From a technical standpoint, a key area for improvement is temporal modeling; future work should explore using transformers to integrate long-term imaging data with EHR sequences. Ethically, it is crucial to investigate issues of causality and bias to ensure the model performs fairly across diverse demographics and varying smoking statuses.

Clinical Workflow Integration: Finally, a significant practical hurdle is workflow integration. For the tool to be useful, it must ensure compatibility with standard medical imaging formats like DICOM and include mechanisms for real-time quality control.

VII.FUTURE DIRECTIONS

Future research will be pursued along two main avenues: enhancing the core technical capabilities of the model and ensuring its seamless and trusted integration into clinical practice. A primary goal is to expand multi-modal data integration by incorporating genomics and additional imaging modalities like PET scans alongside CT and clinical records, which could provide a more comprehensive view of lung cancer. Another key technical direction is the development of real-time prognostic models. Such models could predict disease progression dynamically, enabling timely survival analysis, especially in response to treatments like immunotherapy. To improve the model's adaptability, we will also explore transfer learning, which can be used to adjust models trained on large datasets for use in specific clinical settings, thereby improving their effectiveness for different patient populations.

For the model to be adopted clinically, continued progress in Explainable AI (XAI) is essential. Models must not only provide accurate predictions but also offer understandable insights into their reasoning. Improving transparency with frameworks like SHAP is critical for building trust among clinicians in AI-driven decisions. Finally, bridging the gap to practice requires a focus on the EHR workflow. This involves ongoing collaboration between radiologists and oncologists to ensure that machine learning models are designed to meet clinical needs and can be integrated smoothly into existing healthcare processes.

VIII. CONCLUSION

LungCare-Opt stands as a solid and integrated framework designed for the complex challenges of early lung cancer diagnosis and prognosis. By effectively merging clinical data with both CT and CXR imaging, it achieves a holistic view of the patient's condition. The framework's reliance on multi-objective optimization, uncertainty estimation, and clear, interpretable outputs allows it to maintain stable

performance across its diverse tasks. This makes it a flexible and powerful tool for assisting in clinical decision support and triage. Future work will build upon this foundation, with a focus on three key areas: temporal modeling, prospective external validation, and ensuring fair optimization.

REFERENCE

[1] M. Monteleone et al., "Validating an explainable radiomics approach in non-small cell lung cancer," Physica Medica / European Journal of Medical Physics, 2025.

[2] M. Hammad, M. ElAffendi, A. A. Abd El-Latif, A. Ateya, G. Ali, "Explainable AI for lung cancer detection via a custom CNN on CT images," Scientific Reports, 2025, vol. 15, Article 12707.

[3] "Radiomics for lung cancer diagnosis, management, and …," American Journal of Roentgenology, 2025.

[4] "An Optimized Predictive Machine Learning Model for Lung Cancer Diagnosis," Biomed & Pharma Journal, 2024/2025.

[5] R. Javed, T. Abbas, A. H. Khan et al., "Deep learning for lung cancer detection: a review," Artificial Intelligence Review, 2024, vol. 57, Art. 197.

[6] Mohsen Asghari Ilani, Saba Moftakhar Tehran, et al., "Exploring Machine Learning Models for Lung Cancer Level Classification: A Comparative ML Approach," arXiv, 2024.

[7] Hazrat Ali, Farida Mohsen, Zubair Shah, "Improving diagnosis and prognosis of lung cancer using vision transformers: A scoping review," arXiv, 2023.

[8] M. Monteleone, S. Gennai, P. Govoni, and C. Paganelli, "A novel explainable approach in radiomics pipeline for local recurrence prediction of lung cancer: a feasibility study exploiting high energy physics potential to evaluate the model," in Proc. 10th Int. Conf. Bioinformatics Research and Applications (ICBRA), Barcelona, Spain, Sep. 22–24, 2023, pp. 1–4.

[9] Y. P. Zhang et al., "Artificial intelligence-driven radiomics study in cancer: the role and challenges," Military Medical Research, 2023.

[10] E. Dritsas, M. Trigka, "Lung Cancer Risk Prediction with Machine Learning Models," Big Data and Cognitive Computing, 2022.

[11] A. Leblebici, O. Gesoglu, Y. Basbinar, "AI-Decision Support System Interface Using Cancer Related Data for Lung Cancer Prognosis," arXiv, 2021.

[12] Irvin et al., CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. AAAI, 2019.

[13] Katzman et al., DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network. JMLR, 2018.

[14] Wang et al., ChestX-ray8: Hospital-scale chest X-ray database and benchmarks. CVPR, 2017.

[15] Setio et al., Pulmonary nodule detection in CT images: false positive reduction using multi-view CNNs. IEEE TMI, 2016.

[16] Armato et al., The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. Med Phys, 2011.