Exploring the Value of Lung Texture Features in Distinguishing Usual and Non-specific Interstitial Pneumonia

Xinhui Chen
Affiliated Hospital of Guilin Medical University
Ge Cheng
Affiliated Hospital of Guilin Medical University
Xinguan Yang
Affiliated Hospital of Guilin Medical University
Yuting Liao
GE Healthcare
Zhipeng Zhou (bigbird_zhou@hotmail.com)
Affiliated Hospital of Guilin Medical University

Research Article

Keywords: Interstitial pneumonia, High Resolution CT, Radiomics

Posted Date: June 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-533242/v1

Abstract

Backgground At present, the most common types of interstitial pneumonia are usual interstitial pneumonia (UIP) and non-specific interstitial pneumonia (NSIP), and different types have different prognosis. In addition, if there is a mixture of different classifications, it will be difficult for radiologists to diagnose, and it will make clinical treatment difficult. Therefore, clinicians urgently need new imaging methods to solve such problems. This article aims to explore the CT lung texture images of UIP and NSIP to provide evidence for the identification of UIP and NSIP.

Methods A retrospective analysis of 96 cases of interstitial pneumonia diagnosed by the Department of Pathology and the Affiliated Hospital of Guilin Medical College. Among them, there are 40 cases of UIP and 56 cases of NSIP. All patients are scanned by CT. Lung Intelligence Kit was utilized to perform lung segmentation and texture feature extraction. Variance analysis, least absolute shrinkage and selection operator (Lasso) and multivariate logistic regression were used to select effective features. Finally, a multivariate logistic regression model was constructed to identify two kinds of interstitial pneumonia. Receiver operating characteristic (ROC) curve, area under the curve (AUC), sensitivity, specificity were used to evaluate performance of the constructed model. We used the LK software to segment the two sets of lungs. Feature calculation and selection were performed on the data of the two groups of interstitial pneumonia after lung segmentation, the logistic regression model was established for the selected features, and the ROC curve was drawn.

Results A total of 100 texture features are extracted from the whole lung segmented by LK, and finally 8 features are left after feature reduction. The above-mentioned values of UIP and NSIP of the training group are greater than those of the test group.

Conclusions It is possible to distinguish UIP and NSIP by using LK software to extract lung texture in CT images.

1 Introduction

Interstitial pneumonia is a lung disease characterized by pulmonary interstitial inflammation and progressive pulmonary fibrosis. It has many causes and different types have different prognosis[1]. Currently, the most common types of interstitial pneumonia are usual interstitial pneumonia (UIP) and non-specific interstitial pneumonia (NSIP). However, the two types of interstitial pneumonia are difficult to distinguish on the initial clinical manifestations or CT images, especially when they are atypical. In addition, the treatments for the two types of interstitial pneumonia are different. Misdiagnosis and wrong treatment of patients may lead to unforeseen complications[2].

According to related literature reports, pathological biopsy has good performance in the classification of interstitial pneumonia, but it will cause a series of complications and reduce the survival rate of patients [3]. Therefore, it makes sense to find a non-invasive way to distinguish between the two types of interstitial pneumonia. In recent years, radiomics research as an emerging research field has developed

rapidly. It extracts high-level and unrecognizable features from the images; and has shown potential in variable tumor or non-tumor diseases [4–6]. A large number of radiomics studies have been published in lung diseases. However, so far, no studies have been found to distinguish interstitial pneumonia types by radiomics.

In this study, we extracted the double-lung texture features of patients with interstitial pneumonia based on Lung Intelligence Kit (LK; GE Healthcare), and performed feature selection and model construction, aiming to provide imaging evidence for the accurate diagnosis of UIP and NSIP. In this work, our contributions are as follows: (1) Using LK software to perform reliable lung segmentation and texture features extraction. (2) Constructing a radiomics model to support UIP or NSIP diagnosis.

2 Materials And Methods

2.1 Patients

In this study, the data of patients with interstitial pneumonia who were treated in the Affiliated Hospital of Guilin Medical University from March 2015 to July 2020 were retrospectively collected. Whether the patient has interstitial pneumonia is usually determined by the results of frozen lung biopsy during the treatment process or HRCT examination at during the first visit.

The patient's alveolar lavage, lung function, blood oxygen saturation, etc. were consistent with UIP or NSIP respectively, and preliminary classification was performed. The patient continued to recheck after treatment in HRCT, the assessment is divided into effective and ineffective groups, the effective group is NSIP, and the ineffective group is UIP. Finally, inter-division communication is conducted with the clinic, and the clinical diagnosis is UIP or NSIP, and the corresponding grouping is performed.

Patient selection criteria: 1. All patients with interstitial pneumonia 2. The lesion is obvious and clear on HRCT.

Exclusion criteria for patients: 1. There is no pathological result and no clinical diagnosis type. 2. Pathologically and clinically diagnosed non-UIP and non-NSIP patients.

During this period, a total of 96 patients with interstitial pneumonia were collected, of which 40 were in the UIP group (25 males and 15 females), aged 42–91 years old, with a median age of 67 years and an average age of 66.65 years; 56 in the NSIP group Cases (31 males and 25 females), aged 35–89 years old, with a median age of 58 years and an average age of 60.87 years.

2.2 HRCT scanning

CT scans were performed in all cases with GE light VCT 64, GE Optima600 128 or GE Revolution 256-slice CT scanner. The scanning range is from the apex of the lungs to the line 2 cm below the parallel line of the double costal phrenic angle. Scanning parameters are as follows:128KV, automatic mA, noise index 8,

layer thickness and layer spacing of 5mm. The image under secondary thinning and reorganization, layer spacing and layer spacing are both 1.25mm after reorganization.

2.3 Image analysis

Two associate senior or higher physicians who do not understand clinical and pathological classification will make the diagnosis of interstitial pneumonia type on all patients selected for the study. If two doctors make the same type of diagnosis and clinical diagnosis, the diagnosis is still If they are consistent, they are classified as UIP or NSIP accordingly; if the two doctors' diagnosis on the image is inconsistent, they need to communicate again and communicate with the clinic to make the final diagnosis. They are also divided into UIP or NSIP group.

2.4 Lung segmentation and feature extraction

In this study, lung segmentation and feature extraction were conducted by LK software (Lung Intelligence Kit, GE Healthcare). LK is a lung image analysis software, which can automatically segment lung lobe based on DenseVNet, and extract texture features through the integrated pyradiomics package. The schematic diagram of the segmentation result is shown in Fig. 1. Each patient extracts 100 features in 6 categories, including: histogram features, morphological parameters, gray level symbiosis Matrix features, gray-scale dependent matrix, gray-scale run-length matrix, gray-scale area matrix.

2.5 Feature selection and model construction

The data set is randomly divided into training group and test group at a ratio of 7:3. All cases in the training group are used for model training, and the test group is used to independently evaluate the performance of the model. Replace the missing values and outliers in the data with the mean; and use the Z-score method to standardize and preprocess the data. The features with a variance of 0 are removed; after that, we use the LASSO (Least absolute shrinkage and selection operator) method to reduce dimensionality, and finally build a multiple logistic regression model based on the selected features, and draw Nomogram.

2.6 Statistical analysis

The receiver operating characteristic (ROC) curve was used to evaluate the performance of the multiple logistic regression model. Related accuracy, sensitivity, specificity and area under the curve (AUC) were calculated. All statistical analyses in this study were performed using R language (Version 3.5.1, https://www.r-project.org). Two-tailed p value < 0.05 indicates statistical significance.

3 Results

3.1 Comparison of clinical data between groups

Table 1
Comparison of general information between UIP and NSIP groups

Group	cases	Age	Sex(man/femalr)
UIP	40	66.65 ± 10.13	24/15
NSIP	56	60.87 ± 13.54	31/25
Statistics	96	t = 2.273	X^2 = 0.490
P (value)		0.025	0.484

3.2 Feature selection results

Extract 100 features for each patient, 8 features are still retained after LASSO selection of features (Fig. 2), including original_gldm_DependenceVariance, original_glszm_ZonePercentage, original_glcm_DifferenceVariance, original_glrlm_GrayLevelVariance, original_shape_Maximum2DDiameterColumn, original_gldm_DependenceNonUniformity, original_glszm_SizeZoneNonUniformityNormalized, original_glszm_GrayLevelNonUniformityNormalized.

3.3 Model building results

In the training group and the test group, the logistic regression model was used to evaluate the accuracy, F1 value, area under the curve, sensitivity, specificity, positive predictive value, and negative predictive value of UIP and NSIP indicators were 0.896 and 0.759, 0.868 and 0.696, 0.952 and 0.838, 0.821 and 0.667, 0.949 and 0.824, 0.92 and 0.727, 0.881 and 0.778;the values of the training group are higher than those of the test group, as shown in Fig. 3(A-B) and Table 2 .

Table 2
Evaluation of logistic model in the training and testing samples

Item	Train	Test
Accuracy	0.896	0.759
f1_score	0.868	0.696
AUC	0.952	0.838
Sensitivity	0.821	0.667
Specificity	0.949	0.824
positive prediction	0.92	0.727
negative prediction	0.881	0.778

4 Discussion

At present, the diagnosis and classification of interstitial pneumonia mainly rely on HRCT, and the treatment methods used for different classifications are not all the same. On the other hand, the inflammation and fibrosis of interstitial pneumonia may have the same performance on HRCT[8–9].For atypical interstitial pneumonia, HRCT seems powerless, so it usually requires pathological biopsy for further examination. However, according to reports in the literature, pathological biopsy will cause a series of complications and reduce the survival rate of patients. At the same time, pathological biopsy can only represent the structure of local lung lesions, and cannot fully understand the condition of lung lesions. Therefore, using deep learning for automatic lung segmentation and obtaining high-level texture parameters of the lungs can effectively avoid doctors' subjective interference factors, obtain more comprehensive information about interstitial pneumonia, and achieve more accurate diagnosis.

Pang et al. used imaging omics to segment the HRCT of ILD patients and found that it has higher accuracy and overall performance than conventional methods. The segmentation results prove the necessity of denoising and the practicality of radiological features for segmentation[9].From this, we think of using radiomics to explore the ability to distinguish between UIP and NSIP. Through this study, for the first time, we used LK software to segment the lungs of all patients with interstitial pneumonia, and extracted lung texture features, and established logistic regression to distinguish UIP from NSIP based on these high-level features. Through the ROC curve of the logistic regression model, it can be seen that the classification model of radiomics shows high diagnostic efficiency,and the AUC values of the training group and the test group are 0.952 and 0.838, respectively.

The F1 value is a new machine learning indicator. The F1 value is the weighted average of precision and recall. The best F1 value is 1 and the worst is 0. In this study, the F1 values of the training group and the test group were 0.868 and 0.696, respectively, which showed the good diagnostic performance of this model.

The remaining eight features in this study reflect the gray value distribution, texture feature and spatial distribution of UIP and NSIP patient images, respectively. Among them, "original_gldm_DependenceVariance", "original_gldm_DependenceNonUniformity", and "original_shape_Maximum2DDiameterColumn" reflect the apparent texture thickness, shape, gray scale information of the lung. "Original_glrlm_GrayLevelVariance" reflects the variance of Gray Level in gray level co-occurrence matrix. "Original_glszm_ZonePercentage", "original_glszm_SizeZoneNonUniformityNormalized", "original_glszm_GrayLevelNonUniformityNormalized", and "original_glcm_DifferenceVariance" reflect the complexity of the lesion site, the degree of change in the level of complexity, and the level of change. These indicate that the texture characteristics of the lungs are different in UIP and NSIP patients.

There are some limitations in this study: Firstly, it is a single-center, small-sample study. The results should be verified by multi-center large sample data. Secondly, the grouping of UIP and NSIP is partly based on clinical diagnosis. This method should be feasible, because the clinical manifestations of UIP and NSIP in the middle and late stages are quite different. Thirdly, a machine learning model was

established to distinguish two types in interstitial pneumonia, and deep learning algorithms may provide better discrimination performance, which will be researched in the future.Fourthly, the CT images obtained in different devices may affect the image analysis.Fifthly, Most of the patients with interstitial lung disease are elderly and accompanying pulmonary edema, infection, or emphysema can be seen. I think these findings may affect the results.

In general, this article investigates the potential of lung texture to distinguish between UIP and NSIP, which can bring benefits to the diagnosis and treatment of patients with interstitial pneumonia

Abbreviations

UIP：usual interstitial pneumonia

NSIP：non-specific interstitial pneumonia

ROC：Receiver operating characteristic

AUC：area under the curve

LK：Lung Intelligence Kit

Declarations

Ethics approval and consent to participate

The Institutional Review Board of Guilin medical university Hospital approved this retrospective study and waived the requirement for written informed consent due to its retrospective nature.All methods were performed in accordance with the relevant guidelines and regulations.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed in this study are available from the corresponding author on request.

Competing interests

The authors declare that they have no competing interests.

Funding

Research and Application of Key Core Technologies in the Pharmaceutical and Biological Products Industry (20190202-2).

Authors’ contributions

Xinhui Chen, Ge Chengand Zhipeng Zhou conceived and designed this study. Xinhui Chen xinguan Yang, and Yuting Liao conducted the study and collected important background data. Xinhui Chen and Yuting Liao drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgments

I sincerely thank GE Health Care for its statistical assistance. The publication of this article may be related to GE's products.

Contributor Information

Xinhui Chen,Email:381905274@qq.com

Ge Cheng,Email:chargo_cheng@163.com

Xinguan Yang,Email:yang15007739374@163.com

Yuting Liao,Email:Yuting.liao@ge.com

Zhipeng Zhou E-mail bigbird_zhou@hotmail.com

References

1. Distefano G,Fanzone L,Palermo M,et al (2020) HRCT Patterns of Drug-Induced Interstitial Lung Diseases: A Review.Diagnostics 10: 244.
2. Fragoulis GE, Nikiphorou E, Larsen J, Korsten P, Conway R **(**2019**) Methotrexate-Associated Pneumonitis and Rheumatoid Arthritis-Interstitial Lung Disease: Current Concepts for the Diagnosis and Treatment**. Front Med (Lausanne) 6: 238.
3. Hutchinson JP, Fogarty AW, McKeever TM, Hubbard RB (2016) In-Hospital Mortality after Surgical Lung Biopsy for Interstitial Lung Disease in the United States. 2000 to 2011. Am J Respir Crit Care Med 193: 1161–1167.
4. Hassani C, Varghese BA, Nieva J, Duddalwar V (2019) Radiomics in Pulmonary Lesion Imaging.. AJR Am J Roentgenol 212: 497–504.

5. Chaddad A, Kucharczyk MJ, Daniel P, et al (2019) Radiomics in Glioblastoma: Current Status and Challenges Facing Clinical Implementation.. Front Oncol 9: 374.
6. Colen RR, Fujii T, Bilen MA, et al (2018) Radiomics to predict immunotherapy-induced pneumonitis: proof of concept.. Invest New Drugs 36: 601–607.
7. Zare Mehrjardi M, Kahkouee S, Pourabdollah M. (2017) Radio-Pathological Correlation of Organising Pneumonia (OP): A Pictorial Review.. Br J Radiol 90: 20160723.
8. Maher TM, Wuyts W (2019) Management of Fibrosing Interstitial Lung Diseases. Adv Ther 36: 1518–1531.
9. Pang T, Guo S, Zhang X, Zhao L (2019) Automatic Lung Segmentation Based on Texture and Deep Features of HRCT Images with Interstitial Lung Disease.. Biomed Res Int 2019: 1–8.

Figures

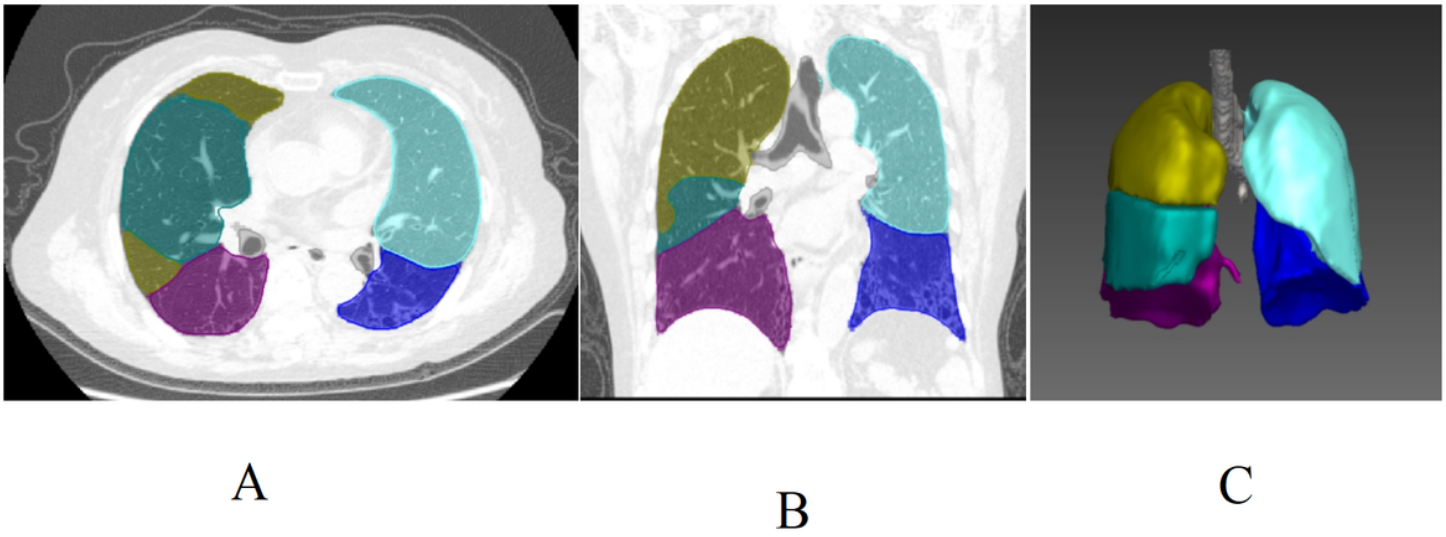

Figure 1

Illustration of the two lungs after being segmented by LK software. (A) the cross section; (B) the coronal plane; (C) Three-dimensional segmentations.

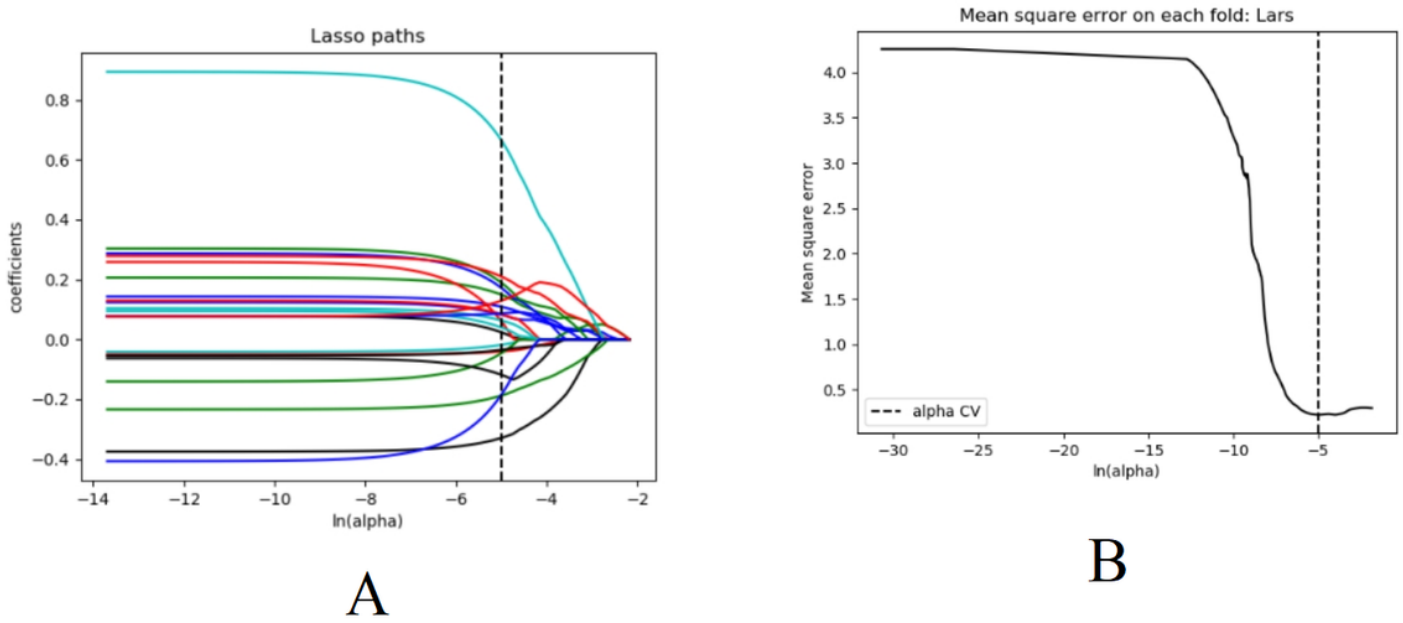

Figure 2

LASSO dimension reduction diagram. (A) the coefficients of each features changes with parameter alpha in Lasso method; (B) the mean square error (MSE) changes with parameters alpha in Lasso method, and the alpha with minimal MSE were selected.

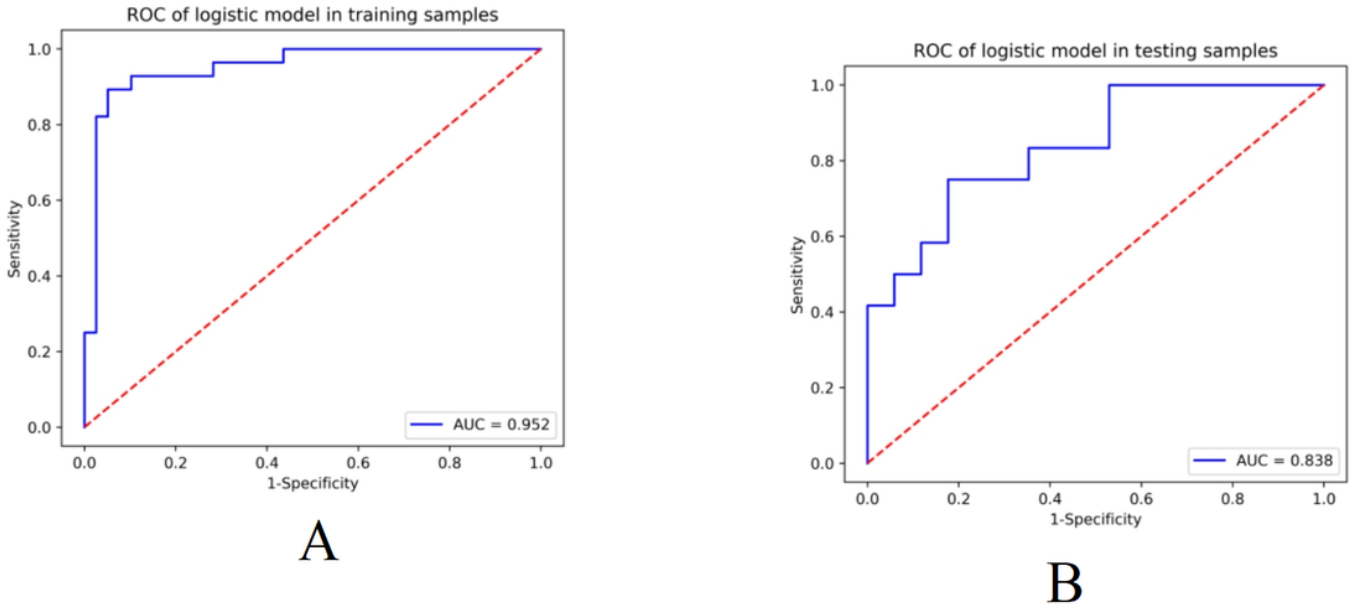

Figure 3

The ROC Curve. (A)the ROC curve of training samples; (B) the ROC curve testing samples