Wavelet transformation can enhance computed tomography texture features: a multicenter radiomics study for grade assessment of COVID-19 pulmonary lesions

Zekun Jiang[1,2,3#^], Jin Yin[1#], Peilun Han[1,2,3], Nan Chen[4], Qingbo Kang[1,2,3], Yue Qiu[1,2,3], Yiyue Li[1,2,3], Qicheng Lao[1,2,3,5], Miao Sun[1,6], Dan Yang[7], Shan Huang[7], Jiajun Qiu[1^], Kang Li[1,2,3,5,8^]

[1]West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu, China; [2]Med-X Center for Informatics, Sichuan University, Chengdu, China; [3]West China Hospital-Sense Time Joint Laboratory, Chengdu, China; [4]Department of Thoracic Surgery, West China Hospital, Sichuan University, Chengdu, China; [5]Shanghai Artificial Intelligence Laboratory, Shanghai, China; [6]College of Electrical Engineering, Sichuan University, Chengdu, China; [7]Department of Radiology, West China Hospital, Sichuan University, Chengdu, China; [8]Sichuan University-Pittsburgh Institute, Sichuan University, Chengdu, China

Contributions: (I) Conception and design: Z Jiang, J Yin, J Qiu; (II) Administrative support: K Li, J Qiu, M Sun; (III) Provision of study materials or patients: J Yin, J Qiu, P Han, N Chen, D Yang; (IV) Collection and assembly of data: J Yin, J Qiu; (V) Data analysis and interpretation: Z Jiang, J Yin, J Qiu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Kang Li; Jiajun Qiu. West China Biomedical Big Data Center, West China Hospital, Sichuan University, No. 37 Guoxue Alley, Wuhou District, Chengdu 610000, China. Email: likang@wchscu.cn; qiujiajun@wchscu.cn.

Background: This study set out to develop a computed tomography (CT)-based wavelet transforming radiomics approach for grading pulmonary lesions caused by COVID-19 and to validate it using real-world data.

Methods: This retrospective study analyzed 111 patients with 187 pulmonary lesions from 16 hospitals; all patients had confirmed COVID-19 and underwent non-contrast chest CT. Data were divided into a training cohort (72 patients with 127 lesions from nine hospitals) and an independent test cohort (39 patients with 60 lesions from seven hospitals) according to the hospital in which the CT was performed. In all, 73 texture features were extracted from manually delineated lesion volumes, and 23 three-dimensional (3D) wavelets with eight decomposition modes were implemented to compare and validate the value of wavelet transformation for grade assessment. Finally, the optimal machine learning pipeline, valuable radiomic features, and final radiomic models were determined. The area under the receiver operating characteristic (ROC) curve (AUC), calibration curve, and decision curve were used to determine the diagnostic performance and clinical utility of the models.

Results: Of the 187 lesions, 108 (57.75%) were diagnosed as mild lesions and 79 (42.25%) as moderate/severe lesions. All selected radiomic features showed significant correlations with the grade of COVID-19 pulmonary lesions ($P<0.05$). Biorthogonal 1.1 (bior1.1) LLL was determined as the optimal wavelet transform mode. The wavelet transforming radiomic model had an AUC of 0.910 in the test cohort, outperforming the original radiomic model (AUC =0.880; $P<0.05$). Decision analysis showed the radiomic model could add a net benefit at any given threshold of probability.

Conclusions: Wavelet transformation can enhance CT texture features. Wavelet transforming radiomics based on CT images can be used to effectively assess the grade of pulmonary lesions caused by COVID-19,

---

^ ORCID: Zekun Jiang, 0000-0002-3178-7761; Kang Li, 0000-0002-8136-9816; Jiajun Qiu, 0000-0002-0972-9338.

which may facilitate individualized management of patients with this disease.

Introduction

In early 2020, the COVID-19 outbreak spread rapidly around the globe. Many patients who became infected with this novel disease developed pneumonia and progressed rapidly into severe acute respiratory failure, with a poor prognosis and high mortality (1,2). The World Health Organization reported that, as of March 17, 2022, there had been 460.3 million confirmed cases of COVID-19, with more than 6 million confirmed deaths (3). Patients with COVID-19 may be at increased risk of developing conditions that can seriously threaten their health and lives, including mental illness (4). Therefore, the early diagnosis and severity assessment of COVID-19 continue to be critical in responding to the global pandemic.

Although reverse transcription-polymerase chain reaction (RT-PCR) remains the gold standard for COVID-19 diagnosis, it is associated with some practical problems. First, the sensitivity of RT-PCR tests is lacking because laboratory errors or low viral loads in test specimens can affect the result (5). Moreover, RT-PCR test kits are in short supply in some developing countries (6). Consequently, in some countries, chest X-ray (CXR) and chest computed tomography (CT) are used as first-line investigation and patient management tools (7,8). In particular, chest CT can show clear early lesions and can achieve high sensitivity if the diagnosing radiologist is experienced. Therefore, chest CT plays an important role in diagnosing COVID-19-positive cases and confirming the severity of pneumonia, and in many epidemic areas, it could even be considered an essential tool (6).

For faster and more accurate examination, many artificial intelligence (AI) techniques for automated detection and quantitative analysis of COVID-19 lesions from CT images have been developed based on deep learning and radiomics (9-18). In addition to detecting lesions, assessing the grade of COVID-19 pulmonary lesions is important for the hierarchical management and treatment of infected patients (19). However, few studies have focused on the use of AI in grading pulmonary lesions (20), and those that exist

have only used "pure" datasets accrued from unified vendor scanners (21,22), which limits the real generalizability of their AI models (23). Thus, it is necessary to develop a more robust AI grading tool based on real-world multicenter datasets.

Radiomics, as a relatively mature medical image analysis technology, can not only be used to build prediction models with a high diagnostic performance, but also to mine valuable imaging features which can provide guidance for clinical practice (24-26). Some previous radiomics studies have suggested that wavelet transformation could be helpful for radiomics analysis (27-29). However, to date, no study has evaluated the effects of various wavelets on radiomic features and models.

In the present real-world multicenter study, we used data from 16 hospitals with 14 different imaging platforms to explore and compare 23 three-dimensional (3D) wavelets and develop an intelligent classifier for the grade assessment of COVID-19 pulmonary lesions based on CT images. It is hoped that the radiomic model developed in this study will aid in reducing radiologists' workloads and cases of misdiagnosis, while improving the accuracy of diagnosis in patients with COVID-19 infection. We present the following article in accordance with the TRIPOD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-22-252/rc).

Methods

Ethics and study design

This retrospective study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of West China Hospital of Sichuan University (No. 2020190), with the need for individual consent waived due to the retrospective nature of the analysis.

Wavelet transforming radiomic models were designed to quickly and accurately assess the grade of COVID-19 pulmonary lesions. In all, 111 patients with 187 pulmonary
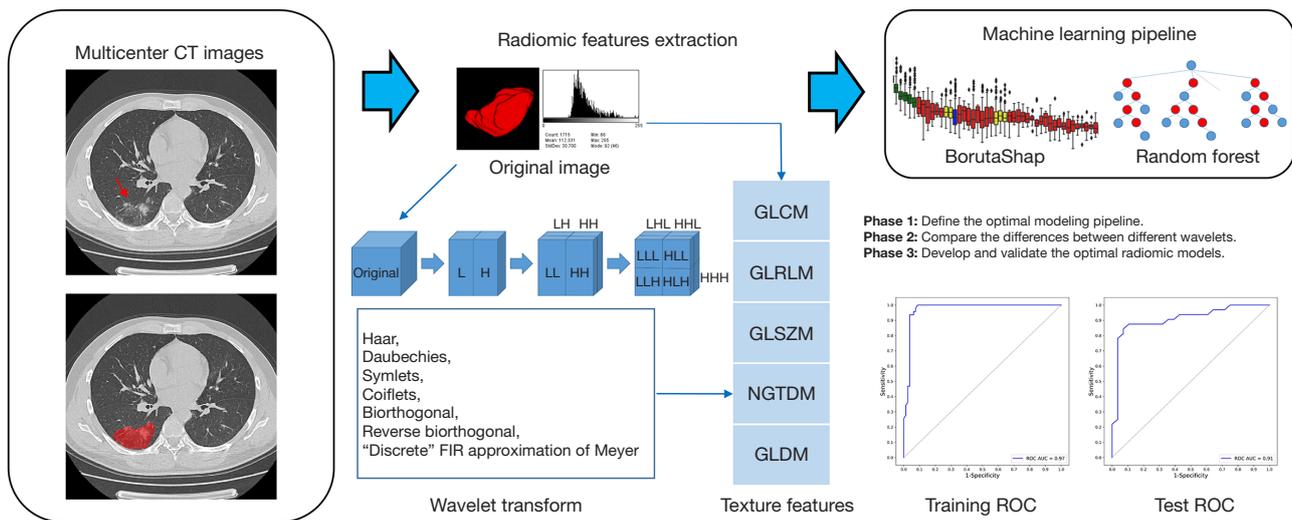
Figure 1 Overview of the study workflow showing the process from CT images to radiomics features to machine learning models. The red arrow indicates the location of the lesion. CT, computed tomography; H, high-pass decomposition filter; L, low-pass decomposition filter; FIR, finite impulse response; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run length matrix; GLSZM, gray-level size zone matrix; NGTDM, neighboring gray tone difference matrix; GLDM, gray-level dependence matrix; ROC, receiver operating characteristic; AUC, area under the receiver operating characteristic curve.

lesions were analyzed retrospectively in this study. Seventy-three quantitative texture features were extracted from the volume of interest (VOI) of the pulmonary lesion based on the original and 23 wavelet-transformed CT images. First, we determined the optimal machine learning pipeline by comparing multiple feature selection algorithms and modeling methods. Then, we selected the valuable features from each image mode. In all, 184 radiomic models of 23 wavelets with eight decomposition modes were built and compared by using the area under the receiver operating characteristic (ROC) curve (AUC) and DeLong test. The final radiomic model was developed using the training cohort and evaluated using the test cohort (see below). Radiomic feature analysis and feature map analysis were also implemented. The study workflow is shown in *Figure 1*.

Study population and grade assessment

Data were collected from 174 patients with RT-PCR-confirmed COVID-19 from 16 hospitals in Sichuan, China, between January 2020 and March 2020. All patients underwent non-contrast chest CT scans. The exclusion criteria were as follows: (I) insufficient CT image quality (n=42); and (II) uncertain lesion grade (n=21). This left 187 pulmonary lesions of 111 patients for analysis. *Figure 2*

shows a flowchart of the inclusion and exclusion process for the patients in this study.

The study data were divided into two cohorts according to the hospital in which the patients underwent CT scanning: (I) a training cohort (72 patients with 127 lesions from nine hospitals); and (II) a test cohort (39 patients with 60 lesions from seven hospitals).

Grade assessment of chest CT images was performed by a radiologist D.Y. with >10 years of experience. Briefly, lesions with scattered ground-glass nodules (GGOs) were graded as mild, whereas high-density lesions with continuous or large areas of GGOs were graded as moderate or severe. Lesions with only continuous GGOs were graded as moderate, whereas those with continuous GGOs, regional texture smoothness, and high CT values were graded as severe (8,10,30). Based on these criteria, 108 lesions were classified as mild and 79 were classified as moderate or severe.

The clinical characteristics of the study population as a whole and according to lesion grade are provided in *Table 1*.

Acquisition and segmentation of CT images

As detailed in Table S1, CT imaging was performed using different scanners according to manufacturers' instructions. The fact that the CT images were obtained with a variety
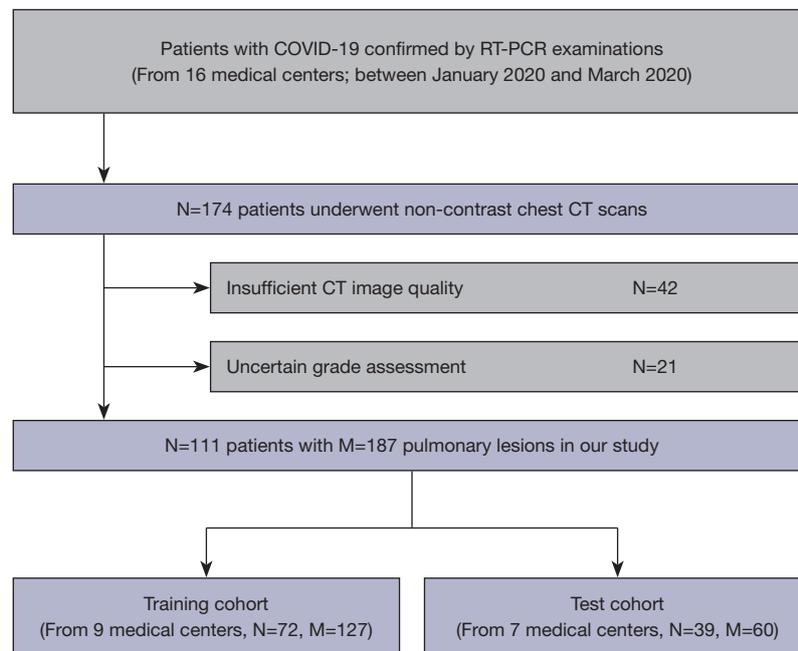
Figure 2 Flowchart showing the application of inclusion and exclusion criteria to the study population. RT-PCR, reverse transcription-polymerase chain reaction; CT, computed tomography; N, the number of patients; M, the number of lesions.

Table 1 Clinical characteristics of the study population

Characteristics	All lesions (n=187)	Mild lesions (n=108)	Moderate/severe lesions (n=79)	P value
Age (years), mean ± SD	44.88±12.15	46.51±12.30	42.65±11.58	0.030
Sex, n (%)				0.112
Male	112 (59.89)	70 (64.81)	42 (53.16)	
Female	75 (40.11)	38 (35.19)	37 (46.84)	
Cohort, n (%)				0.039
Training	127 (67.91)	80 (74.07)	47 (59.50)	
Test	60 (32.09)	28 (25.93)	32 (40.50)	

SD, standard deviation.

of CT scanners reflects the real-world, multicenter design of this study. All CT scans were acquired from the Picture Archiving and Communication System for further processing.

The VOIs of pulmonary lesions were segmented by two radiologists P.H. and S.H. with >5 years of experience each. A third radiologist D.Y. with >10 years of experience in chest CT imaging checked the results and determined the final VOIs through discussion. The VOIs were positioned using ITK-SNAP (version 3.8.0; http://www.itksnap.org/pmwiki/pmwiki.php) (31). During the segmentation

process, the lesion grade was re-evaluated. If a disagreement arose, the final result was determined through discussion.

Radiomic features and wavelet transform

Radiomic features were defined using Python PyRadiomics version 3.0.1 (https://github.com/AIM-Harvard/pyradiomics) (32), with most of them adhering to the definitions of the Imaging Biomarker Standardization Initiative (33). The preprocessing settings were as follows: the interpolator was siktBSpline, the resampled pixel

Table 2 Details of the wavelets used in this study

Wavelets	Abbreviation	N[A]	Decomposition mode
Haar	haar	N/A	LLH, LHL, LHH, HLL, HLH, HHL, HHH, LLL
Daubechies N	dbN	1, 10, 20	
Symlets N	symN	2, 10, 20	
Coiflets N	coifN	1, 2, 3, 4, 5	
Biorthogonal Nr.Nd	bior Nr.Nd	1.1, 2.2, 3.3, 4.4, 5.5	
Reverse biorthogonal Nr.Nd	rbio Nr.Nd	1.1, 2.2, 3.3, 4.4, 5.5	
"Discrete" FIR approximation of Meyer	dmey	N/A	

[A], decomposition progression of wavelets. "N" in dbN, symN, and coifN refers to the number of vanishing moments. In biorNr.Nd and rbioNr.Nd, "Nr" is the number of the order of the functions used for reconstruction and "Nd" is the order of the functions used for decomposition (36). FIR, finite impulse response; N/A, not applicable; L, low-pass decomposition filter; H, high-pass decomposition filter.

spacing was [1, 1, 1], the bin size was 25, and the minimum value in Hounsfield units was –1,000, shift +1,000. Seventy-three texture features were extracted from each VOI in the original CT image, including 22 gray-level co-occurrence matrix (GLCM) features, 16 gray-level run length matrix (GLRLM) features, 16 gray-level size zone matrix (GLSZM) features, 14 gray-level dependence matrix (GLDM) features, and 5 neighboring gray tone difference matrix (NGTDM) features. The details of the extracted texture features are provided in Table S2.

To determine the effect of wavelet transformation on texture features, 23 3D wavelet transform algorithms were used to decompose the original image into eight parts (LLH, LHL, LHH, HLL, HLH, HHL, HHH, and LLL; where L refers to a low-pass decomposition filter and H refers to a high-pass decomposition filter) (34,35). The same texture features were then extracted from each VOI in the translated images. *Table 2* provides details of all the wavelets implemented in our study.

Machine learning pipeline

We defined an optimal pipeline for feature selection and modeling through comparisons of different machine learning pipelines performed on the basis of the original texture features. The performance of different machine learning pipelines is summarized in *Table 3*. The optimal pipeline was set as follows: first, data normalization was implemented, and then the best subset of features was selected by performing a BorutaShap algorithm (37,38), which was a wrapper method combining the Boruta algorithm (39) with SHapley Additive exPlanations (SHAP) (40). Based on the selected features, a random forest model was built in the training cohort with 10-fold cross-validation and then tested in the test cohort.

Comparison of different wavelets

Based on the machine learning pipeline, we determined the valuable texture features in the original image before selecting the same texture features from all the wavelet-transformed images (23 wavelets with eight decomposition modes). Finally, all selected feature groups were modeled and compared. Based on the results of this comparison, the optimal wavelet transforming radiomic model was determined.

Radiomic model and statistical analysis

The final radiomic model was developed in the training cohort and evaluated in the test cohort. To facilitate the clinical use of our model, receiver operator characteristic, calibration, and decision curves were used to evaluate its diagnostic performance and clinical utility.

Statistical analysis was mainly performed using R (version 3.5.3; https://www.r-project.org/). All the machine learning and image processing algorithms were implemented in Python (version 3.7.11; https://www.python.org/). The chi-squared test was used to test the significance of differences in count variables, and the Mann-Whiney U test was used to test the significance of differences in continuous variables. Spearman's test was used to assess associations between different features. Interobserver variability of feature extraction was evaluated using intraclass correlation coefficients (ICCs). Differences in efficacy between different

Table 3 Performance of different machine learning pipelines

Machine learning pipeline	Training AUC	Cross-validation mean AUC	Test AUC
BorutaShap + RF*	0.98*	0.85*	0.88*
BorutaShap + SVM	0.99	0.82	0.84
BorutaShap + LR	0.96	0.82	0.83
BorutaShap + MLP	0.99	0.83	0.85
Boruta + RF	0.98	0.84	0.87
Boruta + SVM	0.97	0.84	0.86
Boruta + LR	0.95	0.83	0.84
Boruta + MLP	0.99	0.85	0.87
LASSO + RF	0.94	0.78	0.80
LASSO + SVM	0.94	0.78	0.80
LASSO + LR	0.98	0.84	0.86
LASSO + MLP	0.97	0.83	0.85
RFE + RF	0.97	0.84	0.86
RFE + SVM	0.96	0.85	0.87
RFE + LR	0.97	0.80	0.86
RFE + MLP	0.97	0.82	0.85

*, the best-performing pipeline. AUC, area under the receiver operating characteristic curve; RF, random forest; SVM, support vector machine; LR, logistic regression; MLP, multilayer perceptron; LASSO, least absolute shrinkage and selection operator; RFE, recursive feature elimination.

ROC curves were determined by using the DeLong test. Two-sided P<0.05 was considered statistically significant.

Results

Clinical characteristics

As shown in *Table 1*, in all, there were 108 (57.75%) mild lesions and 79 (42.25%) moderate or severe lesions. In the training cohort, 80 (74.07%) lesions were defined as mild lesions, and in the test cohort, 28 (25.93%) lesions were defined as mild lesions. In our dataset, age was significantly correlated with lesion grade (P=0.030). Due to the limitations associated with the retrospective nature of this study, many clinical characteristics were not included for analysis.

Optimal machine learning pipeline and selected features

As shown in *Table 3*, BorutaShap + random forest had a cross-validation mean AUC of 0.85 and a test AUC of 0.88, which were higher than the values for other machine learning pipelines. An example of BorutaShap selection in an original image is shown in Figure S1. *Table 4* lists the final selected radiomic features, all of which showed good reproducibility (ICC >0.75). All the features were significantly correlated with the severity grading of COVID-19 pulmonary lesions (P<0.05), with features F1–F6 having correlations with extremely high significance (P<0.001). Of all the features examined, the GLCM features contributed maximum numbers (~44.44%).

Comparison of wavelet transform models

The same radiomic features were selected from all wavelet-transformed images (23 wavelets with eight decomposition modes; 184 image modes in total). Figure S2 shows the main results of the comparisons. Of all the wavelet transforming radiomic models, biorthogonal 1.1 (bior1.1) showed the best performance, and the LLL decomposition mode generally had the highest diagnostic performance among all the image modes. As shown in *Figure 3*, the

Table 4 Radiomic features selected in original images

Feature	Feature type	Feature value	Correlation[A]	P value
F1	GLRLM	Long run high gray-level emphasis	0.441	<0.001
F2	GLCM	Autocorrelation	0.408	<0.001
F3	GLCM	Joint averages	0.414	<0.001
F4	GLCM	Cluster shade	−0.476	<0.001
F5	GLDM	Large dependence high gray-level emphasis	0.395	<0.001
F6	NGTDM	Busyness	−0.345	<0.001
F7	GLDM	Gray-level non-uniformity	−0.227	0.010
F8	GLSZM	Gray-level non-uniformity	−0.207	0.020
F9	GLCM	Correlation	0.168	0.042

[A], Spearman correlation coefficient with two-sided test. GLRLM, gray-level run length matrix; GLCM, gray-level co-occurrence matrix; GLDM, gray-level dependence matrix; NGTDM, neighboring gray tone difference matrix; GLSZM, gray-level size zone matrix.
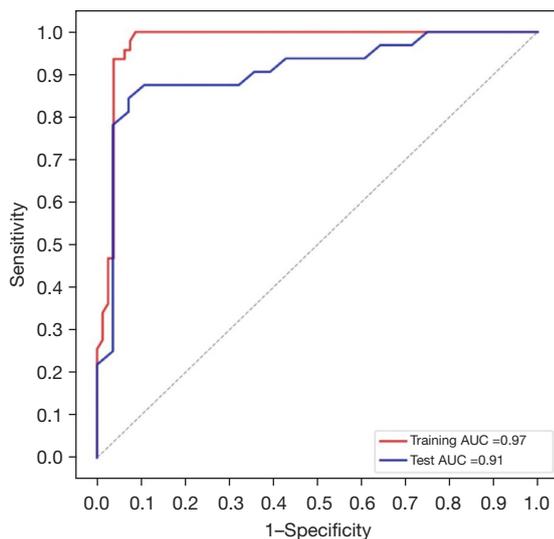

Figure 3 Training and test ROC curves of the bior1.1 LLL wavelet transforming radiomic model. AUC, area under the receiver operating characteristic curve; bior, biorthogonal; L, low-pass decomposition filter; ROC, receiver operating characteristic.

bior1.1 LLL model had AUCs of 0.97 and 0.91 in the training and test cohorts, respectively; thus, it performed significantly better than the original radiomic model (AUC =0.880; P<0.05).

Overall performance of the final radiomic model

Subsequently, the bior1.1 LLL model was chosen as the optimal radiomic classifier, and the feature maps were

calculated. A detailed performance evaluation of the radiomic model is provided in *Table 5*. In the training cohort, the radiomic model had a macro average precision of 0.93, with a sensitivity of 93.75% and a specificity of 93.62%. In the test cohort, the model had a macro average precision of 0.84, with a sensitivity of 96.43% and a specificity of 68.75%.

The calibration and decision curves in the training and test cohorts are shown in *Figure 4*. The calibration curves showed that the mean absolute error of the radiomic model was 0.012 and 0.030 in the training and test cohorts, respectively. Through decision analysis, we found that our radiomic model could add more benefit than could a treat-all or treat-none scheme at any given threshold of probability. Two examples of chest CT diagnosis using the radiomic model with feature maps are shown in *Figure 5*; these examples demonstrate that visualization of radiomic features can aid in clinical decision making.

Discussion

In clinical practice, grade assessment of COVID-19 pulmonary lesions (mild or moderate/severe) is of considerable importance to the further diagnosis of severity and treatment of patients. Intelligent CT-based diagnostic tools may help to overcome the subjectivity and inconsistencies associated with physicians' assessments, thus supporting the use of precision treatment, especially in COVID-19 epidemic areas with limited capacity for diagnosis. Herein, we have reported on the development and validation of a CT-based wavelet transforming

Table 5 Detailed performance evaluation of the radiomic model

Index	Training cohort				Test cohort			
	Precision	Recall	F1	Support	Precision	Recall	F1	Support
Mild lesions	0.96	0.94	0.95	80	0.73	0.96	0.83	28
Moderate/severe lesions	0.90	0.94	0.92	47	0.96	0.69	0.80	32
Accuracy	N/A	N/A	0.94	127	N/A	N/A	0.82	60
Macro average	0.93	0.94	0.93	127	0.84	0.83	0.82	60
Weighted average	0.94	0.94	0.94	127	0.85	0.82	0.81	60

N/A, not applicable.
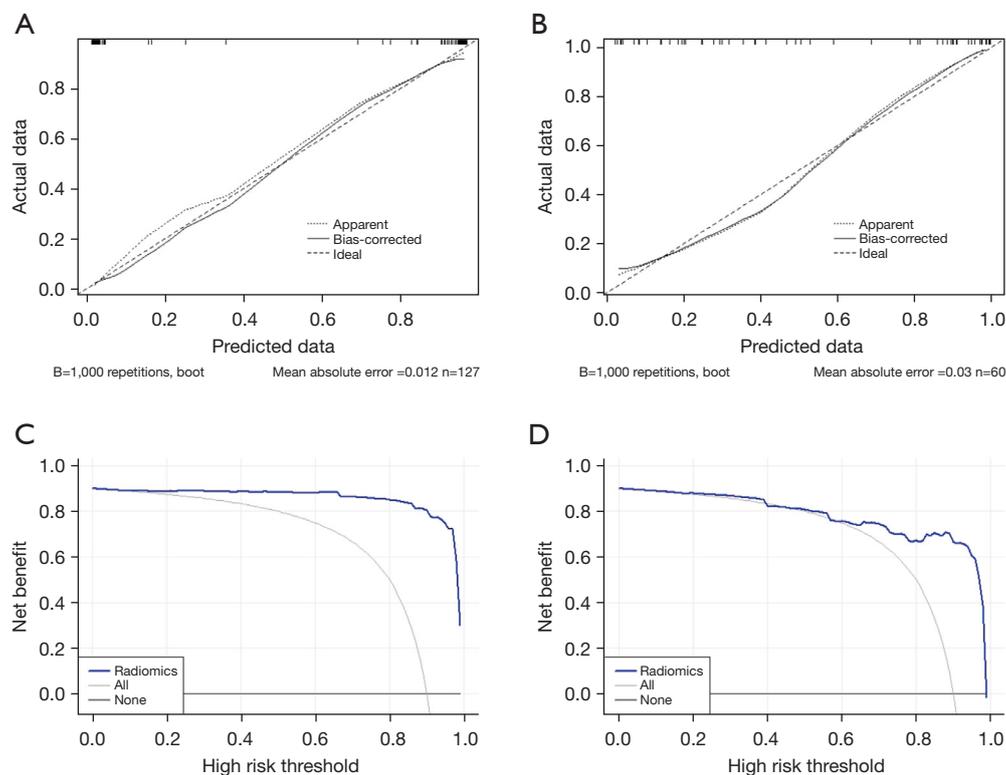
Figure 4 Calibration (A,B) and decision (C,D) curves for our radiomic model in the training (A,C) and test (B,D) cohorts. Apparent curves refer to in-sample calibration of the model, ideal curves refer to the perfect prediction model, and bias-corrected curves show results of overfitting calibrated by bootstrap sampling, where B is the numbers of bootstraps. For decision curves (C,D), gray lines represent the hypothesis that all lesions were moderate or severe and black lines represent the hypothesis that all lesions were mild.

radiomics nomogram based on real-world multicenter data that had a high and robust diagnostic performance (AUC =97.3% in the training cohort; AUC =92.1% in the test cohort).

With the global COVID-19 epidemic, there have been many AI-based COVID-19 diagnostic studies (41,42).

However, most COVID-19-related medical imaging AI technologies focus on object detection, assisting diagnosis, and predicting progress (20). Compared with previously published models to assess the severity of COVID-19 (21,22,43), our radiomics nomogram showed a very high predictive performance and robustness.
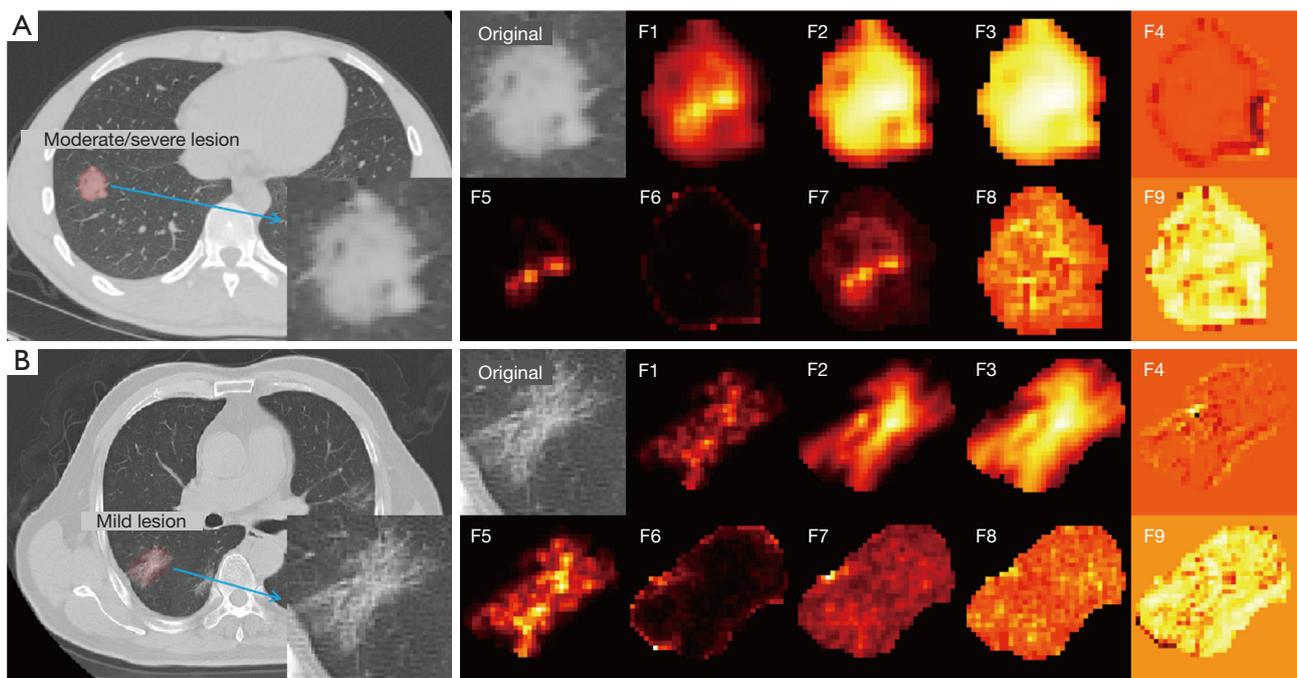
4766

Jiang et al. Wavelet radiomics for COVID-19 grade assessment

Figure 5 CT scans of typical COVID-19 lesions with radiomic feature maps showing the diagnostic performance of our radiomic model. (A) Results for a 49-year-old man with a moderate/severe lesion, showing a prediction probability of 0.98 by using the radiomic model. (B) Results for a 64-year-old man with a mild lesion, showing a prediction probability of 0.08. The details of radiomic features F1–F9 are presented in *Table 4*. Hot color mapping can be used for better visualization. The insets on the CT scans to which arrows point are the regions of interest in the lesion in the original image; the different radiomics feature maps are shown to the right. CT, computed tomography.

Moreover, in addition to building the diagnostic model, this study has also found some interesting insights into CT radiomics.

The realization and promotion of AI are seriously affected by the multivendor nature of real-world data, which contain confounding and discrepant information (23). Under this circumstance, the training of deep learning models is often difficult, resulting in underperformance, and the features of engineering-based radiomics may show a superior performance (44,45). Also, advanced feature selection algorithms can help solve the issue of dimensionality (46). In the present study, we used BorutaShap, which combines Boruta built based on a random forest model with Shapley values (37), and found that this could provide a better features subset and the most accurate global feature rankings. In addition, SHAP improves the interpretability of machine learning (47). As the results of the present study show, BorutaShap is suitable for analyzing and processing complex heterogeneous biomedical data.

Previous studies have shown the value of texture features in the imaging diagnosis of inflammation and tumors (48-50). Some studies have also suggested that wavelet transform may increase the value of texture features (27,51), but specific clinical research is lacking. In the present study, we found that wavelet transforming radiomics showed a better performance than did original radiomics (AUC: 0.921 *vs.* 0.880, respectively; P<0.05). These results indicate that wavelet transform may amplify the heterogeneity information of texture features in medical images to some extent. Zhou *et al.* similarly found that wavelet-transformed textures outperformed original textures in magnetic resonance imaging (28), and Chaddad *et al.* showed that multiscale texture features based on 3D wavelet transform were more sensitive in discriminating colorectal cancer grades (29). Together, these findings provide support for further discussions of the effects of wavelet transform on texture features. Moreover, we found that haar, db1, sym2, coif1, coif2, and bior1.1 wavelets were more valuable in our dataset (see Figure S2). Wavelet transform can

decompose image signals by using low- and high-pass filters. However, it is not usually known which filter is better for amplifying critical information. In the present study, the LLL decomposition modes of 3D wavelets all showed a better diagnostic performance, which may provide reference to other studies. To demonstrate the effects of different decomposition modes on radiomic features, taking GLCM autocorrelation (feature F2) as an example, we provide feature maps in eight decomposition modes in Figure S3; the feature map shows more hierarchical changes and information content in the LLL mode. Radiomic feature maps assist clinical decision making in the form of secondary imaging, which may help to accelerate the promotion and application of radiomics and other AI algorithms in medicine (52,53).

The radiomics quality score (RQS) is an important tool for evaluating the quality of radiomics research (54), and we advocate that every radiomics study should use the RQS for self-examination. The highest possible RQS is 36, and the RQS of the present study was 20 (checkpoint 1: +1; checkpoint 2: +1; checkpoint 3: +18), which is better than most radiologic studies (55).

In our dataset, the mean ± standard deviation (SD) age of patients with mild lesions was significantly higher than that of patients with moderate or severe lesions (46.51±12.3 *vs.* 42.65±11.58 years, respectively; P=0.030). We do not believe that this observation is clinically important, but simply a trend in our dataset. Moreover, it does not affect the main points of this study, because our radiomic model was constructed purely based on CT images, and clinical factors were not incorporated into the model.

This study has some limitations. First, as a retrospective study, related potential bias is inevitable and certain clinical information was not available for inclusion in the study. A prospective study will be able to provide more convincing evidence of the utility of our radiomic model in grading lesions. Second, the segmentation of pulmonary lesions in our study was not fine enough. However, we validated the feasibility of using radiomics based on rough segmentation. Third, 16 hospitals in the study are located in Sichuan province, China. We do not know whether there are regional differences that could affect the broader application of our results and the use of our radiomic model in other populations. Thus, a larger multiregional prospective study is needed to verify our findings. Fourth, due to different research purposes, we did not explore more advanced AI algorithms.

Conclusions

In conclusion, this study has demonstrated that CT-based wavelet transforming radiomics outperformed original radiomics in the grade assessment of COVID-19 pulmonary lesions, and showed high accuracy and robustness in a multicenter validation. Therefore, our radiomic model may be used as a diagnostic tool to help with efficient clinical diagnosis and decision making for patients with COVID-19.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-22-252/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-22-252/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. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The work was approved by the Ethics Committee of West China Hospital of Sichuan University (No. 2020190), with the need for individual consent waived due to the retrospective nature of the analysis.

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

4768

Jiang et al. Wavelet radiomics for COVID-19 grade assessment

References

1. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med 2020;382:1708-20.

2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395:497-506.

3. Coronavirus disease (COVID-19). Available online: https://www.who.int/emergencies/diseases/novel-coronavirus-2019

4. COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021;398:1700-12.

5. Bai HX, Hsieh B, Xiong Z, Halsey K, Choi JW, Tran TML, Pan I, Shi LB, Wang DC, Mei J, Jiang XL, Zeng QH, Egglin TK, Hu PF, Agarwal S, Xie FF, Li S, Healey T, Atalay MK, Liao WH. Performance of Radiologists in Differentiating COVID-19 from Non-COVID-19 Viral Pneumonia at Chest CT. Radiology 2020;296:E46-54.

6. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology 2020;296:E32-40.

7. Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, et al. The Role of Chest Imaging in Patient Management During the COVID-19 Pandemic: A Multinational Consensus Statement From the Fleischner Society. Chest 2020;158:106-16.

8. Wong HYF, Lam HYS, Fong AH, Leung ST, Chin TW, Lo CSY, Lui MM, Lee JCY, Chiu KW, Chung TW, Lee EYP, Wan EYF, Hung IFN, Lam TPW, Kuo MD, Ng MY. Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19. Radiology 2020;296:E72-8.

9. Wang G, Liu X, Shen J, Wang C, Li Z, Ye L, et al. A deep-learning pipeline for the diagnosis and discrimination of viral, non-viral and COVID-19 pneumonia from chest X-ray images. Nat Biomed Eng 2021;5:509-21.

10. Qiu J, Peng S, Yin J, Wang J, Jiang J, Li Z, Song H, Zhang W. A Radiomics Signature to Quantitatively Analyze COVID-19-Infected Pulmonary Lesions. Interdiscip Sci 2021;13:61-72.

11. Panwar H, Gupta PK, Siddiqui MK, Morales-Menendez R, Singh V. Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. Chaos Solitons Fractals 2020;138:109944.

12. Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun 2020;11:4080.

13. Ko H, Chung H, Kang WS, Kim KW, Shin Y, Kang SJ, Lee JH, Kim YJ, Kim NY, Jung H, Lee J. COVID-19 Pneumonia Diagnosis Using a Simple 2D Deep Learning Framework With a Single Chest CT Image: Model Development and Validation. J Med Internet Res 2020;22:e19569.

14. Quattrocchi CC, Mallio CA, Presti G, Beomonte Zobel B, Cardinale J, Iozzino M, Della Sala SW. The challenge of COVID-19 low disease prevalence for artificial intelligence models: report of 1,610 patients. Quant Imaging Med Surg 2020;10:1891-3.

15. Qiu JJ, Yin J, Qian W, Liu JH, Huang ZX, Yu HP, Ji L, Zeng XX. A Novel Multiresolution-Statistical Texture Analysis Architecture: Radiomics-Aided Diagnosis of PDAC Based on Plain CT Images. IEEE Trans Med Imaging 2021;40:12-25.

16. Saha M, Amin SB, Sharma A, Kumar TKS, Kalia RK. AI-driven quantification of ground glass opacities in lungs of COVID-19 patients using 3D computed tomography imaging. PLoS One 2022;17:e0263916.

17. Verma A, Amin SB, Naeem M, Saha M. Detecting COVID-19 from chest computed tomography scans using AI-driven android application. Comput Biol Med 2022;143:105298.

18. Di D, Shi F, Yan F, Xia L, Mo Z, Ding Z, Shan F, Song B, Li S, Wei Y, Shao Y, Han M, Gao Y, Sui H, Gao Y, Shen D. Hypergraph learning for identification of COVID-19 with CT imaging. Med Image Anal 2021;68:101910.

19. Yang R, Li X, Liu H, Zhen Y, Zhang X, Xiong Q, Luo Y, Gao C, Zeng W. Chest CT Severity Score: An Imaging Tool for Assessing Severe COVID-19. Radiol Cardiothorac Imaging 2020;2:e200047.

20. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, He K, Shi Y, Shen D. Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for COVID-19. IEEE Rev Biomed Eng 2021;14:4-15.

21. Cai W, Liu T, Xue X, Luo G, Wang X, Shen Y, Fang Q, Sheng J, Chen F, Liang T. CT Quantification and Machine-learning Models for Assessment of Disease Severity and Prognosis of COVID-19 Patients. Acad

Radiol 2020;27:1665-78.

22. Xu Z, Zhao L, Yang G, Ren Y, Wu J, Xia Y, Yang X, Cao M, Zhang G, Peng T, Zhao J, Yang H, Hu J, Du J. Severity Assessment of COVID-19 Using a CT-Based Radiomics Model. Stem Cells Int 2021;2021:2263469.

23. Doran SJ, Kumar S, Orton M, d'Arcy J, Kwaks F, O'Flynn E, Ahmed Z, Downey K, Dowsett M, Turner N, Messiou C, Koh DM. "Real-world" radiomics from multi-vendor MRI: an original retrospective study on the prediction of nodal status and disease survival in breast cancer, as an exemplar to promote discussion of the wider issues. Cancer Imaging 2021;21:37.

24. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, Allison T, Arnaout O, Abbosh C, Dunn IF, Mak RH, Tamimi RM, Tempany CM, Swanton C, Hoffmann U, Schwartz LH, Gillies RJ, Huang RY, Aerts HJWL. Artificial intelligence in cancer imaging: Clinical challenges and applications. CA Cancer J Clin 2019;69:127-57.

25. Song J, Yin Y, Wang H, Chang Z, Liu Z, Cui L. A review of original articles published in the emerging field of radiomics. Eur J Radiol 2020;127:108991.

26. Attanasio S, Forte SM, Restante G, Gabelloni M, Guglielmi G, Neri E. Artificial intelligence, radiomics and other horizons in body composition assessment. Quant Imaging Med Surg 2020;10:1650-60.

27. Jing R, Wang J, Li J, Wang X, Li B, Xue F, Shao G, Xue H. A wavelet features derived radiomics nomogram for prediction of malignant and benign early-stage lung nodules. Sci Rep 2021;11:22330.

28. Zhou J, Lu J, Gao C, Zeng J, Zhou C, Lai X, Cai W, Xu M. Predicting the response to neoadjuvant chemotherapy for breast cancer: wavelet transforming radiomics in MRI. BMC Cancer 2020;20:100.

29. Chaddad A, Daniel P, Niazi T. Radiomics Evaluation of Histological Heterogeneity Using Multiscale Textures Derived From 3D Wavelet Transformation of Multispectral Images. Front Oncol 2018;8:96.

30. Kazerooni EA, Gross BH. Cardiopulmonary imaging. Philadelphia: Lippincott Williams & Wilkins, 2004.

31. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31:1116-28.

32. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts HJWL. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res

2017;77:e104-7.

33. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology 2020;295:328-38.

34. Procházka A, Gráfová L, Vyšata O, et al. Three-dimensional wavelet transform in multi-dimensional biomedical volume processing. In: Proceedings of the IASTED International Conference on Graphics and Virtual Reality. Cambridge, 2011;263:268.

35. Bhattacharjee S, Kim CH, Park HG, Prakash D, Madusanka N, Cho NH, Choi HK. Multi-Features Classification of Prostate Carcinoma Observed in Histological Sections: Analysis of Wavelet-Based Texture and Colour Features. Cancers (Basel) 2019;11:1937.

36. Salem MA, Ghamry N, Meffert B. Daubechies versus biorthogonal wavelets for moving object detection in traffic monitoring systems. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Informatik, 2009. doi: 10.18452/2487.

37. Keany E. BorutaShap: A wrapper feature selection method which combines the Boruta feature selection algorithm with Shapley values. Zenodo 2020. Available online: https://zenodo.org/record/4247618#.YsKgh3ZBw2w

38. Silva IS, Ferreira CN, Costa LBX, Sóter MO, Carvalho LML, de C Albuquerque J, Sales MF, Candido AL, Reis FM, Veloso AA, Gomes KB. Polycystic ovary syndrome: clinical and laboratory variables related to new phenotypes using machine-learning models. J Endocrinol Invest 2022;45:497-505.

39. Lai WT, Deng WF, Xu SX, Zhao J, Xu D, Liu YH, Guo YY, Wang MB, He FS, Ye SW, Yang QF, Liu TB, Zhang YL, Wang S, Li MZ, Yang YJ, Xie XH, Rong H. Shotgun metagenomics reveals both taxonomic and tryptophan pathway differences of gut microbiota in major depressive disorder patients. Psychol Med 2021;51:90-101.

40. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, 2017;30.

41. Laino ME, Ammirabile A, Posa A, Cancian P, Shalaby S, Savevski V, Neri E. The Applications of Artificial Intelligence in Chest Imaging of COVID-19 Patients: A Literature Review. Diagnostics (Basel) 2021;11:1317.

42. Bouchareb Y, Moradi Khaniabadi P, Al Kindi F, Al Dhuhli H, Shiri I, Zaidi H, Rahmim A. Artificial intelligence-driven assessment of radiological images for COVID-19. Comput Biol Med 2021;136:104665.

4770

Jiang et al. Wavelet radiomics for COVID-19 grade assessment

43. Irmak E. COVID-19 disease severity assessment using CNN model. IET Image Process 2021;15:1814-24.

44. Parekh VS, Jacobs MA. Deep learning and radiomics in precision medicine. Expert Rev Precis Med Drug Dev 2019;4:59-72.

45. Mohammadi A, Afshar P, Asif A, Farahani K, Kirby J, Oikonomou A, Plataniotis KN. Lung Cancer Radiomics: Highlights from the IEEE Video and Image Processing Cup 2018 Student Competition. IEEE Signal Process Mag 2019;36:164-73.

46. Khaire UM, Dhanalakshmi R. Stability of feature selection algorithm: A review. Journal of King Saud University-Computer and Information Sciences 2022;34:1060-73.

47. Molnar C. Interpretable machine learning. 2022. Available online: https://christophm.github.io/interpretable-ml-book/

48. Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT Texture Analysis: Definitions, Applications, Biologic Correlates, and Challenges. Radiographics 2017;37:1483-503.

49. Alves AFF, Miranda JRA, Reis F, de Souza SAS, Alves LLR, Feitoza LM, de Castro JTS, de Pina DR. Inflammatory lesions and brain tumors: is it possible to differentiate them based on texture features in magnetic resonance imaging? J Venom Anim Toxins Incl Trop Dis 2020;26:e20200011.

50. Jiang Z, Wang B, Han X, Zhao P, Gao M, Zhang Y, Wei P, Lan C, Liu Y, Li D. Multimodality MRI-based radiomics approach to predict the posttreatment response of lung cancer brain metastases to gamma knife radiosurgery. Eur Radiol 2022;32:2266-76.

51. Jiang Z, Dong Y, Yang L, Lv Y, Dong S, Yuan S, Li D, Liu L. CT-Based Hand-crafted Radiomic Signatures Can Predict PD-L1 Expression Levels in Non-small Cell Lung Cancer: a Two-Center Study. J Digit Imaging 2021;34:1073-85.

52. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, Cook G. Introduction to Radiomics. J Nucl Med 2020;61:488-95.

53. Crispin-Ortuzar M, Sala E. Precision radiogenomics: fusion biopsies to target tumour habitats in vivo. Br J Cancer 2021;125:778-9.

54. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue RTHM, Even AJG, Jochems A, van Wijk Y, Woodruff H, van Soest J, Lustberg T, Roelofs E, van Elmpt W, Dekker A, Mottaghy FM, Wildberger JE, Walsh S. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749-62.

55. Park JE, Kim D, Kim HS, Park SY, Kim JY, Cho SJ, Shin JH, Kim JH. Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement. Eur Radiol 2020;30:523-36.

Table S1 The details of CT imaging models

Index	Primary cohort	Test cohort
Number of centers	9	7
CT manufacturer's models	GE BrightSpeed, GE LightSpeed VCT, GE Optima CT520 Series, NMS NeuViz 128, Philips Brilliance 64, Siemens Somatom Definition AS+, Siemens Emotion 16(2007), UIH uCT 550, UIH uCT 760	GE BrightSpeed, GE Optima CT520 Series, NMS NeuViz 64 In, Philips Ingenuity CT, Siemens Somatom Definition AS, Siemens Somatom go.Now, UIH uCT 528

CT, computed tomography.

Table S2 The details of radiomic features

Texture analysis	Radiomic features
GLCM	Autocorrelation, joint average, cluster prominence, cluster shade, cluster tendency, contrast, correlation, difference average, difference entropy, difference variance, joint energy, joint entropy, Imc1, Imc2, Idm, Idmn, Id, Idn, inverse variance, maximum probability, sum entropy, sum squares
GLRLM	Gray level non uniformity, gray level non uniformity normalized, gray level variance, high gray level run emphasis, long run emphasis, long run high gray level emphasis, long run low gray level emphasis, low gray level run emphasis, run entropy, run length non uniformity, run length non uniformity normalized, run percentage, run variance, short run emphasis, short run high gray level emphasis, short run low gray level emphasis
GLSZM	Gray level non uniformity, gray level non uniformity normalized, gray level variance, high gray level zone emphasis, large area emphasis, large area high gray level emphasis, large area low gray level emphasis, low gray level zone emphasis, size zone non uniformity, size zone non uniformity normalized, small area emphasis, small area high gray level emphasis, small area low gray level emphasis, zone entropy, zone percentage, zone variance
GLDM	Dependence entropy, dependence non uniformity, gray level non uniformity, dependence non uniformity normalized, dependence variance, gray level variance, high gray level emphasis, large dependence emphasis, large dependence high gray level emphasis, low gray level emphasis, large dependence low gray level emphasis, small dependence emphasis, small dependence high gray level emphasis, small dependence low gray level emphasis
NGTDM	Busyness, coarseness, complexity, contrast, strength

The detailed mathematical definitions were shown in https://pyradiomics.readthedocs.io/en/latest/index.html. GLCM, gray level co-occurrence matrix; GLRLM, gray level run length matrix; GLSZM, gray level size zone matrix; GLDM, gray level dependence matrix; NGTDM, neighboring gray tone difference matrix.
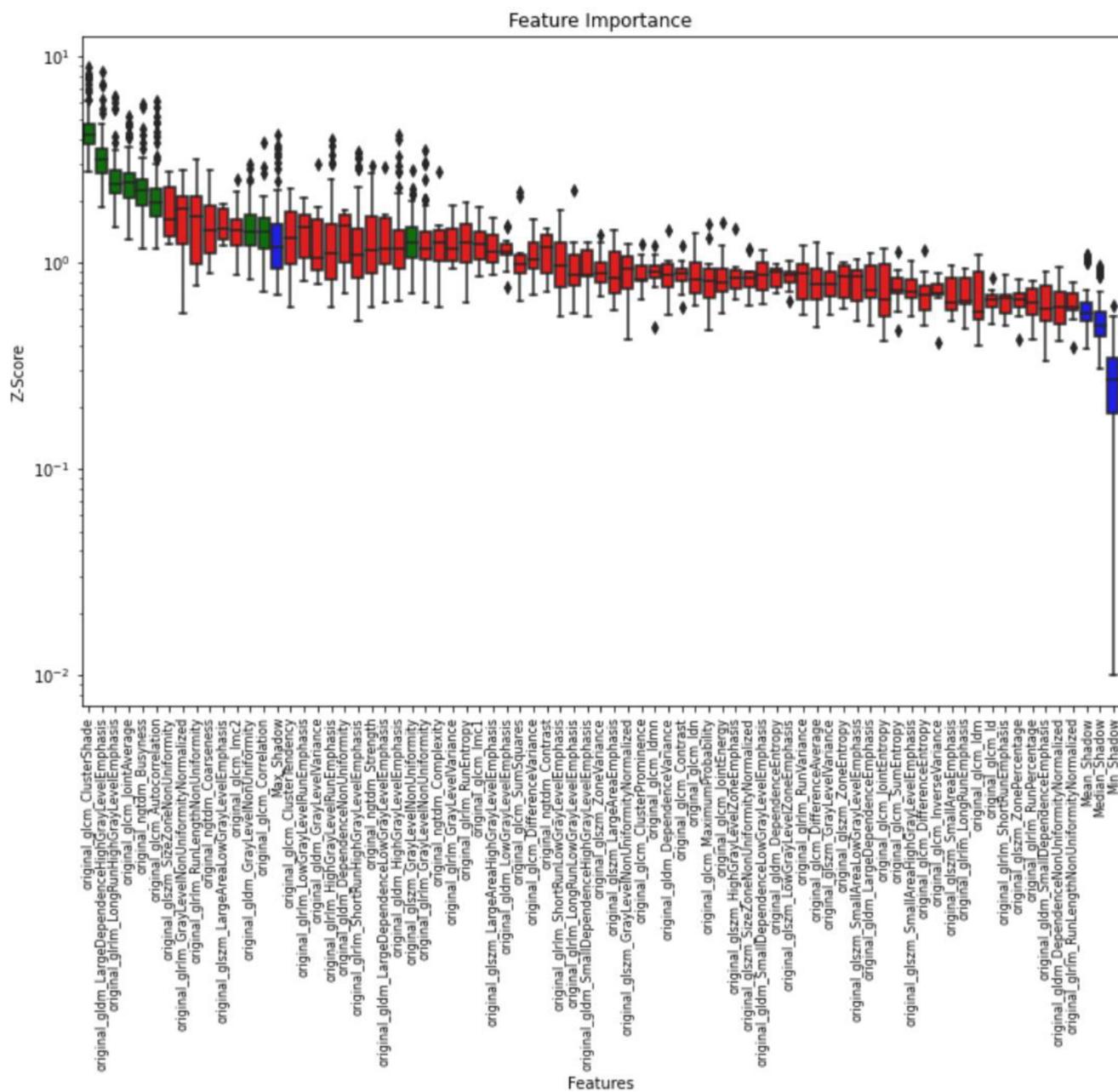
Figure S1 The BorutaShap selection in original image. SHAP, SHapley Additive exPlanations.
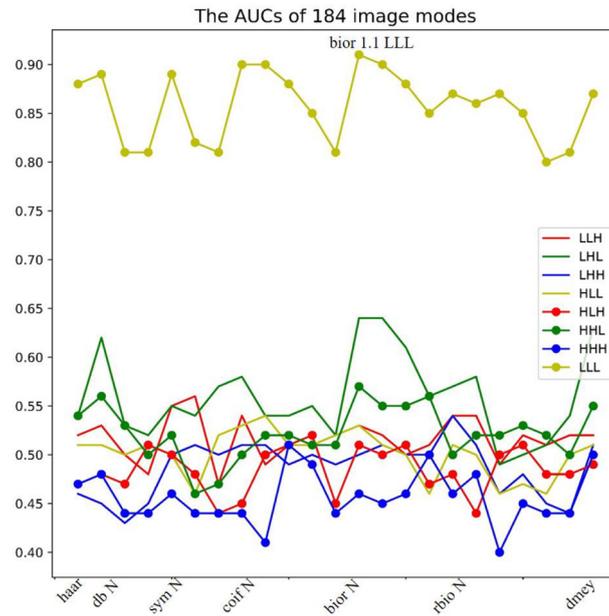
Figure S2 Diagnostic performance of all the wavelet transforming radiomic models. The AUCs of 184 wavelet transforming radiomic models in the test cohort. AUC, area under the receiver operating characteristic curve; L, low-pass decomposition filter; H, high-pass decomposition filter; db, daubechies; sym, symlets; coif, coiflets; bior, biorthogonal; rbio, reverse biorthogonal; dmey, "discrete" FIR approximation of Meyer.
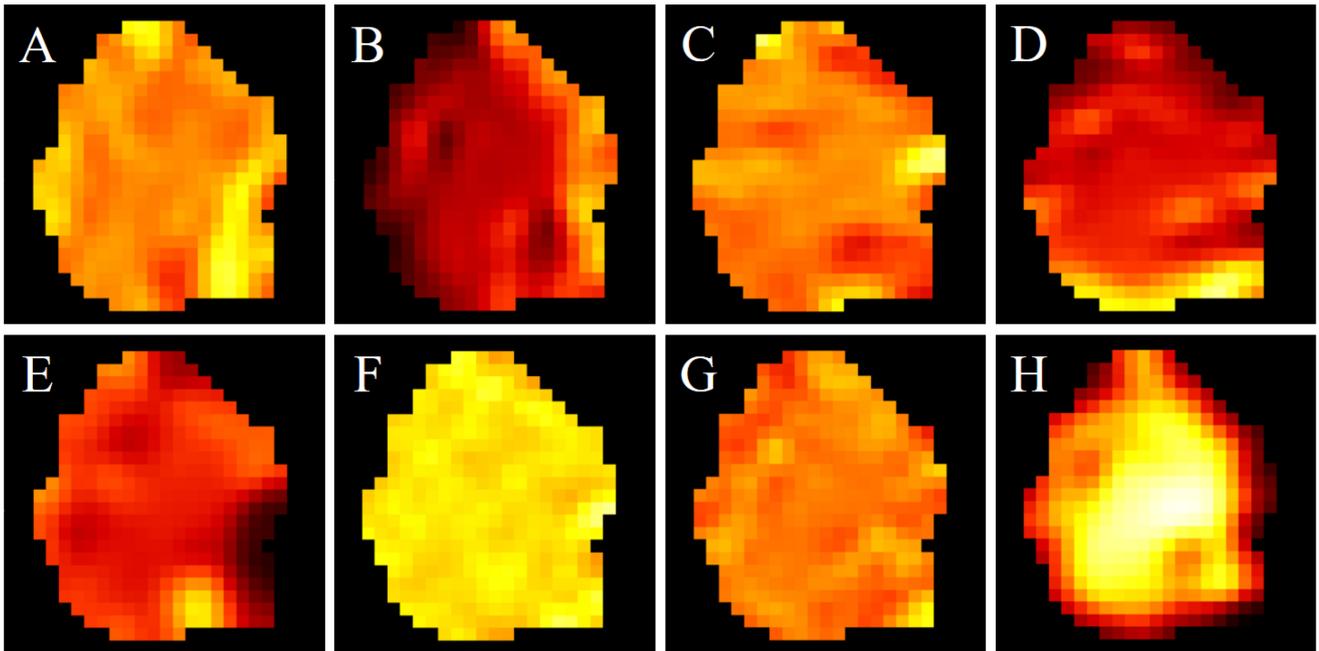

Figure S3 The GLCM autocorrelation feature maps of bior1.1 wavelet transforming. (A) Feature map in bior1.1 HLH, (B) HLL, (C) LHH, (D) LHL, (E) LLH, (F) HHH, (G) HHL, and (H) LLL. GLCM, gray level co-occurrence matrix; bior, biorthogonal; L, low-pass decomposition filter; H, high-pass decomposition filter.