Published in final edited form as:

J Comput Assist Tomogr. 2019 ; 43(2): 300–306. doi:10.1097/RCT.0000000000000836.

Radiomics for Classifying Histological Subtypes of Lung Cancer based on Multiphasic Contrast-Enhanced Computed Tomography

Linning E^a, Lin Lu^{b,*}, Li Li^c, Hao Yang^b, Lawrence H. Schwartz^b, and Binsheng Zhao^b

^a. Department of Radiology, Shanxi DAYI Hospital, 99 Longcheng Street, Taiyuan, Shanxi, 10032, China

^b. Department of Radiology, Columbia University Medical Center, 630 West 168th Street, New York, NY, 10032, USA

^c. Department of Pathology, Shanxi DAYI Hospital, 99 Longcheng Street, Taiyuan, Shanxi, 10032, China

Abstract

Objectives—To evaluate the performance of radiomics method on classifying histological subtypes of lung cancer in multiphasic contrast-enhanced computed tomography (CT) images.

Materials and Methods—A total of 229 patients with pathologically confirmed lung cancer were recruited. All recruited patients underwent non-enhanced and dual phase chest contrast-enhanced CT. 1160 quantitative radiomic features were extracted from CT images to build radiomic classification models. The performance of the classification models was evaluated by the area under the receiver operating characteristic curve (AUC).

Results—The AUC of radiomics models in classifying AD and SCC, AD and SCLC, and SCC and SCLC were 0.801, 0.857, 0.657 at non-enhanced phase; 0.834, 0.855, 0.619 at arterial phase; and 0.864, 0.864, 0.664 at venous phase, respectively.

Conclusion—Our study indicates that radiomics method might be a promising tool for noninvasive predicting histological subtypes of lung cancer based on the multiphasic contrast-enhanced CT images.

Keywords

Lung cancer; Histological Subtype; Computed Tomography; Radiomics; Machine Learning

1. Introduction

Lung cancer remains the most frequently diagnosed cancer and a leading cause of cancer-related mortality worldwide [1]. Lung cancer has two major types, namely, non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC is further classified into

*Corresponding author Lin Lu, Ph.D. Department of Radiology, Columbia University Medical Center, 630 West 168th Street, New York, NY, 10032, USA, Tel: 1 (646) 808-9216, 112860@cumc.columbia.edu.

lung adenocarcinoma (AD) and squamous cell lung cancer (SCC). NSCLC accounts for approximately 85% of all lung cancers, and approximately 10%–15% of lung cancers are SCLC[2]. In recent years, accurate classification of lung cancer has become the initial and a significant step in clinical management and treatment planning for lung cancer patients due to the significant associations between prognosis, recurrence and survival rates, and the histological subtypes of lung cancer [3–5]. Non-invasive and accurate classification of histological subtypes of lung cancer on the basis of imaging methods has important clinical application value.

Previous studies have shown that imaging phenotypic differences exist among the three histological subtypes of lung cancer on the basis of the morphological signs in computed tomography (CT). Moreover, most SCLCs are centrally located and manifest as mediastinal or hilar lymphadenopathy [6]. AD is often characterized by peripheral tumors with air bronchogram and ground glass opacity. The squamous carcinoma is often characterized by tumors with a solid mass density and thick cavity [7]. However, in the current clinical practice, the visual evaluation of CT morphological signs depended mainly on radiologist's knowledge, analytical skills, intuition and diligence.

Radiomics refers to quantitative characterization of tumor phenotypes by using large amounts of quantitative image features extracted from medical images [8]. Radiomics has already been applied in cancer diagnosis, prognosis and treatment response assessment [9–12], and is now gaining increasing attention in oncology. In the field of lung cancer research, studies have shown that radiomics has great potential in classifying lung cancer histological subtypes based on CT image [13–16]. However, different protocols of CT scanning were employed in radiomics studies of lung cancer, including non-enhanced CT and contrast-enhanced CT (CECT) with the starting time varying from 25s to 90s after the injection of contrast agent [10; 17–19]. Previous research [20] indicated that contrast-enhancement could affect the performance of radiomic models in diagnosing solitary pulmonary nodules. For example, Perrin's study[21] also showed that the proportion of reproducible features decreases with the increasing variations in contrast injection rates in liver cancer patients who underwent consecutive CECTs. In diagnostic radiology, a large proportion of CT images are contrast-enhanced. Thus, the reproducibility and variability of radiomic features after the intravenous administration of contrast agent should be investigated due to the dynamic nature of CECT.

Therefore, in this study, we aim to retrospectively assess the performance of radiomics in classifying the histological subtypes of lung cancer based on multiphasic (non-enhanced, arterial phase, and venous phase) CT.

2. Materials and Methods

2.1 Patients and Image Acquisition

Patient data were retrospectively collected from our Hospital. The institutional review board approved the study before its commencement and waived the requirement for informed consent. A radiologist searched the clinical and radiological databases of the institution to identify pathologically confirmed lung cancer patients who underwent preoperative chest CT

between June 2012 and December 2017. Due to the unbalanced population frequency within the three lung histological subtypes, the inclusion of patient into study cohort was consisted of two stages. At the first stage, SCLC patients were searched throughout the entire time-period. At the second stage, AD and SCC patients were searched beginning at April 2012 and stopped when the patient numbers equaling to that of the SCLC patients. In all cases, definitive diagnosis was made through surgical resection or CT-guided percutaneous biopsy. Clinical features, including age at CT examination, and gender were recorded. Cases were excluded for the following reasons: 1) patients received treatment for the pulmonary lesion; 2) image quality was unsatisfactory due to severe artifacts; and 3) no contrast-enhanced CT imaging.

All included patients underwent non-enhanced and dual phase chest contrast-enhanced CT (CECT) using the SOMATOM Definition Flash dual-source CT scanner (Siemens Medical Solutions). CECT was performed after the injection of the iodinated contrast agent (1 ml/kg) (Omnipaque 350, GE Healthcare) at a rate of 3 mL/s via a pump injector. First, the thorax was imaged in the arterial phase of enhancement. This acquisition was triggered by bolus tracking. For this, a region of interest (ROI) was placed within the descending thoracic aorta, and scanning was triggered as soon as peak enhancement (100 HU) was reached. Second, the thorax was imaged at the portal venous phase of enhancement (60-s delay after the injection of contrast agent). The parameters used for CT scanning were as follows: pitch of 1.0, rotation time of 0.5s, tube voltage of 120 kVp, and a tube current-time product of 50 – 250 mAs in automatic dose modulation (Care Dose 4D) and detector collimation: 128×0.6 mm. All CECT scans were reconstructed into an image series by using a thin slice thickness (1 mm) and mediastinal convolution kernels ('B30f'). All non-enhanced CT scans were reconstructed with a slice thickness that varied from 1~5 mm and mediastinal convolution kernels ('B30f').

2.2 Tumor Segmentation

In our study, an experienced radiologist (L.E with 20 years of experience in lung CT interpretation) segmented each primary tumor by using a semi-automated segmentation algorithm integrated into an open source imaging platform, Weasis [22]. Each tumor was segmented on the non-enhanced, arterial phase, and venous phase CT images. A Lung window level (1500 HU and -500 HU) was used during the tumor segmentation. For consistency, segmentation was first performed on venous phase CT images. Then, the contours on the venous phase CT images were copied to the non-enhanced and arterial phase CT images. Finally, the contours on the non-enhanced and arterial phase CT images were reviewed. Care was taken to ensure that the vessels were not included in the contours.

2.3 Feature Extraction

For each tumor, 1,160 well-defined quantitative image features were extracted from each of the three phases using the in-house feature extraction algorithms implemented on the MATLAB 2016b platform (MathWorks). The 1,160 quantitative image features were designed to characterize tumor phenotypes in terms of shape (e.g., eccentricity, compactness), size (e.g., volume, largest diameter), boundary sharpness (e.g., sigmoid slope), and texture patterns (e.g., gray-level co-occurrence matrix). They were extended

from the basic set of 89 image features used in a previous lung radiomic study [23] by increasing the scales of the feature parameters.

2.4 Reproducibility Analysis

The reproducibility of the extracted quantitative image features was independently assessed by using a previously collected NSCLC test-retest dataset[23] that contained 32 NSCLC patients' same-day repeat CT scan images. The repeat CT scan image series of thin slice thickness (1.25 mm) and mediastinal convolutional kernel ("standard" kernel) was used in our study. For the 32 patients, 1,160 radiomic features were extracted from the tumor on both the test and re-test scans. The reproducibility of features was evaluated by using the concordance correlation coefficient (CCC)[24]. Features with a CCC lower than 0.9 were regarded as non-reproducible features and were excluded from the subsequent analyses.

2.5 Model Building

In our study, nine optimal models were developed to classify AD and SCC, AD and SCLC, and SCC and SCLC under the three phases. To build each model, 1,160 candidate features were involved to maximally retain potentially valuable information. In order to deal with such a large pool of the image features, a "coarse" to "fine" strategy was developed to select a small set of informative and non-redundant features for building the classification model to avoid the overfitting effect as much as possible.

The coarse selection procedure involved the application of hierarchical clustering and feature ranking techniques. First, the features were clustered into a hierarchical tree in an unsupervised manner on the basis of their mutual correlations. The number of clusters is determined by a specific correlation threshold. In our study, the correlation threshold was set to 0.8 based on our experience, that is, features with correlations larger than 0.8 were assigned to the same cluster. Then, the minimum redundancy maximum relevance (mRMR) algorithm[25] was applied to rank the correlated features within each cluster. Finally, the top-ranked feature within each cluster was selected for fine selection analysis.

The fine selection procedure involved the incremental forward search (IFS)[25] approach and three supervised machine-learning classification algorithms. IFS is used for evaluating features sequentially. The three classification algorithms, Naïve Bayes[26], Linear Logistic Regression[27] and Random Forest[28] were employed to combine features into models. The candidate classification models built with different numbers of features and different classification algorithms were evaluated in terms of their classification performance. The candidate model with the best performance was selected as the final model.

The "coarse" to "fine" procedures were implemented on the MATLAB 2016b platform (MathWorks).

2.6 Performance Evaluation

The classification models were evaluated using the receiver operating characteristic curve (ROC) analysis. The performance of classification model was indicated by the area under the ROC (AUC), and the sensitivity and specificity were determined by Youden's J statistic.

Tenfold cross-validation was applied for performance validation. In the 10-fold cross-validation, the entire dataset was randomly partitioned into 10 subsets. One subset was used as the testing set, and the other nine subsets were used as the training set. The training-testing procedures were repeated 10 times, until each sample in the dataset was used as training and testing samples once. When used as a test sample, the sample is assigned a prediction score. The final AUC for the classification model, and its confidence interval (95%) were estimated based on the prediction score by using bootstrapping (1,000 times). Besides ROC analysis, the DeLong method[29] was employed to compare the performance of different models; and the two-sample t-test was employed to evaluate the difference between two continuous distributions.

3. Result

In our study, the classification models based on the data of the three histological subtypes, namely, SCLC versus AD, SCLC versus SCC and AD versus SCC, were built and compared at three phases, namely, non-enhanced, arterial, and venous phases.

3.1 Patient Characteristics

In this study, 128 SCLC patients were retrospectively collected from April 2012 to November 2017. Correspondingly, 128 AD and 128 SCC patients were retrospectively collected to match the SCLC patient number beginning from April 2012. For the totally 384 lung cancer patients, 155 patients (including 73 SCLC, 38 AD and 44 SCC) were excluded according to the exclusion criteria. Finally, 55 SCLC, 90 AD and 84 SCC patients were recruited for this study cohort. Among the 229 patients, 157 were males (68.56%) and 72 were females (31.44%). The median ages of the SCLC, AD, and SCC patients were 64, 66, and 69 years (ranges of 59–81, 32–82, and 45–85 years), respectively. The SCLC and SCC groups (88.12% and 89.35%, respectively) included a significantly higher percentage of male patients than the AD group (37.78%, $p = 0.006$).

3.2 Optimal Classification Models for Differentiating Histological Subtypes of Lung Cancer

Out of the 1,160 extracted quantitative image features, 869 features were identified as reproducible features by using the results of a prior repeat CT scan study. Based on the reproducible features and the developed coarse-to-fine strategy, nine optimal classification models were developed. They are models for classifying AD and SCC, AD and SCLC, and SCC and SCLC at three phases. The performances in classifying AD and SCC were AUC (95% CI) = 0.801 (0.737, 0.860), 0.834 (0.798, 0.907), and 0.864 (0.804, 0.912) for the three phases (Fig 1). The performances in classifying AD and SCLC were AUC (95% CI) = 0.857 (0.790, 0.909), 0.855 (0.787, 0.907), and 0.864 (0.782, 0.904) for the three phases (Fig 2). The performances in classifying SCC and SCLC were AUC (95% CI) = 0.657 (0.572, 0.735), 0.619 (0.534, 0.700), and 0.664 (0.580, 0.741) for the three phases (Fig 3). Within all classifications of the three lung cancer histological subtypes, no significant difference was observed between the three phases ($p > 0.05$).

The optimal radiomic features selected for building the nine classification models are listed in Table 1. As shown in Table 1, the models under artery and vein phases shared more common features than under the non-contrast phase, though no significant difference on performance was observed between the three phases. We investigated the mean $\pm$ SD of the HUs of the three histological subtypes under the three phases (Table 2). The HUs under the artery and vein phases were higher than that under the non-contrast phase, and the distributions of HUs were more similar than that at the non-contrast phase. Moreover, we investigated the distribution of “Laws_2”, the top ranked feature in the AD and SCC classification which achieved the highest AUC of 0.864 (See Figure 4). The “Laws_2” in AD was higher than that in SCC, and the trend, values under artery and vein phases were more similar than that under the non-contrast phase.

4. Discussion

In this study, the application of radiomics in classifying histological subtypes of lung cancer utilizing the three-phase CT was investigated. The radiomic models showed a high performance of AUCs ranging from 0.801 to 0.864 (in classifying AD and SCC) and 0.850 to 0.857 (in classifying AD and SCLC) under the three different scanning phases. In the classification of SCC and SCLC, the AUCs under the three phases ranged from 0.619 to 0.664. Our study strongly indicated that phenotypic characteristics could help differentiate AD from SCC and SCLC. However, the differentiation of SCC and SCLC showed relatively low performance.

In the classification of AD and SCC, the highest AUC (0.867) was achieved at the venous phase, although no statistically significant difference was observed among the AUCs of the three phases. In the venous phase, model shown in Table 1, the “Laws_2” feature was selected as the top-ranked feature which contributed the most to the classification model. Technically, the “Laws_2” feature belongs to a kind of textual features that quantifies heterogeneity based on a line-like texture pattern[30]. As shown in Fig. 4, the “Laws_2” value of AD was significantly higher than that of SCC, suggesting that the CT imaging phenotypes of AD were more heterogeneous than those of the SCC, which was consistent with the difference in pathological characteristics between AD and SCC. Fig.5 shows that the pathological structure of AD tissues is loose, and the tumor cells are loosely organized and often show adenoid differentiation and a glandular pattern, while the pathological structure of SCC tissues is compact, and the SCC cells are densely packed with few stromal components. Such pathological difference between AD and SCC was also reflected by their average density on the CT image, that is, the average HUs of SCC were significantly higher than those of AD ($p < 0.01$) under the three scanning phases (See Table 2).

In our study, the radiomic model showed a high performance in classifying AD and SCLC. The “Intensity_Entropy” and “GLCM_Contrast”[31] features were selected as the top-ranked features in classifying AD and SCC. Unlike the line-like feature used in classifying AD and SCC, “Intensity_Entropy” and “GLCM_Contrast” belong to the coarse features kind. The high values of “Intensity_Entropy” and “GLCM_Contrast” indicated heterogeneity within the tumor. These radiomics features could provide additional quantitative pattern information that may not be appreciated by the naked eye.

Radiomic model showed weak performance in classifying SCC and SCLC. The AUCs of the three-phases were 0.657, 0.619 and 0.664, respectively. The lower AUCs of classifying SCC and SCLC can be attributed to two reasons. First, SCC and SCLC are of similar location and morphology. Both SCC and SCLC are common central lung cancers and show a mass morphology with a homogeneous density. Second, SCC and SCLC share similar pathological structures, both have a dense tumor cell arrangement with few stromal components. The average CT values of the tumor tissue in the three phases showed no significant difference between SCC and SCLC.

Moreover, our study showed that contrast-enhanced CT has a definitive impact on the selection of radiomic features. As shown in Table 1, the radiomics models based on the arterial and venous phases of enhancement shared more common features than the model based on the non-contrast phase. The underlying reason for the variability of radiomic features based on contrast-enhanced CT may be correlated with the biological heterogeneity within the tumor tissue (vascularity or lack of it). The heterogeneity on CT can be quantified using radiomics analysis, which reflects the coarseness and regularity that results from local spatial variations in image brightness (an extension from measuring attenuation). The spatial variations in image brightness may be enhanced by the intravenously injected contrast agent, which may then result in the variability of radiomics features. This is supported by the findings of previous studies, which suggested that hepatic texture parameters correlate with hepatic vascularity[32], and the proportion of reproducible radiomics features decreases with the increasing variations in contrast injection rates in liver cancer patients[21]. Therefore, our study indicated that radiomic features may be affected by the phase (arterial or venous) of acquisition, and the reproducibility and stability of radiomics features due to image acquisition should be investigated extensively before using these features for classifying models. Some limitations of this study include, the number of patients is relatively small, and the impact of the volume and rate of injected contrast agent on the reproducibility and stability of radiomics features were not studied.

In conclusion, the use of a radiomics approach for classifying the histological subtypes of lung cancer demonstrates potential for differentiating AD and SCC, and AD and SCLC. Contrast enhancement on the CT has a definitive impact on the extraction of radiomic features. Larger scale studies are suggested to further validate these findings and to investigate the reproducibility and stability of radiomics features from image acquisition on a wide range of lung diseases.

Acknowledgements

Conflicts of Interest and Source of Funding:

This work was supported by the Natural Science Foundation of Shanxi Province, China. [Grant number 201701D121151]. For the remaining authors none were declared.

Abbreviations

CT	computed tomography
CECT	contrast-enhanced computed tomography

QRF	quantitative radiomic feature
AD	adenocarcinoma
SCC	squamous cell carcinoma
NSCLC	non-small cell lung cancer
SCLC	small cell lung cancer
AUC	area under the curve
ROC	receiver operating characteristic
CCC	concordance correlation coefficient
mRMR	minimum redundancy maximum relevance
RFS	incremental forward search

Reference

1. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015; 2:87–108.
2. Ha SY, Roh MS. The new 2011 international association for the study of lung cancer/american thoracic society/european respiratory society classification of lung adenocarcinoma in resected specimens: clinicopathologic relevance and emerging issues. *Korean J Pathol*. 2013; 8:316–325.
3. Tian S, Wang C. An M-W Test on existence of histology subtype-specific prognostic signatures among early stage lung adenocarcinoma and squamous cell carcinoma patients using a Cox-model based filter. *Biology direct*. 2015; 10:15. [PubMed: 25887039]
4. Tian S Identification of subtype-specific prognostic genes for early-stage lung adenocarcinoma and squamous cell carcinoma patients using an embedded feature selection algorithm. *PLoS One*. 2015; 10: e0134630. [PubMed: 26226392]
5. Skrzypski M, Dziadziuszko R, Jassem E, et al. Main histologic types of non-small-cell lung cancer differ in expression of prognosis-related genes. *Clin Lung Cancer*. 2013; 6:666–673.
6. Carter BW, Glisson BS, Truong MT, et al. Small cell lung carcinoma: staging, imaging, and treatment considerations. *Radiographics*. 2014; 6:1707–1721.
7. Kunihiro Y, Kobayashi T, Tanaka N, et al. High-resolution CT findings of primary lung cancer with cavitation: a comparison between adenocarcinoma and squamous cell carcinoma. *Clin Radiol*. 2016; 11:1126–1131.
8. Lee G, Bak SH, Lee HY. CT Radiomics in Thoracic Oncology: Technique and Clinical Applications. *Nucl Med Mol Imaging*. 2018; 2:91–98.
9. Fukui T, Taniguchi T, Kawaguchi K, et al. Comparisons of the clinicopathological features and survival outcomes between lung cancer patients with adenocarcinoma and squamous cell carcinoma. *Gen Thorac Cardiovasc Surg*. 2015; 9:507–513.
10. Ahn SY, Park CM, Park SJ, et al. Prognostic Value of Computed Tomography Texture Features in Non-Small Cell Lung Cancers Treated with Definitive Concomitant Chemoradiotherapy. *Invest Radiol*. 2015; 10:719–725.
11. Chen X, Fang M, Dong D, et al. A Radiomics Signature in Preoperative Predicting Degree of Tumor Differentiation in Patients with Non-small Cell Lung Cancer. *Acad Radiol*. 2018; 10.1016/j.acra.2018.02.019.
12. Chen CH, Chang CK, Tu CY, et al. Radiomic features analysis in computed tomography images of lung nodule classification. *PLoS One*. 2018; 2: e0192002.
13. Zhu X, Dong D, Chen Z. et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. *Eur Radiol*. 2018; 7: 2772–2778.

14. Haga A, Takahashi W, Aoki S. et al. Classification of early stage non-small cell lung cancers on computed tomographic images into histological types using radiomic features: interobserver delineation variability analysis. *Radiol Phys Technol.* 2018; 1:27–35.
15. Hong SJ, Kim TJ, Choi YW, et al. Radiogenomic correlation in lung adenocarcinoma with epidermal growth factor receptor mutations: Imaging features and histological subtypes. *Eur Radiol.* 2016; 10:3660–3668.
16. Wu W, Parmar C, Grossmann P, et al. Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology. *Front Oncol.* 2016; 6:71. [PubMed: 27064691]
17. Suo ST, Cheng JJ, Cao MQ, et al. Assessment of Heterogeneity Difference Between Edge and Core by Using Texture Analysis: Differentiation of Malignant from Inflammatory Pulmonary Nodules and Masses. *Acad Radiol.* 2016; 9:1115–1122.
18. Ravanelli M, Farina D, Morassi M, et al. Texture analysis of advanced non-small cell lung cancer (NSCLC) on contrast-enhanced computed tomography: prediction of the response to the first-line chemotherapy. *Eur Radiol.* 2013; 12:3450–3455.
19. Dennie C, Thornhill R, Sethi-Virmani V, et al. Role of quantitative computed tomography texture analysis in the differentiation of primary lung cancer and granulomatous nodules. *Quant Imaging Med Surg.* 2016; 6:6–15. [PubMed: 26981450]
20. He L, Huang Y, Ma Z, et al. Effects of contrast-enhancement, reconstruction slice thickness and convolution kernel on the diagnostic performance of radiomics signature in solitary pulmonary nodule. *Sci Rep.* 2016; 10:34921.
21. Perrin T, Midya A, Yamashita R, et al. Short-term reproducibility of radiomic features in liver parenchyma and liver malignancies on contrast-enhanced CT imaging. *Abdom Radiol.* 2018; 10.1007/s00261-018-1600-6.
22. Tan Y, Schwartz LH, Zhao B. Segmentation of lung lesions on CT scans using watershed, active contours, and Markov random field. *Med phys.* 2013; 4: 043502.
23. Zhao B, Tan Y, Tsai WY, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep.* 2016; 6:23428. [PubMed: 27009765]
24. Lawrence I, Lin K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989; 1:255–268.
25. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell.* 2005; 8:1226–1238.
26. Murphy KP. Naive bayes classifiers. University of British Columbia. 2006; 10.
27. Tabachnick BG, Fidell LS. Using multivariate statistics. Boston, MA: Allyn & Bacon/Pearson Education; 2007.
28. Breiman L Random forests. *Machine learning.* 2001; 1:5–32.
29. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988; 3:837–845.
30. Laws KI. Rapid texture identification Image processing for missile guidance. *Proc. SPIE.* 1980; 12: 376–382.
31. Haralick RM, Shanmugam K. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics.* 1973; 6:610–621.
32. Bezy-Wendling J, Kretowski M, Rolland Y, et al. Toward a better understanding of texture in vascular CT scan simulated images. *IEEE Trans Biomed Eng.* 2001; 1:120–124.

Figure 1. ROC curve of the models in classifying AD and SCC.

The performances in classifying AD and SCC were AUC (95% CI) = 0.801 (0.737, 0.860), 0.834 (0.798, 0.907), and 0.864 (0.804, 0.912) for the three phases

Figure 2. ROC curve of the models in classifying AD versus SCLC.

The performances in classifying AD and SCLC were AUC (95% CI) = 0.857 (0.790, 0.909), 0.855 (0.787, 0.907), and 0.864 (0.782, 0.904) for the three phases.

Figure 3. ROC curve of the models in classifying SCC versus SCLC.

The performances in classifying SCC and SCLC were AUC (95% CI) = 0.657 (0.572, 0.735), 0.619 (0.534, 0.700), and 0.664 (0.580, 0.741) for the three phases.

Figure 4. Boxplots for the distribution of Laws_2 in AD and SCC.

The “Laws_2” value of AD was significantly higher than that of SCC, suggesting that the CT imaging phenotypes of AD were more heterogeneous than those of the SCC.

Figure 5. Representative photomicrographs (original magnification, $\times 100$; H&E staining) of three histological subtypes of lung cancer.

The AD cells are loosely organized, showing adenoid differentiation and a glandular pattern (A). The internal structure of SCC is dense with few interstitial components (B). The SCLC has a dense tumor cell arrangement with few stromal components, tumor cells are arranged in a disorganized pattern, (C).

Figure 6. Representative photomicrographs (original magnification, $\times 100$; H&E staining) of three histological subtypes of lung cancer.

The pathological structure of three lung cancer histological subtypes is quite different, and these features at the histological level would manifest themselves as different textural gray-level patterns on CT images. The AD cells are loosely organized, showing adenoid differentiation and a glandular pattern (A). The internal structure of SCC is dense with few interstitial components (B). The SCLC has a dense tumor cell arrangement with few stromal components, tumor cells are arranged in a disorganized pattern, (C).

Table 1.

Radiomic features for classification models.

	Non-contrast	Artery phase	Vein phase
AD vs SCC	DWF_Z_LLD	LoG_Mean	Laws_2
	DWF_Z_LLL	Laws_2	DWT1_HH
	GLCM_Sum_Entropy	GLCM_IMC1	Laws_1
			Sigmoid_Slope_Mean
			DWF_Z_H
AD vs SCLC	Intensity_Entropy	GLCM_Contrast	GLCM_Contrast
	Gabor_Min	DWT1_HHL	DWT2_LLH
	DWF_Z_LH	Intensity_Kurtosis	Laws_13
		Laws_13	Sigmoid_Offset_Std
		Laws_1	
SCC vs SCLC		Laws_7	
	GLCM_Homogeneity	Fractal_Std	Fractal_Std
	Fractal_Std	LoG_Z_Uniformity	Sigmoid_Offset_Std
	Round_Factor	Laws_6	Intensity_Kurtosis
	DWF_LD	Sigmoid_Offset_Std	Laws_8
	Shape_SI7	GLCM_IMC1	Gabor_sum
	Laws_7	Round_Factor	Laws_5

Our study showed that contrast-enhanced CT has a definitive impact on the selection of radiomic features. The radiomics models based on the arterial and venous phases of enhancement shared more common features than the model based on the non-contrast phase.

Table 2.

The mean $\pm$ SD of the HUs of the three histological subtypes under the three phases.

	Non-contrast (HU)	Artery phase (HU)	Vein phase (HU)
SCLC	12.1 $\pm$ 28.6	43.4 $\pm$ 23.0	43.6 $\pm$ 30.2
SCC	-3.4 $\pm$ 51.6	34.5 $\pm$ 39.6	37.8 $\pm$ 38.9
AD	-136.6 $\pm$ 157.9	-83.1 $\pm$ 154.4	-80.3 $\pm$ 151.7

The average CT values of the tumor tissue in the three phases showed no significant difference between SCC and SCLC (p value>0.05). There is significant difference in the average CT values of the tumor tissue of three phases between SCC and AD, SCLC and AD (p value<0.05)