

Original Article
Medical Imaging



Quantitative Assessment of Chest CT Patterns in COVID-19 and Bacterial Pneumonia Patients: a Deep Learning Perspective

Myeongkyun Kang ^{1*}, Kyung Soo Hong ^{2*}, Philip Chikontwe ¹, Miguel Luna ¹, Jong Geol Jang ², Jongsoo Park ³, Kyeong-Cheol Shin ², Sang Hyun Park ¹ and June Hong Ahn ²

¹Robotics Engineering, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Korea

²Division of Pulmonology and Allergy, Department of Internal Medicine, Regional Center for Respiratory Diseases, Yeungnam University Medical Center, College of Medicine, Yeungnam University, Daegu, Korea

³Department of Radiology, Seoul National University Hospital, Seoul, Korea



Received: Nov 13, 2020

Accepted: Jan 21, 2021

Address for Correspondence:

June Hong Ahn, MD, PhD

Division of Pulmonology and Allergy,
Department of Internal Medicine, Regional
Center for Respiratory Diseases, Yeungnam
University Medical Center, College of Medicine,
Yeungnam University, 170 Hyeonchung-ro,
Nam-gu, Daegu 42415, Republic of Korea.
E-mail: fireajh@gmail.com

Sang Hyun Park, PhD

Department of Robotics Engineering,
Daegu Gyeongbuk Institute of Science and
Technology, 333 Techno jungang-daero,
Hyeonpung-eup, Dalseong-gun, Daegu 42988,
Republic of Korea.
E-mail: shpark13135@dgist.ac.kr

*Myeongkyun Kang and Kyung Soo Hong
contributed equally to this work.

© 2021 The Korean Academy of Medical
Sciences.

This is an Open Access article distributed
under the terms of the Creative Commons
Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>)
which permits unrestricted non-commercial
use, distribution, and reproduction in any
medium, provided the original work is properly
cited.

ORCID iDs

Myeongkyun Kang
<https://orcid.org/0000-0002-9165-870X>

ABSTRACT


Background: It is difficult to distinguish subtle differences shown in computed tomography (CT) images of coronavirus disease 2019 (COVID-19) and bacterial pneumonia patients, which often leads to an inaccurate diagnosis. It is desirable to design and evaluate interpretable feature extraction techniques to describe the patient's condition.

Methods: This is a retrospective cohort study of 170 confirmed patients with COVID-19 or bacterial pneumonia acquired at Yeungnam University Hospital in Daegu, Korea. The lung and lesion regions were segmented to crop the lesion into 2D patches to train a classifier model that could differentiate between COVID-19 and bacterial pneumonia. The K-means algorithm was used to cluster deep features extracted by the trained model into 20 groups. Each lesion patch cluster was described by a characteristic imaging term for comparison. For each CT image containing multiple lesions, a histogram of lesion types was constructed using the cluster information. Finally, a Support Vector Machine classifier was trained with the histogram and radiomics features to distinguish diseases and severity.

Results: The 20 clusters constructed from 170 patients were reviewed based on common radiographic appearance types. Two clusters showed typical findings of COVID-19, with two other clusters showing typical findings related to bacterial pneumonia. Notably, there is one cluster that showed bilateral diffuse ground-glass opacities (GGOs) in the central and peripheral lungs and was considered to be a key factor for severity classification. The proposed method achieved an accuracy of 91.2% for classifying COVID-19 and bacterial pneumonia patients with 95% reported for severity classification. The CT quantitative parameters represented by the values of cluster 8 were correlated with existing laboratory data and clinical parameters.

Conclusion: Deep chest CT analysis with constructed lesion clusters revealed well-known COVID-19 CT manifestations comparable to manual CT analysis. The constructed histogram features improved accuracy for both diseases and severity classification, and showed correlations with laboratory data and clinical parameters. The constructed histogram features can provide guidance for improved analysis and treatment of COVID-19.


Keywords: COVID-19; Pneumonia, Bacterial; Computed Tomography; Deep Learning; Cluster Analysis

Kyung Soo Hong 

<https://orcid.org/0000-0003-2931-6861>

Phillip Chikontwe 

<https://orcid.org/0000-0002-6995-2312>

Miguel Luna 

<https://orcid.org/0000-0001-6255-8366>

Jong Geol Jang 

<https://orcid.org/0000-0001-8040-5363>

Jongsoo Park 

<https://orcid.org/0000-0001-7204-6700>

Kyeong-Cheol Shin 

<https://orcid.org/0000-0003-1972-1847>

Sang Hyun Park 

<https://orcid.org/0000-0001-7476-1046>

June Hong Ahn 

<https://orcid.org/0000-0001-7104-8325>

Funding

This research was funded by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIT) (No.2019R1C1C1008727).

Disclosure

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: Kang M, Park SH. Data curation: Hong KS, Jang JG, Park J, Ahn JH. Formal analysis: Kang M. Funding acquisition: Park SH. Investigation: Kang M, Chikontwe P, Luna M. Methodology: Kang M, Chikontwe P, Luna M, Park SH. Project administration: Shin K, Park SH, Ahn JH. Resources: Hong KS, Jang JG, Park J, Ahn JH. Software: Kang M. Validation: Kang M, Chikontwe P, Luna M, Hong KS, Jang JG, Park J, Park SH, Ahn JH. Supervision: Shin K, Park SH, Ahn JH. Visualization: Kang M. Writing - original draft preparation: Kang M, Ahn JH. Writing - review & editing: Kang M, Chikontwe P, Luna M, Hong KS, Park SH, Ahn JH.

INTRODUCTION

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic virus originated in Wuhan, China in 2019, has spread rapidly to several countries.¹ Reverse transcription polymerase chain reaction (RT-PCR) of viral nucleic acid is regarded as the reference standard for the diagnosis of coronavirus disease 2019 (COVID-19), but chest computed tomography (CT) examination is mostly recommended for evaluating severity and treatment efficacy given the primary involvement of the respiratory system. Also, CT imaging can be effective for early screening compared to RT-PCR that has shown low sensitivity for early detection.²⁻⁴ Therefore, there is an urgent need for fast and accurate diagnostic tests other than RT-PCR.

There are well-known features of COVID-19 often observed in CT imaging such as ground-glass opacities (GGOs) distributed in the peripheral or posterior lungs.⁵⁻⁷ However, these patterns are often limited and contribute to the challenge of distinguishing COVID-19 from other pneumonia types.^{8,9} Furthermore, since COVID-19 CT manifestations are often observed with mixed or subtle radiological differences, accurate description is challenging even when referencing Fleischner Society listed terms.¹⁰

Recent advances in artificial intelligence with deep learning have shown success in the medical imaging community given the robust feature extraction capability of deep networks.¹¹ We propose a deep learning-based framework to create accurate descriptions of CT manifestations related to COVID-19. Herein, we analyzed chest CT scans from 73 COVID-19 and 97 bacterial pneumonia patients from Daegu, Korea.

METHODS

Study design and subject

We performed a retrospective cohort study of CT scans of 73 patients with COVID-19 infection obtained between February 2020 to March 2020, and 97 patients with bacterial pneumonia between March 2012 to February 2014 at Yeungnam University Medical Center, in Daegu, Korea. During the study period, all consecutive adult patients (age > 18 years) with SARS-CoV-2 infection admitted to the hospital were eligible for inclusion. SARS-CoV-2 infection was confirmed by real-time RT-PCR assay of nasal and pharyngeal swab samples. Severity was defined as a composite outcome of acute respiratory distress syndrome (ARDS), intensive care unit admission, or death. ARDS was diagnosed according to the Berlin definition.¹² The National Early Warning Score (NEWS) is an early warning score, and is composed of seven parameters: pulse oximetry, oxygen, pulse rate, systolic blood pressure, respiration rate, temperature, and central nervous system status. NEWS showed excellent predictive value in predicting critical clinical outcomes of COVID-19. NEWS is stratified into three categories: low risk (0–4), medium risk (5–6), and high risk (≥ 7).^{13,14} Fig. 1 shows the flow chart of data collection, exclusion and splitting ratios applied for training and evaluation.

Imaging pattern analysis using deep learning

To analyze the COVID-19 manifestations using deep learning, the proposed framework consists of 3 key modules: 1) lung and lesion segmentation, 2) deep feature extraction, and 3) K-means clustering modules, respectively. Fig. 2A shows a diagram of the framework.

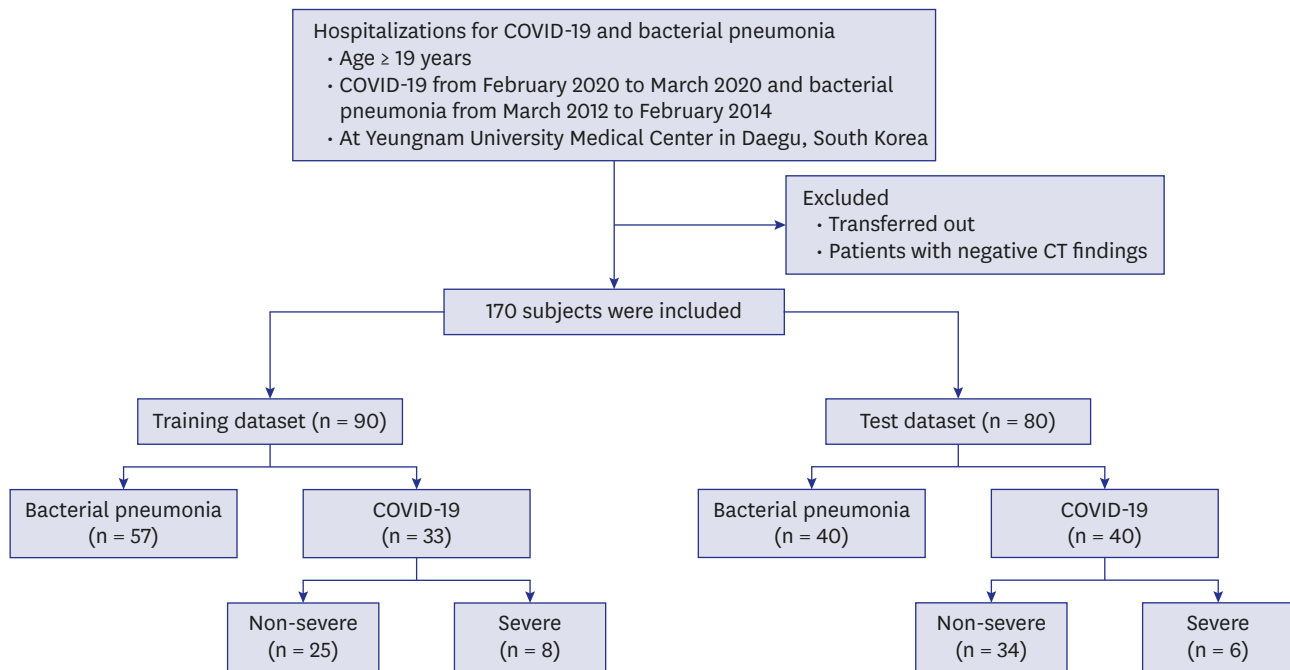


Fig. 1. Flow chart.
COVID-19 = coronavirus disease 2019, CT = computed tomography.

A Mask-cascade-RCNN-ResNeSt-200 with deformable convolution neural network (DCN) architecture was employed in module (a) to extract the lung and lesion regions in the chest CT scans.^{15,16} The lung segmentation model was trained and evaluated on a total 51,978 manually segmented slices (train: 50,756, test: 1,222) from two public datasets (NSCLC, 20 cases).^{17,18} For the lesion segmentation task, 6,971 manually segmented slices (train: 5,854, test: 1,117) from three publicly available datasets were used (20 cases, MosMed, MSD).¹⁸⁻²⁰ Then, for the data from Yeungnam University Medical Center, lesions were extracted using the trained models. The patches smaller than 13 mm were not used to avoid misclassification caused by wrong segmentation or noise such as motion artifacts, as this may have a negative effect on subsequent analysis.

The deep neural network in module (b) employed a ResNet50 model trained to differentiate lesion patches of COVID-19 from those of bacterial pneumonia patients.²¹ The model took a lesion cropped patch as input and returned a label prediction. A total of 12,235 lesion patches (train: 6,181, test: 6,054) from 170 patients were employed. A 2048-dimensional feature vector extracted from the intermediate layer of the ResNet50 model used in the clustering phase, i.e., module (c).

In module (c), the K-means algorithm was applied to cluster the lesion features into 20 groups.²² To profile the typical or relatively atypical imaging features of COVID-19, a total of 12,235 lesion patches from 170 patients were represented in two-dimensional space via a t-distributed Stochastic Neighbor Embedding based reduction of the 2,048-dimensional feature vectors.²³ The lesion images in each cluster initially grouped by K-means were later manually described by one radiologist (Jongsoo Park) using imaging terms, wherein three pulmonologists (Kyung Soo Hong, Jong Geol Jang, and June Hong Ahn) evaluated the descriptions and reached a consensus.

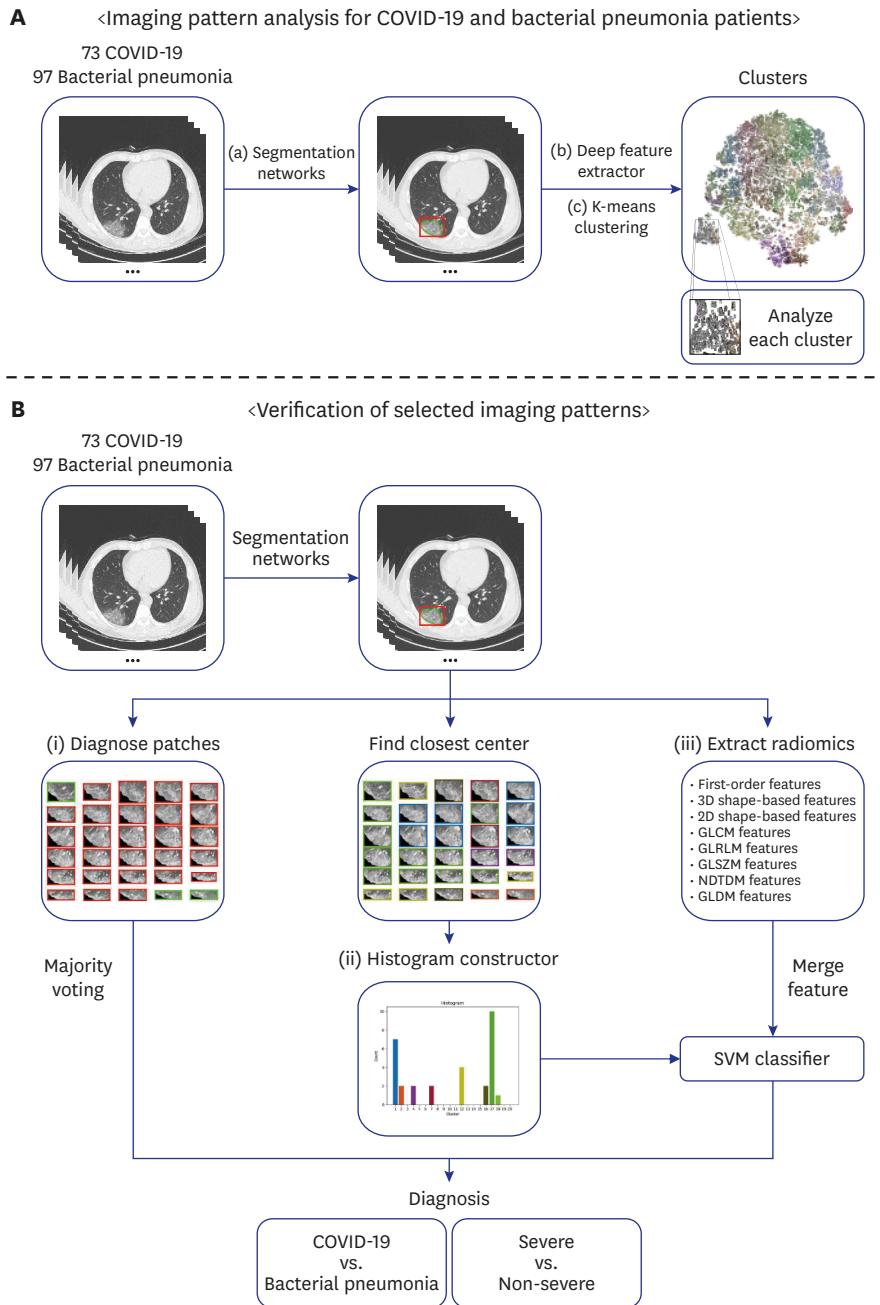


Fig. 2. Diagram of the modeling framework. For analyzing the imaging patterns, the segmentation models were employed to extract regions of interest of the lung and lesions in the chest CT scans. Then, the ResNet50 model takes lesion cropped patches as input and returns a label prediction and a 2048-dimensional feature vector. Next, the K-means clustering algorithm was used to cluster 2D lesion features into 20 groups. To verify the usability of these patterns, (i) majority voting; (ii) constructed histogram with SVM; and (iii) Radiomics features (with/without constructed histogram) with SVM were employed. Disease and severity classification tasks were performed.

COVID-19 = coronavirus disease 2019, CT = computed tomography, SVM = Support Vector Machine.

Verification of selected imaging patterns

Patient-level diagnosis can be achieved by aggregating the lesion-level predictions using majority voting. However, in this way, it is difficult to analyze what patterns the patient has and which combination of patterns are highly correlated with the disease or severity. To further quantify the diagnostic performance of the extracted imaging patterns, we

constructed a histogram of lesion types using the clusters obtained via module (c) for each CT. **Fig. 2B** shows a diagram of the evaluation framework. In addition, the constructed histograms were analyzed using an independent two-sample t-test to evaluate the significance of each cluster regarding: 1) COVID-19 and bacterial pneumonia patients, 2) severe and non-severe cases. A two-tailed $P < 0.05$ was taken to indicate statistical significance. Furthermore, we compared the mean histograms to confirm cluster relevance using the mean of each histogram in each group; i.e., COVID-19, bacterial pneumonia, severe, and non-severe. The cluster which showed high statistical significance ($P < 0.001$) showed significant value differences when performing mean histogram comparison.

Furthermore, we compared the classification accuracy between majority voting of the lesion-level predictions and Support Vector Machine (SVM) based inference which was trained on a combination of the proposed histogram with radiomics features. A total of 107 features were extracted using radiomics from each CT volume, including first-order statistics, shape-based features, etc. Finally, 20 features from our proposed histogram and 107 features from radiomics features were combined. We compared the performance with/without the combination of the radiomic features to verify effectiveness.

Statistical analysis

All statistical analyses were performed using Scipy (1.5.0, <https://www.scipy.org/>). The independent two-sample t-test was used to analyze the differences between groups, and Spearman's correlation test was used to analyze the correlation between laboratory data, clinical parameters and values of a cluster. The cluster which showed high statistical significance ($P < 0.001$) lead to significant value differences when performing mean histogram comparison. In all other analyses, P value < 0.05 was considered to indicate statistical significance.

Ethics statement

This study was conducted in accordance with the tenets of the Declaration of Helsinki and was reviewed and approved by the Institutional Review Board of Yeungnam University Hospital (YUH IRB 2020-05-030). The requirement for informed consent was waived due to the retrospective study design.

RESULTS

Demographic and clinical characteristics

Hospitalized patients with confirmed COVID-19 (73 patients) and bacterial pneumonia (97 patients) were included in this study (**Fig. 1**). Baseline characteristics of all patients are summarized in **Table 1**. The patients in the COVID-19 group were older than the patients in the bacterial pneumonia group (58.70 ± 16.49 vs. 39.86 ± 8.69 , $P < 0.001$). 36 patients (49.3%) of COVID-19 and 55 patients (56.7%) of bacterial pneumonia were male. Body temperature (37.21 ± 0.67 vs. 37.76 ± 0.92 , $P < 0.001$), and heart rate (86.70 ± 13.94 vs. 94.37 ± 16.91 , $P = 0.02$) were significantly lower in patients with COVID-19. Systolic blood pressure (128.55 ± 19.44 vs. 115.71 ± 19.42 , $P < 0.001$), and diastolic blood pressure (80.74 ± 12.15 vs. 71.42 ± 13.52 , $P < 0.001$) were significantly higher for COVID-19 patients. CURB-65 was not significantly different in patients with COVID-19 and bacterial pneumonia (0.69 ± 0.87 vs. 0.57 ± 0.68 , $P = 0.314$). The mean time from the onset of symptoms to CT scan was 7.46 ± 4.04 days.

Table 1. Characteristics of the study participants with COVID-19 and bacterial pneumonia

Variables	COVID-19 (n = 73)	Bacterial pneumonia (n = 97)	P value
Age, yr	58.70 ± 16.49	39.86 ± 8.69	< 0.001
Male, sex	36 (49.3)	55 (56.7)	0.390
Vital signs on admission			
Body temperature, °C	37.21 ± 0.67	37.76 ± 0.92	< 0.001
Heart rate, beats/min	86.70 ± 13.94	94.37 ± 16.91	0.002
Systolic BP, mm Hg	128.55 ± 19.44	115.71 ± 19.42	< 0.001
Diastolic BP, mm Hg	80.74 ± 12.15	71.42 ± 13.52	< 0.001
CURB-65	0.69 (0.87)	0.57 ± 0.68	0.314
Timing of CT scan			
Early phase (≤ 7 days after the onset of symptoms)	47 (64.4)	-	-
Late phase (> 7 days after the onset of symptoms)	26 (35.6)	-	-





















Data are presented as the mean ± standard deviation or number (%).
 COVID-19 = coronavirus disease 2019, BP = blood press.

Deep chest CT manifestations

We summarize the deep chest CT manifestations observed in **Table 2** and **Supplementary Table 1**, with lesion patches visualized in **Fig. 3**, respectively. Typical COVID-19 CT manifestations such as GGOs with interlobular septal thickening in the peripheral lungs were observed in clusters 4 and 17 (**Fig. 4A**). On the other hand, manifestations typical to bacterial pneumonia such as crazy-paving appearance in the posterior lungs were observed in clusters 5 and 10, as presented in **Fig. 4B**. These typical clusters showed *P* values less than 0.001.

Notably, 15 out of the 20 clusters showed key characteristics for the discrimination of COVID-19 from bacterial pneumonia with *P* values less than 0.001. Among 15 clusters, three clusters with *P* values less than 0.001 were shown to be key for severity classification. Moreover, two clusters could classify both COVID-19 and severe patients, with one of the clusters (#8) showing diffuse GGOs in the central and peripheral lungs which represents a typical severe COVID-19 CT manifestation. **Fig. 4C** presents the severe COVID-19 patient's CT manifestations.

Table 2. Summary of lesion cluster with imaging description term and *P* value results

Cluster	Color	Description	Typical	Diagnosis <i>P</i> value	Severe <i>P</i> value
1		Multifocal GGO in the peripheral lung, bilateral.	COVID-19	< 0.001	< 0.001
2		Mixed consolidations and GGOs in the posterior lung, unilateral.	Intermediate	0.006	0.713
3		Focal consolidation in the posterior lung.	Bacterial	< 0.001	0.068
4		GGOs with interlobular septal thickening in the peripheral lungs, bilateral.	COVID-19	< 0.001	0.081
5		Crazy-paving appearance with/without consolidation, diffuse.	Bacterial	< 0.001	0.556
6		Consolidation or clustered micronodules along the bronchovascular bundle, bronchopneumonia pattern.	Bacterial	< 0.001	0.002
7		Multifocal GGO with round morphology, bilateral.	COVID-19	< 0.001	0.011
8		Diffuse GGOs in the central and peripheral lungs, bilateral.	COVID-19	< 0.001	< 0.001
9		Consolidation and GGO with interlobular septal thickening, unilateral.	Bacterial	< 0.001	0.013
10		Crazy-paving appearance in the posterior lungs.	Bacterial	< 0.001	0.437
11		Segmental consolidation in the unilateral or bilateral lungs.	Bacterial	< 0.001	0.810
12		Mixed consolidations and GGOs in the posterior lung, bilateral.	COVID-19	< 0.001	0.466
13		Subtle GGO in the lower lung.	Bacterial	< 0.001	< 0.001
14		Extensive consolidation with air-bronchogram in the bilateral lungs.	Bacterial	< 0.001	0.128
15		GGO with reticular opacity in both lower lungs.	Intermediate	0.097	0.806
16		Mixed consolidation with GGO in the peripheral lung, unilateral.	Intermediate	0.060	0.068
17		GGOs with interlobular septal thickening in the peripheral lungs, bilateral.	COVID-19	< 0.001	0.001
18		Consolidation or clustered nodules along the bronchovascular bundle.	Bacterial	< 0.001	0.534
19		Consolidation or GGO along the bronchovascular bundle.	Intermediate	0.005	0.009
20		Multifocal GGO with interlobular septal thickening, bilateral.	Intermediate	0.163	0.714

The t-test were performed for diseased and severe groups, respectively.
 GGO = ground-glass opacity, COVID-19 = coronavirus disease 2019.

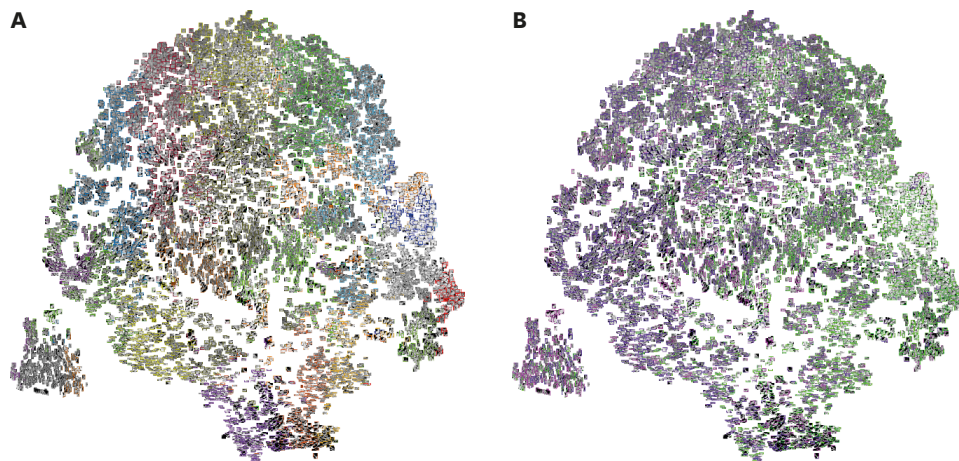


Fig. 3. Visualization of the lesion patches from 170 patients into two-dimensions. **(A)** K-means clusters (20 groups) were used for distinction. Different color edge represents different groups. **(B)** True diagnosis was used for distinction. Purple color represents non-severe COVID-19, magenta represents severe COVID-19, and green for bacterial pneumonia.

COVID-19 = coronavirus disease 2019.

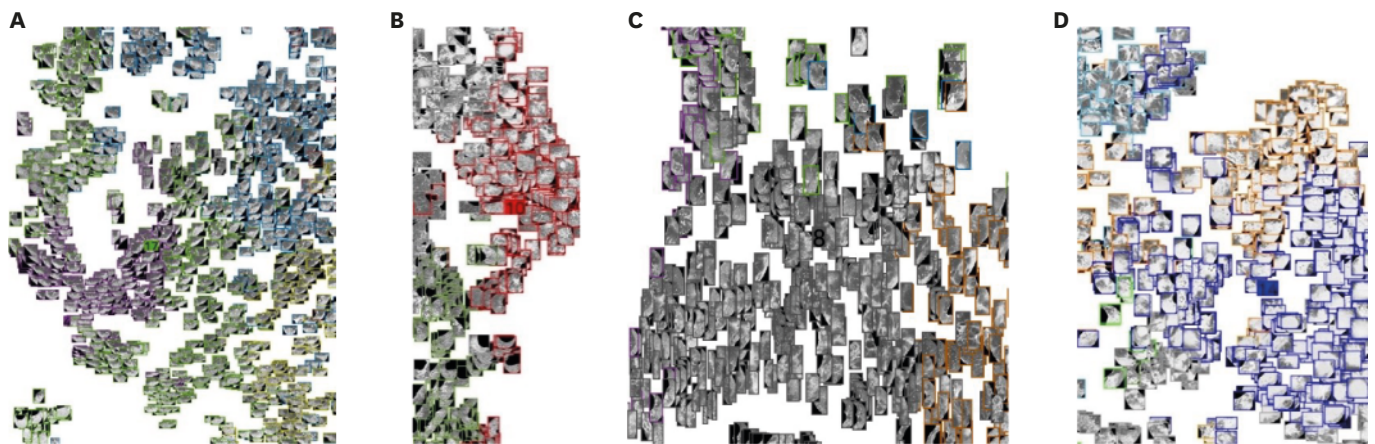


Fig. 4. **(A)** Typical COVID-19 CT manifestations—GGOs with interlobular septal thickening in the peripheral lungs—were observed in cluster 4 and 17. **(B)** bacterial pneumonia CT manifestation—crazy-paving appearance in the posterior lungs—observed in cluster 5 and 10. **(C)** Cluster 8 showed the lowest t-test P value < 0.001 for diseased and severe groups which is the typical pattern for severe COVID-19 patients (i.e., diffuse GGOs in the central and peripheral lungs). **(D)** Bacterial pneumonia CT manifestation—extensive consolidation with air-bronchogram—was observed in cluster 14. These clusters showed a t-test P value < 0.001 . Lesions were colored based on K-means clustering result.

COVID-19 = coronavirus disease 2019, CT = computed tomography, GGO = ground-glass opacity.

In principle, the observed relative differences between lesion clusters are key to better understand the entire lesion distribution. As shown in **Fig. 4D**, extensive consolidation with air-bronchogram was observed in severe bacterial pneumonia patients. Compared with the severe COVID-19 patients who mainly showed diffuse GGOs, we could easily distinguish the difference in patterns observed in the later stages of disease onset.

To confirm cluster relevance, we compared the mean histograms of different tasks as presented in **Fig. 5**. The clusters which showed high significance ($P < 0.001$) in **Table 2** also showed significantly different values in mean histogram comparison. Notably, the mean histogram of COVID-19 can be discriminated from bacterial pneumonia by comparing the relatively increased values in cluster 1, 4, 7, 12, and 17. On the other hand, the mean histogram of bacterial pneumonia can be discriminated from COVID-19 by comparing the significantly increased values of clusters 3, 6, 9, 11, 13, 14, and 18. In addition, the mean histogram of the severe COVID-19

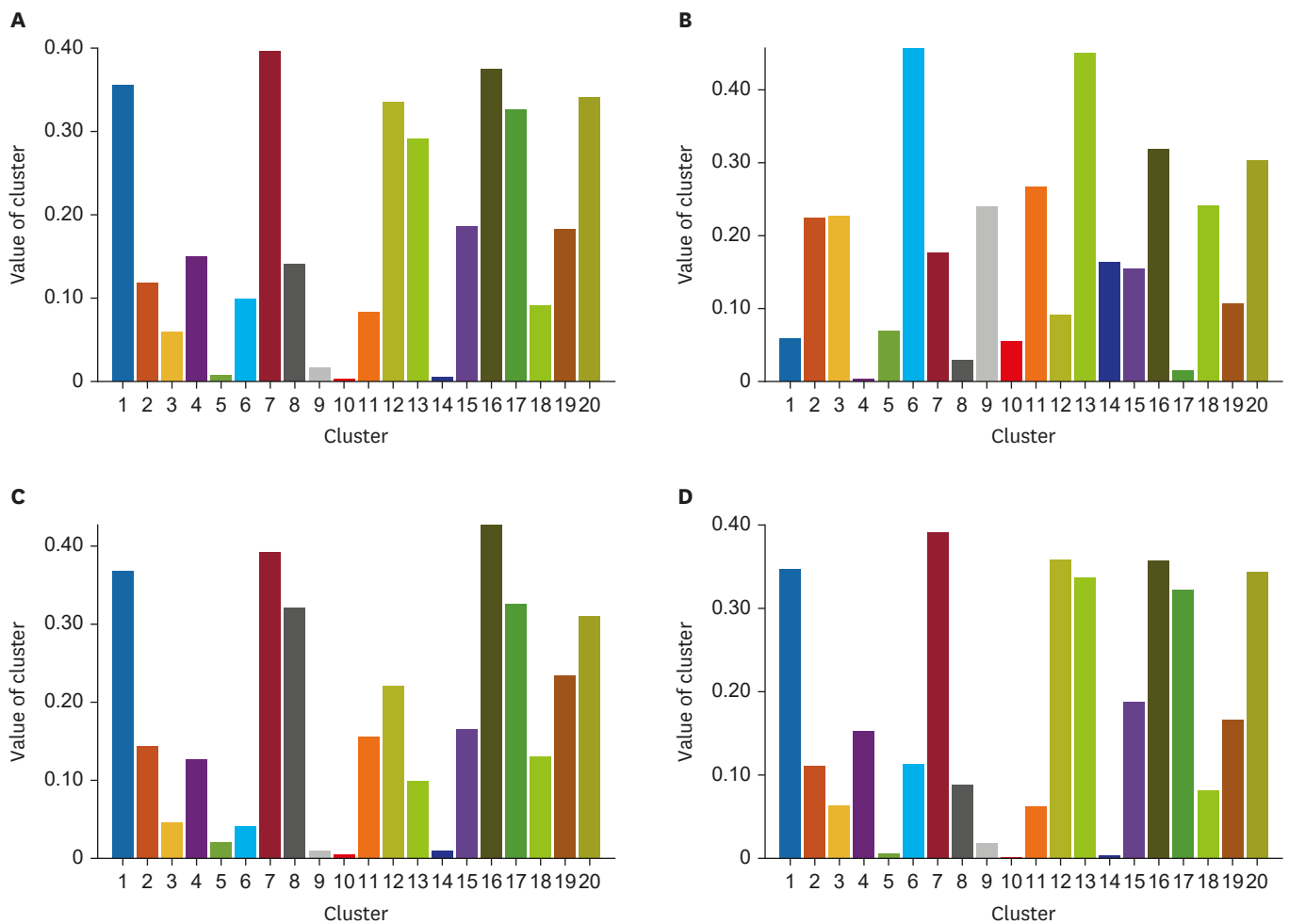


Fig. 5. Mean histograms. (A) Mean histogram of COVID-19; (B) Mean histogram of bacterial pneumonia; (C) Mean histogram of severe COVID-19 patients; and (D) Mean histogram of non-severe COVID-19 patients. COVID-19 = coronavirus disease 2019.

patients showed consistently increased values in cluster 8 compared to the mean histogram of the non-severe COVID-19 patients. This indicates that the diagnosis is highly relevant to the presence of a lesion in a certain cluster, and the patient with a lesion in cluster 8 can be considered as a more severe COVID-19 patient than patients with lesions in other clusters. In summary, the cluster interpreted by the radiologist as being highly relevant to typical COVID-19 and bacterial pneumonia patterns showed high correlation with *P* value significance in statistical analyses, with notable differences observed in mean histogram comparison.

Disease and severity classification

In Table 3, we present disease classification accuracy across several metrics under different settings. First, the baseline method based on majority voting of predictions achieved 87.5% accuracy for COVID-19 patient classification. For the SVM classifier, we consider two scenarios, (i) with histogram features only or (ii) with radiomic features only. Notably, the classifier trained under the first setting achieved 88.7% accuracy; a minor improvement over the baseline, whereas in setting (ii)—81.25% was reported, a considerable decrease from the other models. However, when the SVM classifier was trained with both radiomics and histogram features, accuracy significantly improved to 91.2% from 87.5%. This further shows

Table 3. Test accuracy of COVID-19 diagnosis

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)
Majority voting	87.5 (70 of 80)	92.5 (37 of 40)	82.5 (33 of 40)
Radiomics only	81.25 (65 of 80)	87.5 (35 of 40)	75 (30 of 40)
Histogram only	88.7 (71 of 80)	85 (34 of 40)	92.5 (37 of 40)
Histogram with radiomics	91.2 (73 of 80)	85 (34 of 40)	97.5 (39 of 40)

COVID-19 = coronavirus disease 2019.

Table 4. Test accuracy of severity classification.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)
Radiomics only	82.5 (33 of 40)	66.67 (4 of 6)	85.29 (28 of 34)
Histogram only	82.5 (33 of 40)	50 (3 of 6)	88.24 (30 of 34)
Histogram with radiomics	95 (38 of 40)	83.3 (5 of 6)	97.06 (33 of 34)

that the features learned by deep learning are highly robust for accurate patient diagnosis. In addition, the constructed histograms can accurately express the correlation of lesion features observed in each patient and/or can represent some combination of patterns that could lead to a severe outcome of a disease diagnosis. In comparison to the naïve aggregation method baseline i.e., majority voting, often highly limited in patient-level representations; the proposed deep feature representation highlights several key advantages.

Table 4 shows the severity classification accuracy. In terms of accuracy alone, SVM classifiers that only use either histogram or radiomics only achieved the same performance, i.e., 82.5% for both, with key differences noted in the sensitivity and specificity of the models. On the other hand, when the features were combined considerable improvement (+10%) was noted i.e., 95% accuracy, surpassing previous models. This may be attributed to the fact that the histogram is not able to represent the absolute size difference of lesions, thus the combination mitigates the issue. Rather than using a single feature alone, the combination of features proved to be invaluable for diagnosis.

Lung and lesion segmentation

The lung and lesion segmentation model achieved a dice coefficient score of 97.18%, and 78.06%, respectively. **Fig. 6** shows the lesion segmentation results around the average dice coefficient score of 78.06%, which helps to qualitatively understand the accuracy of lesion segmentation.

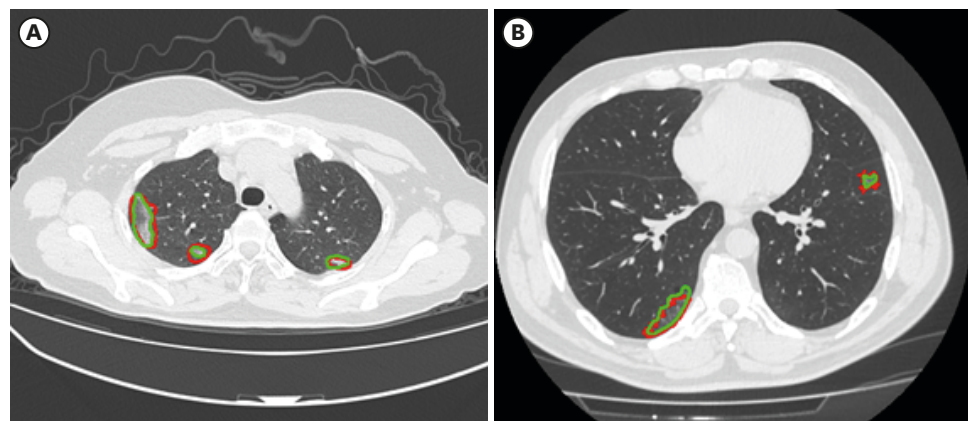


Fig. 6. Representative slices which showed an average dice coefficient score (78.06%) for lesion segmentation. (A) Slice showed a dice coefficient score of 77.08%. (B) Slice showed a dice coefficient score of 78.31%. Red color represents a manual segmentation by human and green color represents an automated segmentation by the trained segmentation model. Due to the ambiguity of lesion boundaries, an average dice score of 78.06% showed satisfactory lesion segmentation results for the deep feature extractor.

Table 5. Correlation between laboratory data, clinical parameters and CT quantitative parameters

Laboratory data and clinical parameters	Correlation between values of cluster 8	P value
White blood cell count, $\times 10^9/L$	0.113	0.341
Neutrophil percentage, %	0.284	0.014
Neutrophil count, $\times 10^9/L$	0.195	0.098
Lymphocyte percentage, %	-0.302	0.009
Lymphocyte count, $\times 10^9/L$	-0.258	0.027
Hemoglobin, g/dL	0.105	0.376
Platelet, $\times 10^9/L$	-0.021	0.863
C-reactive protein, mg/dL	0.018	0.880
Procalcitonin, ng/mL	0.262	0.025
NEWS	0.392	< 0.001
ARDS	0.299	0.010
Septic shock	0.229	0.051
ICU admission	0.209	0.075
IMV	0.229	0.051
ECMO	0.282	0.015
Death	0.110	0.354

CT = computed tomography, NEWS = National Early Warning Score, ARDS = acute respiratory distress syndrome, ICU = intensive care unit, IMV = invasive mechanical ventilation, ECMO = extracorporeal membrane oxygenation.

Correlation between laboratory data, clinical and CT quantitative parameters

The correlation between laboratory data, clinical parameters and CT quantitative parameters are shown in Table 5. The values of cluster 8 were strongly correlated with NEWS ($P < 0.001$), and positively correlated with neutrophil percentage, procalcitonin, ARDS, and extracorporeal membrane oxygenation (ECMO) ($P < 0.05$). On the other hand, a negative correlation with lymphocyte percentage and lymphocyte count ($P < 0.05$) was noted. Comparisons between values of cluster 8 versus NEWS (Fig. 7A) and disease severity (Fig. 7B) in COVID-19 patients are shown. Notably, values were significantly higher among NEWS ≥ 7 patients than in NEWS 0–4 ($P < 0.001$), with severe COVID-19 patients also showing higher values than non-severe COVID-19 patients ($P < 0.05$), respectively.

DISCUSSION

Our deep chest CT analysis is in accordance with various CT findings of COVID-19 patients reported in literature. Chung et al. reported that typical CT findings of COVID-19 include bilateral pulmonary GGO and consolidative opacities which sometimes have a rounded

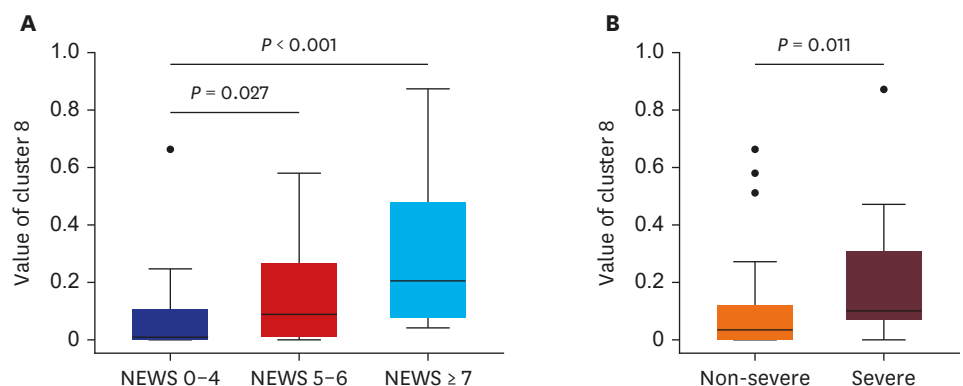


Fig. 7. Comparisons between values of cluster 8 versus NEWS (A) and disease severity (B) in COVID-19 patients. NEWS = National Early Warning Score, COVID-19 = coronavirus disease 2019.

morphology and are distributed in the peripheral lung.²⁴ Song et al.⁵ reported that pure GGO or GGO with reticular and/or interlobular septal thickening with predominant distribution in the posterior or peripheral lung involvements were observed in COVID-19 patients. Caruso et al.²⁵ highlighted the presence of the peripheral GGOs associated with multilobe and posterior lung involvement in COVID-19 patients in Italy. Similar patterns were also observed in our research i.e. GGO distributed in the peripheral lungs in cluster 4 and 17, with multifocal GGOs with round morphology in cluster 1 and 7, respectively. Regarding the common findings of severe COVID-19, Pan et al.²⁶ report that an increase in GGO, consolidative opacities, and interstitial septal thickening was noted, and Song et al.⁵ reported that a significantly more GGOs including pure GGO and a GGO with reticular and/or interlobular septal thickening was observed in the later stages of COVID-19. An increase in GGOs in the central and peripheral lung was observed in our study in cluster 8, therefore, the published CT findings of severe COVID-19 patients show high consistency with our study.

Radiomic features are considered a useful general purpose analysis technique, i.e., for distinguishing the lung nodules (malignant versus benign) or hospital stay (severity) prediction.^{27,28} However, several limitations exist and features alone are often insufficient to distinguish between diseases when subtle radiological differences are observed in the image. Discriminating COVID-19 from bacterial pneumonia is regarded as one of the exemplars of such challenges. Here, our method shows the benefit of using deep learning to obtain more robust representations that are more clinically relevant to key imaging characteristics for COVID-19 diagnosis. We quantitatively show that the constructed histogram better captures the overall statistics of the lesion features. Moreover, the SVM classifier can diagnose diseases or patient's severity more accurately than the radiomics features alone.

Our method has two notable advantages compared to common deep learning algorithms; interpretability and a generalized representation. Common deep learning methods are limited in interpretability even though they can visualize the important regions using heatmaps.^{29,30} Our method can explain the reasons of diagnosis by checking the presence of specific patterns represented in the patient's histogram. We verified the key patterns with mean histograms for each disease and severity group, and found the important key diagnostic imaging patterns in accordance with published literature. This indicates that our method is safer and more transparent for medical assistance. Moreover, although the proposed feature learning model was not trained to classify severity among patients, the obtained features are fairly generalized for severity classification. The imaging features were divided into 20 clusters and verified by radiologists using imaging terms i.e., an independent representation of a specific diagnosis. The constructed histogram can be used for general diagnosis regardless of the trained diseases, and thus showed considerable diagnostic accuracy for severity classification.

This study has several limitations. First, the proposed framework was only trained on a single institute cohort, therefore current models may fail to accurately represent unobserved cohort CT manifestations. External validation on a large cohort would be required to address this issue. Second, though age is significantly different between the two target groups (i.e., COVID-19 and Bacterial pneumonia); severity in pneumonia patients was more of an important factor affecting the CT patterns rather than age. Further, characteristic analysis based on the CURB-65 score revealed little to no differences between the groups, thus we infer the severity of pneumonia in the groups is also not expected to differ. Third, CT scan timing or disease course spectrums were not considered for analysis in this study. We mainly

focused on cluster analysis to compare COVID-19 and bacterial pneumonia characteristics. Future studies analyzing CT patterns according to the timing of CT scans or disease course using deep learning are imperative. Fourth, the K-means algorithm requires an arbitrary number of groups to be chosen, we empirically selected 20 groups in this study though it may be beneficial to further analyze varied cluster groups to draw well informed conclusions across different settings. Applying better clustering algorithms and selecting an optimal number of clusters will be the subject of future research.

In conclusion, the CT images of COVID-19 patients in Daegu, Korea were grouped into 20 clusters. These groups were analyzed and compared with the patterns described in literature. Crazy-paving was extracted as a major pattern for bacterial pneumonia, while GGOs in the peripheral lungs and diffuse GGOs were observed for COVID-19 and severe COVID-19 patients, respectively. To verify the effectiveness of these clusters, we performed two classification tasks by constructing histograms from the clusters. We confirmed the correlations of the image patterns extracted by the proposed method are more relevant to the clinical setting than common methods which use radiomics or naïve deep features. The constructed histogram features improved accuracy for both disease and severity classification, and showed correlations with laboratory data and clinical parameters, thus can provide guidance for improved analysis and treatment of COVID-19.

SUPPLEMENTARY MATERIAL

Supplementary Table 1

Zoomed visualization of the lesion patches from 170 patients with descriptions

[Click here to view](#)

REFERENCES

1. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;382(18):1708-20.
[PUBMED](#) | [CROSSREF](#)
2. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing. *Radiology* 2020;296(2):E41-5.
[PUBMED](#) | [CROSSREF](#)
3. Huang P, Liu T, Huang L, Liu H, Lei M, Xu W, et al. Use of chest CT in combination with negative RT-PCR assay for the 2019 novel coronavirus but high clinical suspicion. *Radiology* 2020;295(1):22-3.
[PUBMED](#) | [CROSSREF](#)
4. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* 2020;296(2):E115-7.
[PUBMED](#) | [CROSSREF](#)
5. Song F, Shi N, Shan F, Zhang Z, Shen J, Lu H, et al. Emerging 2019 novel coronavirus (2019-nCoV) pneumonia. *Radiology* 2020;295(1):210-7.
[PUBMED](#) | [CROSSREF](#)
6. Salehi S, Abedi A, Balakrishnan S, Gholamrezanezhad A. Coronavirus disease 2019 (COVID-19): a systematic review of imaging findings in 919 patients. *AJR Am J Roentgenol* 2020;215(1):87-93.
[PUBMED](#) | [CROSSREF](#)
7. Zhou S, Wang Y, Zhu T, Xia L. CT features of coronavirus disease 2019 (COVID-19) pneumonia in 62 patients in Wuhan, China. *AJR Am J Roentgenol* 2020;214(6):1287-94.
[PUBMED](#) | [CROSSREF](#)

8. Bai HX, Hsieh B, Xiong Z, Halsey K, Choi JW, Tran TM, et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. *Radiology* 2020;296(2):E46-54.
[PUBMED](#) | [CROSSREF](#)
9. Choi H, Qi X, Yoon SH, Park SJ, Lee KH, Kim JY, et al. Extension of coronavirus disease 2019 on chest CT and implications for chest radiographic interpretation. *Radiol Cardiothorac Imaging* 2020;2(2):e200107.
[CROSSREF](#)
10. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. *Radiology* 2008;246(3):697-722.
[PUBMED](#) | [CROSSREF](#)
11. Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60-88.
[PUBMED](#) | [CROSSREF](#)
12. ARDS Definition Task ForceRanieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, et al. Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012;307(23):2526-33.
[PUBMED](#)
13. Jang JG, Hur J, Hong KS, Lee W, Ahn JH. Prognostic accuracy of the SIRS, qSOFA, and NEWS for early detection of clinical deterioration in SARS-CoV-2 infected patients. *J Korean Med Sci* 2020;35(25):e234.
[PUBMED](#) | [CROSSREF](#)
14. Hong KS, Lee KH, Chung JH, Shin KC, Choi EY, Jin HJ, et al. Clinical features and outcomes of 98 patients hospitalized with SARS-CoV-2 infection in Daegu, South Korea: a brief descriptive study. *Yonsei Med J* 2020;61(5):431-7.
[PUBMED](#) | [CROSSREF](#)
15. Zhang H, Wu C, Zhang Z, Zhu Y, Zhang Z, Lin H, et al. ResNeSt: split-attention networks. *arXiv*. Forthcoming 2020.
16. Dai J, Qi H, Xiong Y, Li Y, Zhang G, Hu H, et al. Deformable convolutional networks. In: Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV); 2017 Oct 22-29; Venice, Italy. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2017, 764-73.
[CROSSREF](#)
17. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5(1):4006.
[PUBMED](#) | [CROSSREF](#)
18. Jun M, Cheng G, Yixin W, Xingle A, Jiantao G, Ziqi Y, et al. COVID-19 CT lung and infection segmentation dataset. *Zenodo*. Forthcoming 2020. DOI: 10.5281/zenodo.3757476.
[CROSSREF](#)
19. Morozov SP, Andreychenko AE, Pavlov NA, Vladzimirskyy AV, Ledikhova NV, Gombolevskiy VA, et al. MosMedData: chest CT scans with COVID-19 related findings dataset. *medRxiv*. Forthcoming 2020. DOI: 10.1101/2020.05.20.20100362.
[CROSSREF](#)
20. Simpson AL, Antonelli M, Bakas S, Bilello M, Farahani K, van Ginneken B, et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv*. Forthcoming 2020.
21. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27-30; Las Vegas, NV. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2016, 770-8.
[CROSSREF](#)
22. Lloyd S. Least squares quantization in PCM. *IEEE Trans Inf Theory* 1982;28(2):129-37.
[CROSSREF](#)
23. Lvd M, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res* 2008;9(86):2579-605.
24. Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X, et al. CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology* 2020;295(1):202-7.
[PUBMED](#) | [CROSSREF](#)
25. Caruso D, Zerunian M, Polici M, Pucciarelli F, Polidori T, Rucci C, et al. Chest CT features of COVID-19 in Rome, Italy. *Radiology* 2020;296(2):E79-85.
[PUBMED](#) | [CROSSREF](#)
26. Pan Y, Guan H, Zhou S, Wang Y, Li Q, Zhu T, et al. Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): a study of 63 patients in Wuhan, China. *Eur Radiol* 2020;30(6):3306-9.
[PUBMED](#) | [CROSSREF](#)

27. Qi X, Jiang Z, Yu Q, Shao C, Zhang H, Yue H, et al. Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study. *medRxiv*. Forthcoming 2020. DOI: 10.21037/atm-20-3026.
[CROSSREF](#)
28. Nishino M. Perinodular radiomic features to assess nodule microenvironment: does it help to distinguish malignant versus benign lung nodules? *Radiology* 2019;290(3):793-5.
[PUBMED](#) | [CROSSREF](#)
29. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, et al. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. *Radiology* 2020;296(2):E65-71.
[PUBMED](#) | [CROSSREF](#)
30. Bai HX, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, et al. Artificial intelligence augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other origin at chest CT. *Radiology* 2020;296(3):E156-65.
[PUBMED](#) | [CROSSREF](#)