Journal of Translational Medicine

RESEARCH

Open Access

Deep learning radiopathomics predicts targeted therapy sensitivity in EGFR-mutant lung adenocarcinoma

Taotao Yang[1,2†], Xianqi Wang[1,2†], Yuan Jin[3], Xiaohong Yao[4], Zhiyuan Sun[5], Pinzhen Chen[1], Suyi Zhou[1], Wentao Zhu[3*] and Wei Chen[1,2*]

Abstract

Background Ttyrosine kinase inhibitors (TKIs) represent the standard first-line treatment for patients with epidermal growth factor receptor (EGFR)-mutant lung adenocarcinoma. However, not all patients with EGFR mutations respond to TKIs. This study aims to develop a deep learning radiological-pathological-clinical (DLRPC) model that integrates computed tomography (CT) images, hematoxylin and eosin (H&E)-stained aspiration biopsy samples, and clinical data to predict the response in EGFR-mutant lung adenocarcinoma patients undergoing TKIs treatment.

Methods We retrospectively analyzed data from 214 lung adenocarcinoma patients who received TKIs treatment from two medical centers between September 2013 and June 2023. The DLRPC model leverages paired CT, pathological images and clinical data, incorporating a clinical-based attention mask to further explore the cross-modality associations. To evaluate its diagnostic performance, we compared the DLRPC model against single-modality models and a decision level fusion model based on Dempster-Shafer theory. Model performances metrics, including area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), were used for evaluation. The Delong test assessed statistically significantly differences in AUC among models.

Results The DLRPC model demonstrated strong performance, achieving an AUC value of 0.8424. It outperformed the single-modality models (AUC = 0.6894, 0.7753, 0.8052 for CT model, pathology model and clinical model, respectively. $P < 0.05$). Additionally, the DLRPC model surpassed the decision level fusion model (AUC = 0.8132, $P < 0.05$).

Conclusion The DLRPC model effectively predicts the response of EGFR-mutant lung adenocarcinoma patients to TKIs, providing a promising tool for personalized treatment decisions in lung cancer management.

Keywords CT, WSI, Deep learning, EGFR, Targeted therapy, Multi-modality

†Taotao Yang and Xianqi Wang contributed equally to this work.

*Correspondence:
Wentao Zhu
wentao.zhu@zhejianglab.org
Wei Chen
landcw@tmmu.edu.cn
[1]Department of Radiology, Southwest Hospital, Army Medical University (Third Military Medical University), Chongqing 400038, China
[2]Yu-Yue Pathology Scientific Research Center, Chongqing 400039, China
[3]Zhejiang Lab, Hangzhou 311121, China
[4]Institute of Pathology and Southwest Cancer Center, Southwest Hospital, Army Medical University, (Third Military Medical University), Chongqing 400038, China
[5]Department of Radiology, Jinling Hospital, Medical School of Nanjing University, Nanjing 210002, China

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Background

Lung cancer remains the leading cause of cancer-related deaths worldwide, with lung adenocarcinoma being the most common histological subtype of non-small cell lung cancer (NSCLC) [1, 2]. The epidermal growth factor receptor (EGFR) gene mutation represents the most common oncogenic driver in lung adenocarcinoma is the [3]. For patients with EGFR-mutated lung adenocarcinoma, tyrosine kinase inhibitors (TKIs) have been recommended as the first-line treatment by the national comprehensive cancer network (NCCN) [3]. TKIs have significantly improved progression-free survival and overall survival in these patients [4]. However, up to 30% of patients with EGFR mutations exhibit primary resistance to TKIs, resulting in a limited therapeutic benefit [5]. Therefore, early and accurate prediction of TKIs treatment sensitivity in EGFR-mutant patients is crucial to informing therapeutic strategy.

Currently, clinical predictions of TKIs treatment sensitivity rely primarily on conventional information, such as demographic data and CT morphological characteristics [6]. However, demographic features are low-dimensional and provide limited predictive value. Additionally, the assessment of CT morphological features is often subjective, relying on clinicians' experience, which can lead to low inter-observer agreement and reduced accuracy. Deep learning, a key subfield of artificial intelligence, excels in automatically extracting and learning features from medical images to characterize heterogeneity of tumors [7, 8]. This approach has been widely applied in various aspects of the medical field [9, 10]. Deep learning methods based on CT images have been used for the differential diagnosis, prognosis prediction, tumor subtype prediction, and gene mutation prediction [11–14]. A previous study [15] used deep learning to predict responses to targeted therapy using CT imaging; however, the model's limited predictive performance restricts its clinical utility.

In parallel, deep learning has demonstrated outstanding performance in the analyzing pathological images and has been widely used for tasks such as tumor diagnosis, classification, and prognosis [16–19]. Studies have highlighted the potential of deep learning to identify prognostic factors from whole slide images (WSIs) and predict treatment response [20–22]. A previous study [23] has shown that features extracted from hematoxylin and eosin (H&E)-stained digital pathology images could predict EGFR mutation status using deep learning methods. However, few studies have focused on predicting the sensitivity of targeted therapies in EGFR-positive patients using digital pathology images. Furthermore, histopathological sections are often obtained through localized tumor sampling, which may not fully represent the entire tumor.

Recent studies have shown that integrating multimodal information can better reflect tumor heterogeneity across different scales, thereby enhancing the predictive power of single-modality approaches [24–27]. Huang et al. [28] developed a radiopathomics model that combined ultrasound image data of the whole tumor (macroscopic scale) with digital pathology data from localised areas (microscopic scale) to predict ductal tumors in breast cancer, The integrated model outperformed single-modality models. Building on this, we hypothesized that combining CT imaging with digital pathology data may would improve predictive performance in the early assessment of targeted therapy response. Therefore, in this study, we proposed a deep learning radiological-pathological-clinical (DLRPC) model that integrates pre-treatment CT images and H&E-stained WSIs from aspiration biopsy samples to predict the sensitivity of EGFR-mutant patients to targeted therapy.

Materials and methods

Patients

This retrospective study was approved by the Ethics Committee of the First Affiliated Hospital of the Army Military Medical University (Ethics Approval No. KY2023143), which exempted patients from the requirement of informed consent. Both medical centers followed the same inclusion and exclusion criteria. Patients who meet the following criteria were included in the study: (1) histopathological confirmation of lung adenocarcinoma through lung aspiration biopsy and EGFR gene sequencing results indicating EGFR (+); (2) a baseline chest CT scan conducted within 4 weeks prior to the aspiration biopsy; (3) high-resolution digital images of H&E stained tissue samples; (4) treatment with standard first-generation EGFR-TKIs, including gefitinib, erlotinib, or icotinib; (5) availability of complete clinical and follow-up data. Patient were excluded from the study based on the following criteria: (1) inability to segment CT lesions (supplementary Fig. S1A); (2) poor CT image quality with severe artefacts (supplementary Fig. S1B); (3) discoloration or missing tumor tissue in histological images. Ultimately, 214 patients with lung adenocarcinoma were included: 131 patients from the First Affiliated Hospital of Army Medical University (Hospital A) between January 2014 and June 2023, and 83 patients from Jinling Hospital, affiliated with Nanjing University Medical School (Hospital B), between September 2013 and October 2019 (Fig. 1).

Clinical information

The collected clinical data included gender, age, smoking history, cytokeratin 19 fragment (CYFRA21-1), carcinoembryonic antigen (CEA), neuron-specific enolase (NSE), T stage, N stage, M stage, and clinical stage.

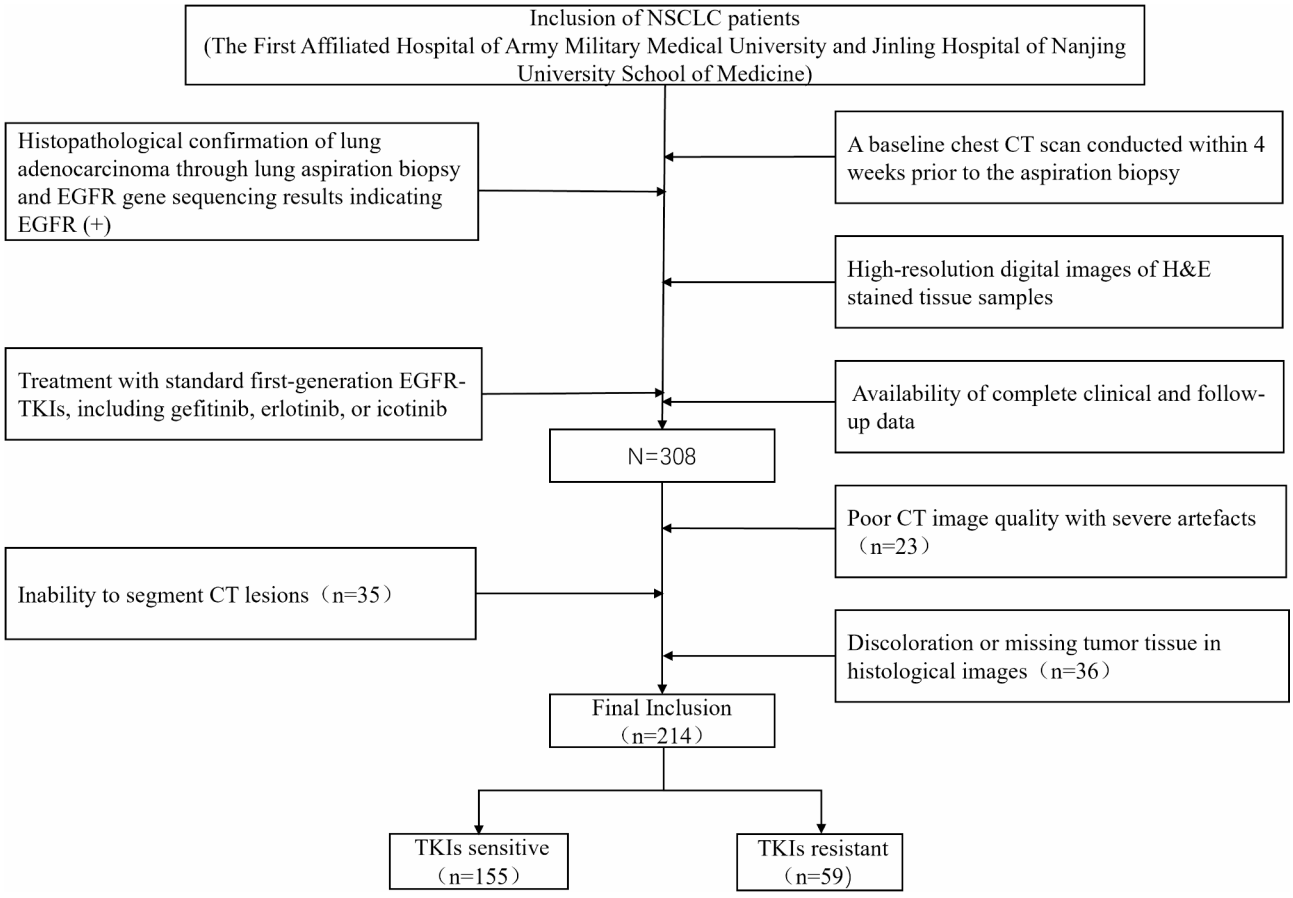

Fig. 1 Flowchart of patient enrolment

We analyzed pre-treatment baseline chest CT images of the patients. The CT morphological characteristics of the patients were assessed and evaluated by two radiologists with 10 and 15 years of experience in thoracic imaging diagnosis, respectively. Both radiologists were blinded to all patient information. The evaluation included tumor location, maximum lesion diameter, shape, interface, lobulation, spiculation, pleural retraction, bronchial cut-off sign, pneumonia, density, necrosis, air-containing cavities, lymphadenopathy, and pleural effusion. In cases of disagreement, the two radiologists discussed and reached a consensus.

Treatment response assessment

After 3 months of standardised TKIs, a physician with 10 years of experience in diagnostic chest imaging at both hospitals assessed treatment response using the Response Evaluation Criteria in Solid Tumors (RECIST 1.1) [29]. Patient were classified as treatment sensitive if they achieved partial remission (PR) or complete remission (CR), while stable disease (SD) or progressive disease (PD) were classified as exhibiting primary resistance.

CT imaging protocol and tumor segmentation

The chest CT scans were performed with the patients in a supine position, extending from the thoracic inlet to the lung base. The CT images were stored and exported in DICOM format. Detailed CT scanning and reconstruction parameters are provided in Table 1.

Tumor segmentation was performed on lung window images using ITK-SNAP software (version 4.0.1). A radiologist with 10 years of experience in thoracic imaging manually delineated the region of interest (ROI) on thin-slice lung window images. The radiologist was required only to outline a general ROI that encompassed the tumor area (Fig. 2A).

WSI acquisition and ROI delineation

The biopsy tissue obtained via puncture was fixed in formalin and embedded in paraffin, followed by sectioning at 4 μm intervals. Each slide was then stained with H&E. Prior to scanning, a senior pulmonary pathology expert selected and confirmed representative slides containing abundant tumor tissue. The selected slides were then digitally scanned using a KF-PRO-400 scanner (Ningbo Jiangfeng Biological Information Technology Co., China) at 40x magnification to obtain high-resolution whole

Table 1 Parameters of CT scan and image reconstruction in twin centres hospitals

Hospital	apparatus	Tube voltage (kV)	Tube current (mA)	Pitch	Rotation Time(s)	Collimation (mm)	Convol-ution kernel	Slice Thickness (mm)	Slice inter-val (mm)
The First Affili-ated Hospital of Army Mili-tary Medical University	SIEMENS Force	120	automatic	1.2	0.5	192×0.6	Br59	1–3	1–3
	SIEMENS Flash	100	automatic	1.5	0.5	128×0.6	B80f	1–3	1–3
	SIEMENS Emotion	110	automatic	1.5	0.6	16×1.2	B90s	1.5-3	1.5-3
	SIEMENS AS	120	automatic	1.2	0.5	128×0.6	I70f	1–3	1–3
	GE Optima	120	automatic	0.984	0.5	64×0.625	standard	1.25-3	1.25-3
Jinling Hos-pital, Nanjing Univer-sity School of Medicin	SIEMENS Perspective	130	automatic	1.2	0.6	64×0.6	I50S	1–3	1–3
	SIEMENS Flash	100	automatic	1.2	0.5	128×0.6	L50f	1–3	1–3
	United-imaging UCT 760	120	automatic	1.0875	0.5	128×0.625	B_SHARP_C	1–3	1–3

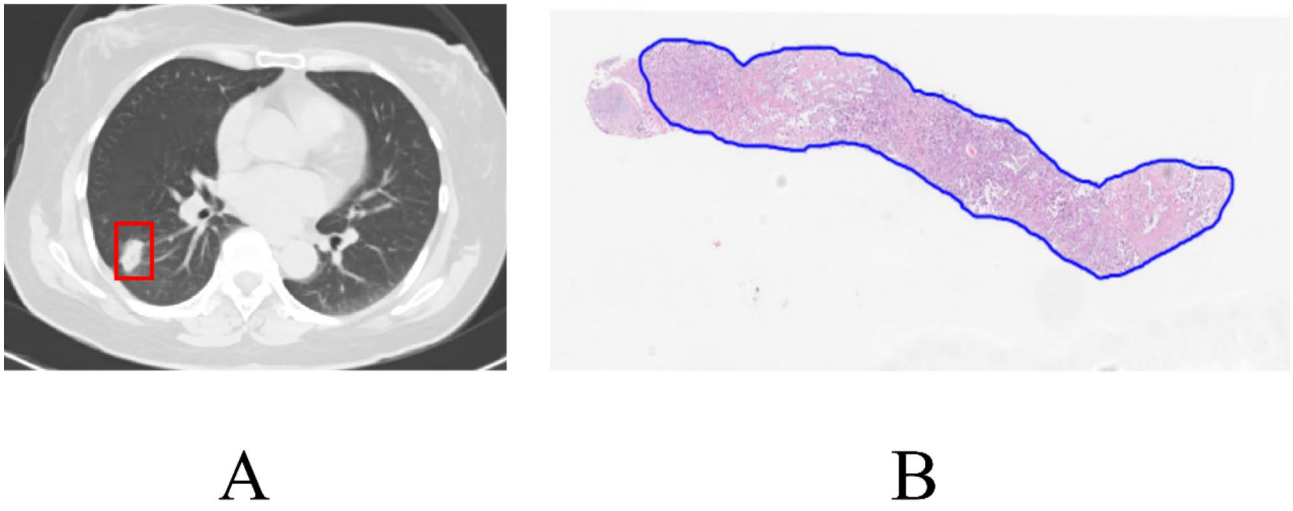

Fig. 2 ROI segmentation. Presentative diagrams of manual annotation of tumor regions in CT (**A**) as well as aspiration biopsy WSI (**B**) in EGFR-mutant lung adenocarcinoma patients

slide digital images [30]. Only one WSI was obtained per patient. The ROI was delineated by a pathologist using the Automated Slide Analysis Platform, which includes all tumor tissues. (Fig. 2B).

Feature extraction

The feature extraction process, which involves pathological, radiological, and clinical feature extractors, is crucial for capturing informative data and performing dimensionality reduction. In this study, we utilized a 2D and 3D vision transformer (ViT) network and ResNet-50 accordingly for CT images, pathological images and clinical data, respectively.

For pathological feature extraction, the high-resolution WSIs are firstly cropped into 256 × 256 patches, which were then fed into the 2D Transformer network module. To enable the model to adaptively attend to discriminative features across heterogeneous tissue contexts, we incorporate the multi-head self-attention mechanism into the network architecture. This mechanism allows the model to jointly capture and learn relevant features from the patches, enhancing its ability to identify important patterns and relationships across the input data. In each iteration, 16 patches from one WSI are randomly selected as inputs, tokenized with positional embeddings, merged along the channel dimension, and fed into the transformer encoder. The output pathological features have the same size of 1024*1.

For radiological features extraction, the raw CT data were reconstructed with consistent parameters,

maintaining an in-plane resolution of 1.0 mm and a slice thickness of 1.0 mm. To ensure comprehensive malignancy coverage, the cancer masks are automatically dilated by three voxels as part of the region delineation process. A final Volume of Interest (VOI) with fixed dimensions of 128 × 128 × 128 was then cropped for each CT image. To standardize data representation, the intensity values of all extracted patches were normalized to a range of zero to one based on the lung window. These pre-processed CT images are input into the 3D Transformer network module, where a multi-head self-attention mechanism extracted the corresponding radiological features, matching the size of the pathological features.

Regarding clinical features, all clinical information was numerically encoded and processed through a fully connected layer to generate feature embeddings comparable to those from pathological and radiological modalities. These embeddings were used to construct a clinical-guided attention mask—a computational mechanism that assigns adaptive weights to specific regions or features within the input data. The weights prioritize interactions between modalities that are clinically meaningful, ensuring cross-modality associations are grounded in biological or clinical relevance rather than being purely data-driven. By focusing the model's attention on clinically plausible and biologically interpretable interactions (e.g., linking histologic patterns to molecular biomarkers with known prognostic roles), the mask enhances the discovery of latent relationships that might remain hidden in unsupervised approaches.

Model development

In this study, the dataset was randomly divided into five subsets with a 4:1 ratio, and a five-fold cross-validation method was used to evaluate the patients in the dataset (supplementary Fig. S2). We have developed a model called DLRPC which integrates information from paired CT, pathological images and clinical data. The overall architecture of the model is shown in Fig. 3. Following the feature extraction process, both CT and pathological features of the same size are obtained. These features are then concatenated, and the clinical feature map is reshaped to align with the concatenated CT and pathological features. A dot-product operation is performed to compute a score matrix based on the clinical feature and the extracted from CT and histological features. This score matrix is normalized using the Softmax function to generate an attention matrix. The attention matrix is

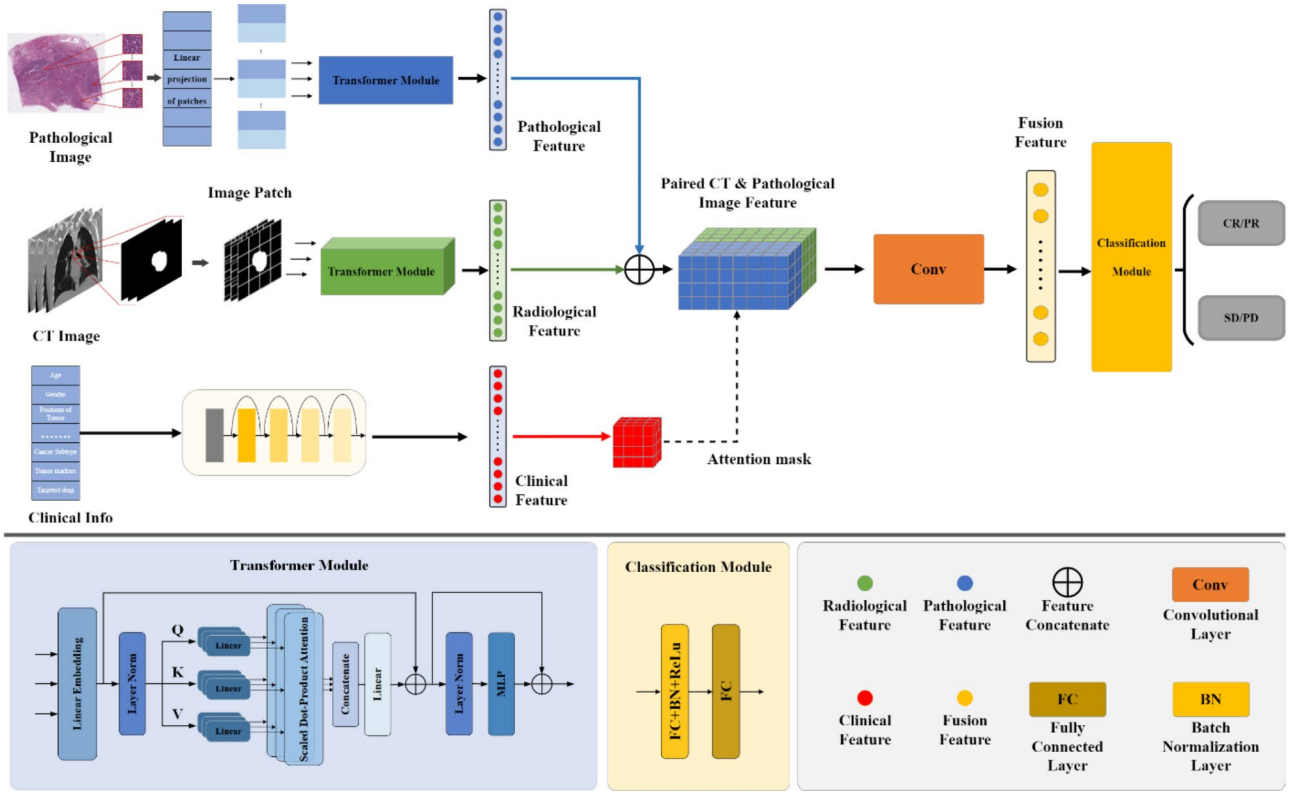

Fig. 3 Workflow Diagram of the DLRPC Model. The preprocessed pathological images, CT images, and clinical information are fed into feature extractors based on 2D ViT, 3D ViT, and ResNet models, respectively. A clinical-based attention mask is generated and applied to fuse the radiological and pathological features, and then used for the final prediction. It is worth noting that although the figure depicts a 3D architecture, the pathological feature extractor is implemented using a 2D ViT architecture

then applied to the feature maps from the CT and histological images, resulting in a fused feature representation, which is subsequently passed through additional convolutional layers. The fused feature is then input into a final convolutional layer, followed by the classification module consisting of a batch normalization layer, a ReLU activation function and two fully connected layers. To mitigate potential biases, we employed class-weighted focal loss as the loss function, which adaptively reduces the influence of well-classified majority samples while amplifying the contribution of minority class samples based on inverse class frequency.

For comparison, we have proposed four models, in which three models are based on single modality and one model is based on multi-modal data. The single modality models include the deep learning CT model, the deep learning pathology model and the clinical model. The deep learning CT model and the deep learning pathology model are constructed using the same feature extraction strategy as the DLRPC model. The clinical model is developed using a Support Vector Machine (SVM) [31] to build a C-Support Vector Classification (C-SVC) model with a radial basis function (RBF) kernel. To optimize the hyperparameters γ and C, we performed a grid search optimization, which determined $\gamma = 0.01$ and $C = 10$ as the optimal values for generating the final results. Additionally, for comparison, the Dempster Shafer Theory was applied to fuse the corresponding modalities in the decision level. The decision level fusion model features were extracted from each source data and decision results were made using local classifiers, which are two fully connected layers and an SVM model for CT images, pathological images and clinical information, respectively.

The proposed DLPRC model, along with the other deep learning baseline methods, was implemented in PyTorch using the CUDA 10.2 toolkit and CUDA Deep Neural Network (cuDNN) 8.0.2 on the Ubuntu 18.04 operating system. The SVM model was implemented using the scikit-learn 1.6.1 toolkit. All training, validation, and testing experiments were conducted on a Tesla V100-SXM2 GPU. During training, the batch size was set to 4, and the initial learning rate was set to 0.0001. Each approach was trained for 400 epochs using the Adam optimizer to update model parameters.

Statistical analysis

Continuous variables between the targeted therapy sensitive group and the primary resistance group were compared using an independent samples t-test or the Mann-Whitney U test, while categorical variables were compared using the chi-square test or Fisher's exact test. To assess the performance of the models in predicting sensitivity to targeted therapy, we quantified the results using several metrics: area under the curve (AUC), accuracy (ACC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The DeLong test was used to compare the differences in AUCs across the models. A *P*-value of less than 0.05 was considered statistically significant. Additionally, Cohen's d was used to measure the effect sizes between the models. All analyses in this study were performed on python 3.8 platform.

Results

Clinical characteristics

A total of 214 patients with EGFR-mutant lung adenocarcinoma from Hospital A and B received first-generation targeted therapies, in which 155 patients were sensitive to TKIs and 59 patients were primary resistant to TKIs. Approximately 28% of patients with EGFR-mutant lung adenocarcinoma did not respond objectively to TKIs, which aligns with findings in a previous study [32]. Detailed clinical information of the patients is shown in Table 2. Among all patients, we found statistical differences between the TKIs sensitive and TKIs resistant groups between maximum diametera ($p = 0.026$), lobulation ($p = 0.005$), and bronchial cut-off sign ($p = 0.034$).

Ablation analysis: Multi-Modal inputs and fusion strategies

The DLRPC model showed good predictive performance in predicting targeted therapy sensitivity in EGFR mutant lung adenocarcinoma patients (AUC 0.8424). The model had good accuracy (0.7980), sensitivity (0.8119) and specificity (0.7644), and the PPV of the DLRPC model was significantly higher (0.9214), while the NPV was lower at 0.6130.

Compared to all single-modality models and the decision level fusion model, the DLRPC model exhibits better predictive performance in predicting the sensitivity to targeted therapy in EGFR-mutant lung adenocarcinoma patients (Table 3). The ROC curves for all models are shown in Fig. 4. Compared with the DLRPC model, the best single-modality model was the clinical model (AUC 0.8052), followed by the pathology model (AUC = 0.7753), with the CT model exhibiting the lowest performance (AUC 0.6894). The decision level fusion model had an AUC of 0.8132, which was higher than any single-modality model but still lower than the DLRPC model. Using the DeLong test to analyze the ROC curves, the DLRPC model significantly improved the AUC for predicting sensitivity to targeted therapy in EGFR -mutant lung adenocarcinoma patients compared to all three single-modality models and the decision level fusion model (all $p < 0.05$, Table 4). Additionally, we quantified the effect size of architectural innovations using Cohen's d, with the ablation models as the reference and our DLRPC model as the target. All negative Cohen's d values (Table 4)

Table 2 Clinical information characteristics of all patients

Features		TKIs ensitive(155)	TKIs resistant (59)	*P*-value
Age (x ± s)		58.7 ± 9.8	60.0 ± 10.4	0.535
Sex	Female	94 (60.6%)	41 (69.5%)	0.231
	Male	61 (39.4%)	18 (30.5%)	
Smoking	No	117 (75.5%	44 (74.6%)	0.891
	Yes	38 (24.5%)	15 (25.4%)	
CYFRA21-1 (ug/L)	≤ 5	109 (70.3%)	38 (64.4%)	0.404
	>5	46 (29.7%)	21 (35.6%)	
CEA (ug/L)	≤ 5	50 (32.3%)	2 (37.3%)	0.486
	>5	105 (67.7%)	37 (62.7%)	
NSE (ug/L)	≤ 25	140 (90.3%)	50 (84.7%)	0.248
	>25	15 (9.7%)	9 (15.3%)	
T stage	T1	24 (15.5%)	10 (16.9%)	0.561
	T2	41 (26.5%)	14 (23.7%)	
	T3	18 (11.6%)	11 (18.6%)	
	T4	72 (46.5%)	24 (40.7%)	
N stage	N0	42 (27.1%)	15 (25.4%)	0.591
	N1	6 (3.9%)	5 (8.5%)	
	N2	39 (25.2%)	15 (25.4%)	
	N3	68 (43.9%)	24 (40.7%)	
M stage	0	22 (14.2%)	9 (15.3%)	0.844
	1	133 (85.8%)	50 (84.7%)	
Stage	I	2 (1.3%)	2 (3.4%)	0.278
	II	1 (0.6%)	2 (3.4%)	
	III	19 (12.3%)	5 (8.5%)	
	IV	133 (85.8%)	50 (84.7%)	
Maximum diametera (cm)		4.0 ± 1.7	3.4 ± 1.4	0.026
Location	Right upper lobe	60 (38.7%)	19 (32.2%)	0.792
	Right middle lobe	9 (5.8%)	5 (8.5%)	
	Right lower lobe	25 (16.1%)	8 (13.6%)	
	Left upper lobe	36 (23.2%)	17 (28.8%)	
	Left lower lobe	25 (16.1%)	10 (16.9%)	
Shape	Regular	74 (47.7%)	30 (50.8%)	0.685
	Irregula	81 (52.3%)	29 (49.2%)	
Interface	Smooth	80 (51.6%)	23 (39.0%)	0.098
	Rough	75 (48.4%)	36 (61.0%)	
Lobulation	No	40 (25.8%)	27 (45.8%)	0.005
	Yes	115 (74.2%)	32 (54.2%)	
Spiculation	No	53 (34.2%)	24 (40.7%)	0.377
	Yes	102 (65.8%)	35 (59.3%)	
Pleural retraction	No	13 (8.4%)	7 (11.9%)	0.435
	Yes	142 (91.6%)	52 (88.1%)	
Bronchial cut-off sign	No	40 (25.8%)	24 (40.7%)	0.034
	Yes	115 (74.2%)	35 (59.3%)	
Pneumonia	No	103 (66.5%)	35 (59.3%)	0.330
	Yes	52 (33.5%)	24 (40.7%)	
Density	No	53 (34.2%)	17 (28.8%)	0.453
	Yes	102 (65.8%)	42 (71.2%)	
Necrosis	No	118 (76.1%)	44 (74.6%)	0.813
	Yes	37 (23.9%)	15 (25.4%)	
Air-containing cavities	No	136 (87.7%)	52 (88.1%)	0.937
	Yes	19 (12.3%)	7 (11.9%)	
Lymphadenopathy	No	96 (61.9%)	37 (62.7%)	0.917

Table 2 (continued)

Features		TKIs ensitive(155)	TKIs resistant (59)	*P*-value
	Yes	59 (38.1%)	22 (37.3%)	
Pleural effusion	No	103 (66.5%)	41 (69.5%)	0.672
	Yes	52 (33.5%)	18 (30.5%)	

Table 3 Comparison of predictive performance between the DLRPC model, single-modality models, and the decision level fusion model

Model	AUC	ACC	Sensitivity	Specificity	PPV	NPV
clinical	0.8052	0.7620	0.7728	0.7325	0.9114	0.5754
CT	0.6894	0.6919	0.7060	0.6560	0.8713	0.4676
pathology	0.7753	0.7389	0.7503	0.7054	0.9029	0.5338
Decision level fusion	0.8132	0.7750	0.7835	0.7509	0.9156	0.5802
DLRPC	0.8424	0.7980	0.8119	0.7644	0.9214	0.6130

indicate that ablation models exhibited substantially lower performance metrics compared to the DLRPC model. Per Cohen's conventions ($|d| > 0.8$ = large effect), these results confirm the material impact of our design choices, with absolute effect sizes exceeding clinical relevance thresholds in computational pathology.

Ablation analysis: model architecture components and clinical prior integration

In our model design, we developed a multi-modality deep learning model to predict sensitivity to targeted therapy in EGFR-mutant lung adenocarcinoma patients. To evaluate the impact of the clinical feature-based attention mask and the contributions of individual benchmark models, we conducted ablation tests.

Our proposed DLRPC model employs a ViT architecture for imaging, utilizes ResNet-50 for the clinical data processing, and incoporates a clinical feature-based attention mask to enhance feature extraction. As illustrated in Table 5, the ViT with ResNet-50 architecture outperforms other SOTA modules. Clinical features also significantly contribute to the final prediction, and the attention mask mechanism effectively extracts multi-modality features, thereby enhancing overall model performance. In addition, we applied the analysis of variance model to select 10 significant measurements from the original clinical data. The feature selection slightly improved the performance of the single-modality clinical SVM model, increasing AUC from 0.8052 to 0.8107. However, the performance of the DLPRC model and the decision level fusion model remained stable, with their AUC values changing from 0.8132 to 0.8120 and from 0.8424 to 0.8389, respectively.

Discussion

In this study, we developed a DLRPC model that integrates deep learning features from CT images, pathological images, and clinical characteristics to predict the sensitivity of EGFR mutated lung adenocarcinoma patients to targeted therapy. The DLRPC model outperforms single-modality models and traditional decision level fusion models, underscoring the potential of integrating multi-modal data for predicting treatment response in cancer patients. The DLRPC model may serve as a valuable tool to aid clinicians in identifying patients who are more likely to benefit from EGFR-TKIs therapy. This information can help guide treatment decisions and potentially avoid unnecessary exposure to ineffective therapy for patients predicted to exhibit primary resistance.

The core contribution of our model lies in its innovative design, particularly in its multimodal fusion algorithm. Unlike traditional models that perform data fusion at the decision level, our DLRPC model integrates data at the feature level by leveraging paired CT images, pathological images, and clinical data. This approach enables a more nuanced combination of information from different modalities, potentially capturing complex interactions and relationships that may be missed when fusing decisions post-feature extraction. Additionally, we introduced a novel attention-based mask generator that utilizes clinical features to create attention masks for pathological and radiological features. This mechanism helps the model focus on classification-relevant features, effectively incorporating clinical priors into the model design, which contributes to its strong performance. The model employs state-of-the-art deep learning architectures, including 2D and 3D ViT networks and ResNet, capable of automatically extracting a vast array of features from images that describe tumor heterogeneity.

Given that traditional clinical and CT morphological characteristics are low-dimensional, there is a critical need to develop methods based on extracting higher-dimensional imaging features for predicting treatment outcomes. Previous studies have explored the use of radiomics features extracted from CT or MRI images to

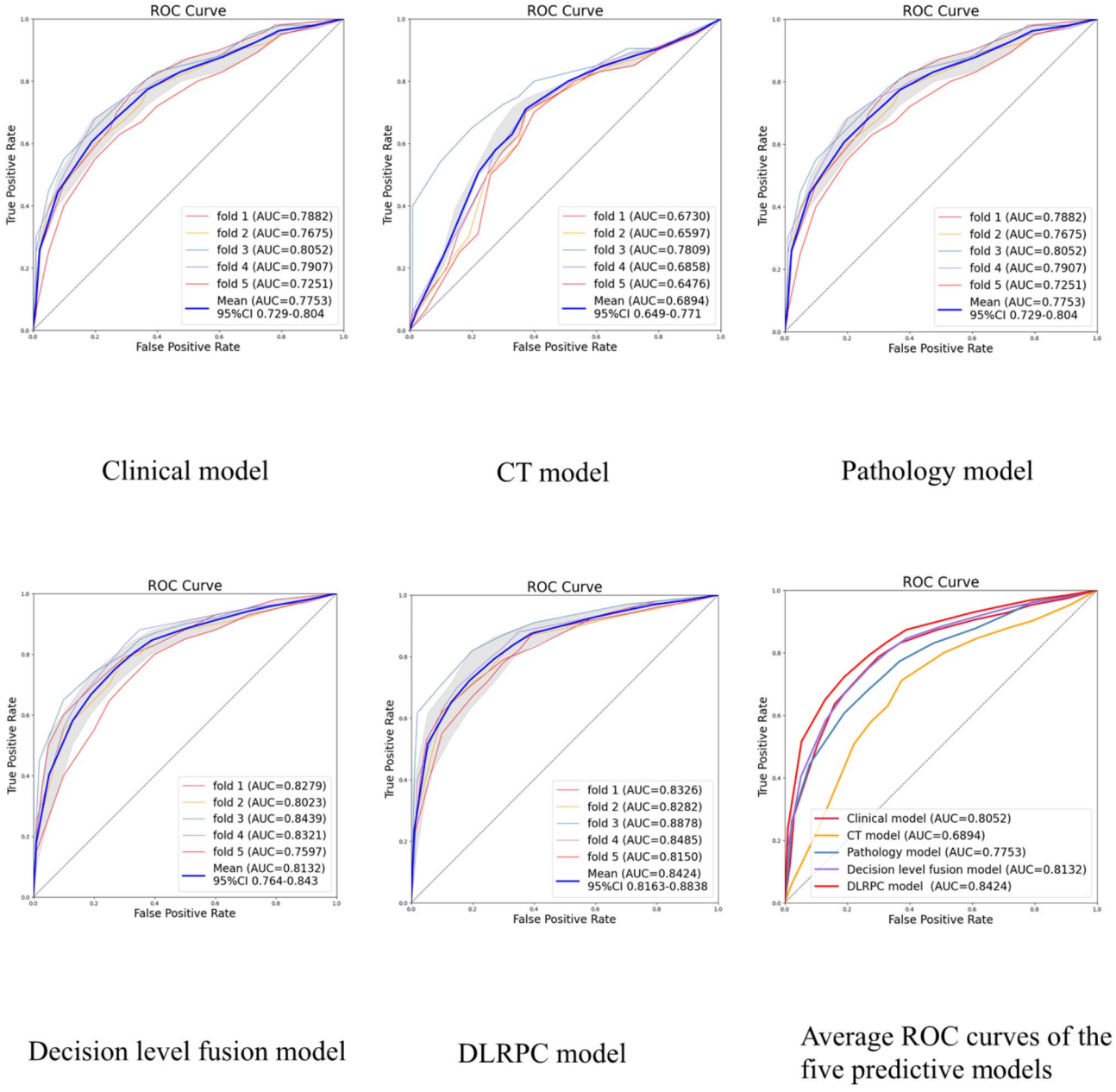

Fig. 4 ROC curves comparing the predictive performance of the DLRPC model with single-modality models and the decision level fusion model

Table 4 Statistical analysis of the DLRPC model compared to single-modality models and the decision level fusion model

Model	*P*-value	Cohen's d
Clinical VS DLRPC	0.0139	-1.08546
CT VS DLRPC	0.0044	-2.82485
Pathology VS DLRPC	0.0079	-2.20206
Decision level fusion VS DLRPC	0.0481	-0.88581

P-values are calculated using the DeLong test, and effect sizes are measured using Cohen's d

predict TKIs therapy response in lung cancer patients [33, 34]. However, conventional radiomics analysis requires precise, manual delineation of lesions, which is a time-consuming and labor-intensive process. In contrast, our deep learning model requires only a rough region of interest covering the tumor area, eliminating the need for detailed tumor boundary delineation and thereby saving time and effort. Although the predictive efficacy of this

Table 5 The ablation test about various model architectures and the presence of the clinical feature-based attention mask

Ablation conditions		AUC	ACC
Image model architecture	ResNet-50	0.8302	0.7874
	DenseNet-121	0.8317	0.7885
	Vision Mamba	0.8386	0.7944
Clinical model architecture	ResNet-18	0.8401	0.7969
	Fully connected layer	0.8359	0.7922
Attention mask	No clinical features	0.8091	0.7689
	No attention mask	0.8222	0.7810
DLRPC		0.8424	0.7980

model is not very high the CT images in this study were obtained using different equipment in different hospitals, which proves that the model has strong adaptability. In this study, the predictive performance of a singular deep learning feature based on CT images proved insufficient, indicating the necessity to integrate CT deep learning features with additional characteristics to achieve better predictive efficacy.

In contrast to CT images, which provide spatial macrostructural information of tumors, histopathological images mainly reflect microscopic morphological features [35]. Computerised image processing techniques can extract massive amounts of information from digital pathology slides that elude human detection, capturing molecular characteristics and prognostic indicators of tumors [36]. Numerous studies [37, 38] have shown that deep learning algorithms can extract features from H&E-stained histopathology slides to predict EGFR mutation status, achieving AUCs between 0.754 and 0.82. Yu et al. [39] used a machine learning approach to predict the prognosis of patients with lung adenocarcinomas and squamous cell carcinomas by extracting objective features from H&E-stained histopathology WSIs. Consistent with these findings, our model achieved an AUC of 0.7753 in predicting the sensitivity to targeted therapy in EGFR-mutant lung adenocarcinoma patients based on H&E-stained histopathological WSIs. Although the pathology deep learning model performed slightly lower than the clinical model (AUC of 0.8052), likely because puncture biopsies capture only a portion of the tumor tissue rather than the entire tumor, it nonetheless provides crucial additional insights into the tumor microenvironment and prognosis for targeted therapy in lung cancer.

The enhanced performance of the DLRPC model stems not only incorporating the clinical features but also from integrating imaging features that capture tumor macrostructures with pathological features that reflect microstructures. This finding is consistent with previous study integrating MRI images with pathology images to predict complete pathological remission [40]. Unlike traditional decision level fusion models, our DLRPC model achieves better clinical prediction performance by fusing different modalities at feature-level with an attention mechanism. In clinical practice, the DLRPC model can quickly predict whether EGFR-mutant lung adenocarcinoma patients are likely to be sensitive to TKIs treatment or primarily resistant. Identifying patients who are sensitive to TKIs supports targeted therapy administration, which can yield real clinical benefits. Conversely, patients predicted to be primarily resistant may be better served with more cost-effective chemotherapeutic agents, potentially reducing unnecessary expenses. Although the DLRPC model in resource-limited environments may incur an initial cost, the long-term improvements in patient outcomes justify this investment. The model's high PPV indicates that it reliably identifies patients who are sensitive to TKIs. In contrast, the relatively lower NPV may be attributed to the imbalance in our study cohort, where 72.43% of patients were sensitive to targeted therapy compared to 27.57% who were resistant, which is consistent with the incidence of sensitivity to targeted therapy in the real world. In this study, we employed a five-fold cross-validation method to optimize data utilization and ensure a robust evaluation of model performance. This method partitions the dataset into five subsets, sequentially using four for training and one for validation, thereby giving each data point the opportunity to contribute to the validation process. Additionally, we balanced the class distribution in each fold to mitigate potential bias and improve validation accuracy.

This study has several limitations. First, its retrospective design introduces selection bias, and prospective studies are necessary to validate the diagnostic efficacy of the model. Then, the sample size is relatively small, only 214 patients from the Chinese population meeting the inclusion criteria, so larger sample sizes and multi-centre studies are warranted. Additionally, the DLRPC model employs a relatively complex design that integrates data from three modalities. In the future, we plan to explore mechanisms for managing missing modalities and implement model compression techniques to improve computational efficiency and enhance model predictive performance. Finally, variations in efficacy among the three first-generation EGFR TKIs and differences in the proportion of patients receiving each drug treatment may affect the study outcomes. A more comprehensive

dataset is needed to better predict sensitivity following treatment with second- and third-generation TKIs.

Conclusion

In this study, we developed a DLRPC model that integrates pre-treatment CT scans, H&E-stained WSIs and patient clinical information to predict TKIs therapy sensitivity in EGFR-mutant lung adenocarcinoma patients. The DLRPC model demonstrated more effective predictive performance compared to single-modality models and decision level fusion models, highlighting the potential of multi-modal data fusion for personalized treatment planning. The DLRPC model may help clinicians identify patients who are more likely to benefit from TKIs therapy, facilitating personalized treatment strategies and potentially reducing unnecessary exposure to ineffective therapies.

Abbreviations

ACC	Accuracy
AUC	Area under the curve
CEA	Carcinoembryonic antigen
CR	Complete remission
CT	Computed tomography
CYFRA21-1	Cytokeratin 19 fragment
EGFR	Epidermal growth factor receptor
H&E	Hematoxylin and eosin
NCCN	National comprehensive cancer network
NPV	Negative predictive value
NSCLC	Non-small cell lung cancer
NSE	Neuron-specific enolase
PD	Progressive disease
PPV	Positive predictive value
PR	Partial remission
ROI	Region of interest
SD	Stable disease
SVM	Support vector machine
TKIs	Ttyrosine kinase inhibitors
ViT	Vision transformer
VOI	Volume of interest
WSIs	Whole slide images

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-06480-9.

Supplementary Material 1

Author contributions

TTY, XQW, WTZ, and WC contributed to the conception, design, and statistical analysis plan of the study. YJ, XHY, and TTY contributed to the methodology, software. ZYS, PZC, SYZ, XQW and XHY participated in data collection, tumor segmentation. TTY, XQW, and YJ analyzed and interpreted the data, TTY, XQW, WTZ and WC prepared the first draft of the manuscript. All authors reviewed the manuscript.

Funding

This study was supported by the 2023 Chongqing Science and Health Joint Key Project (Project No. 2023ZDXM008), Yu-Yue Pathology Science Research Centre Autonomous Project (Project No. YYKYXM202303B05).

Data availability

The data generated in this study are available upon reasonable request from the corresponding author.

Declarations

Ethics approval and consent to participate

This retrospective study was approved by the Ethics Committee of the First Affiliated Hospital of Army Military Medical University (Approval Number: KY2023143).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 14 September 2024 / Accepted: 11 April 2025
Published online: 29 April 2025

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71:209–49.
2. Uras IZ, Moll HP, Casanova E. Targeting KRAS mutant Non-Small-Cell lung cancer: past, present and future. Int J Mol Sci. 2020;21:4325.
3. Lahiri A, Maji A, Potdar PD, Singh N, Parikh P, Bisht B, et al. Lung cancer immunotherapy: progress, pitfalls, and promises. Mol Cancer. 2023;22:40.
4. Apicella M, Giannoni E, Fiore S, Ferrari KJ, Fernández-Pérez D, Isella C, et al. Increased lactate secretion by Cancer cells sustains Non-cell-autonomous adaptive resistance to MET and EGFR targeted therapies. Cell Metab. 2018;28:848–e8656.
5. Maemondo M, Inoue A, Kobayashi K, Sugawara S, Oizumi S, Isobe H, et al. Gefitinib or chemotherapy for Non–Small-Cell lung Cancer with mutated EGFR. N Engl J Med. 2010;362:2380–8.
6. Qin X, Gu X, Lu Y, Zhou W. EGFR-TKI-sensitive mutations in lung carcinomas: are they related to clinical features and CT findings? Cancer Manag Res. 2018;10:4019–27.
7. Jiang Y, Yang M, Wang S, Li X, Sun Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. Cancer Commun. 2020;40:154–66.
8. van der Velden BHM, Kuijf HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Med Image Anal. 2022;79:102470.
9. Kha Q-H, Tran T-O, Nguyen T-T-D, Nguyen V-N, Than K, Le NQK. An interpretable deep learning model for classifying adaptor protein complexes from sequence information. Methods. 2022;207:90–6.
10. Le NQK. Hematoma expansion prediction: still navigating the intersection of deep learning and radiomics. Eur Radiol. 2024;34:2905–7.
11. Lami K, Ota N, Yamaoka S, Bychkov A, Matsumoto K, Uegami W, et al. Standardized classification of lung adenocarcinoma subtypes and improvement of grading assessment through deep learning. Am J Pathol. 2023;193:2066–79.
12. Saad MB, Hong L, Aminu M, Vokes NI, Chen P, Salehjahromi M, et al. Predicting benefit from immune checkpoint inhibitors in patients with non-small-cell lung cancer by CT-based ensemble deep learning: a retrospective study. Lancet Digit Health. 2023;5:e404–20.
13. Shetty MV, Tunga DJ. Optimized deformable Model-based segmentation and deep learning for lung Cancer classification. J Med Invest. 2022;69:244–55.
14. Wang S, Yu H, Gan Y, Wu Z, Li E, Li X, et al. Mining whole-lung information by artificial intelligence for predicting EGFR genotype and targeted therapy response in lung cancer: a multicohort study. Lancet Digit Health. 2022;4:e309–19.
15. Hou R, Li X, Xiong J, Shen T, Yu W, Schwartz LH, et al. Predicting tyrosine kinase inhibitor treatment response in stage IV lung adenocarcinoma patients with EGFR mutation using Model-Based deep transfer learning. Front Oncol. 2021;11:679764.
16. Cheng N, Ren Y, Zhou J, Zhang Y, Wang D, Zhang X, et al. Deep Learning-Based classification of hepatocellular nodular lesions on Whole-Slide histopathologic images. Gastroenterology. 2022;162:1948–61..e7.
17. Wang Y, Acs B, Robertson S, Liu B, Solorzano L, Wählby C, et al. Improved breast cancer histological grading using deep learning. Ann Oncol. 2022;33:89–98.

18. Jiang X, Hoffmeister M, Brenner H, Muti HS, Yuan T, Foersch S, et al. End-to-end prognostication in colorectal cancer by deep learning: a retrospective, multicentre study. Lancet Digit Health. 2024;6:e33–43.
19. Liu Y, Chen W, Ruan R, Zhang Z, Wang Z, Guan T, et al. Deep learning based digital pathology for predicting treatment response to first-line PD-1 Blockade in advanced gastric cancer. J Transl Med. 2024;22:438.
20. Chen Y, Gao R, Jing D, Shi L, Kuang F, Jing R. Classification and prediction of chemoradiotherapy response and survival from esophageal carcinoma histopathology images. Spectrochim Acta Part A Mol Biomol Spectrosc. 2024;312:124030.
21. Park S, Ock C-Y, Kim H, Pereira S, Park S, Ma M, et al. Artificial Intelligence–Powered Spatial analysis of Tumor-Infiltrating lymphocytes as complementary biomarker for immune checkpoint Inhibition in Non–Small-Cell lung Cancer. JCO. 2022;40:1916–28.
22. Fu Y, Karanian M, Perret R, Camara A, Le Loarer F, Jean-Denis M, et al. Deep learning predicts patients outcome and mutations from digitized histology slides in Gastrointestinal stromal tumor. Npj Precis Onc. 2023;7:71.
23. Zhang W, Wang W, Xu Y, Wu K, Shi J, Li M, et al. Prediction of epidermal growth factor receptor mutation subtypes in non–small cell lung cancer from hematoxylin and eosin–stained slides using deep learning. Lab Invest. 2024;104:102094.
24. Wang R, Dai W, Gong J, Huang M, Hu T, Li H, et al. Development of a novel combined nomogram model integrating deep learning-pathomics, radiomics and immunoscore to predict postoperative outcome of colorectal cancer lung metastasis patients. J Hematol Oncol. 2022;15:11.
25. Boehm KM, Aherne EA, Ellenson L, Nikolovski I, Alghamdi M, Vázquez-García I, et al. Multimodal data integration using machine learning improves risk stratification of high-grade serous ovarian cancer. Nat Cancer. 2022;3:723–33.
26. Song H, Xiao X, Han X, Sun Y, Zheng G, Miao Q, et al. Development and interpretation of a multimodal predictive model for prognosis of Gastrointestinal stromal tumor. Npj Precis Onc. 2024;8:157.
27. Ma M, Gu W, Liang Y, Han X, Zhang M, Xu M, et al. A novel model for predicting postoperative liver metastasis in R0 resected pancreatic neuroendocrine tumors: integrating computational pathology and deep learning-radiomics. J Transl Med. 2024;22:768.
28. Huang Y, Yao Z, Li L, Mao R, Huang W, Hu Z, et al. Deep learning radiopathomics based on preoperative US images and biopsy whole slide images can distinguish between luminal and non-luminal tumors in early-stage breast cancers. eBioMedicine. 2023;94:104706.
29. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45:228–47.
30. Wang Y, Lin H, Yao N, Chen X, Qiu B, Cui Y, et al. Computerized tertiary lymphoid structures density on H&E-images is a prognostic biomarker in resectable lung adenocarcinoma. iScience. 2023;26:107635.
31. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20:273–97.
32. Mok TS, Wu Y-L, Thongprasert S, Yang C-H, Chu D-T, Saijo N, et al. Gefitinib or Carboplatin–Paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009;361:947–57.
33. Yang C, Chen W, Gong G, Li Z, Qiu Q, Yin Y. Application of CT radiomics features to predict the EGFR mutation status and therapeutic sensitivity to TKIs of advanced lung adenocarcinoma. Transl Cancer Res. 2020;9:6683–90.
34. Qi H, Hou Y, Zheng Z, Zheng M, Qiao Q, Wang Z, et al. Clinical characteristics and MRI based radiomics nomograms can predict iPFS and short-term efficacy of third-generation EGFR-TKI in EGFR-mutated lung adenocarcinoma with brain metastases. BMC Cancer. 2024;24:362.
35. Erak E, Oliveira LD, Mendes AA, Dairo O, Ertunc O, Kulac I, et al. Predicting prostate Cancer molecular subtype with deep learning on histopathologic images. Mod Pathol. 2023;36:100247.
36. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. Nat Rev Clin Oncol. 2019;16:703–15.
37. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nat Med. 2018;24:1559–67.
38. Zhao D, Zhao Y, He S, Liu Z, Li K, Zhang L, et al. High accuracy epidermal growth factor receptor mutation prediction via histopathological deep learning. BMC Pulm Med. 2023;23:244.
39. Yu K-H, Zhang C, Berry GJ, Altman RB, Ré C, Rubin DL, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun. 2016;7:12474.
40. Feng L, Liu Z, Li C, Li Z, Lou X, Shao L, et al. Development and validation of a radiopathomics model to predict pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a multicentre observational study. Lancet Digit Health. 2022;4:e8–17.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.