



# Detection of stage of lung changes in COVID-19 disease based on CT images: a radiomics approach

Mohammad Mehrpouyan<sup>1,2</sup> · Hamed Zamanian<sup>3</sup> · Ghazal Mehri-Kakavand<sup>4</sup> · Mohamad Pursamimi<sup>4</sup> · Ahmad Shalbaf<sup>3</sup> · Mahdi Ghorbani<sup>3</sup> · Amirhossein Abbaskhani Davanloo<sup>5</sup>

Received: 15 June 2021 / Accepted: 16 May 2022 / Published online: 7 July 2022  
© Australasian College of Physical Scientists and Engineers in Medicine 2022

## Abstract

The aim of this study is to classify patients suspected from COVID-19 to five stages as normal, early, progressive, peak, and absorption stages using radiomics approach based on lung computed tomography images. Lung CT scans of 683 people were evaluated. A set of statistical texture features was extracted from each CT image. The people were classified using the random forest algorithm as an ensemble method based on the decision trees outputs to five stages of COVID-19 disease. Proposed method attains the highest result with an accuracy of 93.55% (96.25% in normal, 74.39% in early, 100% in progressive, 82.19% in peak, and 96% in absorption stage) compared to the other three common classifiers. Radiomics method can be used for the classification of the stage of COVID-19 disease with good accuracy to help decide the length of time required to hospitalize patients, determine the type of treatment process required for patients in each category, and reduce the cost of care and treatment for hospitalized individuals.

**Keywords** COVID-19 disease · Stage · Radiomics · Computed tomography

## Introduction

Since December 2019, an outbreak of rare disease has been reported that has significant inflammatory effects on the respiratory system, particularly the lungs. The disease was first reported from Wuhan in China. The official report of the disease was presented in January 2020, and it was named

COVID-19 disease or, more specifically, coronavirus with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. Similar to other diseases that have emerged in the Middle East and Africa in recent years, such as severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome (MERS), this disease is also associated with acute respiratory distress syndrome (ARDS) [2]. Due to the high number of patients referred to medical centers and the need to perform the diagnosis and treatment process, the real-time polymerase chain reaction test (RT-PCR) is currently considered the gold standard for the identification of COVID – 19 disease. However, because of the possibility of false-negative results, low sensitivity [3, 4], and the long time required to obtain the results, other diagnostic methods are being considered. Also, the severity of the disease is not measureable by PCR. In the meantime, computed tomography (CT) can be considered as a potential tool for the diagnosis of COVID-19 disease [5, 6]. Also, lung CT images contain important information as the stage of COVID-19 disease. Having a radiologist examine CT images of individuals suspected of having the disease allows for quicker decisions and isolates patients from others during the treatment process.

✉ Ahmad Shalbaf  
shalbaf@sbm.ac.ir

✉ Mahdi Ghorbani  
mhdghorbani@gmail.com

<sup>1</sup> Non-Communicable Diseases Research Center, Sabzevar University of Medical Sciences, Sabzevar, Iran

<sup>2</sup> Medical Physics and Radiological Sciences Department, Sabzevar University of Medical Sciences, Sabzevar, Iran

<sup>3</sup> Biomedical Engineering and Medical Physics Department, School of Medicine, Shahid Beheshti University of Medical Sciences, 19857-17443 Tehran, Iran

<sup>4</sup> Department of Medical Physics, School of Medicine, Semnan University of Medical Sciences, Semnan, Iran

<sup>5</sup> Department of Radiology, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

According to the guidelines, COVID-19 disease can be diagnosed on lung images CT with ground-glass opacity (GGO), a crazy-paving pattern, and subsequent consolidation in the lungs [7, 8]. As for the time course of COVID-19 disease, there are four stages: the early stage (0–4 days after the onset of the disease), the progressive stage (5–8 days), the high stage (9–13 days), and the absorption stage, which occurs two weeks after the onset of the first symptoms [9]. At these stages, there are differences in the histological tissue patterns that can be observed on CT images of the patients' lungs. CT images in the early stage showed GGO; CT images in the progressive stage showed an increase in the crazy-paving pattern; CT images in the peak stage showed consolidation, and CT images in the absorption stage showed gradual resolution of consolidation without crazy-paving pattern. The absorption stage shows that the disease is managed and the consolidation is slowly absorbed [9]. The importance of classification based on these stages is to help decide the length of time required to hospitalize patients, to determine the type of treatment process required for patients in each category, and to reduce the cost of care and treatment for hospitalized individuals. The temporal changes in CT lung images of COVID-19 patients have been the subject of several studies [10–13]. Interpretation of these images and their textural changes requires relevant specialists who are not present everywhere and is also tedious and subjective. Therefore, the application of radiomics and machine learning methods is increasingly required.

In recent years, radiomics and machine learning methods have been used to diagnose and even predict a variety of diseases, such as brain tumors [14], breast cancer [15], cardiovascular disease [16, 17], and Leukemia [18]. During the spread of COVID-19 disease, several studies have been conducted to use radiomics and machine learning algorithms for segmentation and categorization of CT images and consequently for detection of COVID-19 infection and prediction of patient's condition [19, 20]. Ouyang et al. [21] introduced a sampling network that differentiated COVID-19 from the community acquired pneumonia (CAP) pulmonary inflammation by focusing on the infected areas of the lungs. In a study by Lin et al. [22] a radiomics model was developed for the prediction of COVID-19 pneumonia and the model was validated to help clinicians quickly and accurately rule out influenza virus pneumonia. They showed that the radiomics model has good performance in distinguishing COVID-19 pneumonia from influenza virus pneumonia. In another study, Yang et al. [23] proposed a classification method based on the traditional machine learning method with the use of the radiomics features of CT chest images to identify and diagnose patients with COVID-19 and non-COVID-19 pneumonia. Their results showed that radiomics features can classify COVID-19 patients and other patients who had other types of pneumonia. Kadry et al. [24]

developed a machine learning system for the classification of CT images of COVID-19. The proposed algorithm uses a set of different methods, including multi-threshold, image separation using threshold filter, feature extraction, feature selection, feature-fusion, and classification. In a study by Fu et al. [25], radiomics and a machine learning-based tool were used to evaluate the prognosis of COVID-19 patients. The data were divided into two groups including a stable group and a progressive group. The imaging features from whole-lung images were extracted. They concluded that radiomics signature of the whole lung which was based on machine learning method could reveal the lung microstructure changes. It also can help to indicate the progress of the disease. In a study by Tan et al. [26], an automatic machine learning method was used to predict the clinical types of COVID-19 pneumonia by analyzing the non-focus region of the lung CT images of patients. In a study by Wang et al. [27] feature engineering methods were compared to know image features in CT images of COVID-19 pneumonia and to predict those features which are significant for this pneumonia. Then, specific deep learning and radiomics features were determined which could be used with these two methods to help the human diagnosis of COVID-19 pneumonia. Guiot et al. [28] developed an automated artificial intelligence framework to extract radiomics features from chest CT images. The obtained results showed that the presented artificial intelligence framework could accurately differentiate COVID-19 cases from other clinical cases. In a study by Wang et al. [29], the efficacy of radiomics was evaluated for the diagnosis of patients with COVID-19 and other types of viral pneumonia, based on CT images from patients. Those radiomic features which were significantly associated with the classification of COVID-19 disease were determined using multiple classification algorithms. Finally, Liu et al. [30] investigated the application of chest CT radiomics for diagnosing COVID-19 pneumonia and an open-source diagnostic tool was developed based on the constructed radiomics model. They concluded that the combined radiomics model could improve the performance of the clinical model and COVID-19 reporting for the diagnosis of COVID-19 pneumonia.

In the recent studies mentioned, classification was limited only to the presence or absence of COVID-19 disease. However, there is a low study to determine the staging or prognosis of this disease using radiomics. In two recent studies, CT lung image segmentation has used to automatically quantify infection regions and then classified the COVID-19 patients into 3-class severity [31, 32]. However, these few methods are based on the segmentation of lesion of lung involvement which is a problemist task. The aim of this study is to classify COVID-19 patients based on lung CT images into normal, early, progressive, peak, and absorption stages by using automatic radiomics based on a combination of

texture statistical features and Random Forest (RF) classifier without any need to image segmentation. This proposed method plays an important role in predicting the course and progression of the disease and the condition of the infected patient and will in turn be used to guide decisions regarding the length of hospital stay, making better decisions to cure and manage the costs of maintenance.

## Materials and methods

The present study attempts to use the radiomics approach to classify the CT images of 683 individuals with suspected COVID-19 into five stages: normal stage, early stage, progressive stage, peak stage, and absorption stage, depending on the involvement of the lung parenchyma. This procedure was carried out on the basis of the extent of the changes noted in the various slices of the images obtained from the patients' lungs. The features of the images relating to each individual were extracted and defined as a basis for assessing the classification of the suspect into a particular stage of COVID-19 disease. In the preprocessing step, CT slices were taken from all thoracic sections, which contained lung organs. Then, different algorithms were used to extract the appropriate features. For this purpose, the first order statistical texture features extracted from the image histogram including variance, skewness and kurtosis [33] and second-order texture features included the gray-level co-occurrence matrix (GLCM [33, 34], 22 features), the gray-level run-length matrix (GLRLM [33–35], 13 features), the gray-level size-zone matrix (GLSZM [33, 34], 13 features), and the gray-level neighborhood difference matrix (NGTDM [34, 36], 5 features) were extracted. Consequently, a total of 56 statistical features (including 3 first-order features and 53 s-order features) were extracted from each CT image. Then, for each participant, these mentioned features are extracted from all CT images. The sum of these features was considered as the feature set of the participant. Finally, based on the extracted features, individuals were classified into five stages: normal, early, progressive, peak, and absorption stages of COVID-19 disease. The results were compared for the same number of features with three other commonly used classifiers: linear discriminant analysis (LDA), artificial neural network (ANN) and support vector machine (SVM). All calculation was done using MATLAB R2019b software.

### Database

Lung CT images were collected from 683 subjects (408 men and 275 women) with suspected COVID-19. Chest CT was performed using Hispeed CT Dual Slice Scanner (GE healthcare, USA) at Qaboos Teb Golestan Medical Imaging Center (Gonbad-e-Kavous, Iran). The images were acquired

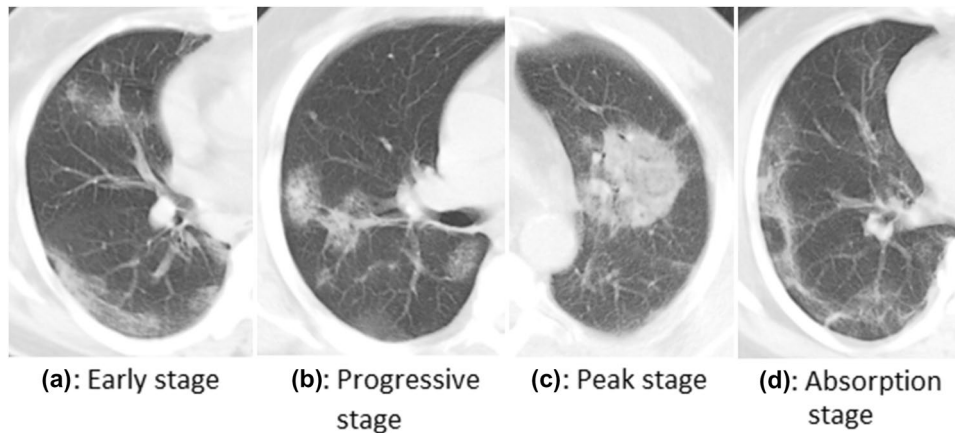
in spiral mode. Slice thicknesses were 10 mm, intervals were 10 mm, and velocity was 15 mm/rotation. The average number of CT axial images for a patient is 24 and was stored in the standard Digital Imaging and Communications in Medicine (DICOM) format with a size of 512×512 pixels. The chest CT is used to assess the stage of lung involvement in COVID-19. The lung findings on the chest CT scans were GGOs that grew in size and developed with a crazy-paving pattern and subsequent consolidation. A quantitative system was used to assess the stage of lung involvement including normal, early, progressive, peak and absorption stages of COVID – 19 disease. The evaluation and labeling of the CT images of different patients in terms of the desired stage of disease was based on the diagnoses of two independent radiologists with at least 10 years of experience in thoracic radiology and thousands of patients examined. The final report is specified by consensus. Evaluation of previous studies has shown that radiologists are widely used for the evaluation of CT images in the diagnosis of COVID-19 disease [31, 32]. The numbers of subjects in normal, early, progressive, peak, and absorption stages of COVID-19 disease were 320, 82, 108, 73, and 100 cases, respectively. Sample CT images of lung slices for the five classes are presented in Fig. 1. CT images in the early stage indicated ground-glass opacities; in the progressive stage indicated an increase in the crazy-paving pattern; in the peak stage indicated consolidation, and in absorption stage indicated gradual resolution of consolidation without crazy-paving pattern. This research was performed based on ethical guidelines of Sabzevar University of Medical Sciences (IR.MEDSAB.REC.1399.152).

### Preprocessing

A set of lung CT images comprising about 20–30 slices were acquired for each participant. Then, the upper and lower part of the lung CT slices was excluded for finding the lung organ completely. Therefore, the images that clearly contained the lung volume were selected. Then both sides of the vertical and horizontal parts of each image were cut off.

### Feature extraction

First- and second-order texture features were used to classify patients into the mentioned five stages. These statistical features were determined on the basis of changes in intensity in points relative to the other points by providing information about the gray levels of the CT image. The statistical methods for extracting first and second-order texture features used in this study for each image include: Variance, skewness and kurtosis [33] as the first order and GLCM [33, 34], GLRLM [33–35], GLSZM [33, 34], and NGTDM [34, 36] as the second order.



**Fig. 1** Lung CT images illustrating histological pattern including **a** early stage (0–4 days) with ground-glass opacities, **b** progressive stage (5–8 days) with an increase in the crazy-paving pattern (interloper septal thickening), **c** peak stage (9–13 days) with consolidation, **d** absorption stage ( $\geq 14$  days) with gradual resolution of consolidation without crazy-paving pattern (reticular). The images are related to the sagittal plane

The GLCM feature extraction method was selected based on the properties of gray level changes in neighboring pixels [33]. This method, the most common feature extraction algorithm, is a technique for extracting second-order statistical features from images. It represents a matrix whose number of rows and columns corresponds to the number of gray levels in the image. The matrix element  $P(i, j | \Delta x, \Delta y)$  is defined as the number of times at which two pixels with distances of  $(\Delta x, \Delta y)$  occur in a defined neighborhood. Various forms of neighbor pixels can be defined, of which the most common are: Single-pixel, double-pixel, and three-pixel steps, and consideration of 0-degree, 45-degree, 90-degree, and 135-degree neighborhoods. For these different neighborhoods, the following properties were evaluated on the pixels: energy, entropy, dissimilarity, variance, contrast, inverse difference, correlation, homogeneity, autocorrelation, cluster shade, cluster prominence, maximum probability, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, information measures of correlation, maximal correlation coefficient, inverse difference normalized, and inverse difference moment normalized. The total number of these features is 22 and was calculated separately for each image. These 22 GLCM features were applied in four main degrees ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ ) in three steps. Then, the GLCM features were averaged in four main degrees and 3 kinds of steps.

GLRLM is another feature extraction method commonly used to extract high-level second statistical texture features from medical images [34]. In this technique, the number of pairs of pixels with gray scale value and their run length are calculated for a given region of interest. This method is also used in four main directions ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$ ). Then the GLRLM features are averaged in four main directions. The following features are routinely extracted for the GLRLM method: short run emphasis

(SRE), long run emphasis (LRE), gray-level nonuniformity (GLN), run-length nonuniformity (RLN), run percentage (RP), low gray-level run emphasis (LGRE), high gray-level run emphasis (HGRE), short run low gray-level emphasis (SRLGE), short run high gray-level emphasis (SRHGE), long run low gray-level emphasis (LRLGE), long run high gray-level emphasis (LRHGE), gray-level variance (GLV), and run-length variance (RLV). The total number of these features is 13 and was calculated separately for each image.

GLSZM is an extended version of the GLRLM method, composed of large areas of equal intensity rather than small groups of pixels or segments in a particular direction [34]. The following features are routinely extracted for the GLSZM method: small zone emphasis (SZE), large zone emphasis (LZE), gray-level nonuniformity (GLN), zone-size nonuniformity (ZSN), zone% (ZP), low gray-level zone emphasis (LGZE), high gray-level zone emphasis (HGZE), small zone low gray-level emphasis (SZLGE), small zone high gray-level emphasis (SZHGE), large zone low gray-level emphasis (LZLGE), large zone high gray-level emphasis (LZHGE), gray-level variance (GLV), zone-size variance (ZSV). The total number of these features is also 13 and which was calculated separately for each image. Finally, NGTDM is another feature extraction method commonly used for extracting the second statistical texture features from medical images [36]. The following five features are routinely extracted in the NGTDM method: Coarseness, Contrast, Busyness, Complexity, and Strength.

A total of 56 features (3 global, 22 GLCM, 13 GLRLM, 13 GLSZM and 5 NGTDM) were extracted for each patient image. Then, for each participant, the sum of mentioned features from all CT images was considered as the feature set of that participant.



## Classification methods

Four classifiers were applied as supervised machine learning algorithms to categorize CT images into five stages: normal, early, progressive, peak, and absorption stage of COVID-19 disease. The first method is LDA as a known linear model for classification [37]. In practice, it projects data from  $D$ -dimensional feature space to  $D'$ -dimensional space ( $D > D'$ ) which maximize the distinctions between classes along with reduction of the variance within them. The second method is a feed-forward neural network (NN) which is designed to solve problems that are not linearly separable [38]. It consists of the input, output, and hidden layers. The input layer receives the input features and the output layer is performed the required classification. An arbitrary number of hidden layers that are placed in between the input and output layer are the main computational machine of the NN. The neurons in this method are trained with the back propagation learning algorithm. The third method is SVM as an accurate and widely used algorithm in classification [39]. SVM tries to find an optimal hyperplane between classes to provide the maximum separating between classes to be distinguished well. The SVM algorithm tries to choose among an infinite number of linear decision boundaries based on the principle of structural risk minimization. The last method is the RF method. The RF algorithm is easy to use and one of the most successful classifiers for multi-class problem classification. In most times, it achieves excellent results compared to other commonly used classification methods [40, 41]. The decision of RF was determined to be the optimal outcome by a majority vote between the results of the different trees, which responded independently.

## Evaluation metrics

Due to the limited number of cases in the dataset, the 10-fold cross-validation technique was used for partitioning the data. In this method, the dataset is divided into 10 folds, from which 10-1 folds are the training data and the left-out fold is the testing data. This is repeated until all data are evaluated

as a test data for one time. This would mean we have 10 different models for each classifier at the end and we have reported the mean and standard deviation of their results. As the experiment step, the total dataset was split into three sets randomly including training, validation, and testing sets which included 80%, 10%, and 10% of sample cases in the total dataset, respectively. The training data was used to develop an appropriate model which was modified using the validation data. In other words, the validation set is used to tune the model parameters like sigma in SVM and numbers of trees in the RF classifier. The performance of the developed model was calculated using the test data based on the accuracy metric. Therefore, in each fold, 90% percentages (80% training and 10% validation) of sample cases are used for determining the best parameters of the model and 10% percentages of sample cases are used for testing the performance of the developed model. It should be noted that the division between sets is made at the participant level. For example, in the Early stage with 82 people, 80% (66 people), 10% (8 people), and 10% (8 people) of datasets are used for training, validation, and testing sets. For each classifier (RF, LDA, ANN, and SVM), average ( $\pm$  standard deviation) performance measure into five stages over 10 different models is reported. Finally, the confusion matrix was calculated to show the performance of the proposed algorithm for each defined class compared to the reference label.

## Results

The results of RF, LDA, ANN, and SVM algorithms for the prepared data are presented in Table 1. The data in this table are in terms of values of accuracy. The parameters of these classifiers were optimized by applying the trial and error method. The results of the selection of the best number of trees in the RF algorithm are presented in Fig. 2. The optimum number of trees is 65. It is demonstrated that the RF classifier performed more powerfully than the others and achieved the highest accuracy with a value of 93.55% in the detection of the five stages of COVID-19 disease from CT images compared to the other mentioned

**Table 1** The results of classification accuracy using RF, SVM, LDA, and ANN algorithms in terms of accuracy for the combination of 56 extracted features (Global, GLCM, GLRLM, GLSZM, and NGTDM)

Algorithm	Overall accuracy (%)	Stage accuracy (%)				
		Normal	Early stage	Progressive stage	Peak stage	Absorption stage
LDA	80.10 ( $\pm$ 01.83)	89.83 ( $\pm$ 01.23)	38.64 ( $\pm$ 03.78)	100.00 ( $\pm$ 00.00)	53.39 ( $\pm$ 03.11)	95.50 ( $\pm$ 01.04)
SVM	90.20 ( $\pm$ 01.27)	93.53 ( $\pm$ 01.03)	69.33 ( $\pm$ 02.56)	100.00 ( $\pm$ 00.00)	82.19 ( $\pm$ 01.75)	93.94 ( $\pm$ 01.04)
NN	91.10 ( $\pm$ 02.03)	98.40 ( $\pm$ 00.95)	63.75 ( $\pm$ 04.45)	100.00 ( $\pm$ 00.00)	74.70 ( $\pm$ 02.96)	94.06 ( $\pm$ 01.83)
RF	93.55 ( $\pm$ 01.08)	96.25 ( $\pm$ 00.89)	74.39 ( $\pm$ 01.57)	100.00 ( $\pm$ 00.00)	82.19 ( $\pm$ 01.56)	96.00 ( $\pm$ 01.42)

For each model, average ( $\pm$  standard deviation) performance measure is reported

algorithms. The data in this table show that the RF classifier had acceptable accuracies with values of 96.25% in normal, 74.39% in early stage, 100% in progressive stage, 82.19% in peak stage, and 96% in the absorption stage of COVID-19 diseases. Finally, the confusion matrix of the RF algorithm was achieved and presented in Table 2. It should be noted, the average time spent on processing the proposed algorithms in MATLAB for each test data on a laptop computer (CPU: i7-3.5 GHz, RAM: 16 GB) is 300 s which is a good time in clinical applications.

## Discussion

In the present study, in addition to diagnosing whether or not volunteers have COVID-19 disease, classification was performed to assess the stage of their disease, based on a radiomics and machine learning approach. This classification was based on CT images collected from the lungs of the people. A patient infected by COVID-19 disease can be classified into one of the five stages. These stages are: normal, early, progressive, peak, and absorption stage. The classification process which was presented in the present study showed that different stages of COVID-19 disease can be classified with good accuracy by radiomics and machine learning methods.

In clinical application, this approach is useful for the early determination of the stage of the disease. The importance of classification based on these stages is to determine the type of treatment process required for patients in the category and to reduce the cost of maintenance and treatment for hospitalized people, given the rapid spread of the disease worldwide. Moreover, the use of this processing approach helps in deciding the time period required for hospitalization of the patient. These assessments can also

be useful tools for medical staff in caring for patients with this epidemic disease.

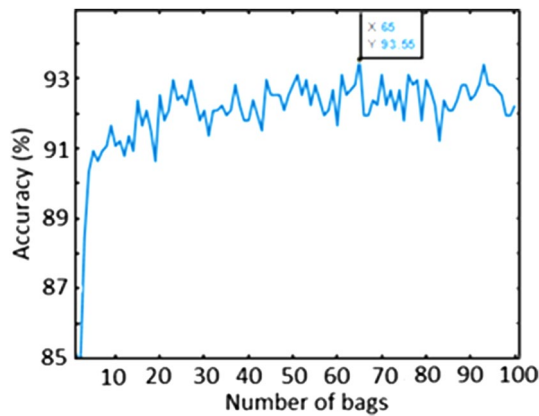
The collected CT images were classified by RF, LDA, ANN, and SVM algorithms in the determination of the stage of COVID-19 disease, and the results were compared with each other for the same selected numbers of features (Table 1). The four evaluated algorithms have resulted in good outcomes for many studies due to their high interpretability and simplicity in understanding performance. However, the comparison results showed that the RF algorithm has more accurate results, in terms of accuracy criteria and is accompanied by higher outputs. The RF algorithm, in general, always plays a dominant role in classifying different data. In other words, the rate of false diagnosis in the RF algorithm is to some extent lower. The reason for this effect can be due to the random batches in different classifiers. In fact, this algorithm, in the form of a continuous “if” and “then” batch algorithm, evaluates the condition of comparison between different features in successive classes, and at the end of the generated tree, leads to one of the desired classes. This process is checked for random sets of features collected in different trees, and the final result is determined by majority voting. However, the SVM algorithm compares the degree of separation of different classes relative to each other. In fact, this algorithm is a two-class algorithm that considers each class relative to the data of the other classes and calculates their separation. Therefore, it cannot accurately distinguish between four different classes in some cases. This issue justifies the lower accuracy of this algorithm compared to the RF algorithm and can be observed based on the data presented in Table 1.

For the RF algorithm, the number of different trees was examined to achieve the optimal number of trees. Based on the results of the selection of the best number of trees (which are presented in Fig. 2), the optimum number of trees in terms of accuracy is 65. The number of trees in the RF algorithm is needed to be tuned. Therefore, the classification error for different numbers of trees was calculated. As the number of trees increases, more decision units participate in voting and the performance of the classifier improves. The optimum number is where adding more trees does not significantly improve the result.

Based on the data in Table 1 and the confusion matrices of the RF algorithm in Table 2, the RF classifier had acceptable accuracies with values of 96.25% in normal, 74.39% in early stage, 100% in progressive stage, 82.19% in peak stage, and 96% in absorption stage of COVID-19 diseases. As it is evident from this table, people in the progressive and early stages are categorized with higher and lower accuracy, respectively. The higher accuracy for the progressive stage may be justified by the fact that the histological patterns in lung tissue are completely different in the progressive state, compared to the other stages.

**Table 2** Comparison of estimation of labels provided by RF algorithm with all combination of features against those assigned by the reference one (radiologist)

	Reference classifier (radiologist)				
	Normal	Early stage	Progressive stage	Peak stage	Absorption stage
Estimated labels by RF algorithm					
Normal	308	4	0	0	1
Early stage	6	61	0	12	1
Progressive stage	0	0	108	0	0
Peak stage	3	16	0	60	2
Absorption stage	3	1	0	1	96



**Fig. 2** The results of the selection of the best number of trees in the RF algorithm with all combinations of features. The optimum number of trees is 65

However, due to the correlation between the physical characteristics of the different stages of the disease, as mentioned earlier, the accuracy which was obtained in some stages (especially the early stage) seems to be slightly lower. In other words, the early stage includes blurriness of ground glass opacities and some squamous and consolidation tissues. In the progressive stage, the squamous tissue is observed to a higher extent and in the peak stage, the consolidation pattern is being dominant. Therefore, a small share of each of these features can lead to discrepancies in the identification of the stages. There is a need for additional information from the patients studied, such as clinical information and history of other diseases (such as diabetes, obesity, etc.) or the results of other tests of patients such as serological tests to determine the stage of the disease more accurately. For example, for a normal person who has been exposed to respiratory pollution for a long time in his normal life, there are observable signs of grounded glass opacity and he can be classified in an irrelevant class. Because, the general features of this normal test subject can be associated with the other stages and cause inaccuracies in the diagnosis process.

The obtained results in the present study show that these algorithms can achieve remarkably good accuracies. Table 3 shows a comparison between the results of this manuscript with other studies [31, 32, 42–44]. Compared to other studies, this work has some advantages such as combination of statistical texture features of all 3D CT slices of a patient and using RF classifier. Increasing the number of subjects and a multi-site investigation to validate and improving the results is important at future work. Therefore, it is suggested that in future studies a higher number of cases be evaluated as a subject for radiomics determination of stage of COVID-19 disease. Additionally, the application of deep-learning

**Table 3** Comparison of the results of other studies on classification of severity of COVID-19 patients in compared with the presented approach from CT images

Study	Methods	Results (accuracy)
Huang et al. [31]	2D U-Net deep learning	N/A.
Shan et al. [32]	Deep learning	85.1%
Tang et al. [42]	Volume of GGO and RF probabilities	87.5%
Shen et al. [43]	Non trainable CV	N/A.
Shi et al. [44]	COV with RF	87.9%
Our work	RF with first and second statistical features	93.55%

methods for this purpose and more algorithms can be useful in future studies.

## Conclusions

In the present study, a classification method based on lung CT images was presented, and it was shown that the different stages of COVID-19 disease can be classified with good accuracy by radiomics and machine learning. The presented classifier, using the algorithm RF and a combination of first- and second-order statistical texture features, had an accuracy of 96.25% in the normal stage, 74.39% in the early stage, 100% in the progressive stage, 82.19% in the peak stage, and 96% in the absorption stage of COVID-19 disease. Therefore, in clinical application, the results of this investigation can be helpful in triaging the patients and influence on clinician's decision on referral patients as soon as possible.

**Acknowledgements** Sabzevar University of Medical Sciences has financially supported this study.

**Funding** This work was approved by the Ethics Committee of Sabzevar University of Medical Sciences (with ethical code number of IR.MEDSAB.REC.1399.152).

## Declarations

**Conflict of interest** There is not any potential conflict of interest.

## References

- Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A et al (2020) World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). *Int J Surg* 76:71–76
- Pan L, Mu M, Yang P, Sun Y, Wang R, Yan J et al (2020) Clinical characteristics of COVID-19 patients with digestive symptoms in Hubei, China: a descriptive, cross-sectional, multicenter study. *Am J Gastroenterol* 115(5):766–773

3. Alsharif W, Qurashi A (2021) Effectiveness of COVID-19 diagnosis and management tools: a review. *Radiography* 27(2):682–687
4. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P et al (2020) Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* 296(2):E115–E17
5. Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Ning Zhang N (2020) Chest CT findings in coronavirus disease-19 (covid-19): relationship to duration of infection. *Radiology* 295(3):200463
6. Pontone G, Scafuri S, Mancini ME, Agalbato C, Guglielmo M, Baggiano A et al (2021) Role of computed tomography in COVID-19. *J Cardiovasc Comput Tomogr* 15(1):27–36
7. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR et al (2017) Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology* 284(1):228–243
8. Koo HJ, Lim S, Choe J, Choi SH, Sung H, Do KH (2018) Radiographic and CT features of viral pneumonia. *Radio Graphics* 38(3):719–739
9. Pan F, Ye T, Sun P, Gui S, Liang B, Li L et al (2020) Time course of lung changes at chest CT during recovery from coronavirus disease 2019 (COVID-19). *Radiology* 295(3):715–721
10. Li M, Lei P, Zeng B, Li Z, Yu P, Fan B et al (2020) Coronavirus disease (COVID-19): spectrum of CT findings and temporal progression of the disease. *Acad Radiol* 27(5):603–608
11. Huang G, Gong T, Wang G, Wang J, Guo X, Cai E et al (2020) Timely diagnosis and treatment shortens the time to resolution of coronavirus disease (COVID-19) pneumonia and lowers the highest and last CT scores from sequential chest CT. *Am J Roentgenol* 215(2):367–373
12. Al-Tameemi -Muayad SO, Alghazali H, Abed M, Abu Ghniem M, Hawiji MN (2020) Pulmonary CT manifestations of COVID-19: changes within 2 weeks duration from presentation. *Egypt J Radiol Nucl Med* 51(1):105
13. Zhang Z, Tang R, Sun H, Dai H, Chen K, Ye X et al (2020) Temporal lung changes on thin-section CT in patients with COVID-19 pneumonia. *Sci Rep* 10:19649
14. Holloway L, Bezak E, Baldock C (2021) Artificial intelligence (AI) will enable improved diagnosis and treatment outcomes. *Phys Eng Sci Med* 44:603–606
15. Ketabi H, Ekhlasi A, Ahmadi H (2021) A computer-aided approach for automatic detection of breast masses in digital mammogram via spectral clustering and support vector machine. *Phys Eng Sci Med* 44:277–290
16. Sani ZA, Shalhaf A, Behnam H, Shalhaf R (2015) Automatic computation of left ventricular volume changes over a cardiac cycle from echocardiography images by nonlinear dimensionality reduction. *J Digit Imaging* 28(1):91–98
17. Shalhaf A, Behnam H, Alizade-Sani Z, Shojaifard M (2012) Left ventricle wall motion quantification from echocardiographic images by non-rigid image registration. *Int J Comput Assist Radiol Surg* 7(5):769–783
18. Mirmohammadi P, Ameri M, Shalhaf A (2021) Recognition of acute lymphoblastic leukemia and lymphocytes cell subtypes in microscopic images using random forest classifier. *Phys Eng Sci Med* 44:433–441
19. Sahoo P, Roy I, Ahlawat R et al (2021) Potential diagnosis of COVID-19 from chest X-ray and CT findings using semi-supervised learning. *Phys Eng Sci Med* 45:31–42
20. Gillman AG, Lunardo F, Prinable J et al (2021) Automated COVID-19 diagnosis and prognosis with medical imaging and who is publishing: a systematic review. *Phys Eng Sci Med* 45:13–29
21. Ouyang X, Huo J, Xia L, Shan F, Liu J, Mo Z et al (2020) Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia. *IEEE Trans Med Imaging* 39(8):2595–2605
22. Lin L, Liu J, Deng Q, Li N, Pan J, Sun H et al (2021) Radiomics is effective for distinguishing coronavirus disease 2019 pneumonia from influenza virus pneumonia. *Front Public Health* 9:682
23. Yang N, Liu F, Li C, Xiao W, Xie S, Yuan S et al (2021) Diagnostic classification of coronavirus disease 2019 (COVID-19) and other pneumonias using radiomics features in CT chest images. *Sci Rep* 11(1):1–7
24. Kadry S, Rajinikanth V, Rho S, Raja NSM, Rao VS, Thanaraj KP (2020) Development of a Machine-Learning System to Classify Lung CT Scan Images into Normal/COVID-19 Class. *arXiv preprint. arXiv:2004.13122*
25. Fu L, Li Y, Cheng A et al (2020) A novel machine learning-derived radiomic signature of the whole lung differentiates stable from progressive COVID-19 infection: a retrospective cohort study. *J Thorac Imaging* 16(6):361–368
26. Tan HB, Xiong F, Jiang YL et al (2020) The study of automatic machine learning base on radiomics of non-focus area in the first chest CT of different clinical types of COVID-19 pneumonia. *Sci Rep* 10:8926
27. Wang H, Wang L, Edward H et al Decoding COVID-19 pneumonia: comparison of deep learning and radiomics CT image signatures. *Eur J Nucl Med Mol Imaging* 2020:1–9
28. Guiot J, Vaidyanathan A, Deprez L et al (2020) development and validation of an automated radiomic CT signature for detecting COVID-19. *Diagnostics (Basel)* 11(1):41
29. Wang L, Kelly B, Lee EH et al (2021) Multi-classifier-based identification of COVID-19 from chest computed tomography using generalizable and interpretable radiomics features. *Eur J Radiol* 136:109552
30. Liu H, Ren H, Wu Z et al (2021) CT radiomics facilitates more accurate diagnosis of COVID-19 pneumonia: compared with CO-RADS. *J Transl Med* 19(1):29
31. Huang L, Han R, Ai T, Yu P, Kang H, Tao Q et al (2020) Serial quantitative chest ct assessment of covid-19: deep-learning approach. *Radiol Cardiothorac Imaging* 2(2):e200075
32. Shan F, Gao Y, Wang J, Shi W, Shi N, Han M et al (2020) Lung infection quantification of COVID-19 in CT images with deep learning. *arXiv preprint. arXiv:2003.04655*
33. Novitasari DCR, Lubab A, Sawiji A et al (2019) Application of feature extraction for breast cancer using one order statistic, GLCM, GLRLM, and GLDM. *Adv Sci Technol Eng Syst J* 4(4):115–120
34. Vallières M, Freeman CR, Skamene SR et al (2015) A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys Med Biol* 60(14):5471–5496
35. Zhang H, Hung C-L, Min G et al (2019) GPU-accelerated GLRLM algorithm for feature extraction of MRI. *Sci Rep* 9(1):1–13
36. Amadasun MR (1989) Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern* 19:1264–1274
37. Balakrishnama S, Ganapathiraju A (1998) Linear discriminant analysis—a brief tutorial. *Inst Signal Inf Process* 18:1–8
38. Bellingegni AD, Gruppioni E, Colazzo G et al (2017) NLR, MLP, SVM, and LDA: a comparative analysis on EMG data from people with trans-radial amputation. *J Neuroeng Rehabil* 14(1):82
39. Mathur A, Foody GM (2008) Multiclass and binary SVM classification: implications for training and classification users. *IEEE Geosci Remot Sens Lett* 5(2):241–245
40. Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
41. Criminisi A, Shotton J, Konukoglu E (2012) Decision forests: a unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Found Trends Comput Graph Vis* 7(2–3):81–227
42. Tang Z, Zhao W, Xie X, Zhong Z, Shi F, Liu J, Shen D (2020) Severity assessment of coronavirus disease 2019 (COVID-19)



- using quantitative features from chest CT images. arXiv preprint. [arXiv:2003.11988](https://arxiv.org/abs/2003.11988)
43. Shen C, Yu N, Cai S, Zhou J, Sheng J, Liu K, Zhou H, Guo Y, Niu G (2020) Quantitative computed tomography analysis for stratifying the severity of coronavirus disease 2019. *J Pharm Anal* 10(2):123–129
  44. Shi F, Xia L, Shan F, Song B, Wu D, Wei Y, Yuan H, Jiang H, He Y, Gao Y, Sui H (2021 Mar) Large-scale screening to distinguish between COVID-19 and community-acquired pneumonia using infection size-aware classification. *Phys Med Biol* 17(6):065031

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