Advances in Consumer Research
https://acr-journal.com/
Volume-2 | Issue-6 | Dec: 2025
Original Researcher Article

FusionNet-LC: A Vision-Clinical Fusion Transformer for Lung Cancer Prediction.

Parvathi.M[1]*, Jaishree T K [2], Mahalakshmi C [3], Monisha S [4],Janani S [5], Kaviyadevi S[6]
[1]Electronics and Communication Engineering V.S.B Engineering College Karur,India
Email ID : parvathiiece@gmail.com
[2]Department of Electronics and Communication Engineering V.S.B Engineering College Karur,India
Email ID : jaishreethamil@gmail.com
[3]Department of Electronics and Communication Engineering V.S.B Engineering College Karur,India
Email ID : mahalakskmi2005@gmail.com
[4]Department of Electronics and Communication Engineering V.S.B Engineering College Karur,India
Email ID : monishasenthil4@gmail.com
[5]Department of Electronics and Communication Engineering V.S.B Engineering College Karur,India
Email ID : jananijaanuselvan@gmail.com
[6]Department of Electronics and Communication Engineering V.S.B Engineering College Karur,India
Email ID : kaviya.devi636@gmail.com

ABSTRACT

Lung cancer remains one of the dominant causes of cancer-related deaths. Precision is needed in the classification of various types of lung cancers in order to properly diagnose and treat the disease. But certain difficulties are being faced by the conventional approaches in incorporating information from various modalities of medicine, including CT scans and other variables.

FusionNet-LC bridges the gap by developing a combination AI diagnostic tool using both image intelligence and clinical intelligence. It includes a Vision Transformer (ViT) module, which examines the CT scan image of the lungs, as well as an MLP module, which examines the clinical data. The two modules are integrated using an intelligent Hybrid Attention Module. In this way, the diagnostic tool is capable of flagging both interesting image and clinical features. This enables the tool to make a diagnosis with a certain level of accuracy.

By embracing the use of explainable AI and adaptive optimization, the value of the FusionNet-LC increases among oncologists with respect to speed, accuracy, and intensive data.The traditional approaches adopted for cancer diagnosis rely completely on imaging or clinical information. The imaging modality, though strong, could potentially ignore small details visible only after considering variables specific to patients like age, markers, or smokers vs. non-smokers. Similarly, statistical modeling based on clinicalinformation would not be able to provide details similar to those provided by imaging modality approaches.

There have been major advancements in the area of deep learning and specifically in transformer models with regard to computer vision. There has been another method known as fusion models that have been quite productive in handling multimodal data in medical fields. However, the issue arises when incorporating this data while maintaining interpretability and results with no bias.

The FusionNet-LC bridges this gap through its fusion platform, enabled by artificial intelligence, to synchronize both visual and non-visual modalities to support explainable inference for lung cancer risk and type..

Keywords: Lung Cancer Prediction, Multimodal Deep Learning, Vision Transformer, Clinical Data Fusion, Explainable AI, Medical Decision Support.
.

1. INTRODUCTION:

Lung cancer is still one of the prominent reasons for cancer-related deaths. Precision in the categorization of different types of lung cancers is required for the effective diagnosis and treatment of the disease. But the traditional systems experience certain difficulties in the incorporation of multimodal medical information, such as CT scans and clinical variables.

FusionNet-LC fills the gap with a hybrid AI diagnostic system that blends image intelligence and clinical intelligence. It utilizes a Vision Transformer (ViT) module that analyzes the lung CT scan and an MLP module that analyzes structured clinical information, merging the two through an intelligent Hybrid Attention Module. In this way, the diagnostic system can point out noteworthy image and clinical features, allowing it to make a diagnosis with a level of precision.

With the incorporation of explainable AI and adaptive optimization, the importance of the FusionNet-LC algorithm increases for oncologists regarding faster, accurate, and data-intensive insights.

The conventional methods used for cancer diagnosis rely utterly on either imaging or clinical data. The imaging technique, although robust, might overlook the fine details observable only after integration with patient-specific variables such as age, biomarkers, or smoking status. Likewise, clinical statistical modeling cannot offer the same level of detail as offered by imaging.

There have been breakthroughs in the field of deep learning and the transformer architecture specifically for computer vision. Another approach called fusion models has also been promising for combining multimodal medical data. However, the challenge comes into play while combining the data without compromising on interpretability and avoiding any biased outcomes.

FusionNet-LC fills this gap with its fusion platform driven by AI to harmonize both visual and non-visual modalities towards enabling explainable inference on lung cancer risk and subtype.

2. RELATED WORK

In lung cancer diagnosis and risk assessment, there was extensive use of CAD systems because of the large mortality rate associated with late-stage cancer patients. During the initial stages of research in this field, there had been some use of manual feature extraction in radiographs and CT scans. These features had included texture features, shape features, and intensity features, with the use of SVM and k-NN as classifiers. There had been consistency in the diagnosis process; however, this process had relied heavily on manual features as well.

The arrival of deep learning technology contributed immensely to the improvement in the evaluation of lung cancer with the ability to learn features from images automatically. The Convolutional Neural Network (CNN) model was on top when it came to the detection, definition, and diagnosis of lung nodules. Design concepts for models comprising both 2D and 3D structures were put forth for the evaluation of CT scans. Although the models performed better than the conventional computer-aided systems, the models' limited Receptive Field made it challenging for the evaluation of the global context.

In order to overcome the limitations introduced by CNNs, recent studies have started investigating the transformer-based models for medical image analysis. Vision Transformers (ViTs), which are based on the self-attention mechanism, have attained massive success in predicting the classification of lung disease based on CT scans. In contrast to characteristics of CNNs, ViTs are able to predict distant areas of the body and thus their applications may prove quite beneficial in malignancy patterns identified in lungs. Yet, the transformer models ignore the medical aspects of patients, and they play an important role in making concrete decisions.

Realizing the importance of non-imaging variables, attempts have been made by several researchers in exploring multimodal learning techniques where CT image processing is simultaneously carried out along with clinical information such as age, smoking status, and symptoms. Significant improvements have resulted from these approaches when compared to previous unimodal methods. Nevertheless, many of these approaches adopted direct concatenation techniques, which are inefficient in processing complex relationships.

Interpretability has emerged to be one of the major obstacles to employing deep learning models for clinical tasks. As part of initiatives to improve the interpretability of deep models, attention mechanisms have been coupled with the multimodal models to highlight areas of the clinical image and assign corresponding clinical variables depending on their significance. It has emerged that these models display greater correlation with clinical opinion and can visualize important areas.

Recently, some research work has also started exploring the concept of using explanations for artificial intelligence (explainable AI or XAI) techniques for making trust and usability better for the models involved. Some techniques, such as attention explanation, attribution, or post-hoc explanation models, have been used for some lung cancer predictive models. Nevertheless, most predictive models have included explanation as some sort of addon module or have not considered explanation at the architecture level. In consolidation, although there have been great advancements for the utilization of deep learning approaches along with multimodal prediction approaches for lung cancer prediction, there are several drawbacks for context modeling, efficacy for fusion, or interpretability. The proposed approach, FusionNet-LC, suggests the integration of transformer learning, modeling by structure, or attention-guided modeling.

3. SYSTEM OVERVIEW AND DESCRIPTION OF DATA

Introduction This is a paper that focuses on the overall

The Proposed FusionNet-LC is an AI-based system, a kind of multimodal decision support system that is targeted to assist in lung cancer risk evaluation based on radiologic images and structured clinical data. This approach has been inspired by an important aspect of medicine, which is that correct diagnosis is not only dependent on radiologic images but also on some individual characteristics of the patients. This needs to make effective use of these two kinds of sources of inputs.

It can be observed that the entire pipeline starts with the acquisition of chest CT scan images and their clinical records. Each image of the CT scan will be tagged to one clinical profile through a patient ID for a perfect match of visual and non-visual data.

3.2 CT Image Data and Preprocessing

The chest CT scans used in this system are non-contrast axial scans, very common in lung cancer screening studies. These CT scans are pre-processed following some standard procedures before extraction is done. The pixel values are normalized using a Hounsfield window, making it very easy to visualize results of the lung tissue. All images are normalized in this manner by resizing them to a uniform resolution so as to get uniform samples. The CT scans are gray-scale images; therefore, duplicate images are provided in all three channels, as required by transformer models in computer vision tasks.

3.3 Vision Transformer-Based Imaging Feature Extraction

In the imaging branch, the pre-processed CT scans are inputted into FusionNet-LC. For feature extraction, a Vision Transformer (ViT) has been employed therein. The process of ViT involves dividing each input image into non-overlapping patches embedded into vector representations and augmented with positional embeddings to conserve spatial information. Based on multi-head self-attention mechanisms, the ViT can encode distant relational information in various regions of the lungs, so that it learns the local and global imaging features related to malignancy. The eventual output is obtained in the form of a compact, global image feature representation of the input CT scan from this imaging branch.

3.4 Clinical Data Description and Modeling

Corresponding to this is the clinical modeling branch processing the structured clinical data. Clinical data involves a clinical dataset of patient demographics, patient external exposure factors, patient symptoms, and findings on the imaging study that are documented in the clinical report. Preprocessing of clinical data before feeding into the model involves validation, normalization, and conversion of clinical data into numbers. Continuous variables are normalized to identical scales and categorical properties are converted to a format processable for neural networks.

The processed clinical information is used to train a Multi-layer Perceptron, comprising several fully connected layers with activation functions that introduce nonlinear properties into the model. In effect, this allows prediction of complex relations among multiple clinical variables to determine the risk of lung cancer. Regular methods such as dropout regularization strategies are applied to eliminate possible instances of model overfitting. The network provides a clinical embedding representation that summarizes patient risk information.

3.5 Multimodal Fusion and Risk Prediction

The point or the pivot part of the FusionNet-LC model is the fusion stage. This stage combines high-level features obtained from the Vision Transformer with clinical embedding features obtained through the MLP by a hybrid attention fusion mechanism. Unlike a simple concatenation of the features, the attention mechanism will allow adaptive weighting of features from the two domains, enabling the model to focus more on the relevant aspects for diagnostics. This allows the model to learn non-linearity between the two domains when studying each domain in isolation.

The resultant multimodal feature fusion is then fed through a classification head with fully connected layers and produces the probability output for lung cancer risk. Probability output is informative in making clinical decisions since it allows the clinicians to give an estimate of likelihood of malignancy instead of making decisions on deterministic inputs. The proposed framework is capable of handling image, clinical, or multimodal data depending on input availability.

3.6 Interpretability and Dataset

Management In an attempt to improve the transparency and clinical credibility of the predictions, the architecture of FusionNet-LC integrates mechanisms for interpretation and explanations into the predictions. These include attention maps generated by the imaging component to indicate the parts of the lungs responsible for the patient's predicted risk and attention scores to reveal the clinical variables responsible for the predictions. For the development and testing of models, the multimodal data is parted into training, testing, and validation sets along with conventional splitting techniques. In cases where access to clinical data is an issue, simulated clinical data sets based on statistics could be used for maintaining the privacy of patients and still have conventional distributions for data. The architecture of this model is source agnostics and could easily be coupled with EHR systems for clinical use.

Item	Method	Output
Data	CT+Clinical	Input features
Model	VIT+mlp	Fused embedding
Result	Attention fusion	Risk score

4.PROPSED METHODOLOGY

The proposed FusionNet-LC framework is a heterogeneous artificial intelligence framework composed of a Vision Transformer (ViT) and a Multi-Layer Perceptron (MLP) architecture designed to complement current early lung cancer risk assessment capabilities by leveraging both radiological image modalities and patient-specific clinical variables. In an effort to promote interpretability, a Hybrid Attention Module is used within the framework to emphasize image and clinical variables of primary diagnostic interest. The model is trained and evaluated on a combination of both synthetic and real-world datasets by comprehensively splitting 70%, 15%, and 15% of the subsets for model training, validation, and testing, respectively. In practice, a Lung Cancer Predictor module is used for inference and clinical reasoning, along with a Gemini-based AI consultation layer used to convert prediction results into meaningful clinical practice-relevant recommendations.

4.1 Vision Transformer for CT Feature Extraction

Computed Tomography (CT) imaging serves as a chief mode of initial screening for lung cancer because it allows detailed visualization of anatomy with high sensitivity to variations in tissue characteristics. Conventional convolutional neural networks excel at capturing local details but struggle with modeling inter-region dependencies within the lungs to provide accurate predictions for this task. To achieve this, the design of

FusionNet-LC utilizes the Vision Transformer model actuated as the feature extraction engine.

Each CT scan also undergoes some preprocessing, which involves intensity normalization and resolution resize to a uniform resolution. Additionally, the grayscale nature of the CT scan images needs to be accommodated because the input to Vision Transformers is multi-channel data. Therefore, the CT scan images are transformed to multi-channel data. Then the images are split into uniform-size patches without overlapping, flattened, and projected to embedding space through linear transformation. Finally, positional encoding is added to the extracted embeddings to preserve the spatial information, and the embeddings are fed to the transformer encoder.

The ViT encoder is made up of blocks of self-attention and feed-forward networks that support the ability to model the context of the whole image. This helps in capturing patterns associated with malignancy and other abnormalities of the lung, which can occur far away from the nodules. The outputs are combined to get a compact representation of the image that is used in the fusion of different models in the next step.

4.2 Clinical Data Modeling

Even though the characteristics of images are useful in giving structural details, there are other details from a clinical perspective that are important and help weigh lung cancer risk greatly. These characteristics are composed of demographics, smoking history, symptom patterns, and disease attributes. Since the characteristics are heterogeneous, their modeling approach is done in its own distinct process.

In the proposed system, the clinical variables receive preprocessing operations such as validation, normalization, and encoding to guarantee numerical consistency. Continuous clinical variables receive standardization processing via scaling, whereas categorical clinical variables receive numerical encoding treatments appropriate for the neural network model. The final clinical vector is fed into the Multi-Layer Perceptron.The architecture of the MLP comprises several fully connected layers with non-linear activation functions; hence, the ability to capture the non-linear relationships between the clinical parameters. Dropout regression is also incorporated to avoid overfitting. This resulted in the production of the latent clinical variable for the patient's risk profile to the multimodal fusion component.

4.3 Vision - Clinical Fusion

In order to seamlessly integrate the complementary details from CT images and clinical metadata, the FusionNet-LC architecture employs an intermediate feature fusion method.

The high-level visual features obtained by the Vision Transformer and the clinical embeddings obtained by the MLP are concatenated to obtain a multimodal representation. These are then propagated through fully connected layers to capture cross-modal relationships. It ensures flexible inference by allowing predictions to be done solely by the image, solely by the clinical inputs, or together. Through the combination of both visual biomarkers and clinical risk indicators, the proposed method is able to consolidate predictive robustness over the unimodal methods. This is also able to offer analysis for contributions within each modality, hence promoting transparency during lung cancer risk prediction. V. Interpretability and Clinical Explanation

The FusionNet-LC model makes use of the hybrid attention mechanism in interpreting regions within the CT scan and the relevant clinical factors that affect the prediction model for lung cancer. Additionally, the model uses the generative AI module for the interpretation of the model outputs for clinical explanations.

4.4 Training Strategy

The FusionNet-LC model is trained using a supervised learning approach for binary lung cancer classification. The dataset is divided into training, validation, and testing subsets using a 70%, 15%, and 15% split, respectively, to ensure unbiased evaluation. Model optimization is performed using the AdamW optimizer, and cross-entropy loss is employed as the objective function. Performance is assessed using accuracy and AUC-ROC metrics to evaluate both classification correctness and discriminative capability. This training strategy ensures stable convergence and reliable generalization across unseen data.

5.INTERPRTABILITY AND CLINICAL EXPLANATION

The FusionNet-LC model makes use of the hybrid attention mechanism in interpreting regions within the CT scan and the relevant clinical factors that affect the prediction model for lung cancer. Additionally, the model uses the generative AI module for the interpretation of the model outputs for clinical explanations.

5.1 Hybrid Attention-Based Interpret

Furthermore, interpretability is a key consideration in artificial intelligence technology used in a hospital setting where transparency and traceability are essential in making a medical decision. To deal with interpretability, FusionNet-LC applies a Hybrid Attention Mechanism on both image modalities and clinical modalities.

Under the CT image branch, a Vision Transformer (ViT) is adopted to interpret images from lung scans. The self-attention mechanism in the Vision Transformers allows the Vision Transformers to Zeroiei noteworthy regions in images, including lung nodules, irregularly shaped lungs, and abnormal tissue. These maps are helpful in confirming whether or not the Vision Transformers are pointing at noteworthy regions in images.

At the same time, the clinical data branch is responsible for the MLP to process structured patient data. Attention weights are also assigned to the specific clinical variables of age, smoking status, and nodule details. The feature attention particularly emphasizes the importance of each feature in the process to determine the risk of lung cancer.

With the help of attention-enhanced representations of the combination of both modalities, the FusionNet-LC network allows for the possibility of dual modality interpretation. This helps to overcome the "black box" problem and improves clinical trust as the interpretation of the contribution of imaging and clinical information to the final diagnosis can be easily understood by allowing the reader.

5.2 Generative AI–Based Clinical Explanation

Even though attention interpretability ensures better decision-making understandings in models, attention visualization scores and values only sometimes result in successful communicative expressions. To address the issue, the model, FusionNet-LC, combines a clinical explanation module based on the Generative AI approach.

After that, the system proceeds with extracting the necessary outputs, such as lung cancer risk scores, classifications, and insights obtained through attention from CT scans and clinical characteristics. These outputs are then passed through a generative AI tool using Gemini, such that the predictions are interpreted and presented in a clinical way.

The reports provide an overview of the risk levels, important imaging regions and factors, and possible courses of action based on predictions made by the algorithm. The role of this AI explanation layer becomes essential for the clinicians and patients as it helps to enhance the interpretability and confidence levels associated with clinical decisions made by humans. The use of generative AI closes the gap between predictions and clinical decisions made by humans.

6.EXPERIMENTAL SETUP AND RESULTS

The proposed model, named FusionNet-LC, was trained using the multimodal images provided by CT imaging, along with their respective clinical data. The data was split into training, validation, and test sets, and optimization was done using the AdamW optimizer. The experimental outcome shows the superiority of the proposed model over traditional single-modality-based methods.

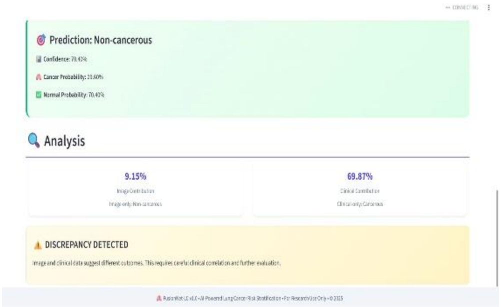
Existing models	FusionNet-LC
CT images only	CT + Clinical data
CNN-based feature extraction	Vision Transformer-based learning
Black-box predictions	Hybrid attention-based interpretability
No clinical explanation	Generative AI-based explanation

6.1 Dataset and Training Parameters

The FusionNet-LC was trained with a multimodal dataset, where a set of chest CT images is paired with relevant clinical information. The clinical information includes demographic information, smoking status, and nodule information that is relevant to lung cancer.

The CT images were preprocessed by resizing and normalizing them. Data augmentations involving rotations and flips were also done. The images were divided into 70% training, 15% validation, and 15% testing sets. This was done to allow fair evaluation of model performances.

In the image branch, a Vision Transformer was used with transfer learning, and in the clinical branch, a Multi-Layer Perceptron was used. Additionally, both branches were trained simultaneously after attention-based feature enhancement. The optimization process used the AdamW optimizer with early stopping and dropout methods. Experiments involving training were done within a GPU-enabled setup, proving its applicability in a cost-effective setup.



6.2 Performance Evaluation and Analysis

The performance of FusionNet-LC was assessed on common metrics such as Accuracy, Precision, Recall, F1-measure, and Area Under Receiver Operating Characteristics Curve (AUC-ROC). The model was found to be consistently accurate for risk-stratification with regard to Lung Cancer.

Comparative analysis indicated that the proposed method of multimodal fusion performs better than models for individual modalities, which consider CT images alone. Including clinical information enhances the dependability of predictions for cases with uncertain imaging results. Additional analysis indicated that measurements taken from attention maps validate the fact that the model focuses appropriately on significant areas of the lungs and clinical information.These results show that by incorporating multimodal learning techniques and interpretability mechanisms, an efficient clinical decision support system is established, making FusionNet-LC useful in medical applications

7. Discussion, Feasibility, and Limitations

The emergence of FusionNet-LC, a multimodal AI-driven diagnostic assistant, marks a significant advancement in supporting early lung cancer detection. By integrating CT imaging data with structured clinical information, the system addresses limitations commonly observed in conventional diagnostic approaches, including inter-observer variability, high clinician workload, and limited interpretability associated with image-only analyses.

7.1 Discussion

FusionNet-LC demonstrates the advantages of multimodal learning in improving diagnostic accuracy. By simultaneously processing imaging features and patient-specific clinical attributes, the system captures complex interactions between radiological findings and established risk factors. This approach mirrors clinical reason ising, wherein physicians consider both imaging results and patient history for comprehensive evaluation. The system's ability to quantify the contribution of each modality enables clinicians to understand how predictions are influenced by imaging versus clinical data, enhancing trust and interpretability.

Attention mechanisms embedded within the model architecture further enhance performance and interpretability. The hybrid attention module emphasizes diagnostically relevant regions in CT scans while weighting informative feature channels. Additionally, the generative AI consultation module translates model outputs into actionable clinical guidance, suggesting appropriate follow-up actions based on calculated risk levels. This combination of predictive performance and explainability addresses a critical barrier to the adoption of AI in clinical workflows.

Operational flexibility is another key benefit of FusionNet-LC. The system can perform image-only, clinical-only, or multimodal inference, allowing it to function effectively even when partial data is available. The use of synthetic clinical data during model development supports experimentation without compromising patient privacy, and the architecture is compatible with integration into electronic health records (EHRs) for real-world deployment.

7.2 Feasibility

The feasibility of deploying FusionNet-LC in clinical practice depends on several factors, including data availability, computational resources, workflow integration, and regulatory compliance. CT scans are routinely acquired in lung cancer screening, and patient-specific clinical variables such as age, smoking history, and symptoms are generally available in clinical records. Preprocessing pipelines and data alignment procedures ensure uniform input formatting, enabling reliable multimodal fusion.

From a computational perspective, FusionNet-LC employs efficient transformer-based imaging models alongside lightweight multi-layer perceptrons (MLPs) for clinical data. This design enables real-time inference on modern workstations equipped with GPUs, allowing seamless integration into existing reporting systems without introducing delays. The modular structure also facilitates future updates, such as the inclusion of additional modalities, including genetic or laboratory data.

Workflow integration is critical for clinical adoption. FusionNet-LC is explicitly designed as a decision-support system rather than a standalone diagnostic tool, ensuring that ultimate responsibility remains with clinicians. By providing interpretable probabilistic outputs and actionable recommendations, the system can be incorporated into radiology review processes, supporting informed decision-making. The dual-modality contribution analysis further allows clinicians to evaluate discrepancies between image-based and clinical-based predictions.

7.3 Limitations

Despite its advantages, FusionNet-LC has inherent limitations. First, the reliance on synthetic clinical data during model development may not fully capture the diversity and complexity of real-world patient populations, potentially affecting generalizability. Validation with real EHR data is essential prior to deployment.

Second, predictive performance depends on the quality and resolution of input CT scans. Variations in imaging protocols or low-quality scans may introduce noise, impacting feature extraction and predictions, though attention mechanisms mitigate this to some extent.

Third, the generative AI consultation module relies on external language models and APIs, which may introduce variability or inaccuracies in recommendations if the underlying knowledge base is incomplete. Human oversight remains crucial to ensure clinically appropriate guidance.

Fourth, FusionNet-LC is intended for risk stratification, not definitive diagnosis. Confirmatory procedures such as biopsies or histopathological evaluations remain essential. Misinterpretation or overreliance on AI outputs could result in inappropriate clinical decisions, highlighting the need for adequate user training.

Finally, ethical and regulatory considerations must be addressed. Patient privacy, data security, informed consent, and compliance with medical device regulations are critical. Continuous monitoring, bias detection, and model updates based on new clinical evidence are necessary to maintain safety, reliability, and fairness.

7.4 Summary

In conclusion, FusionNet-LC integrates multimodal AI, attention-based interpretability, and generative consultation to support early lung cancer detection. Its design aligns with clinical reasoning and workflow requirements, offering flexibility and actionable insights. While the system demonstrates significant promise, careful consideration of limitations, clinical validation, and adherence to ethical standards are imperative. With these measures, FusionNet-LC has the potential to serve as a reliable, interpretable, and practical tool for augmented clinical decision-making in lung cancer diagnostics.

8. CONCLUSION AND FUTURE ENHANCEMENT

This research contributes the design of a hybrid multimodal AI model, named FusionNet-LC, capable of early lung cancer risk assessment. Through the combination of CT scans and clinical inputs by means of a hybrid attention technique, the developed model is able to provide reliable and explainable diagnoses. It is also aided by an explanation component based on generative AI to explain the output of the model. Future directions would involve validating the proposed framework on real-life clinical data and incorporating the model for multi-class cancer staging and survival analysis. Further opportunities exist along the lines of optimizing the solution for real-time use within the hospital setup for long-term monitoring of patients. As a whole, the performance of the proposed solution, namely FusionNet-

LC, is very promising to act as a reliable and interpretable computer-aided solution for lung cancer analysis.

REFERENCES

1) Gao R., Li L., Tang Y., Antic S.L., Paulson A.B., Huo Y., Sandler K.L., Massion P.P., Landman B.A. — *"Deep Multi-task Prediction of Lung Cancer and Cancer-free Progression from Censored Heterogeneous Clinical Imaging"* (arXiv pre-print, 2019)
2) Li (et al.) — *"Fusion of CT images and clinical variables based on deep learning for predicting invasiveness risk of stage I lung adenocarcinoma"* (PubMed, 2023)
3) Wu Y., Ma J., Huang X., Ling S.H., Su S.W. — *"DeepMMSA: A Novel Multimodal Deep Learning Method for Non-small Cell Lung Cancer Survival Analysis"* (arXiv pre-print, 2021)
4) Systematic review: *"Deep Learning Techniques for Lung Cancer Diagnosis with Computed Tomography Imaging: A Systematic Review for Detection, Segmentation, and Classification"* — (Information, 2024)
5)On multimodal dataset: "LUCAS: LUng CAncer Screening" — (MLCDS 2020).

.