# **Determination of Robust Regional CT Radiomics Features for COVID-19**

Mahbubunnabi Tamal <sup>1,\*</sup>

\*Department of Biomedical Engineering

Imam Abdulrahman Bin Faisal University

PO Box 1982, Dammam 31441, Saudi Arabia

mtamal@yahoo.com; mtamal@iau.edu.sa

Declarations of interest: none

# ABSTRACT

### **Background:**

The lung CT images of COVID-19 patients can be characterized by three different regions – Ground Glass Opacity (GGO), consolidation and pleural effusion. GCOs have been shown to precede consolidations. Quantitative characterization of these regions using radiomics can facilitate accurate diagnosis, disease progression and response to treatment. However, according to the knowledge of the author, regional CT radiomics analysis of COVID-19 patients has not been carried out. This study aims to address these by determining the radiomics features that can characterize each of the regions separately and can distinguish the regions from each other.

### Methods:

44 radiomics features were generated with four quantization levels for 23 CT slice of 17 patients. Two approaches were the implemented to determine the features that can differentiate between lung regions -1) Z-score and correlation heatmaps and 2) one way ANOVA for finding statistically significantly difference (p<0.05) between the regions. Radiomics features that show agreement for all cases (Z-score, correlation and statistical significant test) were selected as suitable features. The features were then tested on 52 CT images.

#### **Results:**

10 radiomics features were found to be the most suitable among 44 features. When applied on the test images, they can differentiate between GCO, consolidation and pleural effusion successfully and the difference provided by these 10 features between three lung regions are statistically significant.

#### **Conclusion:**

The ten robust radiomics features can be useful in extracting quantitative data from CT lung images to characterize the disease in the patient, which in turn can help in more accurate diagnosis, staging the severity of the disease and allow the clinician to plan for more successful personalized treatment for COVID-19 patients. They can also be used for monitoring the progression of COVID-19 and response to therapy for clinical trials.

**Keywords:** COVID-19; CT Radiomics; Ground Glass Opacity (GGO); Consolidation; Pleural Effusion.

# INTRODUCTION

COVID-19 caused by severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) is the worldwide leading cause of death at the moment and is a public health emergency concern across the globe [1]. A total of 437,532 deaths with a confirmed infection 8,063,488 has been reported [2].

To control and minimize the spread of COVID-19, it is important to diagnose it as quickly and accurately as possible. Though reverse transcription polymerase chain reaction (RT-PCR) is considered as a gold standard point-of-care diagnostic tool with the highest specificity [3], it requires several hours to get the results, the sensitivity is very low with a very high rate of false negatives and cannot provide the detail condition of the patient [4].

On the other hand, it has been reported that 20% people who are suffering from COVID-19 needs to visit hospital because of respiratory distress and 30% of them require intensive care [5]. In the absence of appropriate drug, two third of the COVID-19 patients who goes to critical care requires prolong period of mechanical ventilation within 24 hours of admission [6]. This group of patients also requires continuous monitoring of their lung health to assess their overall conditions.

Because of its high sensitivity specially within the first two days of infection and capability to provide detail condition of the patient, computed tomography (CT) is routinely being used as a diagnostic and screening tool in several countries including China [7-11]. Quantitative imaging information such as the lesion percentage and the CT mean density values could help the

physician to monitor the progression of the disease and the response to therapy [12]. More in depth analysis of different first, second and third order statistical CT features known as radiomics have been successfully utilized for decoding the radiographic phenotype in cancer [13, 14]. Accurate quantification of tumour radiomic features was made possible by utilizing AI [15]. In case of COVID-19, only a very few studies have been carried out for radiomics analysis. One of the studies has shown that CT radiomics not only can help to screen patients [16] but also can be used as a discriminating tool for diagnosis and monitoring different stages of respiratory abnormalities [17]. Utilizing AI with radiomics can make the process faster and more accurate [18].

All the studies so far extracted the radiomics features from the whole CT lung region [16, 18, 19]. The lung on a CT of a COVID-19 patient can contain either one or any combinations of three different regions - ground-glass opacities (GCO), consolidation and pleural effusion as shown in Figure 1.



**Figure 1:** (a) One CT image slice with three distinctive regions – GCO (red), consolidation (blue) and pleural effusion (green). (b) two distinctive regions GCO (red) and consolidation (blue).

Radiomics features extracted from the whole lung provides the combinatorial outcome of the features of each of the three different regions. On the other hand, not all the features are able to distinguish between regions and are suitable to be used as image based biomarkers for accurate diagnosis, evaluation as well as monitoring disease progression and response to therapy. Since GCOs have been shown to precede consolidations, it is vital to determine the most suitable radiomics features that can distinguish and quantify different regions. This requires detail and careful investigation. According to the knowledge of the author, no such study has been conducted for COVID-19 lung CT regions. This study aims to address these by determining the radiomics features that can characterize each of the regions separately and can distinguish the regions from each other.

#### MATERIALS AND METHODS

COVID-19 CT images were collected from available data derived from Italian Society of Medical and Interventional Radiology website [20, 21]. A total of 100 CT slices of 43 COVID-19 patients were available. GCO, consolidation and pleural effusion were segmented by a radiologists. Out of 100 slices, 96 slices have GCO, 78 slices have consolidation and 25 slices have pleural effusion. Only 23 of slices of 17 patients have all three regions. Rest of the slices had any one or two regions. A total of 44 radiomics features were extracted as shown in Table 1. Out of these 44 radiomics features, 43 features are textural features with the first three features are global features extracted from image histogram [22]. These 43 features can be broadly categorized into six different groups- 1) global texture based on histogram, 2) Grey Level Co-Occurrence Matrix (GLCM), 3) Grey Level Run Length Matrix (GLRLM), 4) Grey Level Size

Zone Matrix (GLSZE) and 5) NGTDM. To extract the rest 40 features, each region needs to be quantized to remove dependency on the image intensity.

No.	Feature Name	Feature Type	No	Feature Name	Feature Type
1	Variance	Histogram (global texture)	23	Long Run High Gray-Level Emphasis (LRHGE)	GLRLM
2	Skewness	24 Gray-Level		Gray-Level Variance (GLV)	
3	Kurtosis		25	Run-Length Variance (RLV)	
4	Energy	GLCM	26	Small Zone Emphasis (SZE)	GLSZM
5	Contrast		27	Large Zone Emphasis (LZE)	
6	Entropy		28	Gray-Level Non-uniformity (GLN)	
7	Homogeneity		29	Zone-Size Non-uniformity (ZSN)	
8	Correlation		30	Zone Percentage (ZP)	
9	Sum Average		31	Low Gray-Level Zone Emphasis (LGZE)	
10	Variance		32	High Gray-Level Zone Emphasis (HGZE)	
11	Dissimilarity		33	Small Zone Low Gray-Level Emphasis (SZLGE)	
12	Autocorrelation		34	Small Zone High Gray-Level Emphasis (SZHGE)	
13	Short Run Emphasis (SRE)	GLRLM	35	Large Zone Low Gray-Level Emphasis (LZLGE)	
14	Long Run Emphasis (LRE)		36	Large Zone High Gray-Level Emphasis (LZHGE)	
15	Gray-Level Non- uniformity (GLN)		37	Gray-Level Variance (GLV)	
16	Run-Length Non- uniformity RLN		38	Zone-Size Variance (ZSV)	
17	Run Percentage (RP)		39	Coarseness	NGTDM
18	Low Gray-Level Run Emphasis (LGRE)		40	Contrast	
19	High Gray-Level Run Emphasis (HGRE)		41	Busyness	
20	Short Run Low Gray- Level Emphasis (SRLGE)		42	Complexity	
21	Short Run High Gray- Level Emphasis (SRHGE)		43	Strength	

**Table 1:** Radiomics Features [22]

22	Long Run Low Gray-	44	Edgeness	Edge
	(LRLGE)			

To investigate if quantization has any impact on the features, four different levels of quantization were used (16, 32, 64 and 128). The final feature is an edge feature where numbers of pixels having edges more than a certain threshold within a region were counted.

Two approaches were implemented to determine the features that can differentiate between lung pathological regions. Firstly, Z-score and correlation heatmaps were generated for all regions for all 100 slices to identify the features that can classify different regions and to remove redundant features respectively. Secondly, one way ANOVA was then performed only on 23 slices that contain all the three regions to determine the features that are statistically significantly different (p < 0.05) between the regions. Radiomics features that show agreement for all cases (Z-score, correlation and statistical significant test) were selected as suitable features and were applied on separate 52 slices containing only GGO and consolidation to test their accuracies by evaluating their *p*-values.

### RESULTS

Figure 2 shows the Z-score heatmap of all 44 features. The Z-score for each patient were calculated and arranged according to different regions. The correlation coefficients along with the *p*-values for testing the hypothesis of no correlation for all the 44 features are shown in Figure 3. Careful investigation of Figure 2 reveals that 10 radiomics features have better classification ability between GGO, consolidation and pleural effusion. These features are also have lower correlation (Figure 3).



**Figure 2:** Z-score heatmap of all 44 features. Each row corresponds to one feature represented by one number (Table 1). Each column represent one regions. Left 96 column represent GGO, middle 78 represent consolidation and last 25 represent pleural effusion. The features that can distinguish between these three pathological regions are shown by the white bounding box.



**Figure 3:** (a) correlation coefficient of 44 radiomics features. (b) p values describing the relationship between the features.

Out of 44 radiomics features, the same 10 features were found to be statistically significantly different between GCO, consolidation and plural effusion and do not depend on the quantization level. The features are shown in Table 2 along with their corresponding p values for three different regions.

No	Foot	Footure Nome	n			nancalidation
110.	reat	Feature Name	p	p(GCO,	p(GCO,	p(consolution,
	ure No			consonuation)	pieurai offusion)	offusion
-	INO.				ellusion)	ellusion)
1	2	Skewness	0.000	0.000	0.000	0.000
2	9	GLCM - Sum Average	0.000	0.000	0.000	0.000
3	12	GLCM - Autocorrelation	0.000	0.000	0.000	0.000
4	19	GLRLM - High Gray-	0.000	0.000	0.000	0.000
		Level Run Emphasis (HGRE)				
5	21	GLRLM - Short Run	0.000	0.000	0.000	0.000
		High Gray-Level				
		Emphasis (SRHGE)				
6	23	GLRLM - Long Run	0.000	0.000	0.000	0.000
		High Gray-Level				
		Emphasis (LRHGE)				
7	24	GLRLM - Gray-Level	0.000	0.001	0.000	0.000
		Variance (GLV)				
8	32	GLSZM - High Gray-	0.000	0.000	0.000	0.000
		Level Zone Emphasis				
0	<b>2</b> (	(HGZE)	0.000	0.000	0.000	0.000
9	34	GLSZM - Small Zone	0.000	0.000	0.000	0.000
		High Gray-Level				
10	27	Emphasis (SZHGE)	0.000	0.040	0.000	0.001
10	37	GLSZM - Gray-Level	0.000	0.048	0.000	0.001
		Variance (GLV)				

**Table 2:** p-values extracted with one way ANOVA

The corresponding boxplots of all the three regions for 10 different radiomics features are shown in Figure 4.



**Figure 4:** Box plot of 23 slices (notches in the boxplot representing group medians) of 10 radiomics features that are significantly different for three different regions (1- GGO, 2- Consolidation and 3- Pleural Effusion).

These 10 radiomics features (Table 2) for GGO and consolidation for the other 52 slices as shown in Figure 5. All these features were statistically significantly different between the two regions with the corresponding p-values were always less than 0.05.



**Figure 5:** Box plot of 52 slices (notches in the boxplot representing group medians) of 10 radiomics features for GGO and consolidation (represented by 1 and 2 on the X-axis respectively).

# DISCUSSION

Because of its high sensitivity specially within the first two days of infection and capability to provide detail condition of the patient, computed tomography (CT) is routinely being used as a diagnostic and screening tool in several countries [23]. It also shows earlier positive sign of improvement for those patients who recover compared to RT-PCR results [23, 24].

The CT images of COVID-19 patients show distinctive patterns of infection with the hallmarks being bilateral and peripheral ground-glass and consolidation [25]. Sometimes CT examination showed a mixed and diverse pattern with the involvement of both lung parenchyma and interstitium [26] that includes bilateral multilobar ground-glass opacification (GGO) with a

peripheral or posterior distribution [27]. It has been reported that temporal changes in CT manifestations can indicate progression and recovery of COVID-19 [28]. A manual chest CT severity score (CT-SS) has been proposed for COVID-19 to rapidly identify patients with severe forms of the disease [29]. 83.3% sensitivity and 94% specificity were achieved with this approach. Lung abnormalities on chest CT scans reported to show greatest severity approximately 10 days after initial onset of symptoms [30].

All these studies were performed retrospectively where radiologist had the time to examine each scan in detail. However, with the increasing number of CT scans performed on COVID-19 patients, it becomes difficult for the radiologists to examine the scan accurately in real time. On the other hand, intra- and inter-observer variability can introduce uncertainty in the staging the disease. Moreover, all the textural pattern that are considered to be the hallmark of COVID-19 may not be visible to the naked human eye especially it becomes more challenging in 3D.

Because of these reason, there is a dire need of a CT diagnostic method that can perform the task of rapid and early detection and monitoring quantitatively with high repeatability. Radiomics features can provide quantitative information regarding tissue heterogeneity and have been correlated with disease progression and prognosis in cancer.

In contrast to the previous radiomics studies where the focus was to extract the radiomics features for the whole lung of the COVID-19 patient [18, 19], this study focused on the regional extraction of the radiomics features from CT images. The argument for such approach is that when radiomics features are extracted for whole lung the regional information is diluted that can

carry vital information regarding the disease stage, progression and response to therapy. So it was important to find the radiomics features that can distinguish different regions.

This investigation confirms that not all the proposed radiomics features can distinguish between GGO, consolidation and pleural effusion. Only a handful of features are able to differentiate between these regions and can be used for accurate quantitation of the disease.

### CONCLUSION

To the best of author's knowledge, this paper presents the first study to determine radiomics features that can differentiate between ground glass opacity, consolidation and pleural effusion from CT data for COVID-19 patients. The study reveals that only ten radiomics features can differentiate between these regions. They are – 1) Skewness, GLCM – Sum Average, 3) GLCM –Autocorrelation, 4) GLRLM - High Gray-Level Run Emphasis (HGRE), 5) GLRLM - Short Run High Gray-Level Emphasis (SRHGE), 6) GLRLM - Long Run High Gray-Level Emphasis (LRHGE), 7) GLRLM - Gray-Level Variance (GLV), 8) GLSZM - High Gray-Level Zone Emphasis (HGZE), 9) GLSZM - Small Zone High Gray-Level Emphasis (SZHGE) and 10) GLSZM - Gray-Level Variance (GLV). These ten features can be useful in extracting quantitative data from CT lung images to characterize the disease in the patient, which in turn can help in more accurate diagnosis, staging the severity of the disease and allow the clinician to plan for more successful personalized treatment for COVID-19 patients. Future work is planned to link these quantitative radiomics information with the clinical outcomes so that they can be used for monitoring the progression of COVID-19 and response to therapy for clinical trials.

# **Funding:**

The author extend his appreciation to the Deputyship for Research & Innovation, Ministry of Education, Saudi Arabia for funding this research work.

### **Disclosure and conflicts of interest**

The author declares that he has no conflict of interest.

### REFERENCES

1 https://<u>www.who.int/</u>

2 https://coronavirus.jhu.edu/map.html

3 Chan, J.F., Yuan, S., Kok, K.H., To, K.K., Chu, H., Yang, J., Xing, F., Liu, J., Yip, C.C., Poon, R.W., Tsoi, H.W., Lo, S.K., Chan, K.H., Poon, V.K., Chan, W.M., Ip, J.D., Cai, J.P., Cheng, V.C., Chen, H., Hui, C.K., and Yuen, K.Y.: 'A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster', Lancet, 2020, 395, (10223), pp. 514-523

Bai, H.X., Hsieh, B., Xiong, Z., Halsey, K., Choi, J.W., Tran, T.M.L., Pan, I., Shi, L.B., Wang, D.C., Mei, J., Jiang, X.L., Zeng, Q.H., Egglin, T.K., Hu, P.F., Agarwal, S., Xie, F., Li, S., Healey, T., Atalay, M.K., and Liao, W.H.: 'Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT', Radiology, 2020, pp. 200823

5 Myers, L.C., Parodi, S.M., Escobar, G.J., and Liu, V.X.: 'Characteristics of Hospitalized Adults With COVID-19 in an Integrated Health Care System in California', Jama, 2020

6 Elisabeth, M.: 'Covid-19: most patients require mechanical ventilation in first 24 hours of critical care', BMJ 2020, 2020, 368:m1201

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., and Xia, L.: 'Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases', Radiology, 0, (0), pp. 200642

8 Fang, Y., Zhang, H., Xie, J., Lin, M., Ying, L., Pang, P., and Ji, W.: 'Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR', Radiology, 2020, pp. 200432

9 Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., and Cao, B.: 'Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China', Lancet, 2020, 395, (10223), pp. 497-506

10 Xu, B., Xing, Y., Peng, J., Zheng, Z., Tang, W., Sun, Y., Xu, C., and Peng, F.: 'Chest CT for detecting COVID-19: a systematic review and meta-analysis of diagnostic accuracy', European radiology, 2020

11 Huang, Y., Cheng, W., Zhao, N., Qu, H., and Tian, J.: 'CT screening for early diagnosis of SARS-CoV-2 infection', The Lancet. Infectious diseases, 2020

12 Qi, X., Lei, J., Yu, Q., Xi, Y., Wang, Y., and Ju, S.: 'CT imaging of coronavirus disease 2019 (COVID-19): from the qualitative to quantitative', Annals of translational medicine, 2020, 8, (5), pp. 256

van Griethuysen, J.J.M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V.,
Beets-Tan, R.G.H., Fillion-Robin, J.C., Pieper, S., and Aerts, H.: 'Computational Radiomics
System to Decode the Radiographic Phenotype', Cancer research, 2017, 77, (21), pp. e104-e107

Aerts, H.J., Velazquez, E.R., Leijenaar, R.T., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M.M., Leemans, C.R., Dekker, A., Quackenbush, J., Gillies, R.J., and Lambin, P.: 'Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach', Nature communications, 2014, 5, pp. 4006

15 Choi, Y., Ha, S., Lee, Y.S., Kim, Y.K., Lee, D.S., and Kim, D.J.: 'Development of tau PET Imaging Ligands and their Utility in Preclinical and Clinical Studies', Nuclear medicine and molecular imaging, 2018, 52, (1), pp. 24-30

Mengjie, F., Bingxi, H., Li, L., Di, D., Xin, Y., Cong, L., Lingwei, M., Lianzhen, Z., Hailin, L., Hongjun, L., and Jie, T.: 'CT radiomics can help screen the Coronavirus disease 2019 (COVID-19): a preliminary study', SCIENCE CHINA Information Sciences, 2020, 63, (7)

17 Xiaolong, Q., Zicheng, J., QIAN, Y., Chuxiao, S., Hongguang, Z., Hongmei, Y., Baoyi, M., Yuancheng, W., Chuan, L., Xiangpan, M., Shan, H., Jitao, W., Dan, X., Junqiang, L., Guanghang, X., Huihong, H., Jie, Y., Jiansong, J., Hongqiu, P., Shengqiang, Z., and Shenghong, J.: 'Machine learning-based CT radiomics model for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: A multicenter study ', medRxiv : the preprint server for health sciences, 2020

18 Cheng, J., Weixiang, C., Yukun, C., Zhanwei, X., Zimeng, T., Xin, Z., Lei, D., Chuansheng, Z., Jie, Z., Heshui, S., and Jianjiang, F.: 'Development and Evaluation of an AI System for COVID-19 Diagnosis', medRxiv : the preprint server for health sciences, 2020

Julien, G., Akshayaa, V., Louis, D., Fadila, Z., Denis, D., Anne-Noelle, F., Marie, T.,
Monique, H., Gregory, C., Stephane, M., Eva, E., Philippe, L., Nathan, T., Benjamin, M., Sean,
W., Michel, M., Renaud, L., Paul, M., Wim, V., Ralph, L., and Pierre, L.: 'Development and

validation of an automated radiomic CT signature for detecting COVID-19 ', medRxiv : the preprint server for health sciences, 2020

20 https://www.sirm.org/en/

21 https://medicalsegmentation.com/covid19/

Vallieres, M., Freeman, C.R., Skamene, S.R., and El Naqa, I.: 'A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities', Physics in medicine and biology, 2015, 60, (14), pp. 5471-5496

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., and Xia, L.: 'Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases', Radiology, 2020, pp. 200642

24 Xie, X., Zhong, Z., Zhao, W., Zheng, C., Wang, F., and Liu, J.: 'Chest CT for Typical 2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing', Radiology, 2020, pp. 200343

Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z.A., Zhang, N., Diao, K., Lin, B.,
Zhu, X., Li, K., Li, S., Shan, H., Jacobi, A., and Chung, M.: 'Chest CT Findings in Coronavirus
Disease-19 (COVID-19): Relationship to Duration of Infection', Radiology, 2020, 295, (3), pp.
200463

Zhou, S., Wang, Y., Zhu, T., and Xia, L.: 'CT Features of Coronavirus Disease 2019 (COVID-19) Pneumonia in 62 Patients in Wuhan, China', AJR. American journal of roentgenology, 2020, 214, (6), pp. 1287-1294

27 Salehi, S., Abedi, A., Balakrishnan, S., and Gholamrezanezhad, A.: 'Coronavirus Disease 2019 (COVID-19): A Systematic Review of Imaging Findings in 919 Patients', AJR. American journal of roentgenology, 2020, 215, (1), pp. 87-93

28 Wang, Y., Dong, C., Hu, Y., Li, C., Ren, Q., Zhang, X., Shi, H., and Zhou, M.: 'Temporal Changes of CT Findings in 90 Patients with COVID-19 Pneumonia: A Longitudinal Study', Radiology, 2020, pp. 200843

29 Ran, Y., Xiang, L., Huan, L., Yanling, Z., Xianxiang, Z., Qiuxia, X., Yong, L., Cailiang,

G., and Wenbing, Z.: 'Chest CT Severity Score: An Imaging Tool for Assessing Severe COVID-

19', Radiology: Cardiothoracic Imaging, 2020, 2, (2)

30 Pan, F., Ye, T., Sun, P., Gui, S., Liang, B., Li, L., Zheng, D., Wang, J., Hesketh, R.L.,

Yang, L., and Zheng, C.: 'Time Course of Lung Changes at Chest CT during Recovery from

Coronavirus Disease 2019 (COVID-19)', Radiology, 2020, 295, (3), pp. 715-721