Fusion of shallow and deep features from ^{18}F-FDG PET/CT for predicting EGFR-sensitizing mutations in non-small cell lung cancer

Xiaohui Yao[1#], Yuan Zhu[1,2#], Zhenxing Huang[2#], Yue Wang[3], Shan Cong[1], Liwen Wan[2], Ruodai Wu[4], Long Chen[3], Zhanli Hu[2]

[1]Qingdao Innovation and Development Center, Harbin Engineering University, Qingdao, China; [2]Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China; [3]Department of PET/CT Center and Department of Thoracic Cancer I, Cancer Center of Yunnan Province, Yunnan Cancer Hospital, The Third Affiliated Hospital of Kunming Medical University, Kunming, China; [4]Department of Radiology, Shenzhen University General Hospital, Shenzhen University Clinical Medical Academy, Shenzhen, China

Contributions: (I) Conception and design: X Yao, Y Zhu, Z Huang, Z Hu; (II) Administrative support: L Chen, R Wu, Z Hu; (III) Provision of study materials or patients: Y Wang, L Chen, Z Hu; (IV) Collection and assembly of data: Y Zhu, Z Huang, S Cong; (V) Data analysis and interpretation: X Yao, Z Huang, L Wan; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Long Chen, MD. Department of PET/CT Center and Department of Thoracic Cancer I, Cancer Center of Yunnan Province, Yunnan Cancer Hospital, The Third Affiliated Hospital of Kunming Medical University, No. 519 Kunzhou Road, Kunming 650118, China. Email: lonechen1983@hotmail.com; Zhanli Hu, PhD. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, No. 1068 Xueyuan Road, Shenzhen 518055, China. Email: zl.hu@siat.ac.cn.

Background: Non-small cell lung cancer (NSCLC) patients with epidermal growth factor receptor-sensitizing (EGFR-sensitizing) mutations exhibit a positive response to tyrosine kinase inhibitors (TKIs). Given the limitations of current clinical predictive methods, it is critical to explore radiomics-based approaches. In this study, we leveraged deep-learning technology with multimodal radiomics data to more accurately predict EGFR-sensitizing mutations.

Methods: A total of 202 patients who underwent both flourine-18 fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) scans and EGFR sequencing prior to treatment were included in this study. Deep and shallow features were extracted by a residual neural network and the Python package PyRadiomics, respectively. We used least absolute shrinkage and selection operator (LASSO) regression to select predictive features and applied a support vector machine (SVM) to classify the EGFR-sensitive patients. Moreover, we compared predictive performance across different deep models and imaging modalities.

Results: In the classification of EGFR-sensitive mutations, the areas under the curve (AUCs) of ResNet-based deep-shallow features and only shallow features from different multidata were as follows: RES_TRAD, PET/CT *vs.* CT-only *vs.* PET-only: 0.94 *vs.* 0.89 *vs.* 0.92; and ONLY_TRAD, PET/CT *vs.* CT-only *vs.* PET-only: 0.68 *vs.* 0.50 *vs.* 0.38. Additionally, the receiver operating characteristic (ROC) curves of the model using both deep and shallow features were significantly different from those of the model built using only shallow features (P<0.05).

Conclusions: Our findings suggest that deep features significantly enhance the detection of EGFR-sensitizing mutations, especially those extracted with ResNet. Moreover, PET/CT images are more effective than CT-only and PET-only images in producing EGFR-sensitizing mutation-related signatures.

Introduction

Lung cancer is the second most common cancer and the leading cause of death in China (1). Non-small cell lung cancer (NSCLC) accounts for 85% of all cases of lung cancer (2,3). In the Asian population, approximately 60% of NSCLC patients exhibit epidermal growth factor receptor (EGFR) mutations, including EGFR-sensitizing and EGFR-resistance mutations (4,5). EGFR-tyrosine kinase inhibitors (EGFR-TKIs) (6) are typically only effective in treating NSCLC patients harboring EGFR-sensitizing mutations. Thus, it is particularly important to predict and identify EGFR-sensitizing mutations accurately and efficiently to enable physicians to make informed clinical decisions.

Traditional detection approaches, such as biopsies and computed tomography (CT) screenings, are not suitable for elderly or physically compromised individuals and have limited specificity (7). However, the integration of artificial intelligence (AI) and radiomics techniques has effectively addressed these limitations, and these novel techniques are now being used to identify patients at risk for various diseases. Recent studies have demonstrated that certain methods have a high level of sensitivity and specificity in detecting and classifying lung tumors and mutations (7-11). For example, Trivizakis *et al.* demonstrated that machine-learning multi-omics analysis methods based on deep features showed superior performance and improved classification metrics compare to traditional detection methods (12). Over the last decade, traditional radiomics techniques using machine-learning methods to extract shallow tumor features have reached a highly mature stage of research (13-15). These shallow features have been applied to characterize tumor heterogeneity (16); however, they are thought to be limited in their ability to fully capture the non-linear information embedded in higher-dimensional tumor images. Recently, with the development of deep-learning algorithms in medicine (17-25), some deep-learning models have been introduced to identify and predict tumors, including those harboring EGFR mutations. Deep-learning models have demonstrated superior performance compared to the traditional radiomics techniques in predicting EGFR mutations; however, it is worth noting that shallow features still provide useful information. To date, few studies have attempted to fuse shallow features with deep features to predict EGFR-sensitizing mutations in NSCLC patients.

Flourine-18 fluorodeoxyglucose (^{18}F-FDG) is a radioactive tracer that plays a pivotal role in positron emission tomography (PET)/CT imaging, offering comprehensive data on metabolism, localization, and pathology (26). Thus, ^{18}F-FDG PET/CT images are widely employed in oncology (15). Despite the widespread use of PET/CT in cancer imaging, existing studies have typically focused on using only CT images for EGFR mutation prediction, and thus do not fully leverage the information available from PET/CT scans (27,28). Recent research has increasingly focused on the use of multimodal cancer images, and success has been achieved in various tasks, such as identifying EGFR mutation subtypes (11), predicting progression-free survival (29), and distinguishing different mutations (10). Few studies have focused on predicting EGFR-sensitizing mutations, which could guide physicians in prescribing appropriate medications. Moreover, only a limited number of studies have investigated the fusion of features from deep-learning networks with shallow radiomics features to predict EGFR-sensitizing mutations.

Given the above observations, this study aimed to leverage deep-learning technology with multimodal imaging data that allows the extraction of both intra- and peri-tumoral radiomics signatures to predict EGFR-sensitizing mutations in NSCLC patients prior to EGFR-TKI treatment. Additionally, multimodal radiomics signatures based on ^{18}F-FDG PET/CT images, especially the fused ResNet-based deep and shallow features, were examined to demonstrate the predictive performance on EGFR-sensitizing mutations. We present this article in accordance with the TRIPOD reporting checklist (available at https://qims. amegroups.com/article/view/10.21037/qims-23-1028/rc).

Table 1 Inclusion and exclusion criteria of the study

Inclusion criteria	Exclusion criteria
No history of other malignancies	Anti-tumor therapy before PET/CT examination
Patients with lung tumors only	Patients with lung and other tumors
Pathological confirmation of NSCLC	pGGN without FDG metabolism

PET/CT, positron emission tomography/computed tomography; NSCLC, non-small cell lung cancer; pGGN, pure ground-glass nodule; FDG, fluorodeoxyglucose.
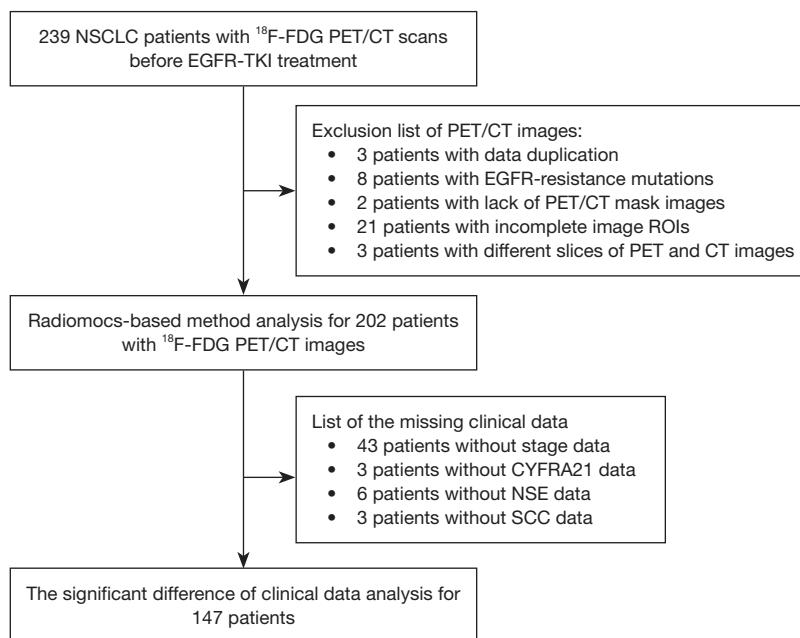

Figure 1 The workflow for NSCLC patients with PET/CT images and clinical data. NSCLC, non-small cell lung cancer; ^{18}F-FDG PET/CT, flourine-18 fluorodeoxyglucose positron emission tomography/computed tomography; EGFR, epidermal growth factor receptor; TKI, tyrosine kinase inhibitor; ROI, region of interest; NSE, neuron-specific enolase; SCC, squamous cell carcinoma.

Methods

Study data

Starting in January 2016, 239 NSCLC patients from the Yunnan Cancer Hospital with thin-slice chest PET/CT images and mask data were included in this study. It should be noted that the Hospital was responsible for applying the inclusion and exclusion criteria for patient selection prior to providing the patient PET/CT images (*Table 1*). Subsequently, we examined the images of these 239 patients (*Figure 1*). Ultimately, the images of 202 patients (117 male and 85 female; median age: 64 years; range, 37–93 years) were eligible for inclusion in this study that used the radiomics-based method to predict EGFR-sensitive mutations. Of these 202 patients, 141 had wild-type EGFR

and 61 had EGFR-sensitizing mutations.

Concurrently, the clinical information from all 239 initial patients was collected by the Hospital. The clinical information included sex, age, weight, stage, carcinoembryonic antigen levels, smoking history, cytokeratin 19 fragment antigen 21 levels, squamous cell carcinoma antigen levels, and neuron-specific enolase levels (*Figure 1*). Due to the difficulty in obtaining clinical data, only 147 of the 202 patients were included in the statistical analysis of the clinical information to explore the potential association between clinical data and EGFR prediction.

This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of Yunnan Cancer Hospital (No. SLKYCS2022217), and written informed

consent was obtained from all the participants or their authorized representatives.

Image acquisition and region of interest (ROI) segmentation

The ^{18}F-FDG PET/CT images were obtained using a SIEMENS Biograph (Munich, Germany), which includes methods for acquiring CT and PET images. The scanning parameters for the chest PET/CT used in our work were as follows: (I) CT: tube voltage: 120 kV, X-ray tube current: 240 mA, exposure time: 500 ms, and slice thickness: 5 mm; (II) PET: acquisition time: 50–60 minutes after the ^{18}F-FDG tracer injection, reconstruction algorithm: Ordered Subset Expectation Maximization (OSEM), and slice thickness: 5 mm.

The ROIs for each patient were independently delineated on their PET and CT images. This procedure was carried out by an experienced from Yunnan Cancer Hospital, and underwent interactive validation by two nuclear medicine physicians with years of expertise in the field. In our work, the size of the CT image density corresponded to the size of the high-metabolism parts of the PET image. First, the nuclear medicine physician delineated the ROIs of the CT images due to the clear anatomical morphology provided therein. Second, using the contours on the CT images as a reference, the corresponding ROIs were established on the PET images, enabling the extraction of the patient's tumor metabolic features. Finally, the ROIs encompassing both the PET and CT images were simultaneously acquired.

Image processing and feature extraction

To normalize the influence of each patient's individual weight and tracer dose, we converted the PET image pixels from gray values to standardized uptake values (SUVs). The CT images were read with mediastinal (level: 40 HU; width: 300 HU) and lung (level: –400 HU; width: 1,500 HU) window settings. We then extracted the lung window from the CT images and finished the normalization processing. The features were extracted after image processing. In this study, traditional radiomics feature extraction was conducted in Python 3.8 with the PyRadiomics software package via machine learning, while deep feature extraction was conducted with the deep neural network ResNet-101 (30). The radiomics workflow is shown in *Figure 2* and includes depictions of the feature extraction, feature fusion, feature selection, and prediction processes. As seen in the first step, the PET/CT images were input

into a machine-learning model (PyRadiomics) and a deep-learning model (ResNet-101). Subsequently, shallow and deep feature matrixes were output, and the features were then integrated in the second step. Finally, the prediction task was performed through the predictor after the feature selection. Next, we introduced each part in detail.

Traditional feature extraction

For the original CT and PET images and those derived based on Laplacian of Gaussian (LoG) and Wavelet transformation (31), the radiomics features were automatically calculated from the tumor ROI using the PyRadiomics package. The following three types of traditional features were included: (I) shape; (II) first order (histogram); and (III) texture [gray-level cooccurrence matrix (GLCM)/gray-level dependence matrix (GLDM)/gray-level run length matrix (GLRLM)/gray-level size zone matrix (GLSZM)/neighbor gray-tone difference matrix (NGTDM)]. The radiomics features were extracted independently from the PET and CT images. During the extraction process, the pixel space resampling value had to be reset for the PET and CT images. Sigma values of 3.0 and 5.0 were selected for the LoG transformation. In total, 1,239 features each were extracted from the PET and CT images. To ensure a fair comparison of the experimental results, both the CT and PET images used the same type of derived images.

Deep feature extraction

Medical images are different from natural images, but due to the sophisticated architecture of convolutional neural networks (CNNs) and abundant training data sets, many studies have confirmed that pretrained CNN models perform well in classification and prediction tasks involving medical images (20,21,32). Thus, the deep features in our study were extracted using a pretrained ResNet-101 (30). Both the CT and PET images had to be interpolated to 49×224×224-pixel input images before training, where 224×224 represents the length and width of each slice of every patient, and 49 represents the number of slices for each patient. It is worth noting that these CT and PET images must completely cover the ROI information. The last layer of the fully connected layer output 1,000 deep features. Finally, a feature matrix (202×1,000) was produced as the output. Importantly, we also used the Visual Geometry Group (VGG)-19 (33) network to extract deep features to compare the predictive abilities of the deep features extracted by the two networks.
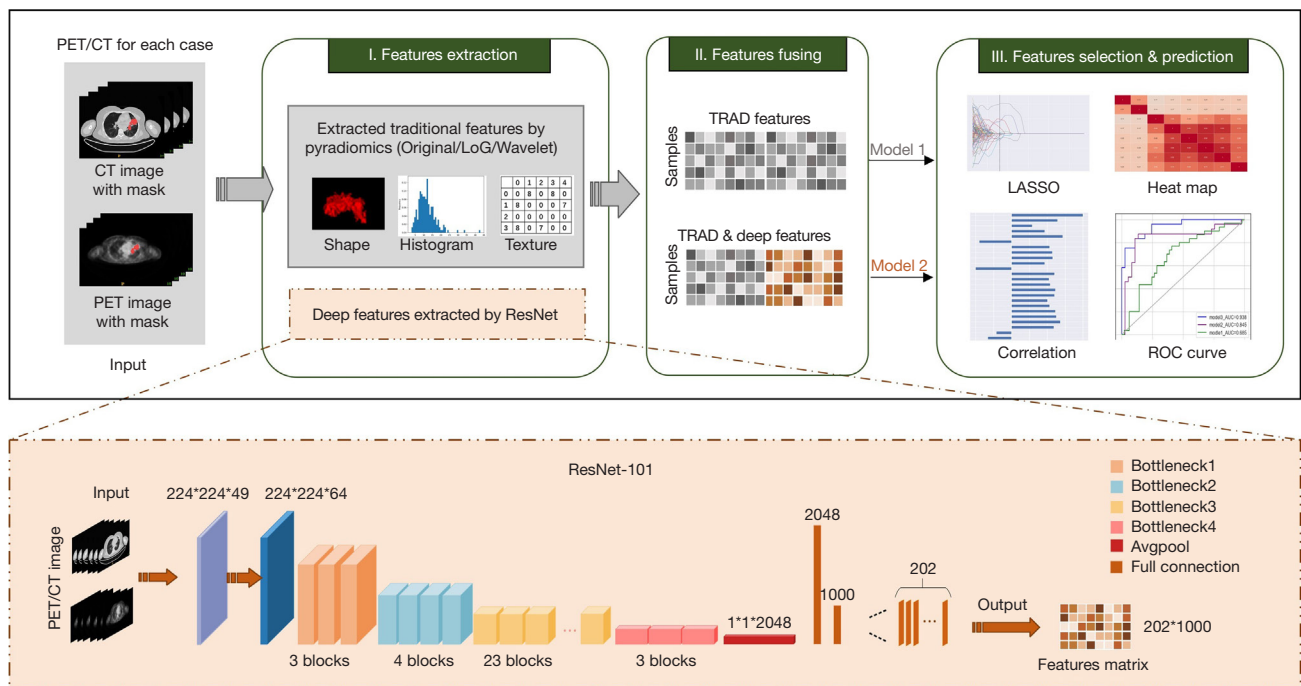
Figure 2 A systematic exposition of the radiomics workflow used in this study for EGFR-sensitizing mutation prediction in NSCLC patients. Model 1, the ONLY_TRAD model uses only traditional features; Model 2, the RES_TRAD model includes all traditional features and deep features extracted with the ResNet-101 network. PET/CT, positron emission tomography/computed tomography; LoG, Laplacian of Gaussian; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; AUC, area under the curve; EGFR, epidermal growth factor receptor; NSCLC, non-small cell lung cancer.

Feature fusion and EGFR-sensitizing mutation-related feature selection

The traditional radiomics features and deep features were concatenated after the feature extraction step. Feature selection was performed in the training cohort to remove redundant features and identify features that had strong potential predictive power. In the feature selection step, we aimed to eliminate redundant features that exhibited high correlations, which might have arisen from the same underlying distribution. We retained only one feature or a subset of representative features. Meanwhile, we aimed to select features with higher correlations to the target variable among the remaining features, as these were presumed to have significant potential for target identification. Thus, two steps were conducted in our work. First, we performed the Mann-Whitney U-test to assess whether the two sample groups were likely to have the same distribution and eliminated features with high confidence, retaining features that had P values lower than the threshold (of 0.05). These remaining features were considered beneficial

for performing the classification and regression tasks. Second, a least absolute shrinkage and selection operator (LASSO) regression algorithm, which is mainly used to solve multicollinearity problems, was built to further refine the remaining features by essentially setting the weights of unimportant features to zero. In our work, we set a large range of alpha values and used regression algorithms to sort alpha and error values, and the best thresholds were selected by cross-validation in the training cohort. The operations were conducted in Python 3.8 with the sklearn package.

Predictive models and statistical analysis

The statistical analysis was performed using Python (version 3.8.0). A P value <0.05 was considered statistically significant. The P values were used to assess the differences in all the variables between the EGFR-sensitizing mutation and wild-type EGFR patients in our work.

Two models were built based on different methods for integrating features. The first model, which included only traditional features, was named ONLY_TRAD. The

second model, which used all the traditional and deep features extracted with the ResNet-101 (30) network, was named RES_TRAD. The deep features extracted by the VGG-19 (33) network were introduced as a control group for comparison with those extracted by ResNet-101 (30). Because some data were linearly inseparable according to the feature analysis conducted earlier, an support vector machine (SVM), a model that has been confirmed to possess excellent classification capabilities (34-37), was introduced into both models, which were then trained with CT-only, PET-only, and PET/CT images as input for the analysis. For both models, the parameters were determined using five-fold cross-validation and grid searching in the training cohort. Classifier performance was evaluated in the independent validation cohort using the area under the curve (AUC) and a decision curve analysis (DCA). A receiver operating characteristic (ROC) curve analysis was used to evaluate the performances of the statistical features and models.

Results

Results for the clinical data

To assess whether factors such as sex and age affected the results, a statistical analysis of these factors was performed. As the results show (*Table 2*), no significant differences were found in these or other features among the groups. Thus, it is unlikely that there was any clinical bias in our radiomics-based analysis of EGFR-sensitive mutations.

Comparison of predictive results between shallow and deep features based on PET/CT

In the training cohort (comprising 140 patients), 56 patients had EGFR-sensitizing mutations (EGRF$^+$) and 84 patients had wild-type EGFR (EGRF$^-$). In the validation cohort (comprising 62 patients), 23 patients had EGFR-sensitizing mutations and 39 patients had wild-type EGFR.

The shallow radiomics features were extracted by the PyRadiomics package. In total, 2,478 features were extracted from the CT and PET images with masks. The Mann-Whitney *U*-test showed that 402 stable features (CT: 323; PET: 79) had latent predictive abilities (*Table 3*). The ONLY_TRAD model was used to predict EGFR-sensitizing mutations (AUC =0.685). Using the different feature selection algorithms and prediction models, after adjusting for the parameters, the AUC was 0.66 (±0.04), and the F1 score was 0.58 (±0.2), which suggests that there was no obvious classification effect.

A total of 3,478 features were extracted with the RES_TRAD model (traditional features: 2,478; deep features: 1,000). After a significance analysis with the Mann-Whitney *U*-test, 934 CT features (traditional: 203; deep: 731) and 829 PET features (traditional: 20; deep: 809) remained. The correlation analysis (threshold >0.2) and the LASSO algorithm indicated that 23 features (CT: 15; PET: 6) were strongly correlated with the EGFR-sensitizing mutations (*Table 3*). From the selected features, it can be seen that the deep features had much stronger potential predictive power than the traditional features regardless of whether they were extracted from the CT or PET images (*Figure 3*). The results showed that the deep-learning algorithms dramatically improved prediction accuracy. Moreover, we conducted the same analysis on the fusion of deep features extracted by the VGG-19 network with traditional features, which yielded 12 features (CT: 7; PET: 5) with strong potential predictive power (*Figure 4*). According to the ROC curve analysis, the RES_TRAD model showed the best performance in distinguishing EGFR-sensitizing mutation and wild-type EGFR patients in the validation cohort based on the PET/CT images (*Figure 5*).

Validation of the predictive performance of the model based on the CT, PET, and PET/CT images

We introduced CT-only and PET-only images into the models and compared the results with those obtained with the PET/CT images. In the validation cohort, the performance of the RES_TRAD model was best for all three types of images, especially PET images (*Figure 5*). The results of the DCA showed that the RES_TRAD model with fused deep-shallow features yielded a higher net benefit compared to other models (*Figure 6*). Further, the results showed that the fusion of shallow and deep features might be sufficient to predict the potential risk of EGFR-sensitizing mutations in NSCLC patients. Additionally, the results based on the PET/CT images were generally superior to those based on the PET-only or CT-only images (*Figures 5,6*). Additionally, we analyzed the P values among the three models and different imaging data methods (*Tables 4*). Both the RES_TRAD and VGG_TRAD models had lower P values when compared to the ONLY_TRAD model (P<0.001). Thus, both the RES_TRAD and VGG_TRAD models possessed greater diagnostic and predictive capability than the ONLY_TRAD model for EGFR-sensitive mutations.

Table 2 Analysis of patients' clinical features

Features	Training cohort (n=103)				Validation cohort (n=44)			
	Total	EGFR⁻	EGFR⁺	P value	Total	EGFR⁻	EGFR⁺	P value
Age (years)	64.8 [40–89]	65 [41–83]	64.5 [40–89]	0.386	61.2 [37–83]	61.6 [37–83]	60.5 [46–81]	0.527
Sex				0.241				0.129
Male	56 (54.4)	34 (60.7)	22 (39.3)		24 (54.5)	16 (66.7)	8 (33.3)	
Female	47 (45.6)	23 (48.9)	24 (51.1)		20 (45.5)	13 (65.0)	7 (35.0)	
Weight (kg)	60.0 [41–90]	60.7 [41–90]	59.1 [42–82]	0.397	61.8 [41–79]	61.7 [41–79]	61.8 [45–76]	0.729
Stage				0.237				0.530
I–II	58 (56.3)	35 (60.3)	23 (39.7)		22 (50.0)	16 (72.7)	6 (27.3)	
III–IV	45 (43.7)	22 (48.9)	23 (51.1)		22 (50.0)	13 (59.1)	9 (40.9)	
CEA (µg/L)				0.238				0.150
≤5	34 (33.0)	23 (67.6)	11 (32.4)		15 (34.1)	10 (66.7)	5 (33.3)	
5–100	52 (50.5)	26 (50.0)	26 (50.0)		22 (50.0)	16 (72.7)	6 (27.3)	
>100	17 (16.5)	8 (47.1)	9 (52.9)		7 (15.9)	3 (42.9)	4 (57.1)	
Smoking				0.839				0.595
Yes	40 (38.8)	23 (57.5)	17 (42.5)		18 (40.9)	12 (66.7)	6 (33.3)	
No	63 (61.2)	34 (54.0)	29 (46.0)		26 (59.1)	17 (65.4)	9 (34.6)	
CYFRA21 (ng/mL)				0.288				0.521
≤3.3	32 (31.1)	15 (46.9)	17 (53.1)		17 (38.6)	10 (58.8)	7 (41.2)	
>3.3	71 (68.9)	42 (59.2)	29 (40.8)		27 (61.4)	19 (0.70)	8 (0.30)	
SCCA (ng/mL)				0.253				0.536
≤1.5	89 (86.4)	47 (52.8)	42 (47.2)		34 (77.3)	22 (64.7)	12 (35.3)	
>1.5	14 (13.6)	10 (71.4)	4 (28.6)		10 (22.7)	7 (70)	3 (30)	
NSE (ng/mL)				0.210				0.177
≤17.0	69 (67.0)	35 (50.7)	34 (49.3)		30 (68.2)	22 (73.3)	8 (26.7)	
>17.0	34 (33.0)	22 (64.7)	12 (35.3)		14 (31.8)	7 (50.0)	7 (50.0)	

Data are presented as median [range] or n (%). The P value represents the univariate association between the EGFR-sensitizing mutations and clinical features. EGFR, epidermal growth factor receptor; CEA, carcinoembryonic antigen; CYFRA21, cytokeratin 19 fragment antigen 21; SCCA, squamous cell carcinoma antigen; NSE, neuron-specific enolase.

Table 3 Distribution of the number of features after the application of the Mann-Whitney *U*-test

Mann-Whitey *U*-test	CT			PET		
	Total	Traditional features (%)	Deep features (%)	Total	Traditional features (%)	Deep features (%)
ONLY_TRAD	402	323	–	79	79	–
RES_TRAD	934	203 (21.7)	731 (78.3)	829	20 (2.4)	809 (97.6)

Data are presented as n (%). ONLY_TRAD denotes the model with only shallow features. RES_TRAD denotes the model with ResNet-deep and shallow features. CT, computed tomography; PET, positron emission tomography.
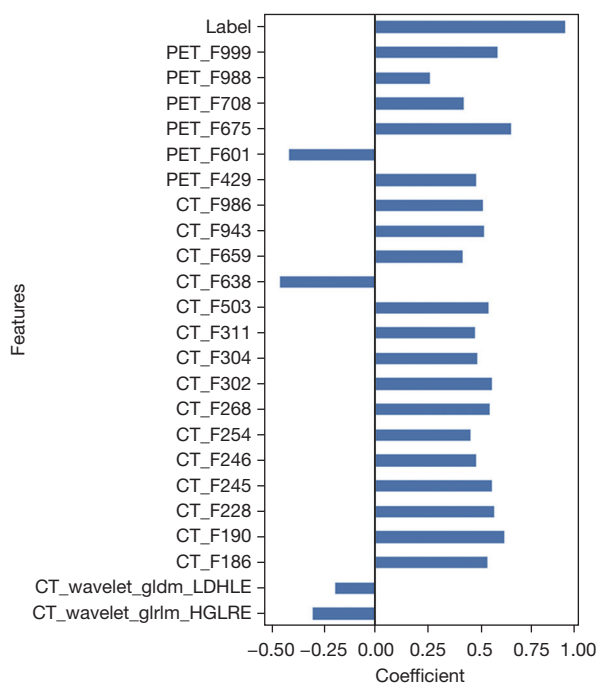
Figure 3 The most predictive features with the corresponding coefficients selected for constructing RES_TRAD model. CT_F denotes CT deep features; PET_F denotes PET deep features. PET, positron emission tomography; CT, computed tomography; LDHLE, large dependence high gray-level emphasis; HGLRE, high gray-level run emphasis.

Discussion

During targeted therapy, EGFR sensitization caused by the T790M mutation occurs in 50% of patients, resulting in drug resistance (38,39). Effectively identifying such sensitive mutations could help doctors make medical decisions. NSCLC patients who are prone to drug resistance will gain little benefits from targeted therapy. The detection of EGFR-sensitizing mutations in NSCLC patients could potentially predict the effects of targeted therapy and provide guidance for clinical treatment. Thus, it is very important to predict EGFR-sensitizing mutations in NSCLC patients before treatment (40). In recent years, research on EGFR genes, EGFR-sensitizing mutations, and targeted therapy for non-small cell cancers has continued unabated (8,41,42). Meanwhile, AI algorithms have been considered to have great potential in medicine (43-51). This study aimed to reveal the relationship between radiomics features and EGFR-sensitizing mutations and to evaluate whether the latter could be predicted in NSCLC patients

before treatment by fusing deep features with traditional radiomics features based on PET/CT images. We also introduced CT-only and PET-only images into the models as control groups.

Before the addition of any deep features, the overall performance of the ONLY_TRAD model was unsatisfactory (*Figure 5*, ONLY_TRAD_AUC =0.685). We speculated that the inclusion of deeper image features extracted by deep-learning algorithms might improve the predictive power of the model. The results of integrating these deep features (*Figure 5*, RES_TRAD_AUC =0.938) showed that they had strong correlations with EGFR-mutation related tumors and potential predictive power in identifying EGFR-sensitizing mutations in a way that was less random than that with the traditional features alone.

Bizzego *et al.* (52) believed that cancer prediction would be more accurate if deep and traditional features were combined, which is supported by our results. We also used the VGG-19 network to extract deep features and fused them with traditional features in our study (to create the VGG_TRAD model). As *Figure 4* shows, the PET deep features extracted by the VGG-19 network for identifying EGFR-sensitizing mutations were less correlated with each other, and the traditional radiomic features extracted from the PET images did not have good predictive power. However, the PET deep features extracted by the ResNet-101 network had strong correlations with the target labels (*Figure 3*). Moreover, as *Figure 5* shows, the features extracted by the RES_TRAD model showed the best performance in terms of the PET image-based ROC curve based on the PET images. These results indicate that the deep features extracted by the residual network had more latent predictive abilities than VGG networks. The reason for this difference between the models might lie in the structure of the corresponding deep algorithms. Compared with the VGG models, the ResNet models introduced a residual block and had a deeper network structure. In addition, the residual network, which was built on a Bayesian network and used polynomial fitting differences, made learning the optimal solution easier, and thus showed great potential predictive power when extracting the deep features.

Among the different imaging methods, CT images convey good structure, shape, size, density, and other information, while PET images, which can better distinguish the lesion from the surrounding normal tissue, always reflect the molecular and metabolic functions of the lesion. Compared to the models with the CT-only and PET-only images, the predictive accuracy of the models with the
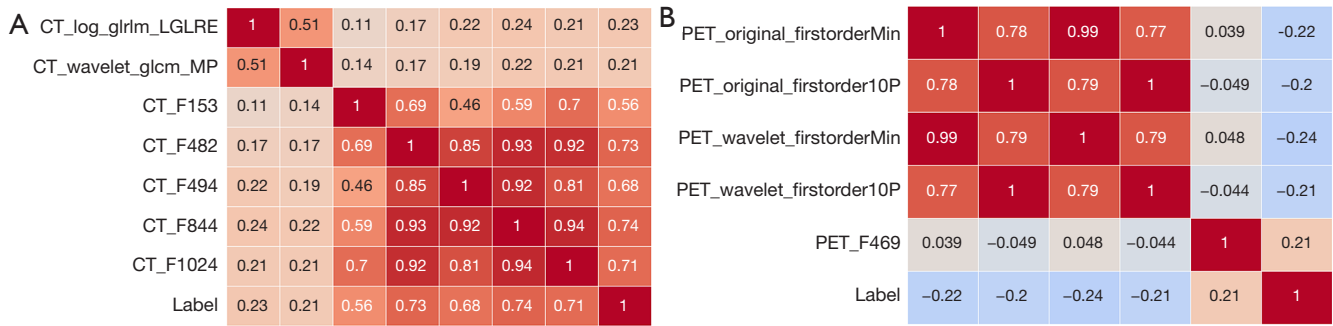
Figure 4 Correlations for features extracted by VGG_TRAD. (A) Heatmap of the correlations between the CT features and labels. CT_F denotes the CT deep features; (B) heatmap of the correlations between the PET features and labels, where PET_F represents the PET deep features. LGLRE, low gray level run emphasis; PET, positron emission tomography; CT, computed tomography; MP, maximum probability; P, percentile.
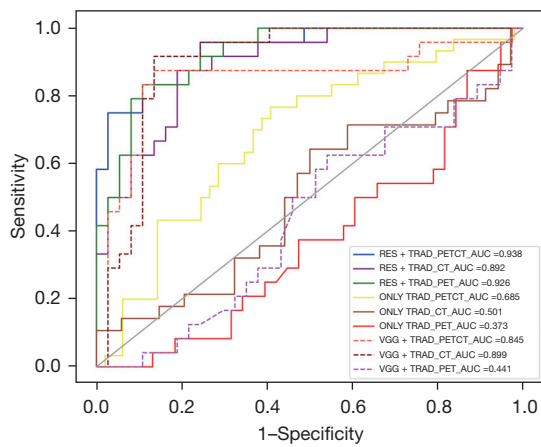

Figure 5 ROC curves and AUC values for evaluating the predictive abilities of the models based on CT-only, PET-only, and PET/CT images. The results of the different multidata in the ONLY_TRAD model are shown as ONLY TRAD_PETCT_ AUC, ONLY TRAD_CT_AUC, and ONLY TRAD_PET_AUC. The results of the different multidata in the RES_TRAD model are shown as RES + TRAD_PETCT_AUC, RES + TRAD _ CT_AUC, and RES + TRAD _PET_AUC. The results of the different multidata in the VGG_TRAD model are shown as VGG + TRAD_PETCT_AUC, VGG + TRAD_CT_AUC, and VGG + TRAD_PET_AUC. ROC, receiver operating characteristic; AUC, area under the curve; PET, positron emission tomography; CT, computed tomography.

PET/CT images was generally higher (*Figures 5,6*). These results showed that the fusion of PET and CT radiomics features simultaneously achieved clarity and had the ability to distinguish the lesion from its surrounding normal tissue.

Despite the valuable results described above, this study still had some limitations. First, while we had complete PET/CT images from 202 patients, we were only able to collect the clinical data of 147 of these patients (*Figure 1*). From a sampling perspective, using clinical data from only 147 patients may appear reasonable. However, there is an incongruity between the number of NSCLC patients included in the clinical information analysis and the number of NSCLC patients included in the PET/CT image analysis. This incongruity may introduce certain unknown sources of bias, the specific effects of which have not been validated or confirmed. Thus, in our future work, we will need to address this issue cautiously to ensure the reliability and credibility of our analysis and conclusions. This also underscores the importance of continuing to collect more clinical data in future work to provide comprehensive support for our research findings. Second, this study did not compare different SUVs, such as SUV_{max}, SUV_{mean}, and SUV_{peak}, from the PET images of the patients. Comparing these values may allow for the extraction of more optimized features. Finally, due to the small sample size of this study, a rigorous external validation design was not implemented. Further research is necessary to address the challenge of predicting disease by using the most appropriate and state-of-art CNN models for medical applications.

Conclusions

In conclusion, fusing deep features with shallow features could improve the ability of clinicians to predict EGFR-sensitizing mutations in NSCLC patients before treatment. Notably, the ResNet-101 network seemed to extract more
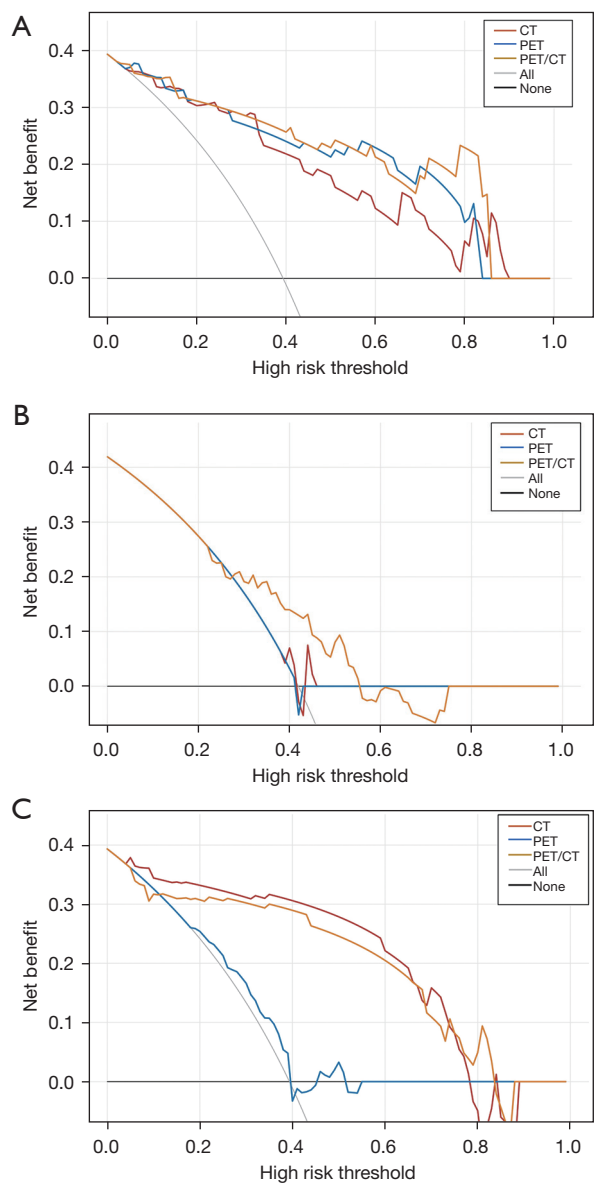
Table 4 P values of the models built with different sets of imaging data

Models	P values
RES_TRAD *vs*. ONLY_TRAD	
PET/CT	<0.001
CT	0.001
PET	0.003
VGG_TRAD *vs*. ONLY_TRAD	
PET/CT	<0.001
CT	<0.001
PET	0.364
RES_TRAD *vs*. VGG_TRAD	
PET/CT	0.116
CT	<0.001
PET	0.907

The P value represents the associations between different models. ONLY_TRAD denotes the model with only shallow features. RES_TRAD denotes the model with ResNet-deep and shallow features. VGG_TRAD denotes the model with VGG-deep and shallow features. CT, computed tomography; PET, positron emission tomography.

Figure 6 The decision curve analysis of three models. (A) The decision curve analysis of the RES_TRAD model; (B) the decision curve analysis of the ONLY_TRAD model; (C) the decision curve analysis of the VGG_TRAD model. RES_TRAD denotes the model with ResNet-deep and shallow features. ONLY_TRAD denotes the model with only shallow features. VGG_TRAD denotes the model with VGGdeep and shallow features. CT, computed tomography; PET, positron emission tomography.

recognizable features that are very likely to have potential predictive abilities than the VGG-19 network due to the better feature extraction capability of the deeper network design of the former. In addition, compared with CT-

only or PET-only images, PET/CT images appear to be a better choice for analyzing EGFR-sensitizing mutations. In the future, the successful integration of different types of features could provide intuitive results to physicians.

5470

Yao et al. Fused shallow-deep features for predicting EGFR mutations

Central Universities (No. 3072022TS2614), and the Key Laboratory for Magnetic Resonance and Multimodality Imaging of Guangdong Province (No. 2023B1212060052).

Footnote

Reporting Checklist: The authors have completed the TRIPOD checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-23-1028/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-23-1028/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of Yunnan Cancer Hospital (No. SLKYCS2022217), and written informed consent was obtained from all the participants or their authorized representatives.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Manafi-Farid R, Askari E, Shiri I, Pirich C, Asadi M, Khateri M, Zaidi H, Beheshti M. [18F]FDG-PET/CT Radiomics and Artificial Intelligence in Lung Cancer: Technical Aspects and Potential Clinical Applications. Semin Nucl Med 2022;52:759-80.
2. Alduais Y, Zhang H, Fan F, Chen J, Chen B. Non-small cell lung cancer (NSCLC): A review of risk factors, diagnosis, and treatment. Medicine (Baltimore) 2023;102:e32899.
3. Smith O. The Epidermal Growth Factor Receptor: A Target for the Treatment of Non-Small Cell Lung Cancer: The University of Manchester (United Kingdom); 2020.
4. Low JL, Lim SM, Lee JB, Cho BC, Soo RA. Advances in the management of non-small-cell lung cancer harbouring EGFR exon 20 insertion mutations. Ther Adv Med Oncol 2023;15:17588359221146131.
5. Marin-Acevedo JA, Pellini B, Kimbrough EO, Hicks JK, Chiappori A. Treatment Strategies for Non-Small Cell Lung Cancer with Common EGFR Mutations: A Review of the History of EGFR TKIs Approval and Emerging Data. Cancers (Basel) 2023.
6. Wang S, Rong R, Yang DM, Fujimoto J, Bishop JA, Yan S, Cai L, Behrens C, Berry LD, Wilhelm C, Aisner D, Sholl L, Johnson BE, Kwiatkowski DJ, Wistuba II, Bunn PA Jr, Minna J, Xiao G, Kris MG, Xie Y. Features of tumor-microenvironment images predict targeted therapy survival benefit in patients with EGFR-mutant lung cancer. J Clin Invest 2023;133:e160330.
7. Vicini S, Bortolotto C, Rengo M, Ballerini D, Bellini D, Carbone I, Preda L, Laghi A, Coppola F, Faggioni L. A narrative review on current imaging applications of artificial intelligence and radiomics in oncology: focus on the three most common cancers. Radiol Med 2022;127:819-36.
8. Yin W, Wang W, Zou C, Li M, Chen H, Meng F, Dong G, Wang J, Yu Q, Sun M, Xu L, Lv Y, Wang X, Yin R. Predicting Tumor Mutation Burden and EGFR Mutation Using Clinical and Radiomic Features in Patients with Malignant Pulmonary Nodules. J Pers Med 2022;13:16.
9. Zhu H, Song Y, Huang Z, Zhang L, Chen Y, Tao G, She Y, Sun X, Yu H. Accurate prediction of epidermal growth factor receptor mutation status in early-stage lung adenocarcinoma, using radiomics and clinical features. Asia Pac J Clin Oncol 2022;18:586-94.
10. Le NQK, Kha QH, Nguyen VH, Chen YC, Cheng SJ, Chen CY. Machine Learning-Based Radiomics Signatures for EGFR and KRAS Mutations Prediction in Non-Small-Cell Lung Cancer. Int J Mol Sci 2021;22:9254.
11. Li S, Ding C, Zhang H, Song J, Wu L. Radiomics for the prediction of EGFR mutation subtypes in non-small cell lung cancer. Med Phys 2019;46:4545-52.
12. Trivizakis E, Souglakos J, Karantanas A, Marias K. Deep Radiotranscriptomics of Non-Small Cell Lung Carcinoma for Assessing Molecular and Histology Subtypes with a Data-Driven Analysis. Diagnostics (Basel) 2021.
13. Hatt M, Cheze Le Rest C, Antonorsi N, Tixier F, Tankyevych O, Jaouen V, Lucia F, Bourbonne V, Schick U, Badic B, Visvikis D. Radiomics in PET/CT: Current

Status and Future AI-Based Evolutions. Semin Nucl Med 2021;51:126-33.

14. Koyasu S, Nishio M, Isoda H, Nakamoto Y, Togashi K. Usefulness of gradient tree boosting for predicting histological subtype and EGFR mutation status of non-small cell lung cancer on (18)F FDG-PET/CT. Ann Nucl Med 2020;34:49-57.

15. Nakajo M, Jinguji M, Ito S, Tani A, Hirahara M, Yoshiura T. Clinical application of (18)F-fluorodeoxyglucose positron emission tomography/computed tomography radiomics-based machine learning analyses in the field of oncology. Jpn J Radiol 2024;42:28-55.

16. Shiri I, Amini M, Nazari M, Hajianfar G, Haddadi Avval A, Abdollahi H, Oveisi M, Arabi H, Rahmim A, Zaidi H. Impact of feature harmonization on radiogenomics analysis: Prediction of EGFR and KRAS mutations from non-small cell lung cancer PET/CT images. Comput Biol Med 2022;142:105230.

17. Huang Z, Chen Z, Chen J, Lu P, Quan G, Du Y, Li C, Gu Z, Yang Y, Liu X, Zheng H, Liang D, Hu Z. DaNet: dose-aware network embedded with dose-level estimation for low-dose CT imaging. Phys Med Biol 2021;66:015005.

18. Huang Z, Chen Z, Quan G, Du Y, Yang Y, Liu X, Zheng H, Liang D, Hu Z. Deep Cascade Residual Networks (DCRNs): Optimizing an Encoder–Decoder Convolutional Neural Network for Low-Dose CT Imaging. IEEE Transactions on Radiation and Plasma Medical Sciences 2022;6:829-40.

19. Huang Z, Liu X, Wang R, Chen Z, Yang Y, Liu X, Zheng H, Liang D, Hu Z. Learning a Deep CNN Denoising Approach Using Anatomical Prior Information Implemented With Attention Mechanism for Low-Dose CT Imaging on Clinical Patient Data From Multiple Anatomical Sites. IEEE J Biomed Health Inform 2021;25:3416-27.

20. Zhang H, Liao M, Guo Q, Chen J, Wang S, Liu S, Xiao F. Predicting N2 lymph node metastasis in presurgical stage I-II non-small cell lung cancer using multiview radiomics and deep learning method. Med Phys 2023;50:2049-60.

21. Wang D, Hu Y, Zhan C, Zhang Q, Wu Y, Ai T. A nomogram based on radiomics signature and deep-learning signature for preoperative prediction of axillary lymph node metastasis in breast cancer. Front Oncol 2022;12:940655.

22. Huang Z, Li W, Wu Y, Guo N, Yang L, Zhang N, Pang Z, Yang Y, Zhou Y, Shang Y, Zheng H, Liang D, Wang M, Hu Z. Short-axis PET image quality improvement based on a uEXPLORER total-body PET system through deep learning. Eur J Nucl Med Mol Imaging 2023;51:27-39.

23. Huang Z, Li W, Wang Y, Liu Z, Zhang Q, Jin Y, Wu R, Quan G, Liang D, Hu Z, Zhang N. MLNAN: Multi-level noise-aware network for low-dose CT imaging implemented with constrained cycle Wasserstein generative adversarial networks. Artif Intell Med 2023;143:102609.

24. Li W, Huang Z, Wang H, Zhou C, Zhang X, Fan W, Hu Z. A 3D noise-level-aware network for low-dose PET imaging. J Nucl Med 2023;64:P791.

25. Lu J, Gao X, Wang S, He Y, Ma X, Zhang T, Liu X. Advanced strategies to evade the mononuclear phagocyte system clearance of nanomaterials. Exploration (Beijing) 2023;3:20220045.

26. Gabelloni M, Faggioni L, Fusco R, Simonetti I, De Muzio F, Giacobbe G, Borgheresi A, Bruno F, Cozzi D, Grassi F, Scaglione M, Giovagnoni A, Barile A, Miele V, Gandolfo N, Granata V. Radiomics in Lung Metastases: A Systematic Review. J Pers Med 2023;13:225.

27. Yang X, Liu M, Ren Y, Chen H, Yu P, Wang S, Zhang R, Dai H, Wang C. Using contrast-enhanced CT and non-contrast-enhanced CT to predict EGFR mutation status in NSCLC patients-a radiomics nomogram analysis. Eur Radiol 2022;32:2693-703.

28. Wu S, Shen G, Mao J, Gao B. CT Radiomics in Predicting EGFR Mutation in Non-small Cell Lung Cancer: A Single Institutional Study. Front Oncol 2020;10:542957.

29. Lu CF, Liao CY, Chao HS, Chiu HY, Wang TW, Lee Y, Chen JR, Shiao TH, Chen YM, Wu YT. A radiomics-based deep learning approach to predict progression free-survival after tyrosine kinase inhibitor therapy in non-small cell lung cancer. Cancer Imaging 2023;23:9.

30. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 27-30 June 2016; Las Vegas, NV, USA. IEEE; 2016.

31. Singh R, Khare A. Fusion of multimodal medical images using Daubechies complex wavelet transform–A multiresolution approach. Information Fusion 2014;19:49-60.

32. Aluka M, Ganesan S, Reddy P. A Comparative Study on Pre-Training Models of Deep Learning to Detect Lung Cancer. International Journal of Intelligent Systems and Applications in Engineering 2023;11:148-55.

33. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014. Available online: https://doi.org/10.48550/arXiv.1409.1556

34. Jiménez Gaona Y, Castillo Malla D, Vega Crespo B, Vicuña MJ, Neira VA, Dávila S, Verhoeven V. Radiomics

5472

Yao et al. Fused shallow-deep features for predicting EGFR mutations

Diagnostic Tool Based on Deep Learning for Colposcopy Image Classification. Diagnostics (Basel) 2022.

35. Liang W, Tian W, Wang Y, Wang P, Wang Y, Zhang H, Ruan S, Shao J, Zhang X, Huang D, Ding Y, Bai X. Classification prediction of pancreatic cystic neoplasms based on radiomics deep learning models. BMC Cancer 2022;22:1237.

36. Xu X, Mao Y, Tang Y, Liu Y, Xue C, Yue Q, Liu Q, Wang J, Yin Y. Classification of Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma Based on Radiomic Analysis. Comput Math Methods Med 2022;2022:5334095.

37. Xu Z, Wang Y, Chen M, Zhang Q. Multi-region radiomics for artificially intelligent diagnosis of breast cancer using multimodal ultrasound. Comput Biol Med 2022;149:105920.

38. Song R, Cheng Y, Zheng T. The Effect of Gefitinib on Treatment Necessity and Prognosis of NSCLC Patients with Early EGFR Mutations. Contrast Media Mol Imaging 2022;2022:2228744.

39. Junaidi A, Ermayanti S, Afriani A. Benefits of Conventional Chemotherapy in Progressive Disease Patients with Tyrosine-Kinase Inhibitors: A Case Report. Bioscientia Medicina: Journal of Biomedicine and Translational Research 2022;6:2625-33.

40. Cao Q, Rui G, Liang Y. Study on PM2.5 pollution and the mortality due to lung cancer in China based on geographic weighted regression model. BMC Public Health 2018;18:925.

41. Zhu JM, Sun L, Wang L, Zhou TC, Yuan Y, Zhen X, Liao ZW. Radiomics combined with clinical characteristics predicted the progression-free survival time in first-line targeted therapy for advanced non-small cell lung cancer with EGFR mutation. BMC Res Notes 2022;15:140.

42. Li Y, Lv X, Wang B, Xu Z, Wang Y, Sun M, Hou D. Predicting EGFR T790M Mutation in Brain Metastases Using Multisequence MRI-Based Radiomics Signature. Acad Radiol 2023;30:1887-95.

43. Huang Z, Liu D, Chen W, Lu P, Chen J. Adversarial Learning for Image Super Resolution Using Auxiliary Texture Feature Attributes. 2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS); 22-24 January 2021; Shenyang, China. IEEE; 2021.

44. Huang Z, Liu H, Wu Y, Li W, Liu J, Wu R, Yuan J, He Q, Wang Z, Zhang K, Liang D, Hu Z, Wang M, Zhang N. Automatic brain structure segmentation for (18) F-fluorodeoxyglucose positron emission tomography/

magnetic resonance images via deep learning. Quant Imaging Med Surg 2023;13:4447-62.

45. Huang Z, Liu Z, He P, Ren Y, Li S, Lei Y, Luo D, Liang D, Shao D, Hu Z, Zhang N. Segmentation-guided Denoising Network for Low-dose CT Imaging. Comput Methods Programs Biomed 2022;227:107199.

46. Huang Z, Liu X, Wang R, Chen J, Lu P, Zhang Q, Jiang C, Yang Y, Liu X, Zheng H, Liang D, Hu Z. Considering anatomical prior information for low-dose CT image enhancement using attribute-augmented Wasserstein generative adversarial networks. Neurocomputing 2021;428:104-15.

47. Huang Z, Wu Y, Fu F, Meng N, Gu F, Wu Q, Zhou Y, Yang Y, Liu X, Zheng H, Liang D, Wang M, Hu Z. Parametric image generation with the uEXPLORER total-body PET/CT system through deep learning. Eur J Nucl Med Mol Imaging 2022;49:2482-92.

48. Weitao T, Grandinetti G, Guo P. Revolving ATPase motors as asymmetrical hexamers in translocating lengthy dsDNA via conformational changes and electrostatic interactions in phi29, T7, herpesvirus, mimivirus, E. coli, and Streptomyces. Exploration (Beijing) 2023;3:20210056.

49. Guo J, Zhao Z, Shang ZF, Tang Z, Zhu H, Zhang K, editors. Nanodrugs with intrinsic radioprotective exertion: Turning the double-edged sword into a single-edged knife. Exploration 2023;3:20220119.

50. Geng Z, Cao Z, Liu J. Recent advances in targeted antibacterial therapy basing on nanomaterials. Exploration (Beijing) 2023;3:20210117.

51. Peng H, Yao F, Zhao J, Zhang W, Chen L, Wang X, Yang P, Tang J, Chi Y, editors. Unraveling mitochondria-targeting reactive oxygen species modulation and their implementations in cancer therapy by nanomaterials. Exploration 2023;3:20220115.

52. Bizzego A, Bussola N, Salvalai D, Chierici M, Maggio V, Jurman G, Furlanello C. Integrating deep and radiomics features in cancer bioimaging. 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB); 09-11 July 2019; Siena, Italy. IEEE; 2019.