*Journal of Industrial and Systems Engineering*
Vol. 13, No. 3, pp. 61-69
Summer (July) 2021

# A quantitative scoring system to compare the degree of COVID-19 infection in patients' lungs during the three peaks of the pandemic in Iran

**Amir Hossein Masumi[1], Rouzbeh Ghousi[*1], Mozhgan Vazifehdoost[2], Faeze Araghi Niknam[1]**

[1]*School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran*

[2]*Faculty of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran*

*amir_masoumi@alumni.iust.ac.ir, ghousi@iust.ac.ir, vazifedoost.m12@gmail.com, s.araghi1357@gmail.com*

**Abstract**

Contrary to most countries, coronavirus has peaked three times during the last eight months in Iran. Unfortunately, increasing the number of positive COVID-19 cases is not the only crucial problem we faced. Indeed, it seems that lung coronavirus infections have also become more severe during these three peaks. Therefore, this study proposes a quantitative scoring system based on medical imaging to score the degree of lung coronavirus infection during each peak. Regarding the degree of lung coronavirus infection for all patients during the last three peaks, we test the mentioned hypothesis by employing statistical methods. Comparing the characteristics of the disease during three different peaks is another goal of the research. To this end, 5265 lung CT scan images from 270 patients with a definite diagnosis of COVID-19 infection were annotated under radiology expert supervision. Then, was used deep learning methods for image segmentation. In the next step, each patient's lung was divided into six sections, and the percentage of infection was calculated in each section. Finally, the Friedman and Games-Howell tests showed that the average degree of COVID-19 infection has increased during the considered period, and the average of infection in men was about twenty percent higher than in women.

**Keywords**: COVID-19, CT scan, statistical analysis, deep learning, image segmentation

## 1-Introduction

WHO (World Health Organization) has declared the outbreak of COVID-19 since 30 January 2020. The virus's genetic mutations may have made it more contagious (Long et al., 2020). Iran's coronavirus death toll has surpassed 50,000, and this number continues to surge. Unfortunately, Iran is facing the third wave of corona; as figure 1 shows, the number of people suffering from COVID 19 peaked in March, June, and September.

CT scan Imaging is playing a crucial role in diagnosing and managing COVID-19, especially with RT-PCR shortages. Radiologists can notice either opacities or bright spots on the patient's lungs with these images (Zhao et al., 2020b). Furthermore, CT can be a reliable instrument for detecting suspected cases that do not show up on PCR tests (Ai et al., 2020). However, the manual segmentation of lung infections is monotonous and time-consuming. Besides, radiologists' risk of error in infection annotation is high.

[*]Corresponding author

ISSN: 1735-8272, 

Today, deep learning techniques are widely used for all medical imaging with diagnosis and treatment aims. Since timing is an essential factor in treating a variety of diseases, timely and accurate diagnosis of lesions through medical images is at a high level of importance.

The current research's main intention is to present a quantitative scoring system to define the degree of lung COVID-19 infection based on CT images using the science of deep learning and neural networks. Comparing the characteristics of the disease during three different peaks in Iran by implementing statistical tools is another goal of the research.

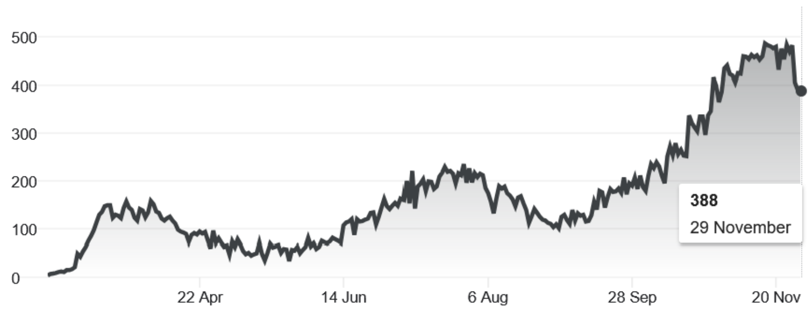


**Fig1.** The trend of coronavirus death toll in Iran (JHU, 2020)

## 2-Related works

More than 185 countries are struggling with the prevalence of COVID-19. This virus has infected more than 60,000,000 individuals and causing almost 1,460,000 deaths by November 2020 worldwide (JHU, 2020). Since the contagious disease rate is high, at least three people on average could be infected by only one person (Alimohamadi et al., 2020). Almost 950,000 patients with COVID- 19 have been identified in Iran from February 2020, and the virus has caused 28,000 deaths. According to (JHU, 2020), Iran ranks first in the Middle East region in fatalities due to COVID- 19.

Contrary to other countries, the number of new cases has fluctuated in Iran and has reached its maximum number three times from the emergence of COVID-19 (WHO: World Health Organization, 2020). Raoofi et al., (2020) have studied the consequences and effects of the corona outbreak in Iran and concluded that combating this virus requires professional advocacy attempts through appropriate inter-sectoral collaboration and whole-government coalitions. The highest number of deaths has occurred in old men with underlying diseases, so it is essential to pay special attention to this group (Nikpouraghdam et al., 2020).

Chest CT scan, Nucleic acid detection, epidemiological history, and clinical treatment form the steps of the most common approach for diagnostic COVID-19. In comparison to antibody tests, the chest CT is more practical to find the infected. Contrary to X-ray imaging, which is not suitable due to the low imaging characteristics, CT scan is a powerful method to identify the infection ((JHU), 2020, Manigandan et al., 2020). Ding et al., (2020) evaluated lung abnormalities on CT scans in patients with COVID-19 and their coloration with the duration of symptoms on 122 patients. They showed that GGO[2], crazy paving, consolidation, and linear opacities are the most frequent lesions, and the infection can be seen in the lower lobes more than other parts of the lungs. According to Pan et al., (2020), CT imaging may reveal abnormalities earlier than RT-PCR testing. They showed that patients might have a single lobe or multiple lobes involvement.

Feature extracting of the lesions such as distribution pattern, quantity, and range, density, shape, interface by using deep learning methods and statistical tools can help radiologists be more accurate in their decision before proposing a biopsy test to patients(Fan et al., 2020b). The purpose of artificial intelligence is not to replace human interactions in medical industries but play as decision support for clinicians based on their models(Phillips-Wren and Ichalkaranje, 2008). Ardakani et al., (2020) introduced a CAD approach based on CT images and suggesting a new model for COVID-19 diagnosis using a deep convolutional network. They created an expert system based on resnet101 backbone and trained their model on 1020 CT images of 108 COVID-19 infected patients, which resulted in 99.51%

[2] Ground-glass opacification

and 99.02% accuracy and specificity, respectively. Ozturk et al., (2020) have designed an automatic deep learning-based model with an end-to-end structure for separating X-ray images with COVID-19 infection. The final accuracy for binary classification was 98.08% on 127 X-ray images. Ribeiro et al., (2020) used clinical information and blood samples besides support vector regression algorithm to predict COVID-19 outbreak in ten Brazilian states. The final accuracy of their short-term forecasting model was 90%. A location-attention-oriented model on CT images was employed to calculate the infection probability of COVID-19 (Xu et al.). Shan et al., (2020) introduced a deep learning network for segmentation and quantification of infections in chest CT scans on 249 COVID-19 patients.

Hu et al., (2020) conducted a meta-analysis to evaluate the risk factors of COVID-19 by the implementation of statistical methods. Considering the physicians' reports about the images can significantly improve image classification accuracy (Litjens et al., 2017). Under Two radiologists' supervision, a semi-quantitative method for evaluating the severity of COVID-19 in the initial chest CT was introduced. Yang et al., (2020) A new flexible extended Weibull distribution can provide the best description of the COVID-19 total death data in the Asian countries (Zuo et al., 2020). Muthusami and Saritha, (2020) used regression and generalized linear models to analyze confirmed infected, deaths, and recovered cases. They also visualized the data to compare the growth rates between the Globe countries. Zhao et al., (2020a) implemented different time series models to predict the number of confirmed cases of coronavirus pneumonia in China and declared that ARIMAX model had the best results in terms of goodness of fit.

Fan et al., (2020a) have also proposed a model, named Inf-Net, by which the identification of infected regions improved. The final MAE for their model was 0.035 (Fan et al., 2020a). Therefore, we employed it as the best deep learning model with the highest accuracy level in infection detection and segmentation and used it on our datasets to segment the infected area of the suspected patient's lung.

## 3-Materials and methods

### 3-1-Data set

We collected data from three pandemic peaks by using simple random sampling to reduce the selection bias. In the first step, a series of lung CT-scans of 450 patients in three different time intervals at Milad Hospital (located in Tehran, Iran) has been collected. CT-scans were recorded with four slices' types of 1.5, 3, 5, and 7 mm. In this study, images with a 5 mm slice were used as the final data.

### 3-2-Data preparation

At this stage, referring to the physician's report, the state of lung health and the corona virus's existence was determined. As a result, 90 patients who suffered from COVID at each peak were collected.
In this research, noise data has been divided into two categories

1. Images that did not include the lungs,
2. The images became blurred due to the patient's motion during the shooting.

The first category of noise data was isolated, and for the second category, the image resolution was improved via using image enhancement techniques; therefore, the images became usable.

#### 3-2-1-Image enhancement

Providing the appropriate input is the first step for using a deep learning network. In this research, input data consisted of lung CT-scans, and the output data is the segmentation of COVID-19 infections. For improving the accuracy of segmentation, the regions of the images not consisting of lungs were extracted from the image using the contour lines concept because the chance of infection existence in this area is zero. We used an image gradient to calculate the contours and identify the border points to cut the useless parts of images. The points with the change in their gradient direction represent the edges, and the input boundary illustrates the omitted region. Thus, the image's vertical and horizontal gradients should be calculated; then, while combining these two gradients, boundary regions should be calculated. If A is the input image, then the size and angle of the image gradient can be calculated as equations (1)-(4):

$$B_v(j,k) = A(j,k+1) - A(j,k-1) \tag{1}$$

$$B_h(j,k) = A(j+1,k) - A(j-1,k) \quad (2)$$

$$B(j,k) = \sqrt{B_h(j,k)^2 + B_v(j,k)^2} \quad (3)$$

$$\theta = \arctan\left(\frac{B_v(j,k)}{B_h(j,k)}\right) \quad (4)$$

The well-known Soble filter for image processing can also be used to obtain gradients in horizontal and vertical directions. After obtaining the lung boundary points in the image, the bounding box should also be calculated. If the set P shows all boundaries of the prostate, the bounding box is calculated as equations (5) and (6).

$$(X_{\min}, Y_{\min}) = (\{\min(x) \mid x \in p\}, \{\min(y) \mid y \in p\}) \quad (5)$$

$$(X_{\max}, Y_{\max}) = (\{\max(x) \mid x \in p\}, \{\max(y) \mid y \in p\}) \quad (6)$$

### 3-2-2-Image Segmentation

This paper has employed a network named Inf-Net to segment the infections. Implicit reverse attention and explicit edge-attention are used to enhance infected regions' identification (Fan et al., 2020a). Two convolutional layers extract the high-resolution, semantically weak features. An edge attention module improves the representation of objective region boundaries. Three convolutional layers extract high-level features for generating a global map and multiple RA[3] module. Finally, the last RA's output is fed to a sigmoid activation function to predict lung infection regions. The segmentation results are shown in figure 3.

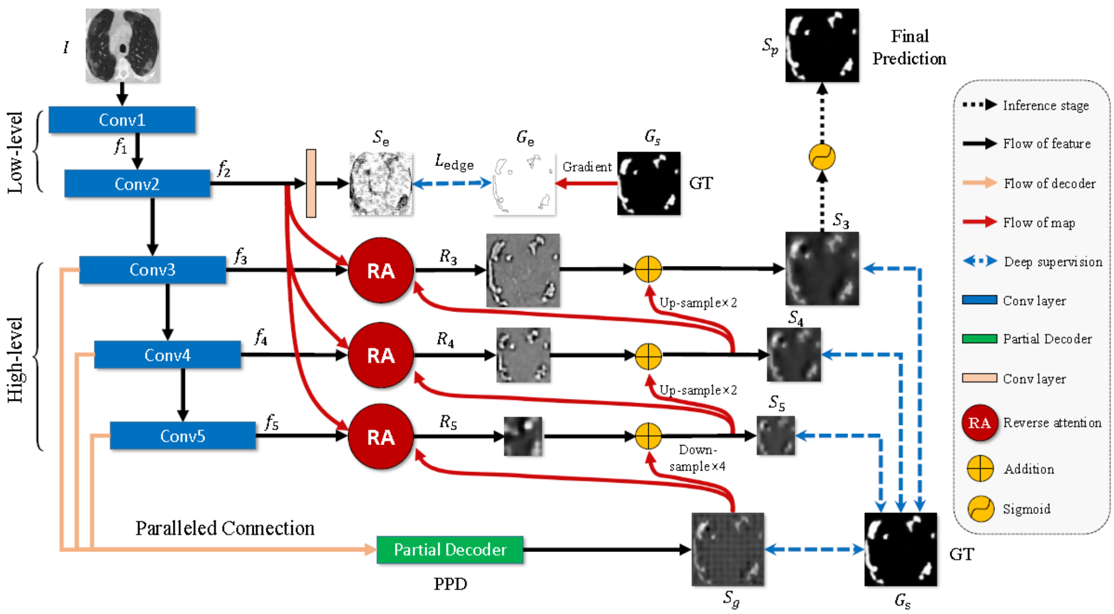


**Fig 2.**The architecture of the Inf-Net model (Fan et al., 2020a)

After segmenting the images, each lung was divided into six parts. Following that, the infection pixels in each section were counted. In the next step, the proportion of infections to the total lung area was calculated. Finally, seven scores were devoted to each patient that represented the total infection rate and the infection rate in each part of their lung. The segmentation results are shown in figure 3.

[3] reverse attention

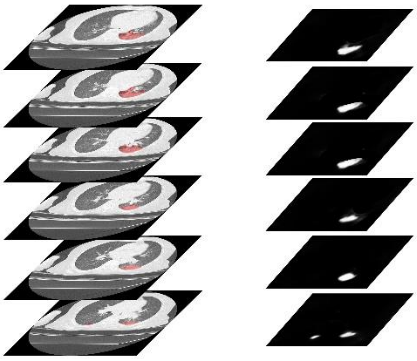

**Fig 3.** Visualization of infection segmentation results

After segmenting the images, each lung was divided into six parts. Following that, the infection pixels in each section were counted. In the next step, the proportion of infections to the total lung area was calculated. Finally, seven scores were devoted to each patient that represented the total infection rate and the infection rate in each part of their lung. The segmentation results are shown in figure 3.

### 3-3-Statistical Analysis

Statistical analysis was performed using SPSS and Minitab. $P<0.05$ was regarded to demonstrate statistical significance. Table 1 provides data about descriptive analysis for each peak and the total of patents. As is presented, the mean standard deviation of infection rate for all the patients was 0.045 and 0.025, respectively.

**Table 1.** Descriptive analysis for total infection rate

| Peak | N (Males, Females) | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| 1 | 90 (50,40) | 0.033 | 0.017 | 0.002 | 0.030 | 0.037 | 0.007 | 0.100 |
| 2 | 90 (46,44) | 0.049 | 0.025 | 0.003 | 0.044 | 0.054 | 0.013 | 0.124 |
| 3 | 90 (51,39) | 0.061 | 0.026 | 0.003 | 0.056 | 0.067 | 0.032 | 0.159 |
| **Total** | **270 (147,123)** | **0.048** | **0.026** | **0.002** | **0.045** | **0.051** | **0.007** | **0.159** |

The Anderson-Darling test was used to identify the distribution of data in each peak. Because the samples' distribution was not normal, we used the Friedman test to detect mean differences among three peaks.

## 4-Total infection analysis

Table 2 illustrates the results of the Friedman test for the total infection index. As can be seen, both assumption and exact sig were less than 0.05, and the null hypothesis was rejected. Test of homogeneity of variances also depicts the difference for them. Hence, the Tamhane test as the multiple comparison tests also shows the difference in means for all three peaks, and the average of COVID-19 infections has increased significantly in the third peak.

**Table 2.** Friedman Test results for the total infection rate

| N | Chi-Square | df | Asymp. Sig. | Exact Sig. | Point Probability |
|---|---|---|---|---|---|
| 90 | 103.2 | 2 | 0 | 0 | 0 |

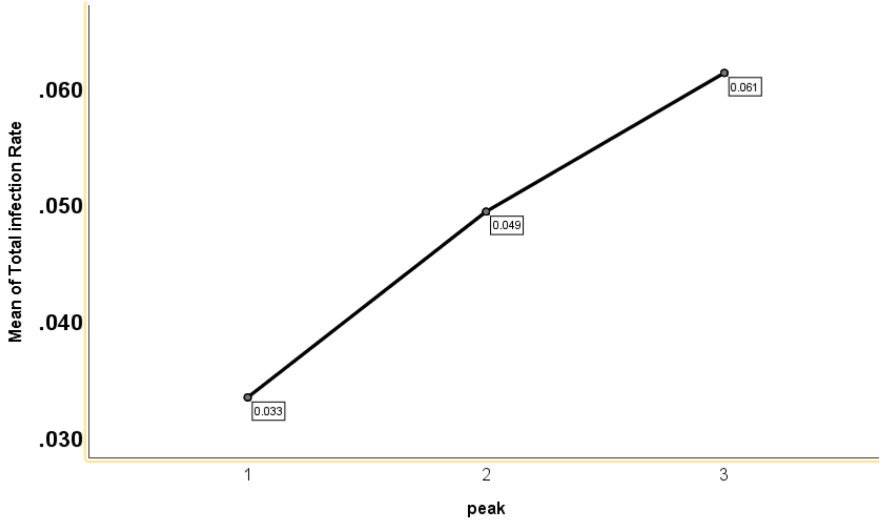


**Fig 4.** Eighty percent increase in the average infection rate during the last nine months

## 4-1-Subsections infection analysis

The table illustrates the result of the Fridman test for subsections. Contrary to other sections, the null hypothesis for the upper left lung and the lower-left lung is not rejected. The average lung involvement with infection is almost constant in these two sections. Like the total infection rate, the other four subsections' average infection rate has increased from the first peak to the third peak.

**Table3.** Friedman Test results for the infection rate in each section

| Section | N | Chi-Square | df | Asymp. Sig. | Exact Sig. | Point Probability |
|---|---|---|---|---|---|---|
| Upper-right | 90 | 6.067 | 2 | 0.048 | 0.059 | 0.004 |
| Upper- left | 90 | 2.222 | 2 | 0.329 | 0.332 | 0.006 |
| Middle-right | 90 | 42.956 | 2 | 0.000 | 0.000 | 0.000 |
| Middle-left | 90 | 20.267 | 2 | 0.000 | 0.000 | 0.000 |
| Lower-right | 90 | 21.422 | 2 | 0.000 | 0.000 | 0.000 |
| Lower-Left | 90 | 4.356 | 2 | 0.113 | 0.116 | 0.006 |

## 4-2-Gender and infection analysis

Friedman's test was used to assess the infection rate of the disease in males and females. As it turns out in figure 5, the null hypothesis was rejected, and the average lesion volume in men is about twenty percent higher than women.

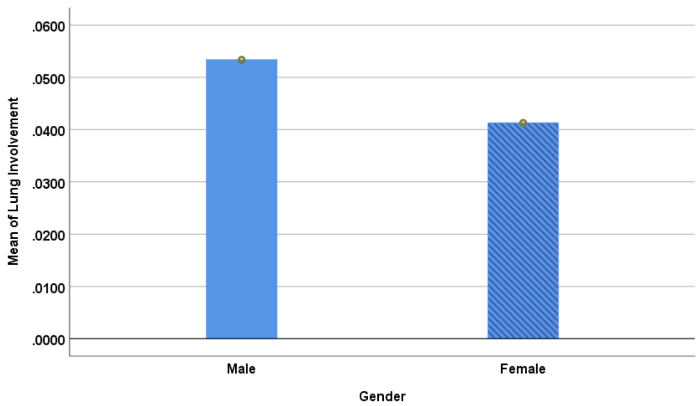


**Fig 5.** Comparison of total infection rate between men women during the period

## 5-Discussion and conclusion

In many countries, the spread of COVID-19 has entered the second wave. In Iran, however, with a marked increase of infections and a rise in the number of deaths, the country is experiencing a third wave of the Coronavirus outbreak. There is a concern about the next waves in other parts of the world. Coronavirus mutation also is a critical concern for scientists and can make the pandemic harder to stop. So it is essential to compare the different features of the virus during the pandemic.

When a patient has fever and respiratory symptoms, CT imaging will be used. Professional radiologists can judge the probability of COVID-19 based on their clinical experiences; however, their determination is easily affected by different factors such as individual proficiency. In comparison, deep learning systems can digitize and standardize the image information and improve clinical decision accuracy.

In this study, a novel method was presented to score the lung infection rate in each patient's lung using a deep learning network to segment CT images. Then, we compared the lung infection rates during the last three peaks using statistical methods. In conclusion, the average infection rate for COVID-19 has been doubled during the last three peaks. Another noticeable point was that 61% of images showed that the lesion volume in the right lung was larger than those on the left, and the average lesion volume in all sections on the right side had been increased. Contrary to them, the lesion volume has remained constant in the same period on the left lung. In terms of gender, lesion volume was about 20% higher in men. It is worth to mention that the quantitative scoring system that calculates the lesion volume in this study can be used as a reliable index for comparing the COVID-19's changes over time.

This study has some limitations because of the strict privacy rules, which hinder us from collecting more data. The accuracy of the proposed model will go up by expanding the number of samples. Second, it is essential to combine the image results with each patient's clinical diagnosis, such as first symptoms and laboratory examination. For future research, we are interested in gathering more data from each patient to investigate the relations between high severity infections and other personal data such as their lifestyle and previous diseases.

## References

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., ... & Xia, L. (2020). Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*, *296*(2), E32-E40.

Alimohamadi, Y., Taghdir, M., & Sepandi, M. (2020). Estimate of the basic reproduction number for COVID-19: a systematic review and meta-analysis. *Journal of Preventive Medicine and Public Health*, *53*(3), 151.

Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., & Mohammadi, A. (2020). Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Computers in Biology and Medicine*, *121*, 103795.

Ding, X., Xu, J., Zhou, J., & Long, Q. (2020). Chest CT findings of COVID-19 pneumonia by duration of symptoms. *European journal of radiology*, *127*, 109009.

Fan, D. P., Zhou, T., Ji, G. P., Zhou, Y., Chen, G., Fu, H., ... & Shao, L. (2020). Inf-net: Automatic covid-19 lung infection segmentation from ct images. *IEEE Transactions on Medical Imaging*, *39*(8), 2626-2637.

Fan, L., Li, D., Xue, H., Zhang, L., Liu, Z., Zhang, B., ... & Jin, Z. (2020). Progress and prospect on imaging diagnosis of COVID-19. *Chinese journal of academic radiology*, *3*(1), 4-13.

Hopkins, J. (2020). COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE). *Coronavirus Resource Center*.

Hu, Y., Sun, J., Dai, Z., Deng, H., Li, X., Huang, Q., ... & Xu, Y. (2020). Prevalence and severity of corona virus disease 2019 (COVID-19): A systematic review and meta-analysis. *Journal of clinical virology*, 104371.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, *42*, 60-88.

Long, S. W., Olsen, R. J., Christensen, P. A., Bernard, D. W., Davis, J. J., Shukla, M., ... & Musser, J. M. (2020). Molecular architecture of early dissemination and massive second wave of the SARS-CoV-2 virus in a major metropolitan area. *MBio*, *11*(6).

Manigandan, S., Wu, M. T., Ponnusamy, V. K., Raghavendra, V. B., Pugazhendhi, A., & Brindhadevi, K. (2020). A systematic review on recent trends in transmission, diagnosis, prevention and imaging features of COVID-19. *Process Biochemistry*.

Nikpouraghdam, M., Farahani, A. J., Alishiri, G., Heydari, S., Ebrahimnia, M., Samadinia, H., ... & Bagheri, M. (2020). Epidemiological characteristics of coronavirus disease 2019 (COVID-19) patients in IRAN: A single center study. *Journal of Clinical Virology*, *127*, 104378.

Ozturka, T., Talob, M., Yildirimc, E. A., Baloglud, U. B., Yildirime, O., & Acharyafgh, U. R. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in Biology and Medicine121 (2020), 103792.*

Pan, Y., Guan, H., Zhou, S., Wang, Y., Li, Q., Zhu, T., ... & Xia, L. (2020). Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): a study of 63 patients in Wuhan, China. *European radiology*, *30*(6), 3306-3309.

Phillips-Wren, G., & Ichalkaranje, N. (Eds.). (2008). *Intelligent decision making: An AI-based approach* (Vol. 97). Springer Science & Business Media.

Raoofi, A., Takian, A., Sari, A. A., Olyaeemanesh, A., Haghighi, H., & Aarabi, M. (2020). COVID-19 pandemic and comparative health policy learning in Iran. *Archives of Iranian medicine*, *23*(4), 220-234.

Ribeiro, M. H. D. M., da Silva, R. G., Mariani, V. C., & dos Santos Coelho, L. (2020). Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. *Chaos, Solitons & Fractals*, *135*, 109853.

Shan, F., Gao, Y., Wang, J., Shi, W., Shi, N., Han, M., ... & Shi, Y. (2020). Lung infection quantification of COVID-19 in CT images with deep learning. *arXiv preprint arXiv:2003.04655*.

World Health Organization. Coronavirus disease (COVID-2019) situation reports. Geneva, Switzerland: World Health Organization; 2020.

Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., ... & Li, L. (2020). A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering*, *6*(10), 1122-1129.

Yang, R., Li, X., Liu, H., Zhen, Y., Zhang, X., Xiong, Q., ... & Zeng, W. (2020). Chest CT severity score: an imaging tool for assessing severe COVID-19. *Radiology: Cardiothoracic Imaging*, *2*(2), e200047.

Zhou, C., Pei, T., DU, Y., CHEN, J., XU, J., WANG, J., ... & JIANG, L. (2020). Big data analysis on COVID-19 epidemic and suggestions on regional prevention and control policy. *Bulletin of Chinese Academy of Sciences*, *35*(2), 200-203.

Zhao, W., Zhong, Z., Xie, X., Yu, Q., & Liu, J. (2020). Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study. *American Journal of Roentgenology*, *214*(5), 1072-1077.

Zuo, M., Khosa, S. K., Ahmad, Z., & Almaspoor, Z. (2020). Comparison of COVID-19 pandemic dynamics in Asian countries with statistical modeling. *Computational and mathematical methods in medicine, 2020.*