Diagnostic classification of coronavirus disease 2019 (COVID-19) and other pneumonias using radiomics features in CT chest images

Ning Yang
Guangdong Second Provincial General Hospital, Guangzhou, China
Faming Liu
Xiao Chang First People's Hospital, Hubei, China
Chunlong Li
Guangdong Medical University , China
Wenqing Xiao
Guangdong Second Provincial General Hospital, Guangzhou, China
Shuangcong Xie
Guangdong Second Provincial General Hospital, Guangzhou, China
Shuyi Yuan
Guangdong Second Provincial General Hospital, Guangzhou, China
Wei Zuo
Guangdong Second Provincial General Hospital, Guangzhou, China
Xiaofen Ma (xiaofenma12@163.com)
Guangdong Second Provincial General Hospital, Guangzhou, China
Guihua Jiang (jianggh@gd2h.org.cn)
Guangdong Second Provincial General Hospital, Guangzhou, China

Research Article

Keywords: COVID-19, pneumonias, radiomics, support vector machine

DOI: https://doi.org/10.21203/rs.3.rs-34648/v1

Abstract

We propose a classification method using the radiomics features of CT chest images to identify patients with coronavirus disease 2019 (COVID-19) and other pneumonias. The chest CT images of two groups of participants (90 COVID-19 patients and 90 other pneumonias patients) were collected, and the two groups of data were manually drawn to outline the region of interest (ROI) of pneumonias. The radiomics method was used to extract textural features and histogram features of the ROI and obtain a radiomics features vector from each sample. Finally, using the radiomics features as an input, a support vector machine (SVM) model was constructed to classify patients with COVID-19 and patients with other pneumonias. This model used 20 rounds of 10-fold cross-validation for training and testing. In the COVID-19 patients, correlation analysis (multiple comparison correction—Bonferroni correction, $p<0.05/7$) was also conducted to determine whether the textural and histogram features were correlated with the laboratory test index of blood, i.e., blood oxygen, white blood cell, lymphocytes, neutrophils, C-reactive protein, hypersensitive C-reactive protein, and erythrocyte sedimentation rate. The results showed that the proposed method had a classification accuracy as high as 88.33%, sensitivity of 83.56%, specificity of 93.11%, and an area under the curve of 0.947. This proved that the radiomics features were highly distinguishable, and this SVM model can effectively identify and diagnose patients with COVID-19 and other pneumonias. The correlation analysis results showed that some texture features were positively correlated with WBC, NE, and CRP and also negatively related to SPO2H and NE.

Introduction

The coronavirus disease 2019 (COVID–19) [1], [2] epidemic began in Wuhan, Hubei Province, China, in December 2019. Other cases in China and other countries soon followed. The main clinical symptoms are fever and fatigue. The respiratory symptoms are mainly dry cough and difficulty in breathing. In severe cases, acute respiratory distress syndrome, septic shock, and metabolic acidosis are difficult to correct and out of coagulopathy. Some patients have mild onset symptoms without fever. Most cases are mild to moderate with a good prognosis; a few patients are critically ill and die. The early manifestations of CT chest imaging [3]-[7] include small-scale interstitial changes, which are evident in the lateral field of the lung. These further develop to multiple ground-glass opacity (GGO) and infiltration in the lung; lung parenchyma may be involved in severe cases.

Nucleic acid testing is the gold standard for the final diagnosis of COVID–2019 in non-invasive diagnosis. However, due to an inadequate supply of kits and complicated sampling methods, there are sure to be false negatives, which causes some patients to delay treatment and control measures. On March 3, 2020, Chinese experts revised the diagnosis and treatment plan on the basis of analysis, research, and summary of the preliminary medical treatment work, and formed the “COVID- 19 Pneumonia Diagnosis and Treatment Plan (Trial Version 7).” This protocol emphasized the importance of CT imaging for the diagnosis of COVID–19. HRCT is currently the main imaging inspection method for screening and diagnosis. Its role is to find lesions and evaluate the size, nature, scope, and dynamic evolution of lesions. The use of CT images combined with the patient’s epidemiological exposure history

is an important basis for screening and diagnosis of COVID–2019. This can be used for the screening, diagnosis, and identification of early COVID- 2019 suspected cases.

However, the diagnostic capability of artificial image reading is affected by factors such as seniority and experience; reading images with a large number is time-consuming. The existence of these factors inevitably has great subjectivity. In recent years, artificial intelligence (AI) has developed in many areas around the world. It is an effective tool for monitoring and responding to this global epidemic. The role of AI in combating COVID- 19 is the earliest detection of potential infectious diseases. Many research teams have used deep learning to process reports from new cases, the Centers for Disease Control, and the World Health Organization, air routes, and other data.

Zhong Nan Shan's team [8] used an AI model (recurrent neural network (RNN)) to infer the epidemic trend of the epidemic in China. This effectively controlled the development of the epidemic. Recently, Zhongnan Hospital of Wuhan University worked with Shanghai United Imaging Intelligence Company to build and operate the United Imaging Cloud uAI platform. This uAI platform uses the VB-Net model [9], [10] to automatically segment and quantify the infected area in the chest CT scan and the entire lung. However, these AI technologies used deep learning technology, which required large data sets (tens of thousands of samples) for model training. Such large data sets are difficult to obtain and are expensive and time-consuming. For smaller data sets, traditional machine learning algorithms are usually better than deep networks. Therefore, this paper proposed a classification method based on the traditional machine learning method, i.e., a support vector machine, that uses the radiomic features of CT chest images to identify and diagnose patients with COVID–19 and non-Corona Virus Disease 2019 pneumonias (other pneumonias).

Materials And Methods

This study was approved by the ethics committee of Guangdong Second Provincial General Hospital and all participants provided written informed consent after they were provided a complete description of the study.

Here, we present a support vector machine (SVM) method for the classification of patients with COVID–19 and patients with other pneumonias via a radiomics framework. The workflow of our proposed method is shown in Figure. 1. First, lung infection areas (region of interest, ROI) in the CT images of each sample were artificially delineated. Second, thirty-two texture features and five histogram features were extracted from ROI data using a quantitative radiomics features model. Finally, a SVM classifier trained using such quantitative radiomics features from training data was used to distinguish COVID–19 patients and other pneumonias patients. The detailed methodology behind each step of the proposed method is described below.

- Participants and Data Acquisition

Ninety patients with COVID–19 (56 males, 34 females; mean ± standard deviation age, 45.36 ± 11.58 years) were recruited, and 90 patients with other pneumonias (COVID–19-negative; 58 males, 32 females; mean ± standard deviation age, 46.54 ± 8.40 years) were recruited as a control group.

Chest CT images of all participants were acquired using a 16- slice CT (Philips). All chest CT images were acquired in about 2 min using a helical scan of the chest as follows: reconstruction slice thickness = 2 mm; reconstruction slice increment = 2 mm. The CT volume was composed of 98–165 slices with 512 × 512 pixels.

- Radiomics Analysis

Radiomics texture analysis has been proposed since the early 1980s as a method for extracting relevant information representing tissue types from various medical images. Previous studies [11], [12] hypothesized that texture features can reflect heterogeneity within tumors, which is of great significance in cancer research. Texture analysis is a key component of radiology [13].

A gray level co-occurrence matrix (GLCM) [14] considers the arrangement of voxel pairs to calculate the texture index. GLCM is calculated from 13 different directions in 3D with a δ-voxel distance ($\|\vec{d}\|$) relationship between adjacent voxels. The index value is the average of the indexes in the 13 directions of the space (X, Y, Z). From this matrix, seven textural indices (homogeneity, energy, contrast, correlation, entropy_log10, entropy_log2, and dissimilarity) are computed. The gray run length matrix (GLRLM) [15] gives the size of the uniform run for each gray level. The matrix is calculated for 13 different directions in 3D (4 in 2D). Eleven texture indices are computed from this matrix: Short-Run Emphasis, Long-Run Emphasis, Low Gray-level Run Emphasis, High Gray-level Run Emphasis, Short-Run Low Gray-level Emphasis, Short- Run High Gray-level Emphasis, Long-Run Low Gray-level Emphasis, Long-Run High Gray-level Emphasis, Gray-Level Non-Uniformity for run, Run Length Non-Uniformity, and Run Percentage.

The neighborhood gray level difference matrix (NGLDM) [16] corresponds to the gray level difference (8 in 2D) of a voxel and its 26 neighborhoods in three dimensions. Three texture indices (coarseness, contrast, and busyness) are computed from this matrix. The Gray Level Zone Length Matrix (GLZLM) [17] provides information about the uniform zone size of each gray level in 3 dimensions (or 2D). Eleven texture indices are computed from this matrix: Short-Zone Emphasis, Long-Zone Emphasis, Low Gray-level Zone Emphasis, High Gray-level Zone Emphasis, Short-Zone Low Gray-level Emphasis, Short- Zone High Gray-level Emphasis, Long-Zone Low Gray-level Emphasis, Long-Zone High Gray-level Emphasis, Gray-Level Non-Uniformity for zone, Zone Length Non-Uniformity, and Zone Percentage.

All texture analysis processes in this article were performed on the LIFEx (Local Image Features Extraction) platform [18]: Three attending physicians with training in imaging delineated the lung infection area (region of interest, ROI) of each slice in the CT image of each sample. The senior physician was responsible for reviewing and modifying; finally, a three- dimensional ROI region was obtained in each CT image (Figure. 2).

The voxel size was then spatially resampled to 1 mm × 1 mm × 0.5 mm for a 3D ROI in each CT image of all participants. The initial voxel values were resampled into 256 grey levels and rescaled between mean–3*Sd - mean+3*Sd of the ROI content, where mean and Sd are the mean and standard deviation of the voxels included in the ROI, respectively. Eventually, the 32 texture features described above were calculated from each ROI of the participants. We also built a histogram of each CT image and calculated five radiomic histogram features related to histogram skewness, kurtosis, and entropy.

- Diagnostic Classification

This study used a machine learning method—support vector machine (SVM). The concept of SVM was first proposed by Vapnik and Cortes [19] in 1995. It is based on the statistical VC dimension theory and the principle of structural risk minimization. It has many advantages in studies of small sample size with nonlinear and high-dimensional pattern recognition problems. The SVM finds a hyperplane that maximizes the distance between the two types of sample points closest to the hyperplane and the hyperplane.

After calculating textural and histogram features in each sample, we obtained a feature matrix (180×37) where 180 is the number of subjects (including 90 patients with COVID–19 and 90 patients with other pneumonias), and 37 is the number of extracted textural and histogram features. Using the feature matrix as input, SVM with different kernels (Linear, Radial Basis Function (RBF), Polynomial (Poly), and Sigmoid) was developed to train a machine learning model for classification in COVID–19 patients and other pneumonias patients. These classification models used a 10-fold cross-validation method for training and testing. The training samples had an inner 10- fold CV for tuning the penalty coefficient C (fault tolerance). This process was repeated 20 times, and the average of 20 rounds of 10-fold CV test results (accuracy, sensitivity, specificity, and area under ROC curve (AUC)) was used as the final SVM classification performance. All machine learning processes for training and testing used PyCharm (http://www.jetbrains.com/pycharm/, JetBrains PyCharm Community Edition 2018.2.4 x64).

- Statistical Analysis and Correlation Analysis

The demographic data for all participants were analyzed using SPSS 22. Differences in age between COVID–19 patients and other pneumonias patients were compared using the Wilcoxon rank-sum tests. Gender differences were assessed via chi- squared tests.

Nonparametric permutation tests estimated the statistical significance of average classification performance by determining whether the average classification performances exceeded the level of opportunity. The class labels of the training data were randomly ranked 1,000 times before training, and the 20 rounds of 10-fold CV procedure were repeated. The P value of the permutation test was defined as: P = (Nexceeds + 1)/(Nsubstitution + 1). Here, Nexceeds represents the number of times the permuted performance exceeded the one obtained for the true labels. The Nsubstitution represents the rounds of permutation.

In the COVID–19 patients, correlation analysis was also conducted to determine whether the textural and histogram features correlated with the laboratory test index of blood, i.e., blood oxygen (SPO2H), white blood cell count (WBC), lymphocytes (LYM), neutrophils (NE), C-reactive protein (CRP), hypersensitive C-reactive protein (hs-CRP), and erythrocyte sedimentation rate (ESR)

Results

- Demographic Data Results

Demographic data results shows that there were no significant differences between the COVID–19 patients and the control group in terms of age (P = 0.61) and gender (P = 0.83).

- Diagnostic Classification Results

The textural and histogram features led to a feature matrix of (180×37), where 180 is the number of subjects (90 COVID–19 patients and 90 other pneumonias patients), and 37 is the number of extracted textural and histogram features. Using the feature matrix as the input, SVM with different kernels (Linear, RBF, Poly, and Sigmoid) was developed to train machine learning models for classification in COVID–19 patients and other pneumonias patients. There were 20 rounds of 10-fold CV procedures for training and testing (Figure. 3). The ROC (receiver operating characteristic) curve of all SVM models prepared based on these different kernels had good performance. The Sigmoid Kernel-SVM model showed the best discrimination with an accuracy of 88.33%, sensitivity of 83.56%, specificity of 93.11%, and AUC of 0.947 (all P < 0.001). RBF Kernel-SVM model also showed good discrimination with an accuracy of 87.56%, sensitivity of 82.78%, specificity of 92.33%, and AUC of 0.939 (all P < 0.001). The results suggested that the textural and histogram features between COVID–19 patients and other pneumonias patients were highly distinguishable, and the machine learning method achieved excellent classification effects. This contributed to the diagnosis of COVID–19 and other pneumonias.

- Correlation Analysis Results

In the COVID–19 group, correlation analysis (multiple comparison correction—Bonferroni correction, p<0.05/7) was also conducted to determine whether the textural and histogram features were correlated with laboratory assays. The results showed that the NE was positively correlated with “GLCM_Dissimilarity” (the variation of grey-level voxel pairs [14]); “GLRLM_LGRE” (Low Gray-level Run Emphasis, the distribution of the low grey-level runs [15]; “GLRLM_LRHGE” (Long-Run High Gray-level Emphasis, the distribution of the long homogeneous runs with high grey-levels [15]); “GLZLM_SZE” (Short-Zone Emphasis, the distribution of the short homogeneous zones in an image [17]); “GLZLM_LZE” (Long-ZoneEmphasis,thedistributionofthelong homogeneous zones in an image [17]); “GLZLM_SZHGE” (Short-Zone High Gray-level Emphasis, the distribution of the short homogeneous zones with high grey-levels [17]). The NE was negatively correlated with “GLRLM_LRLGE” (Long-Run Low Gray-level Emphasis, the distribution of the long homogeneousrunswithlowgrey-levels[15])and “GLZLM_LGZE” (Low Gray-level Zone Emphasis, the distribution of the low grey-level zones [17]). The “GLRLM_LGRE” (Low Gray-level Run

Emphasis, the distribution of the low grey-level runs [15]) was also positively relative to CRP. The following were positively related to WBC and negatively related to SPO2H: “GLZLM_SZE” (Short-Zone Emphasis, the distribution of the short homogeneous zones in an image [17]); “GLZLM_LZE” (Long-Zone Emphasis, the distribution of the long homogeneous zones in an image [17]); and“GLZLM_SZHGE”(Short-ZoneHighGray-level Emphasis, the distribution of the short homogeneous zones with high grey-levels [17]).

Discussion

In this study, we built an effective COVID–19 diagnosis system based on a traditional machine learning method. Our purpose was to examine whether traditional machine learning could be useful in the diagnosis of patients with COVID–19 and non-COVID–19 pneumonias (other pneumonias). We chose the SVM approach over deep learning because deep learning often requires relatively large training samples to avoid over training and because of a small sample size in this study—traditional machine learning methods can sometimes achieve better performance.

One of the challenges in COVID–19 diagnosis is its different diagnosis with other pneumonias. We collected 90 samples with COVID–19 and 90 samples with other pneumonias as a control group. Thirty-two texture features and five histogram features were calculated from the pneumonia lesions in the CT images of each sample. These radiomics features for SVM model training and testing could distinguish COVID–19 from other pneumonias (accuracy of 88.33%, sensitivity of 83.56%, specificity of 93.11%, and AUC of 0.947). The results showed the potential of our SVM in clinical practice by aiding the radiologist in proper diagnosis.

According to “COVID–19 Pneumonia Diagnosis and Treatment Plan (Trial Version 7),” the total number of WBC in the early stage of COVID–19 is normal or decreased. The early warning indicator for medium and critical adult patients with COVID–19 is a progressive increase in CRP; in addition, symptoms of poor breathing and dyspnea will occur due to the virus invading the lungs, and hypoxia will result in decreased SPO2H. In this study, we found that some textural features were positively correlated with WBC and CRP, and also negatively correlated with SPO2H. This explains why textural features are discriminative in the two groups of COVID–19 patients and other pneumonias patients. We also showed how some textural features were correlated with NE; however, the clinical manifestation of NE in COVID–19 patients remains unclear since this disease is new. Knowledge about its appearance is continually updated. Future research will validate this result.

There are several limitations to our study: 1) limited data set, 2) outlines of lesions (ROI) in CT images are manually labeled layer-by-layer, which is time consuming, 3) a lack of feature engineering—This limitation can eliminate some irrelevant and duplicate features while ensuring the validity of the features. This improves the accuracy of the model and reduces the complexity/training time of the model. 4) This single country study might have racial bias. Future work will solve these limitations by collecting more cases, using deep learning to achieve an automatic outline of pneumonia lesions, and performing feature

extraction and feature selection. In summary, despite these limitations, our results showed that radiomic features can classify COVID–19 patients and other pneumonias patients. The SVM model can achieve an excellent diagnosis of COVID–19.

Declarations

Competing interests: The authors declare no competing interests.

References

1. C. Mattiuzzi, and G. Lippi, “Which lessons shall we learn from the 2019 novel coronavirus outbreak?,” Annals of Translational Medicine., vol. 8, no. 3, Feb. 2020.
2. N. Zhu et al., “A Novel Coronavirus from Patients with Pneumonia in China, 2019,” N Engl J Med., vol. 382, pp. 727 –733, Feb. 2020.
3. J. Q. Lei et al., “CT Imaging of the 2019 Novel Coronavirus(2019-nCoV) Pneumonia,” Radiology., vol. 295, no. 1, pp. 200236, Jan. 2020.
4. M. Chung et al., “CT imaging Features of 2019 Novel Coronavirus (2019- nCoV),” Radiology., vol. 295, no. 1, pp. 200230, Feb. 2020.
5. X. Xie et al., “Chest CT for Typical 2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing,” Radiology., to be published. DOI: 10.1148/radiol.2020200343.
6. D. Wang et al., “Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China”,” JAMA., vol. 323, no. 11, pp. 1061–1069, Mar. 2020.
7. F. Pan et al., “Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID–19) pneumonia,” Radiology., to be published. DOI: 10.1148/radiol.2020200370.
8. Z. F. Yang al., “Modified SEIR and AI prediction of the epidemics trend of COVID–19 in China under public health interventions,” Journal of Thoracic Disease., vol 12, no 3, Mar. 2020.
9. F. Shan et al., “Lung Infection Quantification of COVID–19 in CT Images with Deep Learning,” arXiv., Mar 2020.
10. G. R. Mu, Z. Y. Lin, M. F. Han, G Yao and Y. Z. Gao, “Segmentation of kidney tumor by multi-resolution VB-net,” presented at 2019 Kidney Tumor Segmentation Challenge: KiTS19. Jan, 2019.
11. A. Lejla et al., “Quantification of Heterogeneity as a Biomarker in Tumor Imaging: A Systematic Review,” PloS one., vol. 9, no. 10, pp. e110300, Oct. 2014.
12. R. Fisher et al., “Cancer heterogeneity: implications for targeted therapeutics,” British Journal of Cancer., vol. 108, no. 3, pp. 479–485, Jan. 2013.
13. R. J. Gillies et al., “Radiomics: Images Are More than Pictures, They Are Data,” Radiology., vol. 278, no. 2, pp. 563–577, Feb. 2016.
14. R. M. Haralick et al., “Textural features for image classification,” IEEE Transactions on Systems, Man, and Cybernetics., vol. SMC–3, no. 6, pp. 610–621, Nov. 1973.

15. Xu D-H KA, Furst JD and Raicu DS, “Run-Length Encoding for Volumetric Texture,” presented at the 4th IASTED International Conference on Visualization, Imaging and Image Processing—VIP, Marbella, Spain: 452–458.
16. M. Amadasun et al., “Textural features corresponding to textural properties,” IEEE Transactions on Systems, Man, and Cybernetics., vol. 19, no. 5, pp. 1264–1274, Oct. 1989.
17. G. Thibault, B. Fertil, C. Navarro, S. Pereira and P. Cau, “Texture indexes and gray level size zone matrix: application to cell nuclei classification,” presented at the 10th International Conference on Pattern Recognition and Information Processing. 2009.
18. N. Christophe N et al., “LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity,” Cancer Research., vol. 78, no. 16, pp. 4786–4789, Aug. 2018.
19. C. Cortes and V. Vapnik, “Support-Vector Network,” Machine Learning., vol. 20, no. 3, pp. 273–297, Feb 1995.

Figures

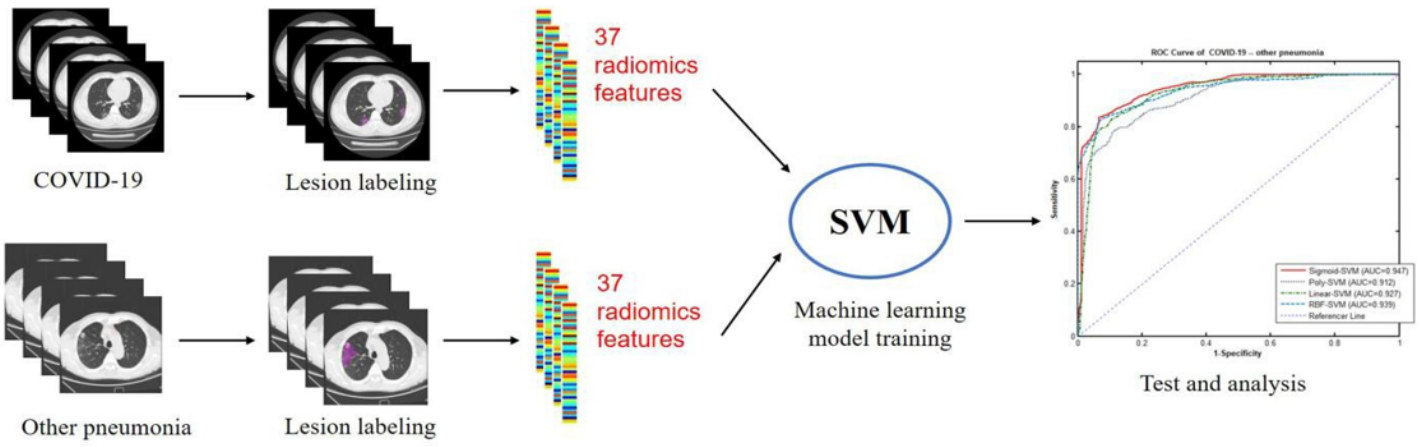

Figure 1

The flow of our proposed machine learning algorithm SVM with radiomics features for classification of COVID-19 and other pneumonias.

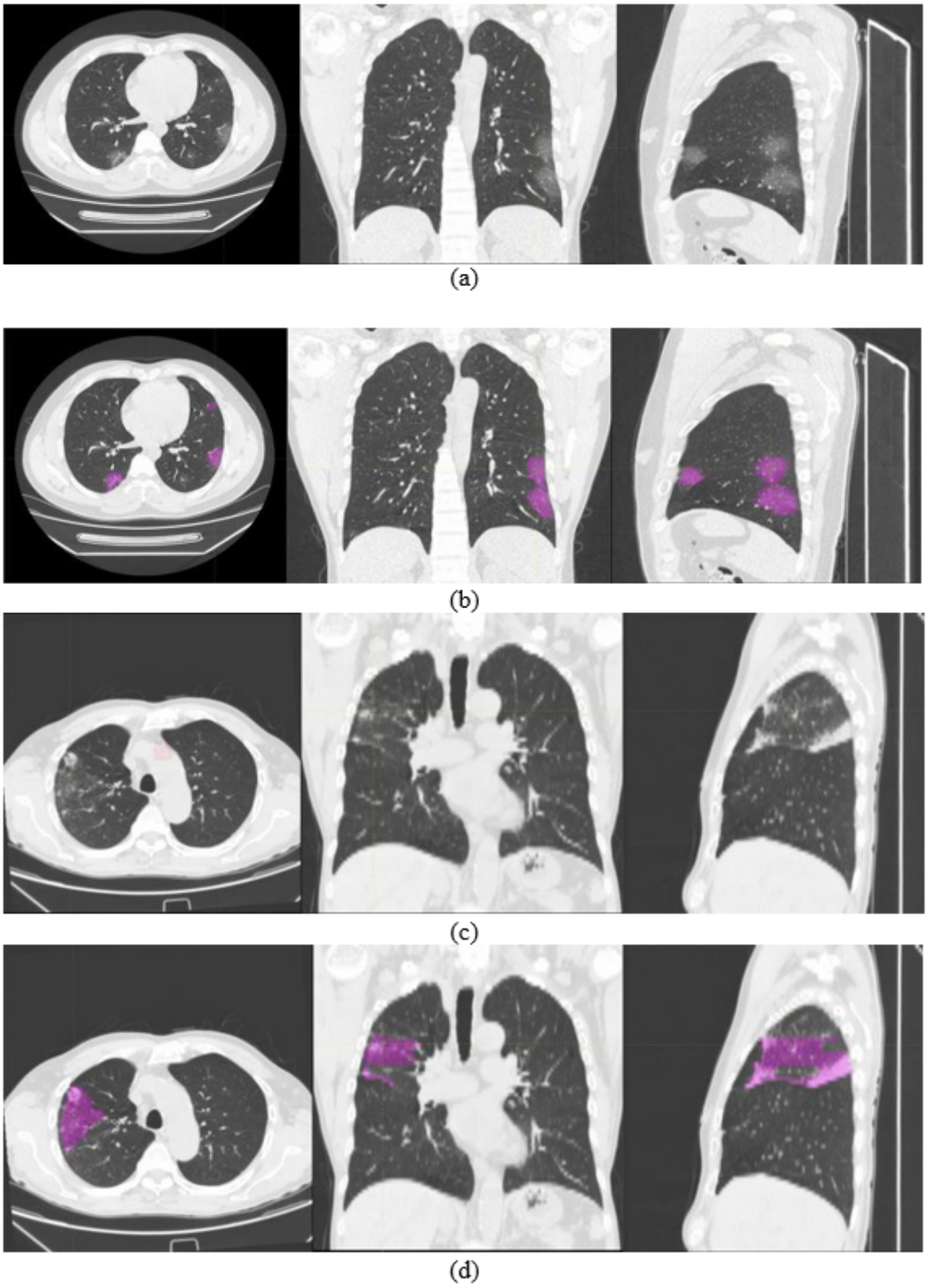

Figure 2

Image examples in different patient groups. From top to bottom: (a) and (b): A 35-year-old male with COVID-19 and COVID-19 lesion labeling (multiple ground-glass opacity, GGO); (c) and (d) A 77-year-old male with other pneumonia and other pneumonia lesion labeling (multiple patches and cloud floccules).

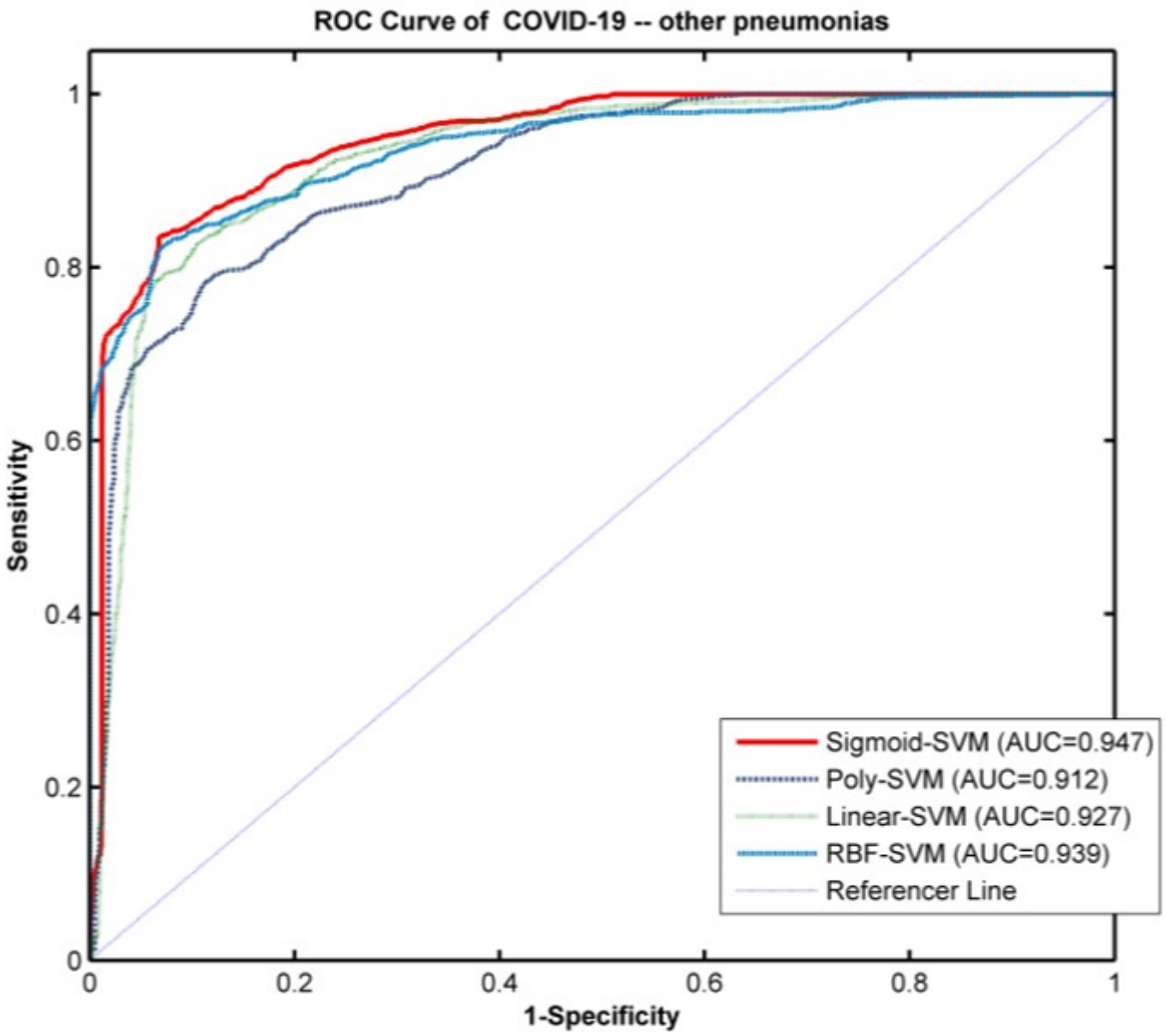

Figure 3

ROC curve of a SVM model based on different kernels for classification of COVID-19 and other pneumonias.