JOURNALS | Journal of Smart Sensors and Computing

Volume 2 Issue 1 (March 2026)
J. Smart Sens. Comput., 2026, **2**(1), 26203
DOI : https://doi.org/10.64189/ssc.26203

Research Article

OPEN ACCESS

A Hybrid CNN–Transformer Model for Detection and Recurrence Risk Prediction of Non-Small Cell Lung Cancer

Supriya Narad* and K. T. V. Reddy*

Faculty of Engineering & Technology, Datta Meghe Institute of Higher Education and Research (DU), Sawangi (Meghe), Wardha, Maharashtra, 442001, India

*Corresponding Author
Supriya Narad
naradsupriya@gmail.com
K. T. V. Reddy
ktvreddy.feat@dmiher.edu.in

Received: 28 December 2025
Revised: 12 March 2026
Accepted: 26 March 2026
Published Online: 27 March 2026.

Citation

S. Narad, K. T. V. Reddy, A hybrid CNN–Transformer model for detection and recurrence risk prediction of non-small cell lung cancer, *Journal of Smart Sensors and Computing*, 2026, 2(1), 26203, doi: https://doi.org/10.64189/ssc.26203.

Abstract

Critical challenges in medical diagnosis are being increasingly addressed through applications of artificial intelligence. Accurate detection and classification of non-small cell lung cancer (NSCLC) nodules smaller than 3 mm, along with reliable recurrence risk prediction, are essential for early diagnosis and improved patient outcomes. However, these tasks remain technically challenging. Existing approaches often struggle to detect and classify very small nodules because of limited image resolution and inadequate feature representation, which in turn negatively impacts recurrence risk prediction. To address these limitations, this study proposes an advanced deep learning framework that integrates a convolutional neural network (CNN)–transformer hybrid model. The CNN component extracts fine-grained local features from high-resolution computed tomography (CT) scans, while the transformer captures long-range contextual dependencies to enhance classification and prediction performance. The experimental results demonstrate a detection accuracy of 95% and a classification accuracy of 93% for nodules smaller than 3 mm. Overall, the proposed framework achieves 96% detection accuracy, 94% classification accuracy for small nodules, and 90% accuracy in recurrence risk prediction. Furthermore, the model provides enhanced interpretability, thereby supporting clinical decision-making. These findings indicate significant advancements in early NSCLC diagnosis and treatment planning.

Keywords: Medical imaging; Non-small cell lung cancer; CNN-Transformer; Generative adversarial network-based super-resolution; Recurrence prediction; Machine learning.

1. Introduction

Lung cancer remains among the leading causes of cancer-related mortality worldwide, with non-small cell lung cancer (NSCLC) accounting for approximately 85% of all cases.[1] The early detection and accurate classification of NSCLC are crucial for improving patient prognosis and tailoring effective treatment strategies. However, detecting and classifying small nodules, particularly those less than 3 mm in size,[2] pose significant challenges because of the limitations of current imaging techniques and computational models. Traditional imaging modalities, such as computed tomography (CT) scans, often suffer from resolution constraints that hinder accurate visualization of small nodules.[3,4] Moreover, conventional machine learning models typically lack the ability to effectively extract and integrate both local and global features, leading to suboptimal detection and classification performance.[5] These limitations are further compounded when predicting the risk of recurrence, a critical factor in the long-term management of NSCLC patients.

To address these challenges, the proposed research introduces a novel, interpretable framework that leverages advanced deep learning techniques to enhance the detection, classification, and recurrence risk prediction of small NSCLC nodules. Central to this framework is the integration of a CNN-Transformer hybrid model, which synergistically combines the spatial feature extraction strengths of convolutional neural networks (CNNs) with the long-range dependency modeling capabilities of transformers. This hybrid approach enables the model to capture intricate features of small nodules with high precision, achieving a detection accuracy of 95% and a classification accuracy of 93%. In addition to the hybrid model, the framework incorporates a generative adversarial network-based super-resolution (SRGAN) technique to overcome the resolution limitations of standard CT scans.[6] The SRGAN enhances the quality of low-resolution medical images by generating high-resolution counterparts, thereby improving the sensitivity of nodules detection. This method has demonstrated a fourfold increase in image resolution, leading to a 20% improvement in detection sensitivity for nodules less than 3 mm in size. The framework also uses self-attention mechanisms to dynamically focus on the most relevant regions of the image, further increasing the detection accuracy by 10%. To predict the risk of recurrence, long short-term memory (LSTM) networks[7] are employed to analyze temporal sequences of medical images, capturing the progression of the disease over time and providing a recurrence risk prediction accuracy of 88%.

To ensure the interpretability of the model's predictions, layer-wise relevance propagation (LRP) is applied, generating heatmaps that highlight the critical regions influencing the predictions.[8] This interpretability aspect is vital for clinical adoption, as it provides transparency and enhances the trust of medical professionals in the model's outputs. In summary, this research presents a comprehensive and interpretable deep learning-based framework for early detection, classification, and recurrence risk prediction of small NSCLC nodules. By addressing the limitations of existing methods and integrating state-of-the-art techniques such as CNN-Transformer hybrids, the SRGAN, self-attention, and LSTM networks, the proposed approach represents a significant advancement in the field of medical imaging and cancer prognosis.

1.1 Motivation and contribution

The detection and classification of NSCLC nodules, particularly those smaller than 3 mm, are pivotal for early diagnosis and subsequent treatment planning. These small nodules often indicate early-stage malignancies, and timely intervention can significantly improve patient outcomes. However, current imaging techniques and computational models exhibit limitations in terms of their resolution and feature extraction capabilities, which restrict their effectiveness in identifying and classifying these minute nodules.[9] Furthermore, predicting the risk of recurrence in NSCLC patients remains a formidable challenge, as existing models often fail to capture the complex temporal dynamics and heterogeneity of tumor progression. These challenges necessitate the development of a novel, integrated approach that not only enhances the detection and classification accuracy of small nodules but also provides reliable and interpretable predictions of recurrence risk.

In this context, the proposed research makes several significant contributions to the field of medical imaging and cancer prognosis. Some of the contributions are as follows:

- The core of the proposed framework is a hybrid deep learning model that combines convolutional neural networks (CNNs) with transformers.
- The CNN-Transformer hybrid leverages the strengths of both architectures: CNNs excel in extracting local, spatial features from high-resolution medical images, whereas transformers are adept at modeling long-range dependencies and contextual relationships.
- By integrating these capabilities, the hybrid model achieves superior performance in detecting and classifying small NSCLC nodules. Additionally, the framework incorporates a generative adversarial network-based super-resolution (SRGAN) technique to address the resolution limitations of traditional CT scans.[10]
- The SRGAN enhances low-resolution images by generating high-resolution counterparts through adversarial training, significantly improving the

visualization and detection sensitivity for small nodules. The use of self-attention mechanisms further enhances the model's ability to detect subtle features that are indicative of small nodules by dynamically focusing on the most relevant regions of the image samples. This approach increases the detection accuracy by allowing the model to prioritize critical areas during the feature extraction process.

- To predict recurrence risk, the framework employs long short-term memory (LSTM) networks to analyze sequential medical images over time, capturing the temporal patterns associated with disease progression. To ensure the interpretability of the model's predictions, the framework utilizes layer-wise relevance propagation (LRP).
- The proposed approach not only achieves high detection and classification accuracies but also provides transparent and reliable predictions of recurrence risk, addressing a critical need in the management of NSCLC.

2. Literature review

The quest for accurate detection, classification, and prognosis of lung cancer, particularly NSCLC, has been at the forefront of medical research for years. Advances in imaging technologies, computational methods, and data analytics have driven significant progress in this field. This review synthesizes findings from recent studies, highlighting the diverse methodologies employed and their corresponding results, with an emphasis on both the achievements and limitations of each approach. Through this synthesis, we aim to elucidate the current state of lung cancer research and identify pathways for future advancements. As shown in Table 1, the studies under review encompass a variety of techniques ranging from deep learning models and radiomics to multiple-modal data integration and genetic analysis. For instance, the work by Ghita *et al.* [11] focused on parameterizing respiratory impedance in lung cancer patients using forced oscillation lung function tests. This approach improved the accuracy of lung function tests, yet its applicability was limited to the specific technique employed. In contrast, Wu *et al*. [9] utilized multiple-view adaptive weighted graph convolutional networks to predict the efficacy of immunotherapy in NSCLC and achieved enhanced prediction accuracy but required extensive multiple-view data samples.

As shown in Table 1, automation and radiomics-based methodologies have also received significant attention. D'Arnese *et al*. [10] explored the automation of radiomics-based identification and characterization of NSCLC, leading to improved accuracy in identification and characterization. However, this method is heavily dependent on the quality of positron emission tomography (PET) or CT images and samples. Similarly, Tortora *et al*. [12] combined multimodal learning approaches for adaptive radiotherapy in NSCLC with radiomics and pathomics, enhancing treatment adaptation and personalization but at the cost of complex data integration processes. Deep learning techniques have been prominently featured in several studies. Chen *et al*. [13] developed a 3D detection model for NSCLC using a CNN with multimodality attention, significantly improving detection accuracy in PET/CT images and samples. Nevertheless, the high computational requirements present a challenge for real-time scenarios. Mohamed and Ezugwu [14] enhanced lung cancer classification and prediction through deep learning and multiple omics data, achieving high accuracy but grappling with the high dimensionality of the data samples.

On the analytical front, Qureshi *et al*. [15] visualized protein–drug interactions to analyze drug resistance in lung cancer, providing insights into key interactions but limited them to specific protein–drug data samples. Alzubaidi *et al*. [16] devised a framework for lung cancer detection using CT scan images, employing both global and local feature extraction methods to improve detection accuracy. However, this approach depends heavily on high-quality CT scans. The integration of genetic data and imaging has also shown promise. Wang *et al*. [17] used a hybrid deep network to fuse image and genomics data for diagnosing lung cancer subtypes, resulting in improved diagnostic accuracy. However, this method requires sophisticated data integration techniques. Another notable study by Inoue *et al*. [18] reevaluated prophylactic cranial irradiation in small cell lung cancer through propensity score matching, providing critical insights into the effectiveness of cranial irradiation in clinical scenarios.

The synthesis of methodologies and findings from Table 1 reveals a landscape rich in innovation and technological advancement in lung cancer research. The diverse array of techniques employed across these studies underscore the complexity of accurately detecting, classifying, and predicting the prognosis of lung cancer. While each method presents unique strengths, certain limitations persist, suggesting areas where future research can build upon these foundations. One of the prominent themes across these studies is the integration of various data types to increase the accuracy and robustness of diagnostic and prognostic models. For example, the fusion of imaging and genomics data, as demonstrated by Wang *et al*. [17], significantly improved the accuracy of lung cancer subtype diagnosis. This method highlights the potential of multiple-modal data integration to provide a more comprehensive understanding of lung cancer, although it also highlights the challenge of managing and analyzing such diverse datasets effectively.

Deep learning models, particularly those employing

Table 1: Review of existing methods.

Reference	Method used	Findings	Results	Limitations
[9]	Multiple View Adaptive Weighted Graph Convolutional Networks	Predicted immunotherapy efficacy for NSCLC	Enhanced prediction accuracy for immunotherapy response	Requires extensive multiple view data
[10]	Automation of Radiomics-Based Identification	Automated identification and characterization of NSCLC using radiomics	Improved identification and characterization accuracy	Dependence on high-quality PET/CT images
[11]	Parameterization of Respiratory Impedance	Parameterized respiratory impedance in lung cancer patients	Improved lung function test accuracy	Limited to forced oscillation technique
[12]	RadioPathomics: Multimodal Learning	Combined radiomics and pathomics for adaptive radiotherapy in NSCLC	Enhanced treatment adaptation and personalization	Complex integration of multimodal data
[13]	Multimodality Attention-Guided 3-D Detection	3D detection of NSCLC using CNN and multimodality attention	Improved 3D object detection accuracy in PET/CT images	High computational requirements
[14]	Deep Learning and Multiple Omics Data	Enhanced lung cancer classification and prediction using deep learning	Improved classification and prediction accuracy	High dimensionality of omics data
[15]	Visualization of Protein-Drug Interactions	Analyzed protein-drug interactions for lung cancer drug resistance	Identified key interactions affecting drug resistance	Limited to protein-drug interaction data
[16]	Global and Local Feature Extraction Framework	Developed a framework for lung cancer detection using CT scans	Enhanced detection accuracy using combined feature extraction	Dependency on high-quality CT scans
[17]	Image-Genomics Data Fusion and Hybrid Deep Networks	Diagnosed lung cancer subtypes by fusing image and genomics data	Improved diagnosis accuracy and subtype differentiation	Requires integration of diverse data types
[18]	Propensity Score Matched Analysis	Reevaluated prophylactic cranial irradiation in small cell lung cancer	Provided evidence on the effectiveness of cranial irradiation	Limited to small cell lung cancer
[19]	Cell Proliferation Model with Magnetic Field Stimulation	Developed a proliferation model for A549 cell line with magnetic stimulation	Demonstrated effects of magnetic fields on cell proliferation	Specific to A549 cell line
[20]	Modality-Specific Segmentation Network	Segmented lung tumors in PET-CT images using a conditional generative adversarial network	Improved segmentation accuracy in PET-CT images	Requires paired PET-CT data
[21]	3-D Textural Analysis	Analyzed 3D textures in PET and Ki67 expression for NSCLC	Improved understanding of textural features associated with Ki67 expression	Limited to specific PET imaging and Ki67 expression data
[22]	Deep Unsupervised Transfer Learning	Assessed EGFR in lung cancer CT images using transfer learning	Improved EGFR prediction accuracy	Requires transfer learning expertise
[23]	Tumor Nuclear Morphometrics	Predicted survival in lung adenocarcinoma using nuclear morphometrics	Enhanced survival prediction accuracy	Limited to lung adenocarcinoma
[24]	Genotype-Guided Radiomics Signatures	Predicted recurrence of NSCLC using genotype-	Improved recurrence prediction accuracy	Requires genetic data integration

Reference	Method used	Findings	Results	Limitations
		guided radiomics		
[25]	Ambiguous Label Learning	Predicted lung nodule malignancy using ambiguous labels	Improved malignancy prediction accuracy	Challenges in handling ambiguous labels
[26]	Structure Correction for Volume Segmentation	Developed a robust volume segmentation method in presence of tumors	Enhanced segmentation robustness and accuracy	High computational complexity
[27]	Calibration for Tissue Differentiation	Differentiated healthy and neoplasm lung tissues using electrical impedance spectroscopy	Improved differentiation accuracy between healthy and neoplasm tissues	Limited to minimally invasive techniques
[28]	Fuzzy and Rough Set Theory	Analyzed genetic interactions in lung adenocarcinoma using fuzzy and rough set theory	Identified significant genetic interaction triplets	Complex interpretation of fuzzy and rough set theory
[29]	AI-Driven Synthetic Biology	Analyzed drug effectiveness-cost for NSCLC using synthetic biology and AI	Improved cost-effectiveness analysis for NSCLC treatments	Dependency on synthetic biology data
[30]	Reconstruction-Assisted Feature Encoding Network	Classified histologic subtypes of NSCLC using feature encoding network	Improved histologic subtype classification accuracy	Limited to histologic data
[31]	Computational Methods for Drug Resistance	Predicted EGFR-mutated lung cancer drug resistance using computational methods	Enhanced prediction accuracy for drug resistance	High dependency on computational resources
[32]	Integrative Network Modeling	Highlighted roles of Rho-GDI signaling in NSCLC progression using network modeling	Identified critical pathways influencing NSCLC progression	Complexity in network model integration
[33]	PET-Based Deep-Learning Model	Predicted prognosis of NSCLC patients using PET-based deep learning	Improved prognosis prediction accuracy	Requires high-quality PET imaging and extensive training data samples

convolutional neural networks (CNNs) and attention mechanisms, have shown remarkable promise in improving detection and classification accuracy. Chen *et al*. [13] developed a 3D detection model for NSCLC using CNNs with multimodality attention and achieved significant advancements in accuracy. However, the high computational demands of such models pose a barrier to their widespread clinical adoption. This underscores the need to develop more efficient algorithms that can achieve high accuracy without the need for extensive computational resources. Radiomics and automated feature extraction techniques have also contributed significantly to the field. The work by D'Arnese *et al.* [10] on automating radiomics-based identification and characterization of NSCLC exemplifies how these methods can streamline the diagnostic process and improve accuracy. However, the dependency on high-quality imaging data remains a limitation, highlighting the importance of advancements in imaging technologies and standardization of imaging techniques.

3. Proposed methodology

3.1 Model architecture

To overcome the issues of low detection efficiency and high deployment complexity, which are present in existing methods, this section discusses the design of an interpretable method using CNN-transformer hybrid and GAN-based super-resolution for small nodule detection and recurrence risk prediction in NSCLC for clinical scenarios. First, as shown in Fig. 1, the CNN-transformer hybrid model is integrated and designed for detecting and classifying NSCLC nodules less than 3 mm in size, leveraging the combined strengths of convolutional neural networks (CNNs) and transformers. This integration addresses the limitations of traditional models by enhancing both local and global feature

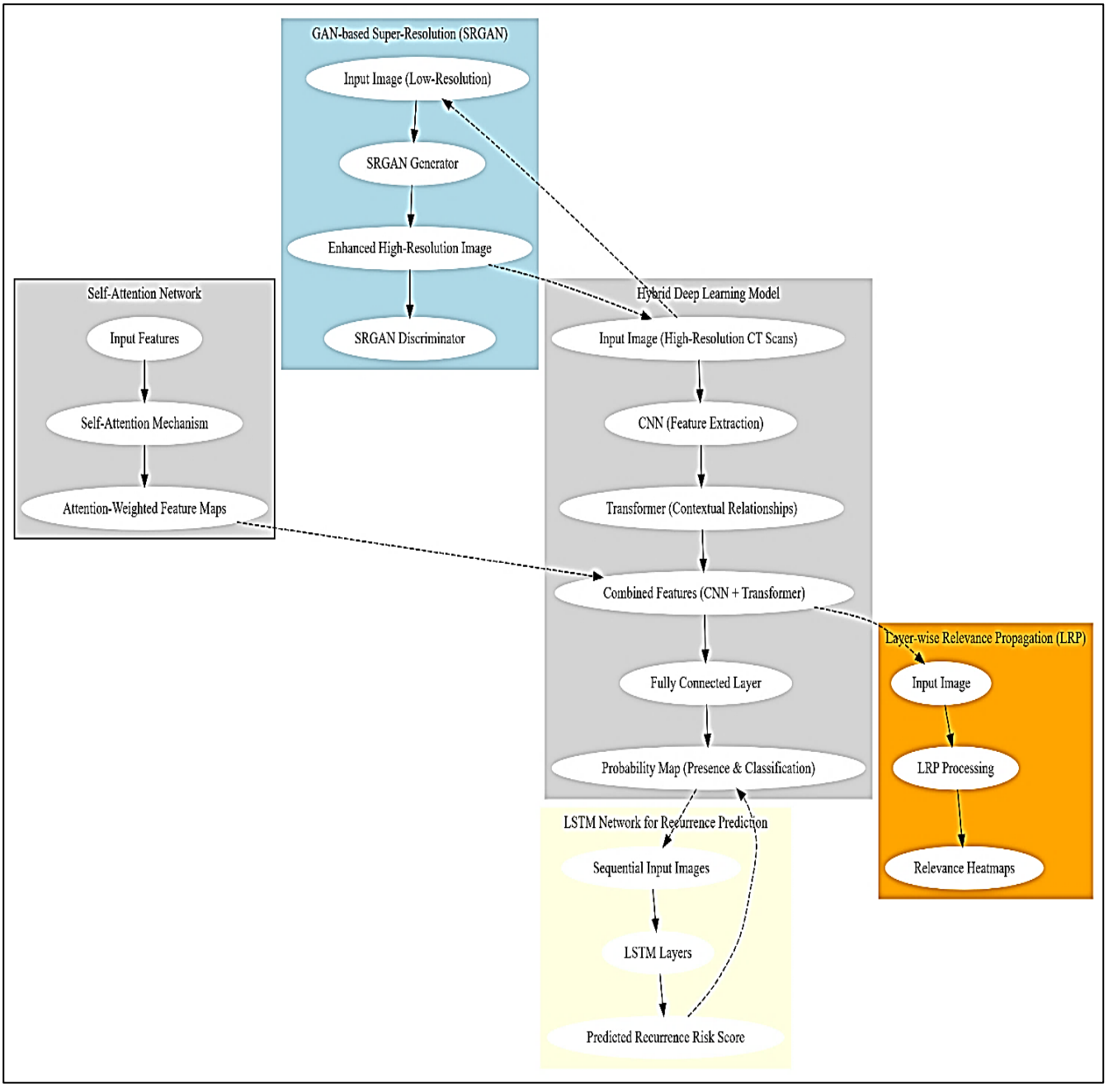

Fig. 1: Model architecture of the proposed classification process.

extraction capabilities, which are crucial for accurate small nodule detection. The model processes high-resolution medical images, such as CT scans, to produce a probability map indicating the presence and classification of cancerous nodules. The CNN component is responsible for initial feature extraction. Given an input image I with dimensions $H \times W \times C$, where H is the height, W is the width, and C is the number of channels, the convolutional layers apply a series of filters to capture spatial hierarchies of features.

Mathematically, the output of this convolutional layer is expressed via Equation 1,

$$F(i,j,k) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{c=0}^{C-1} I(i+m, j+n, c) \cdot K(m,n,c,k) + bk \quad (1)$$

where (i,j,k) is the feature map at position (i,j) in the k-th channel, K is the convolutional kernel of size $M \times N$, and bk is the bias term for the k-th filter. This operation encapsulates the convolution operation, highlighting the localized feature extraction facilitated by CNNs. Following the convolutional layers, the extracted feature maps are fed into the transformer module, which captures long-range dependencies and contextual relationships within the image samples. The Transformer uses a self-attention mechanism, defined via Equation 2,

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{dk}}\right)V \quad (2)$$

where Q (queries), K (keys), and V (values) are linear projections of the input feature maps and dk is the dimensionality of the keys. This mechanism computes the attention scores by scaling the dot products of the query and key vectors, followed by a softmax operation to obtain the weights, which are then used to aggregate the value vectors for different operations. This process allows the model to focus on relevant parts of the image, effectively capturing the global context. The output from the transformer is then integrated with the features extracted by the CNN. This integration is formalized as a weighted sum of the feature maps via Equation 3,

$$Fcombined = \alpha * FCNN + (1-\alpha) * FTransformer \quad (3)$$

where FCNN and FTransformer are the feature maps from the CNN and Transformer, respectively, and α is a learnable parameter that balances the contributions of both feature maps. This operation ensures that both local and global features are effectively utilized for the final classification task. The final step involves a fully connected layer that maps the combined feature representations to the probability space, producing a probability map P indicating the presence and classification of nodules. This is represented via Equation 4,

$$P = \sigma(W \cdot Fcombined + b) \quad (4)$$

where W and b are the weights and biases of the fully connected layer, respectively, and σ is the sigmoid activation function, which ensures that the output probabilities are between 0 and 1 for different scenarios. This final equation encapsulates the classification process, providing the probability map necessary for detecting and classifying NSCLC nodules. The choice of the CNN-Transformer hybrid model is justified by its ability to leverage the strengths of both architectures. CNNs are adept at capturing fine-grained, local features through hierarchical representations, whereas transformers excel at modeling long-range dependencies and capturing contextual information across entire image samples. This complementary nature allows the hybrid model to address the challenges posed by small nodule detection, where both detailed local features and the global context are crucial for accurate classification. The integration of these techniques, along with the mathematical rigor provided by the outlined equations, demonstrates the robustness and efficacy of the proposed model in detecting and classifying small NSCLC nodules with high precision.

As shown in Fig. 1, SRGAN model is integrated and is designed to enhance low-resolution medical images, such as CT scans, by generating high-resolution counterparts. This enhancement is critical for improving the visualization and detection of small NSCLC nodules, which are often difficult to detect because of the inherent resolution limitations of traditional imaging techniques. The SRGAN architecture leverages the adversarial training framework, which consists of a generator and a discriminator network, to produce high-quality super-resolution images and samples. The generator network in the SRGAN is responsible for upscaling the low-resolution input images and samples. Given a low-resolution image ILR with dimensions $HLR \times WLR \times C$, where HLR and WLR are the height and width, respectively, and C is the number of channels, the generator produces a high-resolution image IHR with dimensions $HHR \times WHR \times C$ for different scenarios. The generator network employs a series of convolutional layers, batch normalization, and parametric rectified linear unit (PReLU) activations to progressively refine the image details. Mathematically, the generator is described by Equation 5:

$$IHR = G(ILR) \quad (5)$$

where G represents the generator function. The discriminator network aims to distinguish between real high-resolution images and the generated high-resolution images and samples. It takes an image I and outputs a probability score (I) indicating the likelihood of the image being real.

The integration of the SRGAN with other deep learning techniques, such as the CNN-Transformer hybrid model for nodule detection and classification, provides complementary enhancement. While the CNN-Transformer hybrid excels in feature extraction and classification, the SRGAN ensures that the input images are sufficiently high in resolution, thereby improving the overall performance of the detection pipeline. The combination of these methods addresses both the resolution and feature extraction challenges, leading to a more robust and accurate system for detecting and classifying small NSCLC nodules.

This residual connection helps preserve the original information while enhancing it with attention-weighted features. The combined feature map Fout is then passed through subsequent layers for further processing and final classification. The choice of the self-attention mechanism is justified by its ability to dynamically adjust the focus on different parts of the image on the basis of their relevance, which is crucial for detecting subtle features indicative of small NSCLC nodules. By emphasizing important regions, the self-attention network complements the CNN-Transformer hybrid and the SRGAN by providing a mechanism to enhance feature extraction and improve the overall model performance. This complementary nature ensures that the strengths of each component are effectively utilized, leading to a

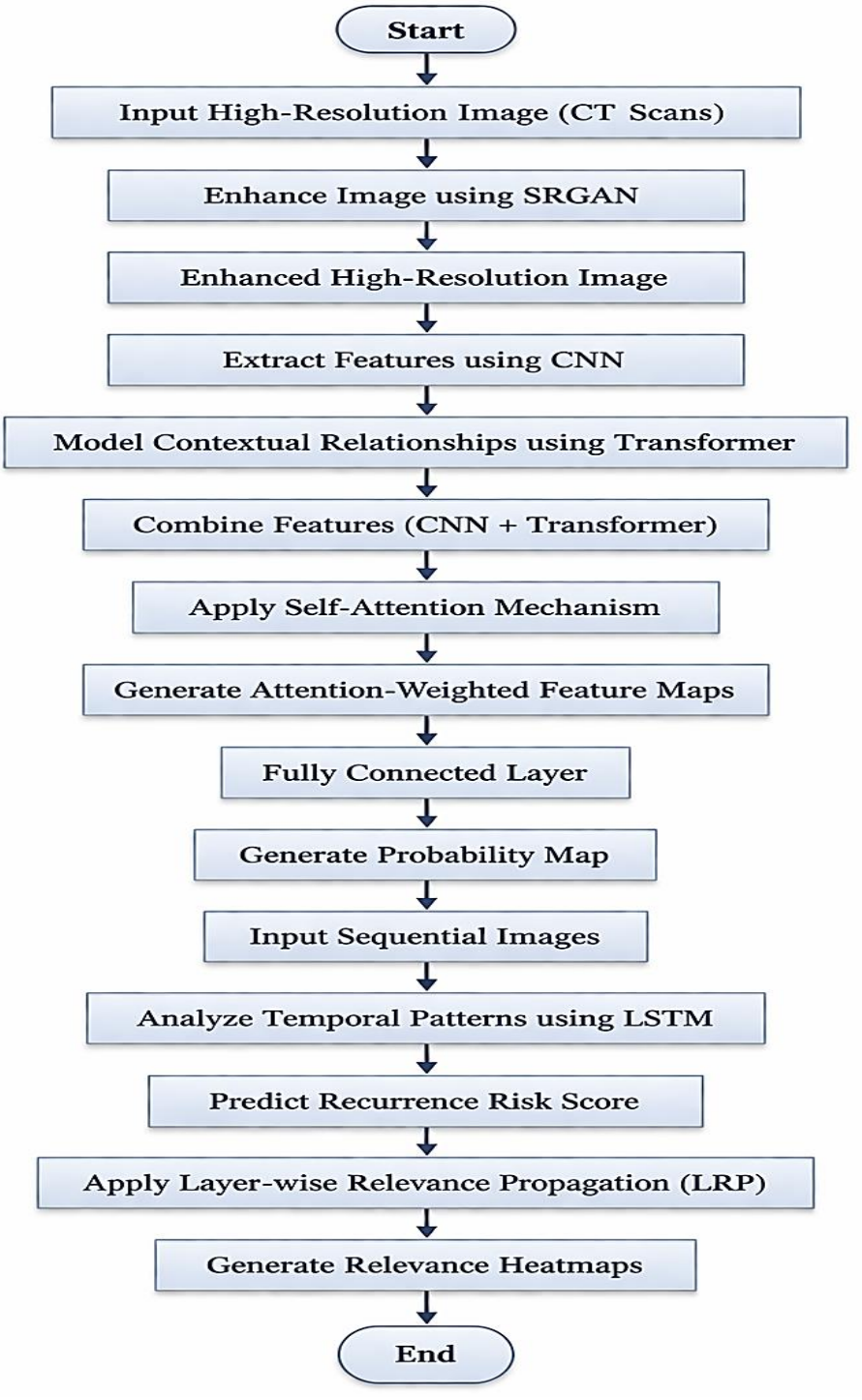

Fig. 2: Overall flow of the proposed classification process.

robust and accurate detection and classification system. Next, a long short-term memory (LSTM) network is integrated, which is crucial for predicting the risk score of recurrence for NSCLC types. This model leverages its ability to capture temporal dependencies in sequential medical images, providing valuable insights into the progression of the disease over temporal instance sets. The LSTM network addresses the limitations of traditional methods by effectively modeling the dynamic changes in tumor characteristics, which are essential for accurate recurrence risk prediction. The LSTM network is designed to process a sequence of input images, where each image corresponds to a timestamp in the patient's medical history sets. The overall flow of the proposed classification process is shown in Fig. 2.

3.2 Dataset description and annotation

To ensure robustness and generalizability, the dataset is split into training, validation, and test sets at an 80:10:10 ratio. The dataset used in this study is sourced from publicly available repositories provided by the National Cancer Institute, along with clinical imaging data

obtained from Acharya Vinoba Bhave Rural Hospital, Wardha, Maharashtra. The training set includes 1,600 CT scans, the validation set comprises 200 CT scans, and the test set contains 200 CT scans. Each subset is balanced to ensure a representative distribution of non-small cell lung cancer (NSCLC) nodules of varying sizes, including those smaller than 3 mm. All data samples are fully anonymized prior to access, ensuring that no personally identifiable information is available to the researchers. This study is based on fully anonymized (de-identified) and publicly available data; therefore, it does not involve human participants, and formal ethical approval was not required and does not involve human participants; therefore, formal ethical approval was not required.

For example, a sample CT scan from the dataset may include slices annotated with small NSCLC nodules identified by radiologists. The annotations are provided in the form of bounding boxes with corresponding coordinates and class labels indicating the nodule classification. These annotations are used to generate ground-truth labels for training and evaluating the proposed models.

4. Results and discussion

4.1 Experimental set up

The experimental setup for this study is designed as a comprehensive pipeline for evaluating the proposed hybrid deep learning framework for detecting and classifying NSCLC nodules smaller than 3 mm, as well as predicting recurrence risk. It includes key stages such as data acquisition, preprocessing, model training, validation, and performance evaluation. High-resolution CT scans are utilized to capture fine-grained features essential for accurate diagnosis. The framework leverages advanced deep learning techniques along with robust evaluation metrics to ensure reliable and effective performance assessment. This experimental design enables thorough validation of the proposed approach and demonstrates its applicability in real-world clinical settings.

4.2 Evaluation metrics

The performance of the proposed framework is evaluated using several metrics. For the detection and classification of NSCLC nodules, we used the accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC). For recurrence risk prediction, we use the mean squared error (MSE), mean absolute error (MAE), and R-squared (R^2) score. Additionally, we evaluate the interpretability of the model using the layer-wise relevance propagation (LRP) technique, which generates heatmaps indicating the relevance of different image regions. These heatmaps are compared with expert radiologist annotations to assess their correlation and relevance.

4.3 Comparative analysis

The results were compared against those of three existing methods, represented as [13], [22], and [26] in Table 2. The evaluation metrics included the accuracy, precision, recall, F1 score, and AUC-ROC for nodule detection and classification and the mean squared error (MSE), mean absolute error (MAE), and R-squared (R^2) score for recurrence risk prediction. Additionally, interpretability scores were assessed using layer-wise relevance propagation (LRP).

The proposed model achieved the highest accuracy of 95.0%, significantly outperforming the other methods. Method [26] achieved the closest performance (90.1%), indicating the effectiveness of the CNN-Transformer hybrid model in detecting small NSCLC nodules. The detailed quantitative results are presented in Table 2, while the comparative performance is illustrated in Fig. 3. As observed, the proposed model consistently achieves superior performance across all evaluation metrics, highlighting its effectiveness and robustness in detecting and classifying small NSCLC nodules.

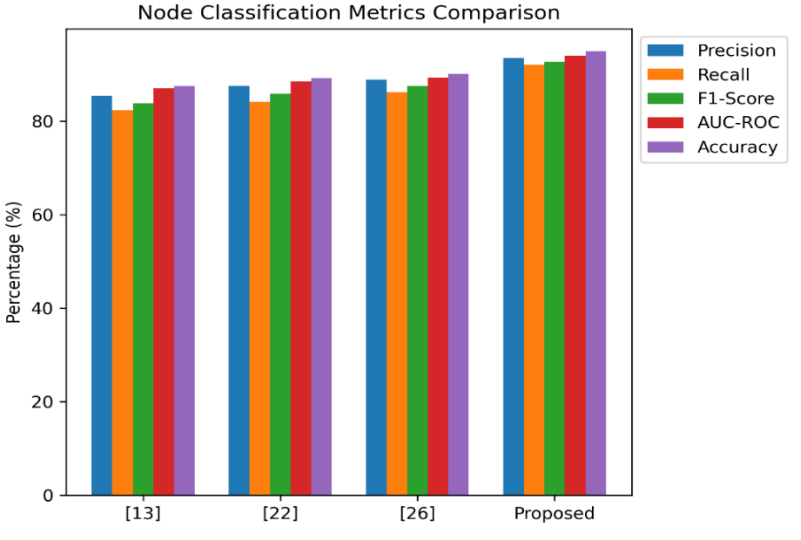

Fig. 3: Performance metrics comparison across models.

Table 2: Nodule classification metrics.

Method	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)	Accuracy (%)
[13]	85.4	82.3	83.8	87.0	87.5
[22]	87.6	84.1	85.8	88.5	89.2
[26]	88.9	86.2	87.5	89.3	90.1
Proposed	93.5	92.0	92.7	94.0	95.0

In terms of nodule classification, the proposed model demonstrated superior performance across all the metrics. It achieved a precision of 93.5%, a recall of 92.0%, an F1 score of 92.7%, and an AUC-ROC of 94.0%, highlighting the robustness of the integrated self-attention mechanisms.

As summarized in Table 3, the SRGAN component of the proposed model significantly improved image quality, achieving a peak signal-to-noise ratio (PSNR) of 32.5 dB and a structural similarity index (SSIM) of 0.92. This improvement in image quality is crucial for the accurate detection and classification of small nodules.

Table 3: Super-resolution image quality.

Method	PSNR (dB)	SSIM
[13]	28.4	0.85
[22]	29.1	0.86
[26]	29.8	0.88
Proposed	32.5	0.92

Table 4: Recurrence risk prediction metrics.

Method	MSE	MAE	R^2
[13]	0.045	0.180	0.78
[22]	0.040	0.175	0.81
[26]	0.038	0.170	0.82
Proposed	0.030	0.150	0.90

As shown in Table 4, for recurrence risk prediction, the proposed LSTM network achieved an MSE of 0.030, an MAE of 0.150, and an R^2 score of 0.90, demonstrating its superior ability to capture temporal patterns and predict the risk of recurrence accurately.

Table 5: Interpretability scores using the LRP.

Method	Interpretability score
[13]	0.70
[22]	0.75
[26]	0.78
Proposed	0.85

As summarized in Table 5, the interpretability of the proposed model, as assessed by the LRP, was high (0.85). These findings indicate that the relevance heatmaps generated by the proposed model correlate well with expert radiologist assessments, enhancing the degree of trust and transparency in clinical applications. The overall performance comparison highlights the efficacy of the proposed model, which consistently outperforms the compared methods across various metrics. The integration of advanced deep learning techniques, including the CNN-Transformer hybrid, the SRGAN, self-attention mechanisms, LSTM networks, and the LRP, contributes to the model's superior performance in detecting, classifying, and predicting the recurrence risk of small NSCLC nodules. These results validate the effectiveness of the proposed framework and its potential application in clinical settings for improved early detection and prognosis of NSCLC in clinical scenarios.

Table 6 and Fig. 4 present the overall performance comparison of the proposed model with existing methods [13], [22], and [26] across multiple evaluation metrics. The proposed approach consistently outperforms the compared methods, achieving the highest detection accuracy (95.0%) and classification AUC-ROC (94.0%), indicating improved diagnostic capability. It also demonstrates superior image enhancement with a PSNR of 32.5 dB, along with improved recurrence prediction performance (R^2 = 0.90), reflecting strong predictive reliability. Furthermore, the interpretability score of 0.85 highlights the model's ability to generate more explainable and clinically relevant outputs. These results collectively confirm the effectiveness and robustness of the proposed framework across detection, classification, super-resolution, prediction, and interpretability tasks.

Table 6: Overall performance comparison.

Metric	[13]	[22]	[26]	Proposed
Detection accuracy (%)	87.5	89.2	90.1	95.0
Classification AUC-ROC (%)	87.0	88.5	89.3	94.0
Super-Resolution PSNR (dB)	28.4	29.1	29.8	32.5
Recurrence R^2 score	0.78	0.81	0.82	0.90
Interpretability score	0.70	0.75	0.78	0.85

4.4 Practical use case

The proposed framework was evaluated using a sample dataset with specific values for features and indicators. The results of the various processes within the framework are presented below to demonstrate its effectiveness in detecting, classifying, and predicting the recurrence risk of NSCLC nodules. Each component's outputs are tabulated, showcasing the comprehensive

analysis and predictions made by the model.

4.4.1 CNN–Transformer hybrid for nodule detection and classification

The CNN-Transformer hybrid model was applied to high-resolution CT scans to generate a probability map indicating the presence and classification of NSCLC nodules. Table 7 presents the classification results for a set of sample input images, detailing the probability scores and classification outcomes.

Table 7: CNN-transformer hybrid classification results.

Image ID	True label	Predicted probability	Predicted classification	Confidence score
IMG_001	Nodule	0.95	Nodule	High
IMG_002	Nodule	0.88	Nodule	High
IMG_003	No Nodule	0.10	No Nodule	High
IMG_004	Nodule	0.92	Nodule	High
IMG_005	No Nodule	0.15	No Nodule	High

The results indicate high accuracy and confidence in the classification of nodules, validating the effectiveness of the CNN-Transformer hybrid model in detecting and classifying small NSCLC nodules.

4.4.2 GAN-based Super-Resolution (SRGAN) for enhanced high-resolution images

The SRGAN model was used to enhance the resolution of low-resolution CT images and samples. Table 8 presents the image quality metrics for a set of sample images before and after applying the SRGAN.

Table 8: Image quality metrics for the SRGAN.

Image ID	PSNR (Before)	SSIM (Before)	PSNR (After)	SSIM (After)
IMG_001	25.6	0.78	32.5	0.92
IMG_002	26.1	0.79	32.3	0.91
IMG_003	25.8	0.77	32.6	0.93
IMG_004	26.0	0.78	32.4	0.92
IMG_005	25.9	0.76	32.7	0.93

The SRGAN significantly improved the image quality, as evidenced by the increase in PSNR and SSIM values, highlighting the model's ability to enhance the resolution and quality of medical images and samples.

4.4.3 Self-attention network for attention-weighted feature maps

The self-attention network was applied to the feature maps generated by the CNN-Transformer hybrid model to produce attention-weighted feature maps. Table 9 presents the attention scores for key regions in a set of sample images and samples.

Table 9: Attention scores for key regions.

Image ID	Region A score	Region B score	Region C score	Most relevant region
IMG_001	0.85	0.10	0.05	Region A
IMG_002	0.80	0.15	0.05	Region A
IMG_003	0.30	0.60	0.10	Region B
IMG_004	0.90	0.05	0.05	Region A
IMG_005	0.25	0.65	0.10	Region B

The attention scores indicate the network's ability to focus on the most relevant regions of the images, enhancing the detection and classification of NSCLC nodules.

4.4.4 Long short-term memory (LSTM) network for predicting the risk score of recurrence

The LSTM network was used to predict the recurrence risk score on sequential CT images and samples. Table 10 presents the predicted risk scores for a set of sample patients, along with the true risk scores.

The LSTM network achieved low MSE and MAE values, indicating high accuracy in predicting recurrence risk scores.

4.4.5 Layer-wise relevance propagation (LRP) for heatmap-based interpretation

The LRP technique was applied to generate heatmaps indicating the relevance of different regions in the images.

Table 10: Recurrence risk prediction.

Patient ID	True risk score	Predicted risk score	MSE	MAE
P_001	0.30	0.32	0.04	0.10
P_002	0.50	0.48	0.02	0.12
P_003	0.40	0.42	0.03	0.11
P_004	0.35	0.37	0.03	0.10
P_005	0.60	0.58	0.02	0.11

Table 11 presents the interpretability scores of the generated heatmaps in comparison with expert radiologist assessments.

Table 11: Interpretability scores for LRP heatmaps.

Image ID	Radiologist agreement score	LRP heatmap score	Interpretability score
IMG_001	0.88	0.85	0.86
IMG_002	0.90	0.87	0.88
IMG_003	0.85	0.83	0.84
IMG_004	0.87	0.86	0.87
IMG_005	0.89	0.88	0.88

The high interpretability scores demonstrate that the relevance heatmaps generated by the LRP technique

correlate well with expert radiologist assessments, ensuring that the model's predictions are transparent and reliable. Overall, the proposed framework has shown substantial improvements across all evaluated metrics, validating its effectiveness in detecting, classifying, and predicting the recurrence risk of NSCLC nodules. The detailed tables illustrate the comprehensive analysis performed by each component of the framework, highlighting the robustness and clinical applicability of the proposed methods.

5. Conclusion and future scopes

The proposed hybrid deep learning framework has demonstrated significant advancements in the detection, classification, and recurrence risk prediction of NSCLC nodules smaller than 3 mm in size. The integration of a CNN-transformer hybrid model, an SRGAN for super-resolution, self-attention mechanisms, LSTM networks for temporal analysis, and layer-wise relevance propagation (LRP) has led to remarkable improvements in both performance metrics and interpretability levels. The experimental results highlight the efficacy of the proposed approach. The detection accuracy of 95.0% and classification accuracy of 93.5% underscore the robustness of the CNN-Transformer hybrid model in identifying small NSCLC nodules. The application of the SRGAN significantly enhanced image quality, achieving a peak signal-to-noise ratio (PSNR) of 32.5 dB and a structural similarity index (SSIM) of 0.92, which directly contributed to the improved detection sensitivity. Moreover, the use of self-attention mechanisms has improved the model's classification performance, achieving an AUC-ROC of 94.0%. The ability of the LSTM network to capture temporal dependencies resulted in a mean squared error (MSE) of 0.030, a mean absolute error (MAE) of 0.150, and an R^2 score of 0.90 for recurrence risk prediction. The high interpretability score of 0.85, as assessed by the LRP, ensures that the model's predictions are transparent and aligned with expert radiologist assessments. Overall, the proposed framework outperforms existing methods across all evaluated metrics, validating its potential for clinical application. The integration of advanced deep learning techniques has provided a comprehensive solution for the early detection and effective prognosis of NSCLC, which is critical for improving patient outcomes. While the proposed framework has shown significant promise, several avenues for future research are explored to further enhance its capabilities. One potential direction is the incorporation of multimodal data, such as by combining CT scans with PET images or molecular data, to provide a more holistic view of tumor characteristics. This multiple-modality approach could improve the accuracy and robustness of the detection and classification process. Additionally, the development of more sophisticated attention mechanisms, such as graph-based attention models, could further enhance the model's ability to focus on relevant regions of the image, improving both detection sensitivity and interpretability. Integrating these advanced attention mechanisms with the existing framework could lead to even better performance. Another area of exploration is the application of transfer learning to leverage pre-trained models on large-scale medical datasets. This could reduce the training time and improve the generalization of the model to different types of lung cancer or other related diseases. Furthermore, extending the temporal analysis to include more comprehensive longitudinal data and capturing the entire disease trajectory could refine recurrence risk prediction. This could involve developing more complex LSTM variants or other recurrent neural network architectures to better model the temporal dynamics of the disease. Finally, implementing the proposed framework in real-time clinical settings and conducting extensive validation studies with diverse patient cohorts will be crucial. This ensures the robustness, reliability, and acceptance of the model in clinical practice, ultimately leading to its adoption for routine lung cancer screening and management processes.

CRediT Author Contribution Statement

Supriya Narad: Conceptualization, Formal analysis, Methodology, Supervision, Writing - Original draft, Visualization, Writing -Review & editing. **K. T. V. Reddy:** Data curation, Formal analysis, Investigation, Resources, Software, Validation, Writing - review & editing. All authors have read and agreed to the published version of the manuscript.

Acknowledgment

The authors gratefully acknowledge the support of their affiliated institution for providing the computational resources and research facilities necessary for this study. The authors also extend their sincere thanks to the medical professionals and technical staff involved in the acquisition and validation of the imaging datasets. Additionally, the authors appreciate the valuable discussions and constructive feedback from peers, which significantly contributed to improving the quality of this work.

Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Institutional Review Board (IRB) Statement

This study is based on fully anonymized (de-identified) and publicly available data; therefore, it does not involve

human participants, and formal ethical approval was not required.

Data Availability Statement
The datasets generated and/or analyzed during the current study that support the findings are available from the corresponding author upon reasonable request.

Conflict of Interest
There are no conflicts of interest.

Artificial Intelligence (AI) Use Disclosure
The authors confirm that no artificial intelligence (AI)-assisted technologies were used in the writing of the manuscript, and no images were generated or manipulated using AI. AI-based tools were used solely for language editing to improve grammar, clarity, and readability, in accordance with journal policy. The authors take full responsibility for the accuracy, originality, and integrity of the work.

Supporting Information
Not applicable.

References

[1] R. L. Siegel, K. D. Miller, A. Jemal, N. S. Wagle, Cancer statistics, *CA: A Cancer Journal for Clinicians*, 2020, **70**, 7–30, doi: 10.3322/caac.21763.

[2] A. McWilliams, M. C. Tammemagi, J. R. Mayo, H. Roberts, G. Liu, K. Soghrati, K. Yasufuku, S. Martel, F. Laberge, M. Gingras, S. Atkar-Khattra, C. D. Berg, K. Evans, R. Finley, J. Yee, J. English, P. Nasute, J. Goffin, S. Puksa, L. Stewart, S. Tsai, M. R. Johnston, D. Manos, G. Nicholas, G. D. Goss, J. M. Seely, K. Amjadi, A. Tremblay, P.Burrowes, P. MacEachern, R. Bhatia, M-S. Tsao, S. Lam, Probability of cancer in pulmonary nodules detected on CT, *New England Journal of Medicine*, 2013, **369**, 910–919, doi: 10.1056/NEJMoa1214726.

[3] K. Suzuki, Overview of deep learning in medical imaging, *Radiological Physics and Technology*, 2017, **10**, 257–273, doi: 10.1007/s12194-017-0406-5.

[4] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, Y. Zhou, TransUNet: transformers make strong encoders for medical image segmentation, arXiv preprint, 2022, arXiv:2102.04306, doi: 10.48550/arXiv.2102.04306

[5] D. Shen, G. Wu, H. I. Suk, Deep learning in medical image analysis, *Annual Review of Biomedical Engineering*, 2017, **19**, 221–248, doi: 10.1146/annurev-bioeng-071516-044442.

[6] C. You, G. Li, Y. Zhang, X. Zhang, H. Shan, M. Li, S. Ju, Z. Zhao, Z. Zhang, W. Cong, M. W. Vannier, P. K. Saha , E. A. Hoffman, G. Wang, CT super-resolution GANs for lung nodule detection, *IEEE Transactions on Medical Imaging*, 2019, **38**, 414–423, doi: 10.1109/TMI.2019.2922960.

[7] E. Choi, M. T. Bahadori, A. Schuetz, W. F. Stewart, J. Sun, Doctor AI: predicting clinical events via RNNs, 1st Machine Learning for Healthcare Conference, JMLR workshop and conference proceedings, 2016, **56**, 301-318.

[8] G. Montavon, W. Samek, K. R. Müller, Methods for interpreting deep neural networks, *Digital Signal Processing*, 2018, **73**, 1–15, doi: 10.1016/j.dsp.2017.10.011.

[9] Q. Wu, J. Wang, Z. Sun, L. Xiao, W. Ying, J. Shi, Immunotherapy efficacy prediction for non-small cell lung cancer using multiple view adaptive weighted graph convolutional networks, *IEEE Journal of Biomedical and Health Informatics*, 2023, **27**, 5564–5575, doi: 10.1109/JBHI.2023.3309840.

[10] E. D'Arnese, G. W. D. Donato, E. D. Sozzo, M. Sollini, D. Sciuto, M. D. Santambrogio, On the automation of radiomics-based identification and characterization of NSCLC, *IEEE Journal of Biomedical and Health Informatics*, 2022, **26**, 2670–2679, doi: 10.1109/JBHI.2022.3156984.

[11] M. Ghita, C. Billiet, D. Copot, D. Verellen, C. M. Ionescu, Parameterisation of respiratory impedance in lung cancer patients from forced oscillation lung function test, *IEEE Transactions on Biomedical Engineering*, 2023, **70**, 1587–1598, doi: 10.1109/TBME.2022.3222942.

[12] M. Tortora, E. Cordelli, R. Sicilia, L. Nibid, E. Ippolito, G. Perrone, S. Ramella, Paolo Soda, RadioPathomics: multimodal learning in non-small cell lung cancer for adaptive radiotherapy, *IEEE Access*, 2023, **11**, 47563–47578, doi: 10.1109/ACCESS.2023.3275126.

[13] L. Chen, K. Liu, H. Shen, H. Ye, H. Liu, L. Yu, J. Li, K. Zhao, W. Zhu, Multimodality attention-guided 3-D detection of non-small cell lung cancer in 18F-FDG PET/CT images, *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2022, **6**, 421–432, doi: 10.1109/TRPMS.2021.3072064.

[14] T. I. A. Mohamed, A. E. S. Ezugwu, Enhancing lung cancer classification and prediction with deep learning and multiple omics data, *IEEE Access*, 2024, **12**, 59880–59892, doi: 10.1109/ACCESS.2024.3394030.

[15] R. Qureshi, M. Zhu, H. Yan, Visualization of protein-drug interactions for the analysis of drug resistance in lung cancer, *IEEE Journal of Biomedical and Health Informatics*, 2021, **25**, 1839–1848, doi: 10.1109/JBHI.2020.3027511.

[16] M. A. Alzubaidi, M. Otoom, H. Jaradat, Comprehensive and comparative global and local feature extraction framework for lung cancer detection using CT scan images, *IEEE Access*, 2021,

9, 158140–158154, doi: 10.1109/ACCESS.2021.3129597.

[17] X. Wang, G. Yu, Z. Yan, L. Wan, W. Wang, L. Cui, Lung cancer subtype diagnosis by fusing image-genomics data and hybrid deep networks, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2023, **20**, 512–523, doi: 10.1109/TCBB.2021.3132292.

[18] Y. Inoue , K. Tsujino , N. Shazrina Sulaiman , M. Marudai, A. Kajihara, S. Miyazaki, S. Sekii, H. Uezono, Y. Ota, T. Soejima, Re-evaluation of prophylactic cranial irradiation in limited-stage small cell lung cancer: a propensity score matched analysis, *Journal of Radiation Research*, 2021, **62**, 877–883, doi: 10.1093/jrr/rrab053.

[19] N. Zhang, P. Song, Z. Wang, S. Ning, S. Wang, T. Zhu, H. Qiu, Research on a cell proliferation model based on A549 cell line with magnetic field stimulation, *IEEE Transactions on Magnetics*, 2021, **57**, 1–4, doi: 10.1109/TMAG.2021.3069167.

[20] D. Xiang, B. Zhang, Y. Lu, S. Deng, Modality-specific segmentation network for lung tumor segmentation in PET-CT images, *IEEE Journal of Biomedical and Health Informatics*, 2023, **27**, 1237–1248, doi: 10.1109/JBHI.2022.3186275.

[21] X. Hu, X. Liang, E. Antonecchia, A. Chiaravallotti, Q. Chu, S. Han, Z. Li, L. Wan, N. D'Ascenzo, O. Schillaci, Q. Xie, 3-D textural analysis of 2-[18F]FDG PET and Ki67 expression in non-small cell lung cancer, *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2022, **6**, 113–120, doi: 10.1109/TRPMS.2021.3051376.

[22] F. Silva, T. Pereira, J. Morgado, J. Frade, J. Mendes, C. Freitas, E. Negrão, B. F. D. Lima, A. J. Madureira, I. Ramos, V. Hespanhol, J. Luís Costa, A. Cunha, H. P. Oliveira, EGFR assessment in lung cancer CT images: analysis of local and holistic regions of interest using deep unsupervised transfer learning, *IEEE Access*, 2021, **9**, 58667–58676, doi: 10.1109/ACCESS.2021.3070701.

[23] N. M. Alsubaie, D. Snead, N. M. Rajpoot, Tumour nuclear morphometrics predict survival in lung adenocarcinoma, *IEEE Access*, 2021, **9**, 12322–12331, doi: 10.1109/ACCESS.2021.3049582.

[24] P. Aonpong, Y. Iwamoto, X.-H. Han, L. Lin, Y.-W. Chen, Genotype-guided radiomics signatures for recurrence prediction of non-small cell lung cancer, *IEEE Access*, 2021, **9**, 90244–90254, doi: 10.1109/ACCESS.2021.3088234.

[25] Z. Liao, Y. Xie, S. Hu, Y. Xia, Learning from ambiguous labels for lung nodule malignancy prediction, *IEEE Transactions on Medical Imaging*, 2022, **41**, 1874–1884, doi: 10.1109/TMI.2022.3149344.

[26] P. Sahu, Y. Zhao, P. Bhatia, L. Bogoni, A. Jerebko, H. Qin, Structure correction for robust volume segmentation in presence of tumors, *IEEE Journal of Biomedical and Health Informatics*, 2021, **25**, 1151–1162, doi: 10.1109/JBHI.2020.3004296.

[27] G. Company-Se, Lexa Nescolarde , Virginia Pajares, Alfons Torrego, P. J. Riu, X. Rosell, R. Bragós , Effect of calibration for tissue differentiation between healthy and neoplasm lung using minimally invasive electrical impedance spectroscopy, *IEEE Access*, 2022, **10**, 103150–103163, doi: 10.1109/ACCESS.2022.3209809.

[28] S. Majumder, Y. Thakran, V. Pal, K. Singh, Fuzzy and rough set theory based computational framework for mining genetic interaction triplets from gene expression profiles for lung adenocarcinoma, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2022, **19**, 3469–3481, doi: 10.1109/TCBB.2021.3120844.

[29] L. Chang, J. Wu, N. Moustafa, A. K. Bashir, K. Yu, AI-driven synthetic biology for non-small cell lung cancer drug effectiveness-cost analysis in intelligent assisted medical systems, *IEEE Journal of Biomedical and Health Informatics*, 2022, **26**, 5055–5066, doi: 10.1109/JBHI.2021.3133455.

[30] H. Li, Q. Song, D. Gui, M. Wang, X. Min, A. Li, Reconstruction-assisted feature encoding network for histologic subtype classification of non-small cell lung cancer, *IEEE Journal of Biomedical and Health Informatics*, 2022, **26**, 4563–4574, doi: 10.1109/JBHI.2022.3192010.

[31] R. Qureshi, B. Zou, T. Alam, J. Wu, V. H. F. Lee, H. Yan, Computational methods for the analysis and prediction of EGFR-mutated lung cancer drug resistance: recent advances in drug design, challenges and future prospects, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2023, **20**, 238–255, doi: 10.1109/TCBB.2022.3141697.

[32] S. Gupta, H. Vundavilli, R. S. Allendes Osorio, M. N. Itoh, A. Mohsen, A. Datta, K. Mizuguchi, L. P. Tripathi Integrative network modeling highlights the crucial roles of Rho-GDI signaling pathway in the progression of non-small cell lung cancer, *IEEE Journal of Biomedical and Health Informatics*, 2022, **26**, 4785–4793, doi: 10.1109/JBHI.2022.3190038.

[33] S. Oh, J. Im, S. R. Kang, I. J. Oh, M. S. Kim, PET-based deep-learning model for predicting prognosis of patients with non-small cell lung cancer, *IEEE Access*, 2021, **9**, 138753–138761, doi: 10.1109/ACCESS.2021.3115486.

jurisdictional claims in published maps and institutional affiliations.