Deep-learning characterization and quantification of COVID-19 pneumonia lesions from chest CT images

D. Bermejo-Peláez[1,7], R. San José Estépar[2], M. Fernández-Velilla[3], C. Palacios Miras[4], G. Gallardo Madueño[5], M. Benegas[6], M.A. Luengo-Oroz[7], J. Sellarés[6], M. Sánchez[6], G. Bastarrika[5], G. Peces Barba[4], L.M. Seijo[5], M.J. Ledesma-Carbayo[1]

[1]Biomedical Image Technologies, Universidad Politécnica de Madrid & CIBER-BBN, Madrid, Spain [2]Applied Chest Imaging Laboratory, Brigham and Women's Hospital, Boston, MA, USA [3]Hospital Universitario La Paz, Madrid, Spain [4]Hospital Universitario Fundación Jiménez Díaz & CIBER-ES, Madrid, Spain [5]Clínica Universidad de Navarra & CIBER-ES, Madrid, Spain [6]Hospital Clinic de Barcelona, Barcelona, Spain [7]Spotlab, Madrid, Spain

ABSTRACT

A relevant percentage of COVID-19 patients present bilateral pneumonia. Disease progression and healing is characterized by the presence of different parenchymal lesion patterns. Artificial intelligence algorithms have been developed to identify and assess the related lesions and properly segment affected lungs, however very little attention has been paid to automatic lesion subtyping. In this work we present artificial intelligence algorithms based on CNN to automatically identify and quantify COVID-19 pneumonia patterns. A Dense-efficient CNN architecture is presented to automatically segment the different lesion subtypes. The proposed technique has been independently tested in a multicentric cohort of 100 patients, showing Dice coefficients of 0.988±0.01 for ground glass opacities, 0.948±0.05 for consolidations, and 0.999±0.0003 for healthy tissue with respect to radiologist's reference segmentations, and high correlations with respect to radiologist severity visual scores.

Keywords: computed-aided diagnosis, COVID-19, lesion subtyping, CT, CNN

1. INTRODUCTION

Between 17% and 29% of patients with COVID-19 are affected by severe or even fatal respiratory disease [1]. In this situation an appropriate classification of patients admitted to hospitals according to their prognosis and severity is needed to define their management.

Currently, prediction and quantification of patient severity is still limited. There are several studies that have tried to determine severity based on different clinical and biological parameters, or on the assessment of lesions on CT images [2].

Computed Tomography (CT) was used early in the pandemic in China as the imaging technique of choice for the diagnosis of patients with suspected COVID-19 (even with negative RT-PCR) and is useful for monitoring disease as well as response to treatment [4]. Many algorithms have been published that successfully identify patients with COVID-19 from CT images [5] or automatically score the severity and prognosis from the CT information [6,7], demonstrating the potential use of artificial intelligence in this field. However, despite the important research effort in artificial intelligence applied to the acquisition, segmentation and diagnosis of data in COVID-19 [8], little attention has been paid to automatic subtyping of lesions. Parenchymal tissue lesions change during disease progression [8], so an automatic characterization of these lesion types could serve for better severity assessment, patient stratification, and treatment monitoring.

In this context, this work proposes a segmentation algorithm based on convolutional neural networks (CNN) for the automatic identification of COVID-19 lesion subtypes from CT images and the quantification of the extent of pulmonary involvement, to report the severity of pneumonia to serve in the stratification of patients. These models have been evaluated in a completely independent cohort of 100 patients.

Medical Imaging 2022: Computer-Aided Diagnosis, edited by Karen Drukker, Khan M. Iftekharuddin, Hongbing Lu, Maciej A. Mazurowski, Chisako Muramatsu, Ravi K. Samala, Proc. of SPIE Vol. 12033, 120331V · © 2022 SPIE · 1605-7422 · doi: 10.1117/12.2613086

2. METHODS

2.1 Training database

The dataset used to train the segmentation algorithm comes from a public database published by the Italian Society of Medical and Interventional Radiology. We used 100 axial CT slices of 60 patients confirmed by COVID-19 that were manually segmented by an expert radiologist considering three different types of lesions, including ground-glass opacities, consolidation pattern and pleural effusion. The images were transformed from JPG format to Hounsfield Units (HU) by setting an intensity transformation function from the original pixel values to HU considering the intensity values of the air and fat areas.

2.2 Dense-Efficient segmentation architecture

The proposed CNN architecture is illustrated in Figure 1. The encoder of the network comprises encoder blocks, which consists of a convolutional dense block and down-sampling Efficient-Net blocks. Dense connections allow for greater feature propagation through the network architecture and alleviates the vanishing gradient problem. [10]. Efficient-Net blocks avoid representation bottlenecks and loss of information by combining traditional max-pooling operations with strided convolutional operations [11]. It has been previously shown that the combination of dense and Efficient-Net blocks into segmentation architectures outperforms state-of-the-art CNN such as the well-known U-Net [12,13].
The decoding pathway is comprised by convolutional dense blocks, which processes transposed convolutional (deconvolutional) features together with higher resolution features maps from the encoder. Finally, the architecture ends with a convolutional operation (kernel of size 1x1) to produce 5 features maps corresponding to the background, healthy tissue, ground glass opacities, consolidation pattern and pleural effusion. This last convolutional operation has a softmax activation function to produce the final probability image map. The final image label map is computed by assigning to each pixel in the image the label which has the highest probability.
The training of the network was performed by minimizing a class-weighted Dice-based loss function which considers class imbalance using an Adam Stochastic Optimizer with an initial learning rate 1e-4 and plateau learning rate decay with a factor of 0.2 when the validation loss is not improved after 5 epochs. The training was ended when the validation loss was not improved after 10 epochs.
The network was implemented in Keras with Tensorflow and using a PC with a GPU NVIDIA Quadro P6000 (24GB).

3. EXPERIMENTS AND RESULTS

3.1 Independent test database

The proposed network for the segmentation of COVID-19 CT patterns was tested on a completely independent test set with data different from the used for training the method. This test set was composed by 100 CT scans from 100 patients positive for COVID-19 confirmed by PCR which were retrospectively collected from 4 different hospitals, Hospital Universitario La Paz, Hospital Universitario Fundación Jiménez Díaz, Clínica Universidad de Navarra, and Hospital Clinic de Barcelona after the ethical approval of their institutional review boards. All test CT scans were preprocessed by segmenting and masking only the lung region using a robust and publicly available model for lung parenchyma segmentation [14].

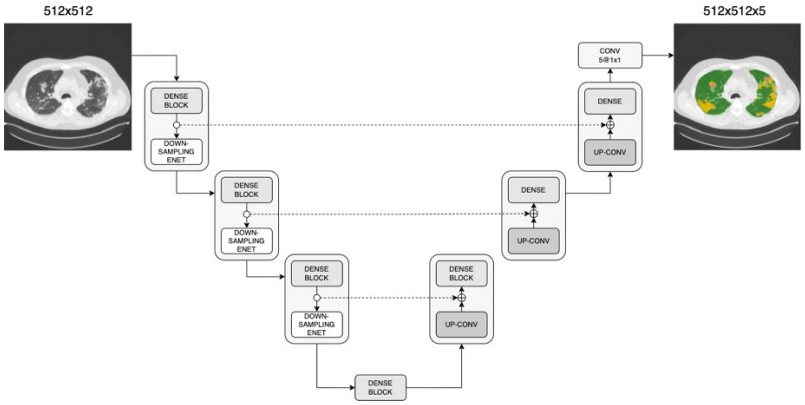

Figure 1. Dense-efficient CNN architecture for the segmentation of COVID-19 lesion subtypes.

3.2 Disease quantification

The proposed method has been developed with the purpose of being a valid and usable tool for radiologists. With this purpose, the algorithm was implemented so it automatically generates a report in an easy-to-read PDF format, reporting the lung involvement of disease subtypes in COVID-19. This solution provides detailed quantification of disease subtypes by reporting the percentage of affected volume for each CT and per lung zone. The report also includes a visual glyph representation which sums up the volume metrics helping to reduce reading variability and subjectivity. Finally, three axial, one coronal and one sagittal image with its corresponding segmentation results are also provided, as can be seen in Figure 2.

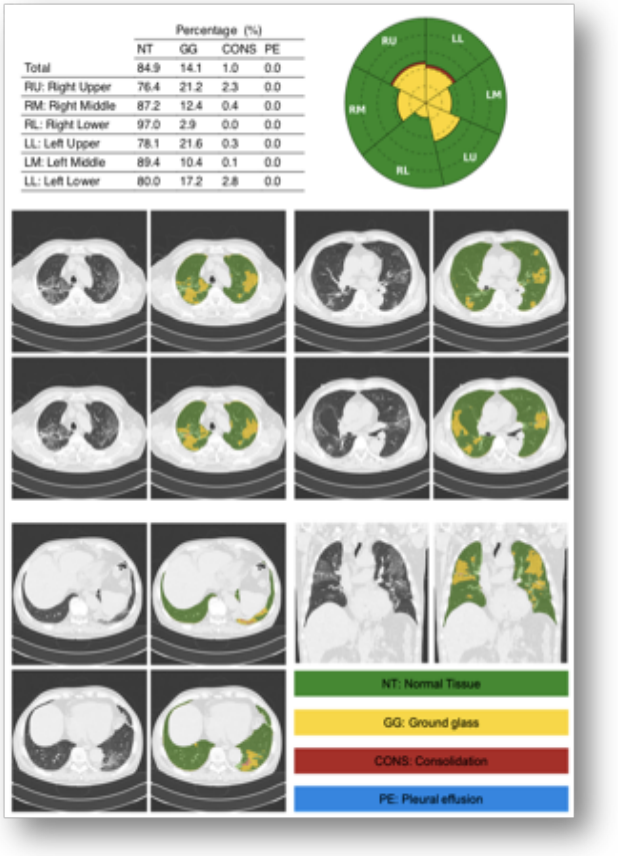

Figure 2. Example of the PDF automatically generated for one subject with mild lung involvement.

3.3 Quantitative evaluation

The method was executed for the automatic segmentation of disease subtypes on the 100 independent CT scans. From these CT scans, four different radiologists with expertise in COVID-19 diagnosis from four institutions revised a total 2716 axial CT images from 10 CT scans together with its corresponding segmentation results. The 10 CT scans were randomly selected ensuring that all disease severity ranges were included. These experts manually supervised and corrected all the segmentation label maps produced by the proposed algorithm, and the Dice coefficient of the segmentation result with respect to the corrected annotations were computed in order to quantitatively assess the model performance. The Dice coefficient was 0.988±0.01 for ground glass opacities, 0.948±0.05 for consolidations, and 0.999±0.0003 for healthy tissue. The most common errors identified were the overestimation of ground glass opacities in basal zones and mostly caused by motion artifacts, as well as the confusion between pleural effusion and consolidation pattern.

3.4 Assessment of disease severity

All 100 CT scans were visually inspected and semi-quantitatively scored by expert radiologists. This scoring was made to assess the extent of lung injury using a 5-point scale (1=<5%, 2=5%-25%, 3=25%-50%, 4=50%-75%, 5=>75%), considering the relative extent of each parenchymal injury pattern for each pulmonary lobe as well as for each lung.
Figure 3 shows the relation between the AI-predicted percentage of each disease subtype and the visually determined severity score for those subtypes (Figure 3a,b), as well as the percentage of the total affected lung tissue with respect to the total severity score assigned to each patient (Figure 3c). The total severity score was computed as the sum of the scores for each subtype. Note that pleural effusion subtype is not present in the results due to its marginal presence in the test cohort.

The figures show a good agreement and high correlation between the disease severity based on the AI-predicted percentage of lung affected by COVID-19 and the severity score visually determined by experts, having a correlation coefficient (R) of 0.77 (p-value<0.001) when considered all disease subtypes.

Additionally, we wanted to test if the proposed method could directly identify those individuals with significant extent of COVID-19 pneumonia. We defined that a patient has a significant lung injury if the percentage of lung affected is more than 5% considering all disease subtypes, i.e. having a mean visual severity score greater or equal than 2 considering both lungs. We calculated the receiver operating characteristic curves (ROC) and the area under the curve (AUC) using a logistic regression model via cross-validation for the prediction of those individuals with visually identified presence of COVID-19 pneumonia. All subjects from the independent test set were included in the analysis (25 without significant extent of COVID-19 pneumonia and 75 with visually defined significant lung injury). The proposed method had and AUC of 0.97 for the detection of patients affected by COVID-19 (see Figure 4).

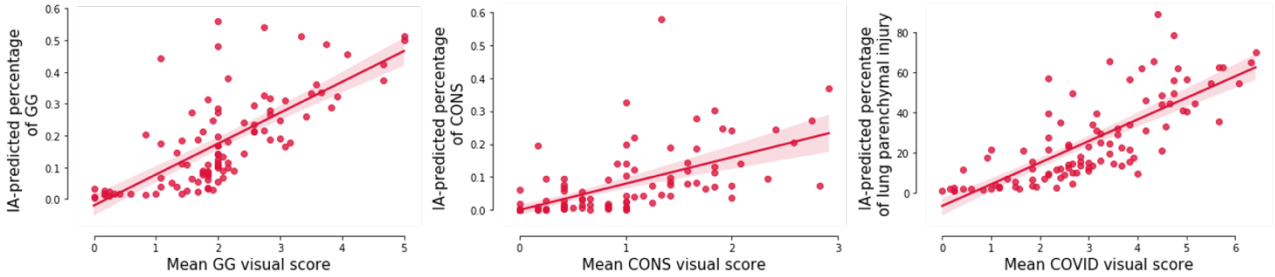

Figure 3. Relation between visually defined CT severity score and the IA-predicted percentage of lung affected by (a) ground glass opacities, (b) consolidation, and (c) considering all disease subtypes.

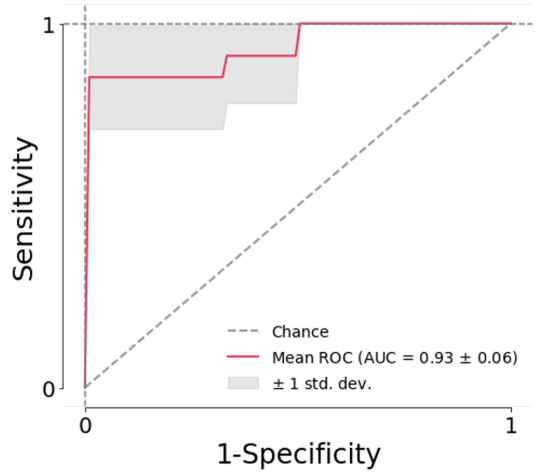

Figure 4. ROC curve for the prediction of individuals with significant involvement of lung injury caused by COVID-19 pneumonia.

4. DISCUSSION AND CONCLUSION

In this work we have presented an artificial intelligence algorithm for the segmentation and differentiation of different CT-findings in COVID-19 pneumonia. CT-findings in the lung of infected patients have been confirmed as one of the earliest indicators of the disease, and the quantification of each disease subtypes may play an important role in the management of the patients and in the monitorization of the course of the disease.

The results obtained proves that the proposed method is a valid tool for the identification and quantification of disease subtypes in COVID-19 (Dice coefficients of 98.8, 94.8 and 99.9% for ground glass opacities, consolidation pattern and healthy tissue respectively), and that is associated with the visually determined presence of parenchymal injuries and disease severity.

The technologies developed could enable an objective and quantitative understanding of disease progression and response to treatment, as well as the objective evaluation of efficacy in clinical trials of new treatments.

ACKNOWLEDGEMENTS

This work has been partially funded by projects RTI2018-098682-B-I00 (MCIU/AEI/FEDER,UE) by the Spanish ministry of Science and Innovation, cofunded by the European Union ERDF (European Regional Development Fund), and the Leonardo grant to researchers and cultural creators 2019 from Fundación BBVA. Raúl San José Estépar has been supported by NHI grant R21LM013670. Research work of Dr. Jacobo Sellarés has been financed by the grant SLT008/18/00176 and the support of the Department of Health of the Generalitat de Catalunya, in the call for grants 2019–2021, as part of the Strategic Research and Innovation Plan in Health 2016–2020. It has also been financed by FEDER Funds (PI19/01152), SEPAR (PI17/562 & PI18/792), SOCAP, FUCAP and Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS).

REFERENCES

[1] Wang D., 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, Feb. 2020.

[2] Zhou F., et al. "Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study". Lancet, vol. 28, vol. 395, no. 10229, pp. 1054-1062, Mar 2020.

[3] Zhou F., et al. "Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study". Lancet, vol. 28, vol. 395, no. 10229, pp. 1054-1062, Mar 2020.

[4] Zhao W., et al. "Relation Between Chest CT Findings and Clinical Conditions of Coronavirus Disease (COVID-19) Pneumonia: A Multicenter Study". AJR Am J Roentgenol, vol.214, no. 5, pp.1072-1077, May 2020.

[5] Li L., et al. "Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy" Radiology, vol. 296, no.2, pp. E65-E71, 2020.

[6] Lessmann N., et al. "Automated Assessment of CO-RADS and Chest CT Severity Scores in Patients with Suspected COVID-19 Using Artificial Intelligence" Radiology, 2020 Jul 30:202439

[7] Zhang K., et al. "Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography" Cell. 2020; vol. 181, no. 6, pp. 1423-1433.e11, 2020.

[8] Shi F., et al. "Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19" IEEE Rev Biomed Eng. 2020 Apr 16;PP.

[9] Pan F. et al. "Time Course of Lung Changes at Chest CT during Recovery from Coronavirus Disease 2019 (COVID-19)". Radiology, vol. 295, no. 3, pp. 715-721, Jun 2020.

[10] Huang G. et al., "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 2261-2269.

[11] Paszke A. et al., "ENET: A deep neural network architecture for real-time semantic segmentation," CoRR, vol. abs/1606.02147, 2016.

[12] López-Linares K., et al., "3D pulmonary artery segmentation from CTA scans using deep learning with realistic data augmentation," in *Image Analysis for Moving Organ,Breast, and Thoracic Images*. 2018, pp. 225–237, Springer I P.

[13] Bermejo-Peláez D. et al., "A SR-NET 3D-to-2D Architecture for Paraseptal Emphysema Segmentation" in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, Venice, Italy, 2019, pp. 303-306.

[14] Hofmanninger, J., et al. "Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem". Eur Radiol Exp 4, 50 (2020)