JOURNAL OF
EMERGING INVESTIGATORS

Article

Lung cancer AI-based diagnosis through multi-modal integration of clinical and imaging data

Arjun S. Ulag[1], Ricardo A. Gonzales[2]

[1] Lakeside Upper School, Seattle, Washington

[2] Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom

SUMMARY

Lung cancer remains the most lethal form of cancer, primarily due to late-stage diagnoses. Early detection significantly improves survival rates, yet it remains challenging. This study aims to enhance early lung cancer diagnosis by developing and evaluating three models: a Multi-Layer Perceptron (MLP) for clinical data, a Convolutional Neural Network (CNN) for imaging data, and a hybrid model combining both data types. We hypothesized that integrating clinical and imaging data would yield higher diagnostic accuracy than single-modality approaches. Using the U.S. National Institute of Health's Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial dataset, which includes over 100,000 chest X-rays and associated clinical records, we preprocessed the data and balanced the distribution of positive and negative samples to train and evaluate our models. The hybrid model achieved the highest accuracy (71.58%), performing slightly better than the MLP (70.88%) and more notably better than the CNN (58.25%), suggesting that multi-modality integration may offer added value under certain conditions. Future research should consider adding other data sources such as genetic and environmental factors to enhance the model's performance further. These findings underscore the promise of multi-modality approaches in transforming lung cancer diagnostics, potentially leading to earlier detection and improved patient outcomes.

INTRODUCTION

Lung cancer begins when abnormal cells in the lungs start to grow uncontrollably, forming tumors, and may spread to other parts of the body through metastasis (1). It is the most fatal form of cancer, leading to more deaths annually in both men and women than any other type of cancer (1). In 2024, lung cancer is projected to account for over 20% of an estimated 611,720 cancer-related deaths in the United States (2). Conversely, lung cancer diagnoses are expected to comprise only 12% of the roughly 2 million new cancer cases anticipated this year in the United States (2). This disparity highlights a mortality rate of 32.4% for lung cancer, which is the highest among all types of cancer.

The elevated mortality rate of lung cancer is primarily due to the disease most often not being diagnosed until it has reached advanced stages. Recent studies have shown that only 30.8% of lung cancer cases are detected at stage I, at which time patients have a significantly higher five-year survival rate than stage IV detection (3). In contrast, a substantial 48% of cases are diagnosed at stage IV, which is often too late for effective intervention, resulting in much lower survival rates (3). For patients diagnosed at stage I, the five-year survival rate can exceed 90%, while those diagnosed at stage IV have survival rates below 10% (4). Clearly, the stage at which lung cancer is diagnosed is closely tied to the prognosis, underscoring the critical need for early detection. Despite the importance of early diagnosis for lung cancer, achieving it is exceedingly challenging. The majority of lung cancer patients do not exhibit obvious symptoms in the early stages of the disease (5). As a result, many are only definitively diagnosed once the cancer has progressed to an advanced stage.

Lung carcinogenesis is influenced by a variety of factors, with smoking being the primary risk factor, accounting for approximately 80% of lung cancer deaths (4). Additional key risk factors include exposure to secondhand smoke, occupational risks such as asbestos and radon, air pollution, genetic factors, and prior chronic lung conditions (4). These risk factors, along with systemic issues, contribute to the observed disparities in lung cancer outcomes. Individuals from minority groups diagnosed with lung cancer, for example, often face poorer outcomes compared to their white counterparts (5). This is due to several factors, including differences in access to healthcare, rates of early diagnosis, and likelihood of receiving surgical treatments (3). Such disparities further complicate the efforts to improve early detection and treatment of lung cancer across all populations. Addressing these systemic and medical challenges is essential for improving survival rates and healthcare equity.

Current diagnostic practices for lung cancer involve a combination of imaging techniques and invasive procedures. Chest X-rays, which are widely used for initial screenings due to their accessibility, minimal intrusion, and cost-effectiveness, unfortunately often miss early-stage tumors, leading to a high rate of false negatives (4). Computed tomography (CT) scans, while more sensitive, are more expensive and carry a higher risk of radiation exposure (4). Recent advancements in low-dose CT scans have shown promise, but they also result in significant incidental findings—potentially important but unrelated abnormalities such as cardiovascular, renal, or thyroid issues—that can escalate healthcare costs due to additional diagnostic testing and follow-up procedures (6). Invasive methods like biopsies are able to provide definitive diagnoses, but they pose risks of complications such as infections and bleeding (7, 8).

In recent years, machine learning approaches have gained prominence in the analysis of chest X-rays for disease detection (9). Multi-Layer Perceptrons (MLPs) are a class of

https://doi.org/10.59720/24-190

feedforward neural networks designed to process structured, tabular data. They consist of multiple layers of interconnected nodes, where each layer transforms input features—such as pixel-based or textural characteristics extracted from X-ray images—through learned weights and activation functions. In prior work, MLPs have been used to classify lung conditions by combining feature engineering techniques like histogram equalization, segmentation, and texture analysis to convert images into numerical vectors suitable for MLP input (10). In contrast, Convolutional Neural Networks (CNNs) are specialized for image data and operate directly on pixel arrays. CNNs use layers of convolutional filters to automatically learn spatial features, such as edges, patterns, and shapes, enabling them to detect abnormalities without manual feature extraction. A recent CNN-based approach introduced a lightweight architecture with custom modules for multiscale feature learning and efficient gradient propagation, enabling end-to-end training on chest X-ray datasets (11).

The aim of this study was to address the need for early diagnosis of lung cancer by proposing the development and evaluation of a multi-modal model that integrates artificial intelligence with current diagnostic methods. We hypothesized that a hybrid approach, combining machine learning algorithms with clinical data on patient history and risk factors (patient demographics, medical and family history, lifestyle factors, etc.) alongside imaging data, would achieve higher accuracy and broader applicability in the early detection of lung cancer. By leveraging the strengths of multiple diagnostic modalities, this model seeks to enhance the accuracy of lung cancer screening and reduce the incidence of false negatives. Furthermore, the proposed model aims to make lung cancer diagnosis more accessible and repeatable on a large scale, potentially transforming the landscape of cancer diagnostics. Although the improvement from combining clinical and imaging data was modest, the results suggest that even limited integration of modalities can enhance diagnostic insight, offering a promising direction for future multi-modal approaches in early cancer detection.

RESULTS

Our research objective was to enhance early cancer diagnosis by integrating multiple data types, specifically clinical and imaging data, thereby leading to better patient outcomes. We developed and evaluated three models for lung cancer diagnosis: an MLP for clinical data processing, a CNN for imaging analysis, and a hybrid model combining these two data types (**Figure 1**). These models incorporate progressively deeper architectures: the MLP processes 24 tabular clinical features, the CNN uses a ResNet152V2 backbone to analyze chest X-rays, and the hybrid model merges both data streams for joint feature learning (**Figure 2**). The developed models were trained and evaluated using data from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial from the U.S. National Institute of Health (NIH), which included over 100,000 chest X-rays and associated clinical records covering patient demographics, body mass index (BMI), smoking history, medication use, family history, and comorbidities (12, 13). The lung cancer cases in the dataset covered a range of clinical stages, with early-stage cases defined as stage I and II and advanced-stage cases as stage III and IV. To ensure all models were trained on equivalent inputs, we filtered the dataset to include only patients with both imaging (as is) and complete clinical records, yielding 12,494 samples. The dataset was then balanced by under sampling the majority class (non-cancer), resulting in a final dataset of 1,422 samples (711 cancer, 711 non-cancer). This dataset was split into 80% for training and 20% for validation. We evaluated model accuracy on the held-out validation set.

We assessed the performance of each model based on accuracy metrics, which measure the proportion of accurate predictions made by the model compared to the diagnosis previously made by clinicians. The MLP model achieved an accuracy of 70.88%, indicating that clinical data alone provides valuable insights for lung cancer diagnosis. The CNN model, trained on imaging data, attained a validation accuracy of 58.25%, suggesting that imaging data alone may have limited

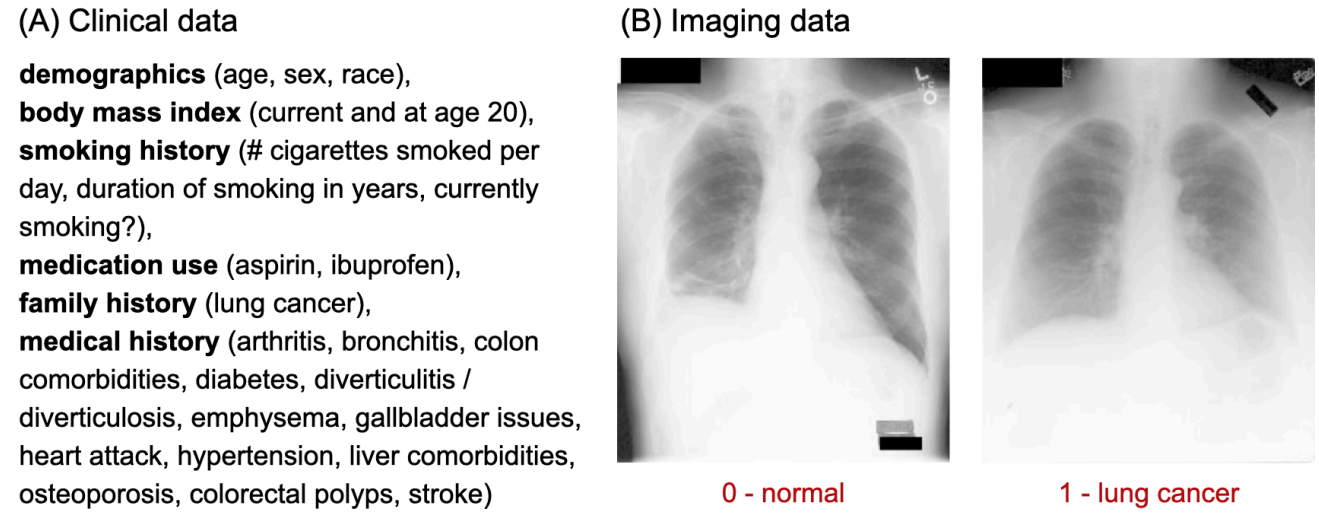

Figure 1: Clinical and imaging data used for lung cancer diagnosis. (A) Clinical data comprising 24 features, grouped into demographics, body mass index, smoking history, medication use, family history, and medical history. (B) Imaging data includes chest X-rays with examples of 'normal' (0; left) and 'lung cancer' (1; right) cases. These data were sourced from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial dataset from the U.S. National Institute of Health (12, 13).

https://doi.org/10.59720/24-190

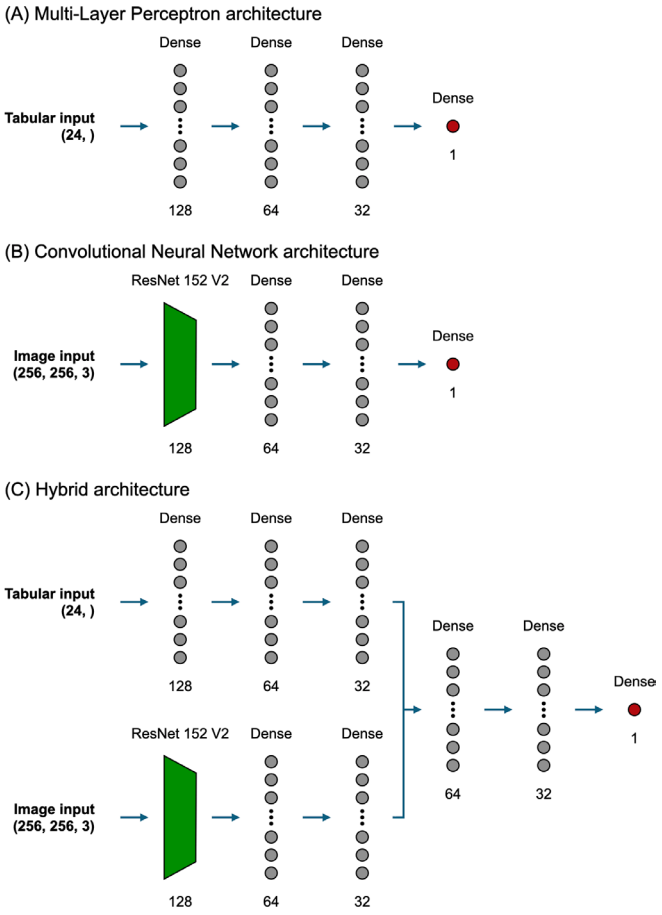

Figure 2: Architectures of the developed models for lung cancer diagnosis. (A) Multi-Layer Perceptron architecture: Processes 24 clinical features through dense layers (128, 64, 32 neurons). (B) Convolutional Neural Network architecture: Uses ResNet152V2 for chest X-ray input, followed by dense layers (64, 32 neurons). (C) Hybrid architecture: Combines clinical and imaging data, each stream passing through dense layers (128, 64, 32 neurons), with a concatenated output processed through additional dense layers (64, 32 neurons). All models culminate in a final 1-neuron classification layer for binary output (0 = normal, 1 = lung cancer).

diagnostic capability. In contrast, the hybrid model, which integrated both clinical and imaging data, reached the highest validation accuracy of 71.58%. However, this improvement over the MLP model was marginal, and without statistical testing, we cannot definitively conclude that the difference is meaningful rather than due to random variation. Given that many X-ray images were rotated or inconsistently formatted, the hybrid model likely did not benefit as fully from the imaging modality as expected.

The confusion matrices comparing actual and predicted results further clarify these findings (**Figure 3**). The MLP model had a false negative rate of 24.6% and a false positive rate of 33.6%, while the CNN model showed a false negative rate of 47.9% and a false positive rate of 35.7%. The hybrid model had a false negative rate of 31.0% and a false positive rate of 25.9%. While it reduced false positives compared to both the MLP and CNN models, the hybrid model increased false negatives relative to the MLP. In the context of early cancer detection—where minimizing false negatives is especially critical—this trade-off underscores the limitations of the hybrid model as implemented.

DISCUSSION

This study highlights the potential of multi-modality approaches in large datasets for early lung cancer detection. We evaluated three models: an MLP using clinical data, a CNN using imaging data, and a hybrid model combining both. The hybrid model outperformed the single-modality models in accuracy, showcasing the advantage of integrating clinical and imaging data. This approach enables the identification of patterns and correlations not visible with a single data type, thus enhancing early detection capabilities essential for improving patient outcomes.

The results do not show clear superiority but rather modest improvement of the hybrid model over the single-modality models. While the hybrid model achieved a slightly higher accuracy than the MLP model, its false negative rate (31.0%) was higher than that of the MLP (24.6%). However, the hybrid model did achieve a lower false positive rate than the MLP (25.9% compared to 33.6%). This trade-off suggests that while combining modalities may help reduce overdiagnosis, it may also increase the risk of missing cancer cases—an important consideration in clinical settings where early detection is critical.

Clinical data provides patient history and risk factors, while imaging data offers visual information about lung abnormalities. Although the hybrid model may capture structural patterns in X-rays—such as irregular opacity, asymmetry, or tissue density shifts—its ability to do so was likely constrained by the quality and quantity of available imaging data. Imaging data were available for only about 1% of the full cohort, and many chest X-rays were rotated or inconsistently collected. In particular, many images were rotated by 90 degrees, which may have affected the CNN's performance, as CNNs are sensitive to orientation (14). However, in specific cases where both data types align, integrating imaging and clinical features may still support a more holistic diagnostic profile, warranting further investigation with higher-quality imaging data and a larger multimodal sample.

The use of the extensive NIH PLCO dataset is an advantage of this study. This dataset, with records from over 155,000 lung cancer patients, offers a rich source of imaging and clinical data. The diversity of the dataset, which includes a wide range of patient demographics, varied clinical presentations, along with access to baseline chest X-ray imaging, helped us to develop models that were robust and applicable across different population segments, enhancing their real-world relevance. Moreover, using a public dataset promotes reproducibility and transparency, allowing other researchers to replicate and build upon our findings, thus accelerating advancements in lung cancer diagnostics (15).

However, the study faced limitations, notably class imbalance within the dataset. The number of positive lung cancer cases was relatively small compared to the total number of records, leading to biased training. To address this, we reduced the sample size of the majority class and adjusted the class weights, but these methods have drawbacks including data loss and increased model complexity (16). Future research should explore data augmentation or synthetic data

https://doi.org/10.59720/24-190

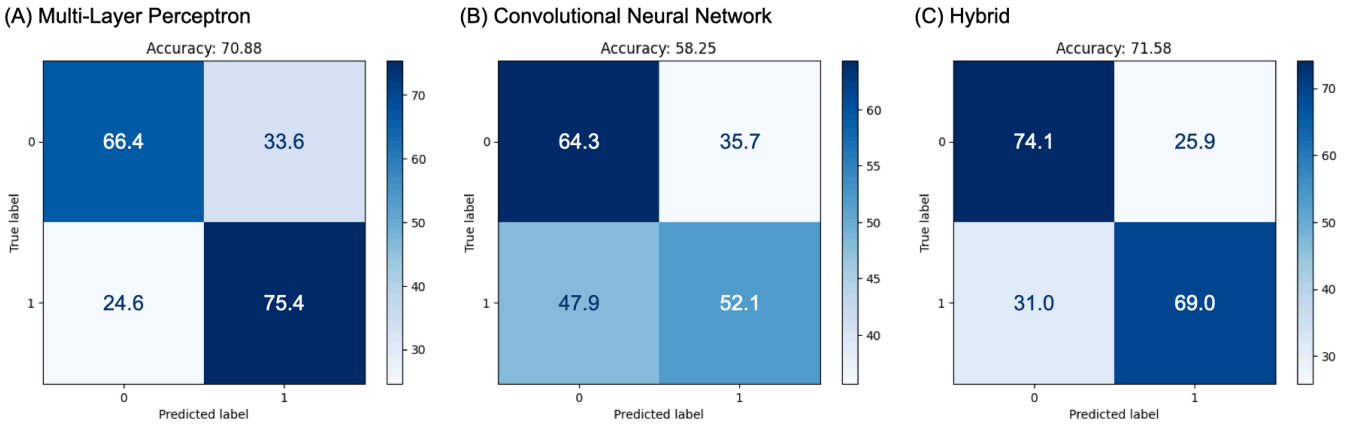

Figure 3: Confusion matrices of model performance. (A) Multi-Layer Perceptron architecture. (B) Convolutional Neural Network architecture. (C) Hybrid architecture. The matrices show the distribution of true positives, true negatives, false positives, and false negatives for each model, illustrating the comparative diagnostic performance. Darker cells indicate higher counts.

generation, such as using Generative Adversarial Networks, to create a more balanced dataset and improve model training (17).

Despite the hybrid model's demonstrated diagnostic accuracy gains over single modality models, there is room for improvement. Future research should focus on adjusting the model architecture, such as incorporating attention mechanisms or domain adaptation layers, which could help the model better learn from heterogeneous data sources and reduce sensitivity to noisy or misaligned inputs (18). Additionally, more data types such as genetic and environmental information might strengthen the model. Genetic data could highlight hereditary lung cancer risks, while environmental data could account for pollutant exposure. Preprocessing methods could also be improved by standardizing image orientation, enhancing contrast, or applying automated quality control to reduce variation in the imaging data before training. Integrating these additional data types and improving data quality can lead to more comprehensive models, capturing a broader spectrum of lung cancer risk factors, ultimately bringing us closer to clinically applicable diagnostic tools. Scaling multimodal learning to larger datasets and more complex architectures can help evaluate whether the benefits of combining clinical and imaging data persist at higher capacity, a key step toward making such models suitable for clinical deployment. More broadly, this approach may inform future efforts in AI-assisted diagnostics across cancer types where risk factors and imaging together provide a more complete view of disease, potentially improving early detection and decision support tools in oncology.

MATERIALS AND METHODS

Data acquisition

The imaging and clinical data for this study were sourced from the NIH PLCO Dataset. This dataset comes from the PLCO Cancer Screening Trial, which is a carefully controlled study that is evaluating screening tests for prostate, lung, colorectal, and ovarian cancers (12, 13). The lung dataset specifically is extensive, covering nearly all the data collected from the PLCO study related to lung cancer screening, its occurrence, and mortality. It includes a single record for each of the roughly 155,000 people who participated in the PLCO trial (12, 13). The dataset spans a diverse demographic, including different ages, genders, and ethnicities. Each patient record includes chest X-ray imaging and detailed clinical data, such as patient history, symptoms, previous diagnoses, and treatment outcomes. Patient profiles cover a wide spectrum of lung tumor types, from benign nodules to advanced malignant tumors. Only samples with confirmed lung cancer diagnoses or no cancer (controls) were used in this study; cases with indeterminate or unrelated diagnoses were excluded.

Data pre-processing

Preprocessing of both imaging and tabular clinical data was a critical step to ensure data integrity and compatibility for this research. The clinical features used were: age, sex, race, current BMI, BMI at age 20, cigarettes per day, years smoked, current smoking status, family history of lung cancer, aspirin use, ibuprofen use, and comorbidities including arthritis, bronchitis, colon issues, diabetes, diverticulitis, emphysema, gallbladder disease, heart attack, hypertension, liver issues, osteoporosis, polyps, and stroke (**Figure 1A**). The selected clinical features were type-cast to appropriate formats (e.g., integers for categorical variables and floats for continuous ones), but no additional normalization was applied. For imaging data, chest X-rays were resized to a uniform dimension of 256x256 pixels to meet the input requirements of the CNN models. This resizing ensures consistent image size and efficient batch processing during model training. The pixel values of the images were normalized to a range of 0 to 1, a standard practice in deep learning that aids in faster convergence during model training. Using patient IDs, the imaging and tabular data were linked to associate each image with its corresponding clinical data. This integration resulted in a comprehensive dataset comprising both image and tabular features, ready to support a multi-modality model. The dataset was balanced by undersampling the majority class (non-cancer) to match the number of cancer cases, resulting in 1,422 samples. The dataset was split into training and a held-out validation sets in an 80:20 split.

https://doi.org/10.59720/24-190

MLP architecture

The MLP model was selected to process tabular data (**Figure 2A**). The MLP architecture began with an input layer designed to handle the specified number of features (n=24). The input layer was followed by several dense layers, each configured with ReLU (Rectified Linear Unit) activation functions to introduce non-linearity, which is essential for capturing complex patterns within the data (19). The architecture included an initial dense layer with 128 neurons, followed by layers with 64 and 32 neurons respectively. Each dense layer was equipped with L2 regularization to prevent overfitting by penalizing large weights. Additionally, dropout layers with a rate of 0.2 were interspersed between the dense layers to further mitigate overfitting by randomly deactivating neurons during training (20). The final layer of the MLP model was a single-neuron dense layer with a sigmoid activation function, providing a binary classification output indicating the presence or absence of lung cancer.

CNN architecture

The CNN model was utilized to process chest X-ray images, leveraging the ResNet152V2 architecture, known for its effectiveness in image recognition tasks (**Figure 2B**) (22). The model began with an input layer designed to handle images resized to 256x256 pixels with three color channels. The input was passed through a data augmentation pipeline, which included random contrast, rotation, brightness adjustments, and zoom, enhancing the model's ability to generalize. Following augmentation, images were rescaled and preprocessed using the ResNetV2 preprocessing function. The base model, ResNet152V2, was pretrained on the ImageNet dataset and used with its weights frozen to leverage transfer learning (22). The output from the base model was passed through dense layers with 64 and 32 neurons, each equipped with ReLU activation and L2 regularization to prevent overfitting. Dropout layers with a rate of 0.2 were also included. The final output layer was a single-neuron dense layer with a sigmoid activation function, providing a binary classification of the images for lung cancer presence.

Hybrid architecture

The hybrid model combined both the MLP and CNN architectures to leverage the strengths of both clinical and image data, hypothesizing that this multi-modality approach would provide superior diagnostic accuracy (**Figure 2C**). The hybrid architecture consisted of two parallel streams: one for processing tabular data and another for processing images. The tabular data stream used the MLP architecture, with dense layers configured as previously described, while the image stream employed the ResNet152V2-based CNN architecture, also as described. The outputs from both streams were concatenated to form a combined feature vector. This combined vector was then processed through additional dense layers with 64 and 32 neurons, each featuring ReLU activation and L2 regularization to ensure robust learning and prevent overfitting. Dropout layers with a rate of 0.2 were included to further enhance generalization. The final layer of the hybrid model was also a single-neuron dense layer with a sigmoid activation function, providing a binary classification output from both clinical and imaging data for lung cancer diagnosis.

Model hyperparameters and optimization

The models were compiled and optimized using TensorFlow and Python, primarily in the Google Colab environment (23, 24). The binary cross-entropy loss function was utilized due to the binary nature of the classification task. The output threshold of the sigmoid activation function was set at 0.5. The Adam optimizer, chosen for its efficiency and adaptability, was used with an initial learning rate set at 0.0001 (25). The training process involved fitting the models over 200 epochs, using a batch size of 16 to balance computational efficiency and training stability. Model weights were saved at each epoch, and the optimal weights were selected based on the convergence point between the training and validation curves, specifically where the difference was minimal. This approach ensured that the models were trained effectively while avoiding overfitting, achieving high performance through careful tuning of hyperparameters and regular monitoring of training and validation metrics. All code for data preprocessing, model training, and evaluation, along with the cleaned tabular dataset, is publicly available at GitHub (https://github.com/arjunulag/plco-multimodal) while access to the full imaging dataset must be requested directly from the PLCO via the National Cancer Institute's data portal.

Model evaluation

The models were evaluated using accuracy metrics and confusion matrices to compare their performance. Accuracy was the primary measure for assessing the models' ability to correctly predict lung cancer presence. Confusion matrices provided detailed insights into the models' performance, showing the number of true positives, true negatives, false positives, and false negatives, which were crucial for understanding the models' diagnostic capabilities and areas needing improvement.

ACKNOWLEDGMENTS

We would like to acknowledge mentorship support from Veritas AI and we would like to thank the National Institute of Health for approving our projects (PLCO-1514 and PLCOI-1517).

Received: July 2, 2024
Accepted: November 5, 2024
Published: June 27, 2025

REFERENCES

1. "Lung Cancer." World Health Organization, 2023, www.who.int/news-room/fact-sheets/detail/lung-cancer.
2. "SEER Cancer Stat Facts." SEER, www.seer.cancer.gov/statfacts.
3. Cainap, Calin, et al. "Early Diagnosis and Screening in Lung Cancer." *American Journal of Cancer Research*, vol. 10, no. 7, 2020, pp. 1993–2009, www.ncbi.nlm.nih.gov/pmc/articles/PMC7407360/.
4. Ning, Jing, et al. "Early Diagnosis of Lung Cancer: Which Is the Optimal Choice?" *Aging* (Albany NY), vol. 13, no. 4, 2021, pp. 6214–6227, https://doi.org/10.18632/aging.202504.
5. Schabath, Matthew B., et al. "Racial and Ethnic Differences in the Epidemiology and Genomics of Lung Cancer." *Cancer Control*, vol. 23, no. 4, 2016, pp. 338–346, https://doi.org/10.1177/107327481602300405.

https://doi.org/10.59720/24-190

6. Gareen, Ilana F., et al. "Medical Care Costs Were Similar across the Low-Dose Computed Tomography and Chest X-Ray Arms of the National Lung Screening Trial (NLST) despite Different Rates of Significant Incidental Findings." *Medical Care*, 2018, p. 1, https://doi.org/10.1097/mlr.0000000000000900.
7. Modi, Pranav, and Abhay Uppe. "Lung Biopsy Techniques and Clinical Significance." PubMed, StatPearls Publishing, 2022, www.ncbi.nlm.nih.gov/books/NBK563153/.
8. Leiten, Elise Orvedal, et al. "Complications and Discomfort of Bronchoscopy: A Systematic Review." *European Clinical Respiratory Journal*, vol. 3, no. 1, 2016, p. 33324, https://doi.org/10.3402/ecrj.v3.33324.
9. Meedeniya, Dulani, et al. "Chest X-Ray Analysis Empowered with Deep Learning: A Systematic Review." *Applied Soft Computing*, vol. 126, 2022, p. 109319, https://doi.org/10.1016/j.asoc.2022.109319.
10. Çelik, Ahmet, and Semih Demirel. "Enhanced Pneumonia Diagnosis Using Chest X-Ray Image Features and Multilayer Perceptron and K-NN Machine Learning Algorithms." *Traitement Du Signal*, vol. 40, no. 3, 2023, pp. 1015–1023, https://doi.org/10.18280/ts.400317.
11. Yen, Chih-Ta, and Chia-Yu Tsao. "Lightweight Convolutional Neural Network for Chest X-Ray Images Classification." *Scientific Reports*, vol. 14, no. 1, 30 2024, https://doi.org/10.1038/s41598-024-80826-z.
12. Kramer, Barnett S., et al. "A National Cancer Institute Sponsored Screening Trial for Prostatic, Lung, Colorectal, and Ovarian Cancers." *Cancer*, vol. 71, no. S2, 1993, pp. 589–593, https://doi.org/10.1002/cncr.2820710215.
13. Prorok, Philip C., et al. "Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial." *Controlled Clinical Trials*, vol. 21, no. 6, 2000, pp. 273S309S, https://doi.org/10.1016/s0197-2456(00)00098-2.
14. Mo, Hanlin, and Guoying Zhao. "RIC-CNN: Rotation-Invariant Coordinate Convolutional Neural Network." *Pattern Recognition*, vol. 146, 2024, p. 109994, https://doi.org/10.1016/j.patcog.2023.109994.
15. Kocheturov, Anton, et al. "Massive Datasets and Machine Learning for Computational Biomedicine: Trends and Challenges." *Annals of Operations Research*, vol. 276, no. 1-2, 2018, pp. 5–34, https://doi.org/10.1007/s10479-018-2891-2.
16. He, Haibo, and E.A. Garcia. "Learning from Imbalanced Data." *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, 2009, pp. 1263–1284, https://doi.org/10.1109/tkde.2008.239.
17. Motamed, Saman, et al. "Data Augmentation Using Generative Adversarial Networks (GANs) for GAN-Based Detection of Pneumonia and COVID-19 in Chest X-Ray Images." *Informatics in Medicine Unlocked*, vol. 27, 2021, p. 100779, https://doi.org/10.1016/j.imu.2021.100779.
18. Li, Jun, et al. "Transforming Medical Imaging with Transformers? A Comparative Review of Key Properties, Current Progresses, and Future Perspectives." *Medical Image Analysis*, vol. 85, 2023, p. 102762, https://doi.org/10.1016/j.media.2023.102762.
19. Lakhani, Paras, et al. "Machine Learning in Radiology: Applications beyond Image Interpretation." *Journal of the American College of Radiology*, vol. 15, no. 2, 2018, pp. 350–359, https://doi.org/10.1016/j.jacr.2017.09.044.
20. Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." *The Journal of Machine Learning Research*, vol. 15, no. 1, 2014.
21. He, Kaiming, et al. "Identity Mappings in Deep Residual Networks." *Computer Vision – ECCV 2016*, 2016, pp. 630–645, https://doi.org/10.1007/978-3-319-46493-0_38.
22. Deng, Jia, et al. "ImageNet: A Large-Scale Hierarchical Image Database." *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, https://doi.org/10.1109/cvpr.2009.5206848.
23. Abadi, Martin, et al. "TensorFlow: A System for Large-Scale Machine Learning." *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 2016, https://doi.org/10.48550/arXiv.1605.08695
24. Ekaba Bisong. *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*. New York, Apress, 2019.
25. Kingma, Diederik P., and Ba, Jimmy. "Adam: A method for stochastic optimization." 3rd International Conference on Learning Representations (ICLR 2015), 2015, https://doi.org/10.48550/arXiv.1412.6980.