



# Dynamic 3D radiomics analysis using artificial intelligence to assess the stage of COVID-19 on CT images

Shengping Cai<sup>1</sup> · Yang Chen<sup>2,3</sup> · Shixuan Zhao<sup>4</sup> · Dehuai He<sup>2,3</sup> · Yongjie Li<sup>4</sup> · Nian Xiong<sup>5</sup> · Zhidan Li<sup>4</sup> · Shaoping Hu<sup>1</sup>

Received: 13 April 2021 / Revised: 26 July 2021 / Accepted: 31 July 2021 / Published online: 29 January 2022  
© The Author(s), under exclusive licence to European Society of Radiology 2022

## Abstract

**Objective** To develop a dynamic 3D radiomics analysis method using artificial intelligence technique for automatically assessing four disease stages (i.e., early, progressive, peak, and absorption stages) of COVID-19 patients on CT images.

**Methods** The dynamic 3D radiomics analysis method was composed of three AI algorithms (the lung segmentation, lesion segmentation, and stage-assessing AI algorithms) that were trained and tested on 313,767 CT images from 520 COVID-19 patients. This proposed method used 3D lung lesion that was segmented by the lung and lesion segmentation algorithms to extract radiomics features, and then combined with clinical metadata to assess the possible stage of COVID-19 patients using stage-assessing algorithm. Area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity were used to evaluate diagnostic performance.

**Results** Of 520 patients, 66 patients (mean age, 57 years  $\pm$  15 [standard deviation]; 35 women), including 203 CT scans, were tested. The dynamic 3D radiomics analysis method used 30 features, including 27 radiomics features and 3 clinical features to assess the possible disease stage of COVID-19 with an accuracy of 90%. For the prediction of each stage, the AUC of stage 1 was 0.965 (95% CI: 0.934, 0.997), AUC of stage 2 was 0.958 (95% CI: 0.931, 0.984), AUC of stage 3 was 0.998 (95% CI: 0.994, 1.000), and AUC of stage 4 was 0.975 (95% CI: 0.956, 0.994).

**Conclusion** With high diagnostic performance, the dynamic 3D radiomics analysis using artificial intelligence could represent a potential tool for helping hospitals make appropriate resource allocations and follow-up of treatment response.

## Key Points

- The AI segmentation algorithms were able to accurately segment the lung and lesion of COVID-19 patients of different stages.
- The dynamic 3D radiomics analysis method successfully extracted the radiomics features from the 3D lung lesion.
- The stage-assessing AI algorithm combining with clinical metadata was able to assess the four stages with an accuracy of 90%, a macro-average AUC of 0.975.

**Keywords** Artificial intelligence · COVID-19 · Stage

---

Shengping Cai and Yang Chen contributed equally to this work.

✉ Zhidan Li  
627296145@qq.com

✉ Shaoping Hu  
hushaopingdoctor@163.com

<sup>1</sup> Department of Radiology, Wuhan Red Cross Hospital, Wuhan 430015, China

<sup>2</sup> West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu 610000, China

<sup>3</sup> Med-X Center for Informatics, Sichuan University, Chengdu 610000, China

<sup>4</sup> MOE Key Lab for Neuroinformatics, University of Electronic Science and Technology of China, Chengdu 610054, China

<sup>5</sup> Department of Neurology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China

## Abbreviations

3D	Three dimensions
AI	Artificial intelligence
AUC	Area under the receiver operating characteristic curve
CI	Confidence interval
COVID-19	Coronavirus disease 2019

## Introduction

By February 16, 2021, there have been over 100 million confirmed cases of COVID-19, and 2,403,641 patients had died worldwide. More seriously, the disease is increasing at a rate of over 400,000 per day [1]. Diagnosis methods include respiratory sample transcription-polymerase chain reaction (RT-PCR) and chest imaging. RT-PCR has high specificity and low sensitivity, which has been reported to be as low as 60–70% [2]. Therefore, revised version 7 of China's COVID-19 Diagnosis and Treatment Protocol indicates that clinically suspected cases with imaging characteristics of pneumonia can be diagnosed as COVID-19 patients. In addition, chest X-rays are of little value in early diagnosis, whereas CT images can detect abnormalities before symptoms appear [3]. Therefore, chest CT examination is strongly recommended during the initial assessment, follow-up of suspected COVID-19 cases.

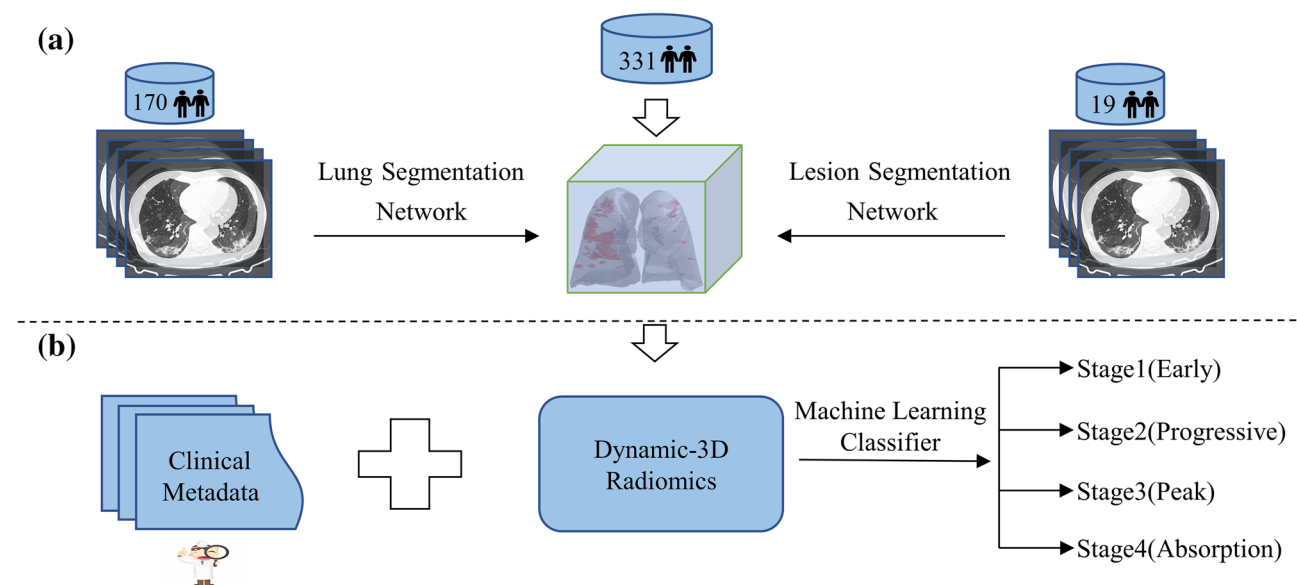
In clinical work, it is stressful for doctors to read thousands of CT images. Artificial intelligence technology may be able to solve this problem. CT radiomics and artificial

intelligence have been used to distinguish COVID-19 from other pneumonias [4]. Some scholars divided COVID-19 into four stages based on the different dynamic imaging manifestations of chest CT images [5]. Considering the evolution of pulmonary lesions by course and the effectiveness of multidirectional and multiangle image observations, this study first proposed an AI framework using dynamic 3D radiomics and clinical metadata to assess the stage of COVID-19 patients (Fig. 1); it will help hospitals with the planning and management of medical resource.

## Materials and methods

### Data and pre-processing

The data used in this study consists of three parts. The first part for lung segmentation was obtained from two public datasets with manually segmented lung boundary [6, 7]. There are 5750 CT slices from 170 COVID-19 patients in this part. The second part for lesion segmentation was collected from The Second Xiangya Hospital [8]; this part consists of 19 patients including 1117 CT images with lesion region delineated by two radiologists (6 and 10 years of experience). The third part for staging is from Wuhan Red Cross Hospital; it contains 331 patients (1023 CT scans with clinical metadata) who underwent continuous chest CT examinations between January 1, 2020, and March 9, 2020, throughout the treatment (Supplemental Material



**Fig. 1** Flowchart for the dynamic 3D radiomics analysis method using artificial intelligence. **a** Using artificial intelligence models to segment the lung and lesion of COVID-19 patients. **b** Extracting the

dynamic 3D radiomics features of 3D lung lesion and combining with the clinical metadata to assess the stage of COVID-19 patients

**Fig. 2** Four stages of COVID-19 on CT images from four patients. **a** Transverse unenhanced images. **b** The coronal reconstructing images. The first row represents early stage, the second row represents progressive stage, the third row represents peak stage, and the last row represents absorption stage

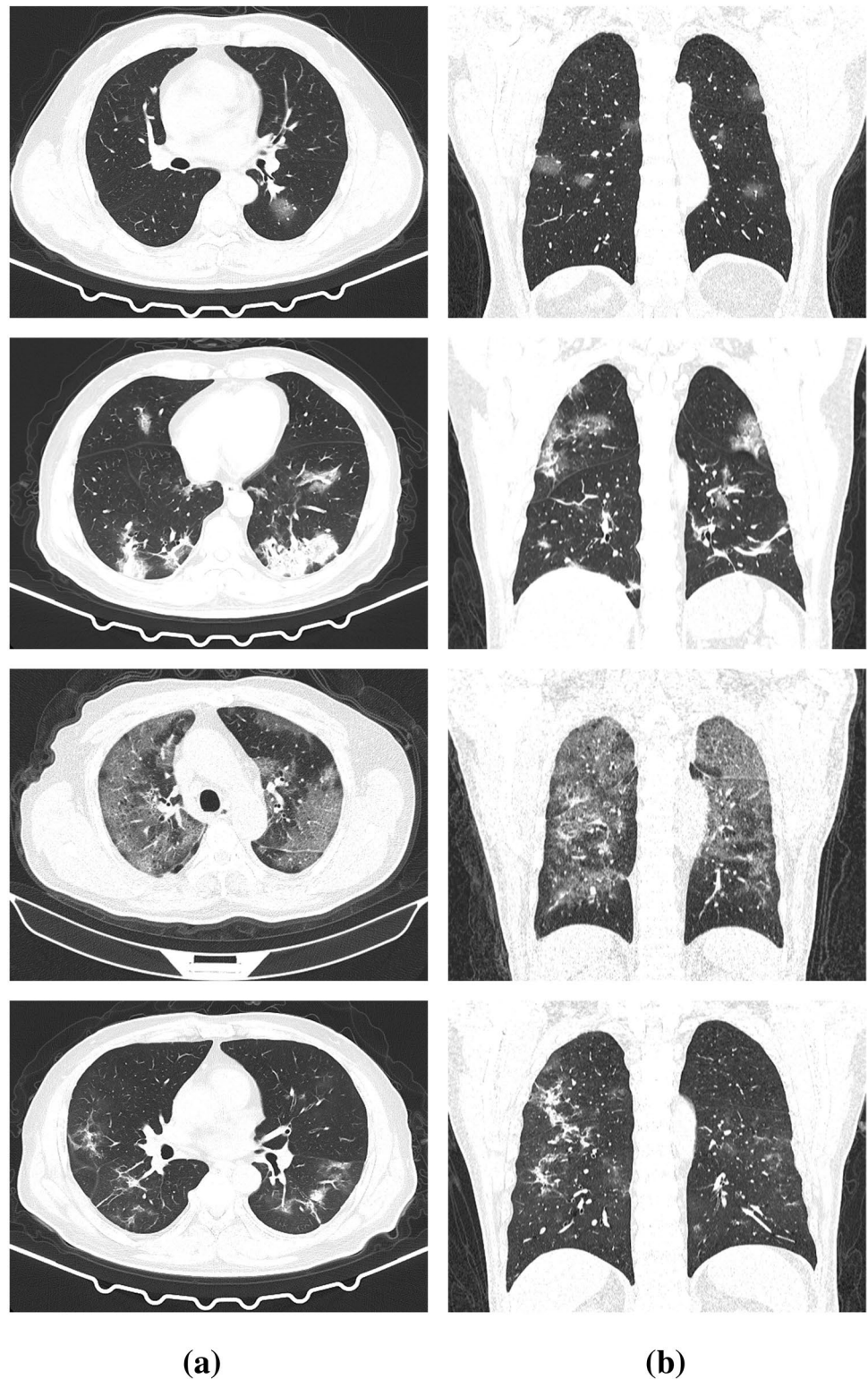


Fig. S1). These CT scans were labeled by two radiologists (14 and 31 years of experience) who were blind to the clinical metadata according to the Guideline for Medical Imaging in Auxiliary Diagnosis of Coronavirus Disease in 2019 [9]. The reference standards about staging in this guideline are

as follows: COVID-19 is divided into four stages (early, progressive, peak, and absorption stages) on CT images (Fig. 2),

- (1) The early lung manifestations are often atypical, and the lesions are light, patchy, and ground-glass opacity

(GGO) with many limitations and scatter in the two middle and lower lung fields, mainly in the subpleural region.

- (2) Multiple lesions are identified in the progressive stage, manifesting as GGO exudation, crazy-paving pattern, fusion, or consolidation, which are more common in both lung foreign zone distributions, with a small amount of pleural effusion.
- (3) The peak stage (critical illness) is equivalent to the advanced stage of the disease, and the diffuse and generalized lung density further increased, which is called “white lung.” This stage of lesion development rapidly increases by greater than 50% in 48 h. Treatment is difficult, and mortality is high.
- (4) During the absorption stage, the lesions are reduced or absorbed, and some cases exhibit changes in pulmonary interstitial fibrosis.

For the annotating procedure, we first let the two radiologists make all the annotations, respectively. Then, we checked for the discrepancies and let the two radiologists discussed for making the final decision. For the data split, we randomly divided each dataset into two independent sets (training and testing sets) with a ratio of 4:1 at the patient level (Fig. S2). We want to emphasize that all the CT images are not overlapped not only among these three datasets but also between the training set and testing set.

Regarding the data pre-processing, we used the fixed lung window (−1200, 0) to adjusted all the raw CT images and normalized them into the range (0, 255). We did not use any resizing technique, and all the CT images are in the same size of 512 x 512 pixels.

### Lung and lesion segmentation

Here, 3D radiomics features were derived from the segmented 3D lung lesion. First, we used the first and second parts of data to train two segmentation models based on the spatial- and channel-wise coarse-to-fine attention network (SCOAT-Net) [8]; the SCOAT-Net is a novel U-Net++ architecture that has a channel-wise attention module and a spatial-wise attention module to attract the self-attention learning of the network, which serves to segment the target area at the channel level and pixel successfully (Fig. S3). Lung segmentation network was trained and tested on the first part of dataset; lesion segmentation network was trained and tested on the second part of dataset. Then, we used the trained networks to segment the lung and lesion of the third part of dataset. The next step was to reconstruct the 3D lung lesion based on the results of lung and lesion segmentation.

### Dynamic 3D radiomics feature extraction and feature selection

Quantitative radiomics approaches have been applied a lot in medical image analysis since Aerts et al [10] used radiomics features to decode tumor phenotype. In this paper, we first extracted common radiomics (intensity features) based on the reconstructed 3D lesion. Then, we decomposed the 3D lesion into nine fixed-view slices (Fig. S4) for extracting the common radiomics (texture features) on each slice. And we added the lung volume, lesion volume, and ratio of lesion-to-lung volume as shape features, because Zhang et al [7] found that the lesion ratio was a significant contributor in the clinical prognosis estimation for COVID-19. Therefore, we extracted 314 regular 3D radiomics features including 3 shape, 14 intensity, and 297 texture features. Moreover, given that the radiologists would consider the lung changes when they assess the stage, we also added the variation values of two adjacent CT scans’ regular 3D features as the dynamic 3D radiomics features. Apart from the radiomics features, we combined with four clinical features including age, sex, time of onset, and time of progress. Time of onset (in days) represents the time after initial onset of symptoms, while time of progress (in days) is the time interval between two adjacent CT scans. The dynamic 3D features and time of progress were set to 0 when the patient made the first CT scan. Thus, we extracted a total of 632 features including 628 radiomics features and 4 clinical features (Table S1, column 1).

For feature selection, we used three feature selection methods, i.e., Random Forest (RF-FS) [11], Relief-F [12], Local Learning-based Clustering Feature Selection (LLCFS) [13]. Each feature selection method assigned weights to these 632 features according to the different evaluation functions (Table S1), and sorted these features in a descending mode. Then, each classifier selected the maximum  $n$  ( $1 \leq n \leq 632$ ) features to calculate the classification accuracies using ten-fold cross-validation on the training set (Fig. S5). Finally, for each classifier, we achieved the optimal feature selection method with the optimal number of features according to the greatest accuracy. We want to emphasize that this part about feature ranking and feature selection procedure was only implemented on the training set.

### Experimental settings and evaluation metrics

We trained the lung and lesion segmentation models based on the Pytorch framework, and we used the common optimizer (SGD) to optimize the loss function (dice coefficient loss) that updated the networks’ parameters that were initialized by Kaiming method. Moreover, we set the number of epochs to 100 for stopping training, and the initial learning

rate was 0.01 that was multiplied by 0.1 every 10 epochs. We finally selected the models that were after 100 epochs of training as segmentation models. For the staging procedure, we implemented it on MATLAB software. We used FEATURE SELECTION TOOLBOX V 6.2 for feature selection and MATLAB's own machine learning classifiers for training staging models. The codes about segmentation and staging are available at <https://github.com/Phanzsx/SCOAT-Net> and <https://github.com/Phanzsx/Assess-the-COVID-19>.

We applied six metrics to evaluate the segmentation and staging performance. The dice similarity coefficient (DSC) and intersection over union (IOU) were used to evaluate the segmentation performance. Accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC) were used to evaluate the performance

of the discrimination classifiers. Moreover, Student's *t*-test was used to test the difference between independent groups; a two-sided *p* value < 0.05 was deemed to be statistically significant.

## Results

### Patient characteristics and laboratory findings

One thousand twenty-three scans from 331 COVID-19 patients were included in the staging dataset. As shown in Table 1, the most prevalent symptoms at presentation were cough (225 of 331 patients [70%]) and fever (220 of 331 patients [67%]). Table 2 shows that most laboratory results were often normal, and a small number were elevated. C-reactive protein levels, D-dimer levels, and erythrocyte sedimentation rates increased in all four stages and peaked in stage 3. In addition, lactate dehydrogenase levels and serum creatinine levels only increased in stage 3. And the time of onset for four stages were obviously different; they were  $4.4 \pm 6.1$ ,  $11.0 \pm 7.1$ ,  $15.4 \pm 8.3$ , and  $26.3 \pm 12.0$  days, respectively.

### Segmentation results and selected features

The segmentation performance of lung and lesion networks on the first and second parts of data is shown in Table S1; we used the trained segmentation models to segment the lung and lesion of the third part of data, then reconstructed the 3D lung lesion to extract radiomics features. The segmentation and reconstruction results are shown in Fig. 3. For feature selection, we first used three feature selection methods

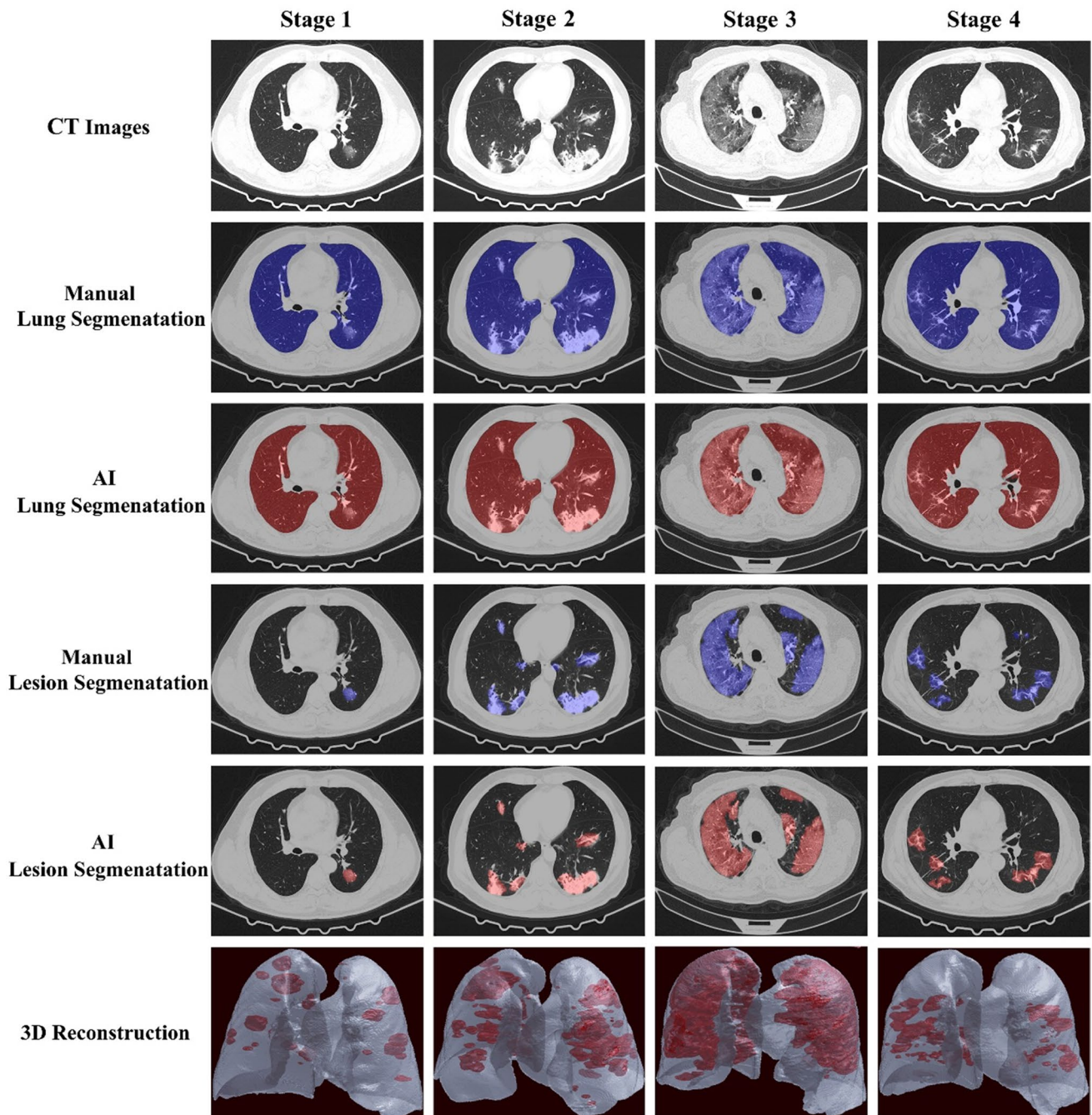
**Table 1** Characteristics of the patient cohort

Age	$57.83 \pm 15.83$ (range, 21–97)		
Sex	Men 159 (48%)	Women 172 (52%)	
Fever	Low-grade fever (37.5– 38.0 °C) 69 (21%)	Moderate-grade fever (38.1– 39.0 °C) 123 (37%)	High-grade fever (> 39.1 °C) 28 (9%)
Cough	225 (70%)		
Expectoration	93 (28%)		
Fatigue	127 (38%)		
Throat pain	24 (7%)		
Myalgia	43 (13%)		
Chest distress	126 (38%)		
Diarrhea	35 (11%)		

**Table 2** Laboratory findings of the patient cohort

Parameter	Stage 1	Stage 2	Stage 3	Stage 4
White blood cell	$5.1 \pm 2.4$	$5.5 \pm 2.8$	$8.6 \pm 3.6$	$5.9 \pm 2.3$
Neutrophil	$3.5 \pm 2.3$	$4.1 \pm 3.0$	$7.3 \pm 3.5$	$4.0 \pm 2.2$
Lymphocyte	$1.2 \pm 0.5$	$1.1 \pm 0.5$	$0.8 \pm 0.4$	$1.4 \pm 0.5$
C-reactive protein	$22.5 \pm 35.4$	$37.6 \pm 43.4$	$96.0 \pm 77.1$	$12.2 \pm 21.7$
D-Dimer	$1.0 \pm 1.7$	$2.7 \pm 6.4$	$13.8 \pm 23.9$	$2.9 \pm 5.0$
Alanine aminotransferase	$22.2 \pm 21.1$	$25.7 \pm 21.5$	$36.0 \pm 24.4$	$32.4 \pm 26.3$
Aspartate aminotransferase	$27.7 \pm 19.8$	$29.8 \pm 18.7$	$37.7 \pm 24.0$	$26.0 \pm 12.9$
Blood urea nitrogen	$5.1 \pm 4.1$	$4.7 \pm 3.9$	$7.5 \pm 6.6$	$4.5 \pm 2.3$
Serum creatinine	$95.2 \pm 120.2$	$82.9 \pm 104.1$	$114.5 \pm 201.4$	$73.5 \pm 30.6$
Serum uric acid	$259.3 \pm 103.8$	$231.3 \pm 92.7$	$229.3 \pm 122.5$	$261.7 \pm 89.4$
Erythrocyte sedimentation	$18.9 \pm 12.4$	$43.5 \pm 25.1$	$61.3 \pm 25.2$	$36.1 \pm 22.3$
Lactate dehydrogenase	$168.6 \pm 57.7$	$234.5 \pm 84.1$	$404.6 \pm 184.4$	$184.6 \pm 50.4$
Time of onset	$4.4 \pm 6.1$	$11.0 \pm 7.1$	$15.4 \pm 8.3$	$26.3 \pm 12.0$

White blood cell, neutrophil, and lymphocyte are in G/L for the unit; C-reactive protein and D-dimer are in mg/L; alanine aminotransferase, aspartate aminotransferase, and lactate dehydrogenase are in U/L; blood urea nitrogen is in mmol/L; serum creatinine and serum uric acid are in  $\mu\text{mol/L}$ ; erythrocyte sedimentation is in mm/h; Time of onset is in days

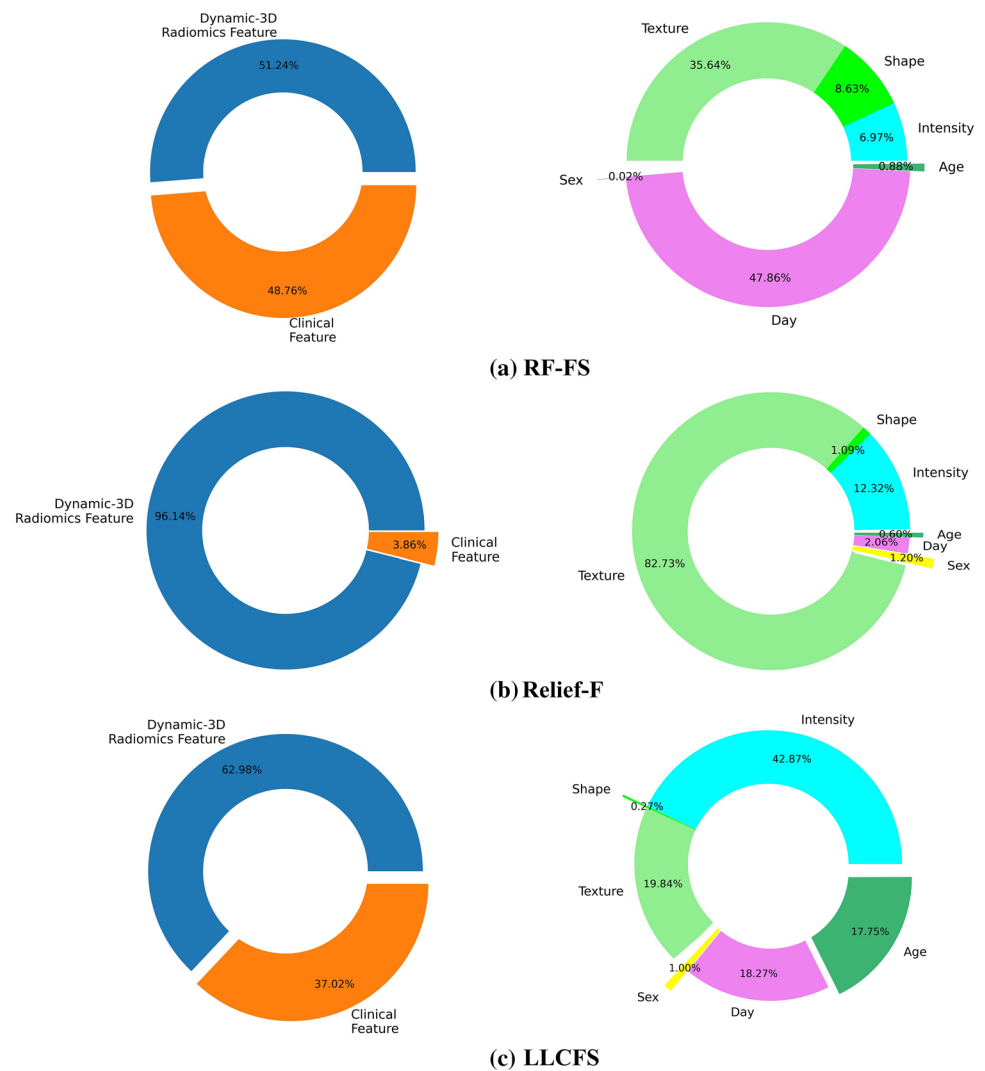


**Fig. 3** The lung-lesion segmentation results by the proposed artificial intelligence models and the corresponding 3D lung-lesion reconstruction results

for assigning weights to these features based on the training set. The results (Table S1) showed that each method could rank these features differently. Moreover, we calculated the sum of weights of different types of features (Fig. 4) and found that both dynamic 3D radiomics features and clinical features were significant factors for stage assessment on the training set. We also listed each feature selection method's top 30 feature items based on weights and found that some

features (time of onset, age, intensity features) were equally significant (Tables S2–S4). After ranking these features, each classifier selected the top  $n$  ( $1 \leq n \leq 632$ ) features to calculate the accuracy using tenfold cross-validation on the training set. Finally, as depicted in Table S5, for each classifier, we achieved the optimal feature selection method with optimal number of features. And the top 30 feature items in RF-FS feature selection method are listed in Table 3. We found

**Fig. 4** The weights of various features in the three feature selection methods. Dynamic 3D radiomics features consists of texture, shape, intensity features, clinical features consists of age, sex, day (time of onset, time of progress). **a** The RF feature selection method (RF-FS). **b** The Relief-F feature selection method. **c** The LLCFS feature selection method



that the top two feature items were time of onset (Table 2) and ratio of lesion-to-lung volume (Fig. 5). The ratios of lesion-to-lung volume at the four stages were  $2.2\% \pm 3.6\%$ ,  $14.8\% \pm 14.1\%$ ,  $46.0\% \pm 16.5\%$ , and  $7.8\% \pm 9.5\%$ , respectively.

### Stage assessing for COVID-19

We used the selected features of two adjacent CT scans as input to train our staging models on the training set, and then used the trained model to evaluate the staging performance on the testing set. The four-way classification results are shown in Table 4. Moreover, given that KNN classifier did not perform well in terms of the accuracy metric, we only applied RF and SVM classifiers for the detailed analysis. More specifically, we calculated the accuracy, sensitivity, and specificity for four two-way classifications (i.e., stage 1/stages 2–3–4, stage 2/stages 1–3–4, stage 3/stages 1–2–4, stage 4/stages 1–2–3) on the testing set; we also calculated

the confusion matrices, ROC curves on the testing set. The results are shown in Table 5 and Fig. 6.

Both Table 5 and Fig. 6 show that these two machine learning classifiers achieved comparable performances in terms of accuracy, specificity, and AUC metrics. However, RF and SVM classifiers performed differently in terms of sensitivity. More specifically, RF classifier outperformed SVM classifier for the absorption stage, and SVM classifier outperformed RF classifier for the early stage. Moreover, we found that these two classifiers could both achieve 100% sensitivity for the peak stage, but would both misclassify the early, progressive stages into the absorption stage; this is probably because the image manifestations of early, progressive stages are similar to the image manifestations of absorption stage without considering the lung changes over time. Furthermore, these misclassified CT scans probably have something to do with the lesion segmentation performance given that our lesion segmentation model could only achieve DSC of 0.885.

**Table 3** The top 30 feature items in the Random Forest Feature Selection Method (RF-FS)

Rank	Kind of feature	Feature detail	Weight
1	Clinical feature	Time of onset	0.4657
2	Dynamic 3D feature	Ratio of lesion-to-lung volume	0.0748
3	Dynamic 3D feature	Texture_view8: autocorrelation	0.0131
4	Clinical feature	Time of progress	0.0129
5	Regular 3D features	Texture_view5: run percentage	0.0122
6	Clinical feature	Age	0.0088
7	Regular 3D feature	Texture_view7: run percentage	0.0088
8	Regular 3D feature	Intensity: median	0.0078
9	Regular 3D feature	Intensity: entropy	0.0076
10	Regular 3D feature	Intensity: skewness	0.0074
11	Regular 3D feature	Intensity: uniformity	0.0066
12	Regular 3D feature	Texture_view8: informational measure of correlation 2	0.0065
13	Regular 3D feature	Lung_area	0.0058
14	Regular 3D feature	Texture_view4: cluster prominence	0.0058
15	Regular 3D feature	Texture_view4: maximum probability	0.0053
16	Regular 3D feature	Intensity: standard deviation	0.0052
17	Regular 3D feature	Texture_view7: inverse difference moment normalized	0.0052
18	Regular 3D feature	Intensity: mean absolute deviation	0.0049
19	Regular 3D feature	Texture_view4: energy	0.0049
20	Regular 3D feature	Intensity: variance	0.0048
21	Regular 3D feature	Texture_view3: cluster prominence	0.0048
22	Regular 3D feature	Texture_view4: variance	0.0046
23	Regular 3D feature	Texture_view4: difference entropy	0.0045
24	Regular 3D feature	Texture_view4: contrast	0.0044
25	Regular 3D feature	Texture_view5: run length non-uniformity	0.0042
26	Regular 3D feature	Texture_view9: run length non-uniformity	0.0042
27	Regular 3D feature	Texture_view8: run length non-uniformity	0.0041
28	Regular 3D feature	Texture_view9: cluster prominence	0.0040
29	Regular 3D feature	Texture_view7: run length non-uniformity	0.0038
30	Regular 3D feature	Texture_view8: contrast	0.0035

Based on the above results, the diagnostic efficiencies were respectable for all the classifiers. More specifically, RF classifier was the most effective in terms of the total accuracy (90%). And SVM classifier yielded the best sensitivity compared with RF classifier when diagnosing the early stage.

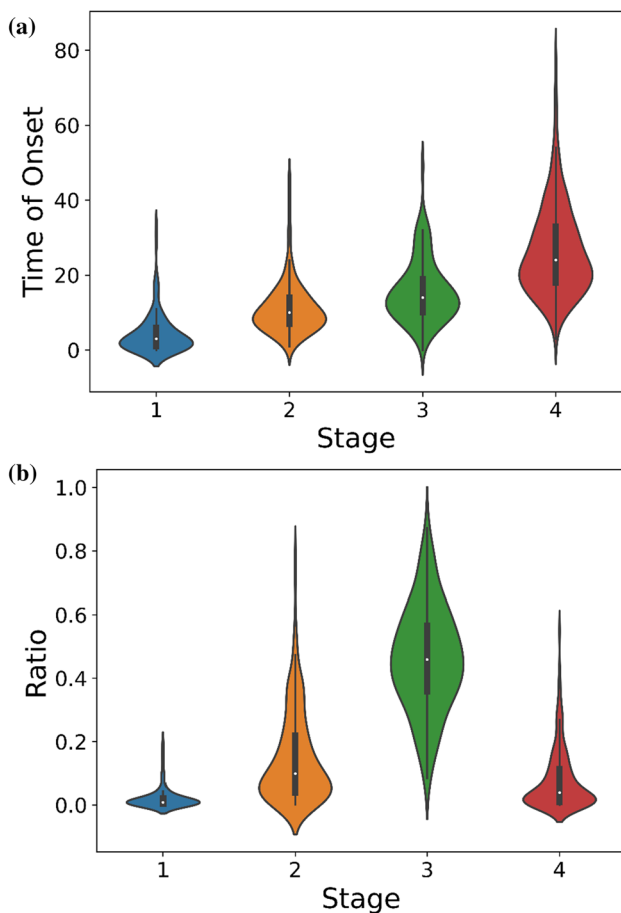
## Discussion

In this study, the characteristics of COVID-19 were analyzed from the clinical symptoms, laboratory examination, and dynamic 3D radiomics features of the patients, and the 30 key features including clinical metadata, and dynamic 3D radiomics features were selected to assess the stage of COVID-19 patients.

We found that patients often had cough (225 of 331 patients [70%]) and fever (220 of 331 patients [67%]) symptoms; this result was the same as that of another meta-analysis [14]. Recently, results of a UK cohort study revealed that

ten symptoms, including cough and fever, were associated with COVID-19 infection [15]. Many studies have found that C-reactive protein levels, D-dimer levels, and erythrocyte sedimentation rate increase [14]. However, other detailed studies have not been performed to assess how these laboratory indicators evolve with disease. Our study revealed this pattern over time. In addition, the role of C-reactive protein levels in prognosis has been studied. The C-reactive protein level could be used as an independent factor to predict the outcome of COVID-19. Higher levels were more likely to be associated with complications [16]. Some scholars have noted that lactate dehydrogenase, which is a metabolic marker, is an independent risk factor for severe COVID-19 patients [17]. An increase or decrease in lactate dehydrogenase levels was indicative of radiographic progress or improvement, which was consistent with our results [18]. An increase in serum creatinine in the severe stage was also reported in another study [19], indicating acute kidney injury [20].





**Fig. 5** **a** Relationships between the stages of COVID-19 and the time of onset (in days). **b** Relationships between the stages of COVID-19 and ratio of lesion-to-lung volume

**Table 4** The optimal feature selection method and feature quantity in terms of accuracy for the three machine learning classifiers on the training and testing sets

Classifier	Feature selection method	Feature quantity	Training Acc	Testing Acc
RF	RF-FS	30	1.0000	0.9018
SVM	RF-FS	17	0.8672	0.8769
KNN	RF-FS	176	1.0000	0.8079

Acc represents accuracy

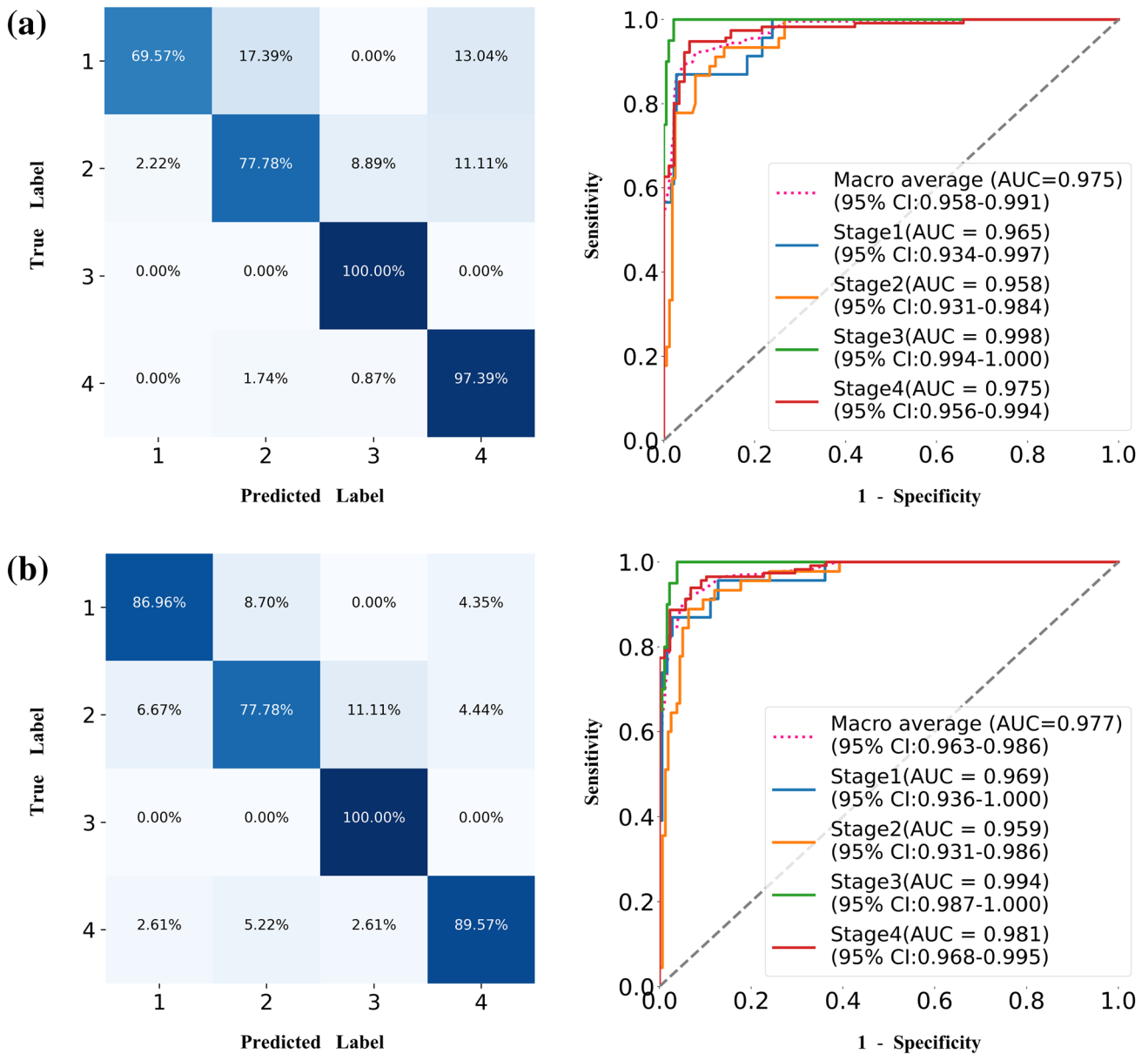
To date, there has been no application of dynamic 3D radiomics technology for COVID-19 analysis. Conventional radiomics techniques were based on 2D medical images for feature extraction. Three-dimensional feature extraction is not common. Dynamic analysis in researches typically refers to the changes of drug metabolism characteristics after injection of enhanced drugs [21]. Our study not only provided detailed feature analysis, including shape, texture, and intensity in 3D space, but also considered the influence

**Table 5** Performance metrics of RF and SVM classifiers on the testing set

Classifier	Metrics	Stage 1	Stage 2	Stage 3	Stage 4
RF	Accuracy (95% CI)	96% (92–98%)	92% (88–95%)	98% (94–99%)	95% (91–97%)
	Sensitivity (95% CI)	70% (47–87%)	78% (63–89%)	100% (83–100%)	97% (93–99%)
	Specificity (95% CI)	99% (97–100%)	96% (92–99%)	97% (94–99%)	91% (83–96%)
SVM	Accuracy (95% CI)	96% (92–98%)	91% (86–95%)	96% (92–98%)	93% (88–96%)
	Sensitivity (95% CI)	87% (66–97%)	78% (63–89%)	100% (83–100%)	90% (82–94%)
	Specificity (95% CI)	97% (93–98%)	95% (90–98%)	96% (92–98%)	97% (90–99%)

of course on the disease of COVID-19, which is a very innovative research approach. We made such great efforts because the characteristics of the lesion, age, sex, and time course of disease are important factors for the diagnosis and prognosis of COVID-19, which have been reported in many studies [22–24]. In the analysis of lesion characteristics, our study suggested that texture and intensity features were the most important factors, which was the same as other studies [25]. In that study, it extracted image features in the 2D space, and the texture features played the greatest role in determining whether it was an early ground-glass opacity, but the AUC was only 0.67. Our study took nine view slices that were extracted in the 3D space, which was the reason why the final ROC achieved a better effect. In our study, the onset time of four stages were different, which was similar to the findings of Pan et al [5], but the onset time of stages 2, 3, and 4 in our study was a little greater than the findings of Pan et al. Our study was based on 1023 scans of 331 patients, and the differences in the number of patients might account for the difference of outcome. The ratio of lesion-to-lung volume represented the degree of lung involvement. Zhang et al also analyzed the correlation between this feature and the clinical parameters of COVID-19 [7]. Autocorrelation of texture\_view8 was a feature of the grey level cooccurrence matrix. The autocorrelation coefficient of the eighth view reflected the changes of image manifestations from the sagittal alignment of the lung from anterior-to-posterior with an elevation of 45°; it played an important role in distinguishing different stages. This finding was consistent with our clinical observation: lesions often involved the lower lobe dorsal segment of both lungs [14].

The COVID-19 is becoming a major challenge to medical resources as a large number of diagnosed patients



**Fig. 6** Normalized confusion matrices and ROC curves of four-way classifications. In the ROC curves, the red dot curve denotes the macro-average area under the curve (AUC). **a** RF classifier. **b** SVM classifier

continue to be hospitalized. Computer-aided diagnosis (CAD) system is an effective tool to make an automatic and rapid diagnosis. Even with the limited training datasets, our dynamic 3D radiomics analysis method using artificial intelligence could provide a potential tool to help hospitals quickly triage patients at admission using only CT images and basic clinical metadata. Moreover, our AI system can enhance the clinical follow-up of disease development and treatment response for the COVID-19 patients. There were also some limitations in our study, such as the lack of application of the model in external

validation datasets and the lack of further study in CT images of patients with concurrent lung tumors. Moreover, our study has a potential limitation because both labels and predictors depend on the imaging data, which will produce incorporation bias (PROBAST criteria 3.3 and 3.5) [26]. The next step will be to collect more data for external validation, and look for a new methodology (e.g., time of onset as label) that ensures the labels are independent of predictors. Furthermore, we also plan to improve the performance of our lesion segmentation model for extracting more accurate radiomics features.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00330-021-08533-1>.

**Acknowledgements** We sincerely thank Jun Ma et al and Kang Zhang et al for providing the public COVID-19 dataset, and Jiao Xu for helping with data pre-processing.

**Funding** This work is supported by Department of Science and Technology of Sichuan Province under grant number 2021YJ0245.

## Declarations

**Guarantor** The scientific guarantor of this publication is Shaoping Hu, MD.

**Conflict of interest** The authors declare no competing interests.

**Statistics and biometry** Yang Chen kindly provided statistical advice for this manuscript.

**Informed consent** Written informed consent was waived by the Institutional Review Board.

**Ethical approval** Institutional Review Board approval was obtained.

## Methodology

- retrospective
- diagnostic or prognostic study
- multicenter study

## References

1. Coronavirus disease (COVID-19) Situation dashboard. <https://covid19.who.int>
2. Fang Y, Zhang H, Xie J et al (2020) Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* 296:E115–E117
3. Kim JY, Choe PG, Oh Y et al (2020) The first case of 2019 novel coronavirus pneumonia imported into Korea from Wuhan, China: implication for infection prevention and control measures. *J Korean Med Sci* 35:e61
4. Fang M, He B, Li L et al (2020) CT radiomics can help screen the coronavirus disease 2019 (COVID-19): a preliminary study. *Sci China Inf Sci*. <https://doi.org/10.1007/s11432-020-2849-3>
5. Pan F, Ye T, Sun P et al (2020) Time course of lung changes at chest CT during recovery from coronavirus disease 2019 (COVID-19). *Radiology* 295:715–721
6. Ma J, Wang Y, An X et al (2020) Towards efficient COVID-19 CT annotation: a benchmark for lung and infection segmentation. *arXiv preprint arXiv:2004.12537*
7. Zhang K, Liu X, Shen J et al (2020) Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* 181(1423–1433):e1411
8. Zhao S, Li Z, Chen Y et al (2020) SCOAT-Net: a novel network for segmenting COVID-19 lung opacification from CT images. *medRxiv*. <https://doi.org/10.1101/2020.09.23.20191726>
9. Li H, Liu S, Xu H, Cheng J (2020) Guideline for medical imaging in auxiliary diagnosis of coronavirus disease 2019. *Chin J Med Imaging Technol* 36:321–331
10. Aerts HJWL, Velazquez ER, Leijenaar RTH et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 5:4006
11. Sylvester EV, Bentzen P, Bradbury IR et al (2018) Applications of random forest feature selection for fine-scale genetic population assignment. *Evol Appl* 11:153–165
12. Liu H, Motoda H (eds) (2007) Computational methods of feature selection (1st ed.) Chapman and Hall/CRC. <https://doi.org/10.1201/9781584888796>
13. Zeng H, Cheung Y-m (2010) Feature selection and kernel learning for local learning-based clustering. *IEEE Trans Pattern Anal Mach Intell* 33:1532–1547
14. Zhu J, Ji P, Pang J et al (2020) Clinical characteristics of 3062 COVID-19 patients: a meta-analysis. *J Med Virol* 92:1902–1914
15. Han R, Huang L, Jiang H, Dong J, Peng H, Zhang D (2020) Early clinical and CT manifestations of coronavirus disease 2019 (COVID-19) pneumonia. *AJR Am J Roentgenol*. <https://doi.org/10.2214/AJR.20.22961>
16. Liu F, Li L, Xu M et al (2020) Prognostic value of interleukin-6, C-reactive protein, and procalcitonin in patients with COVID-19. *J Clin Virol* 104:370
17. Han Y, Zhang H, Mu S et al (2020) Lactate dehydrogenase, an independent risk factor of severe COVID-19 patients: a retrospective and observational study. *Aging (Albany NY)* 12:11245
18. Wu M-y, Yao L, Wang Y et al (2020) Clinical evaluation of potential usefulness of serum lactate dehydrogenase (LDH) in 2019 novel coronavirus (COVID-19) pneumonia. *Respir Res* 21:1–6
19. Chen C, Yan J, Zhou N, Zhao J, Wang D (2020) Analysis of myocardial injury in patients with COVID-19 and association between concomitant cardiovascular diseases and severity of COVID-19. *Zhonghua Xin Xue Guan Bing Za Zhi* 48:E008–E008
20. Cheng Y, Luo R, Wang K et al (2020) Kidney disease is associated with in-hospital death of patients with COVID-19. *Kidney Int* 97:829–838
21. Guan Y, Peck KK, Lyo J et al (2020) T1-weighted dynamic contrast-enhanced MRI to differentiate nonneoplastic and malignant vertebral body lesions in the spine. *Radiology* 297:382–389
22. Klein SL, Pekosz A, Park H-S et al (2020) Sex, age, and hospitalization drive antibody responses in a COVID-19 convalescent plasma donor population. *J Clin Investig* 130:6141–6150
23. Palaodimos L, Kokkinidis DG, Li W et al (2020) Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a cohort of patients with COVID-19 in the Bronx New York. *Metabolism* 108:154262
24. Iaccarino G, Grassi G, Borghi C, Ferri C, Salvetti M, Volpe M (2020) Age and multimorbidity predict death among COVID-19 patients: results of the SARS-RAS study of the Italian Society of hypertension. *Hypertension* 76:366–372
25. Homayounieh F, Ebrahimian S, Babaei R et al (2020) CT radiomics, radiologists and clinical information in predicting outcome of patients with COVID-19 pneumonia. *Radiol Cardiothorac Imaging* 2:e200322
26. Moons KGM, Wolff RF, Riley RD et al (2019) PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 170:W1–W33

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.