

Development and validation of an automated radiomic CT signature for detecting COVID-19

GUIOT J.^{1,*}, VAIDYANATHAN A.^{2,4,*}, DEPREZ L.³, ZERKA F.^{2,4}, DANTHINE L.³,
FRIX A.N.¹, THYS M.⁶, HENKET. M.¹, CANIVET G.⁷, MATHIEU S.⁷, EFTAXIA E.³,
LAMBIN P.⁴, TSOUTZIDIS N.², MIRAGLIO B.², WALSH S.², MOUTSCHEN M.⁵,
LOUIS R.¹, MEUNIER P.³, VOS W.², LEIJENAAR R.T.H.^{2,‡} & LOVINFOSSE P.^{8,‡}

*Guiot J. & Vaidyanathan A. have equally contributed as first author.

‡Lovinfosse P. & Leijenaar R.T.H. have equally contributed as last author.

1. Department of Pneumology, University Hospital of Liège, Liège, Belgium
2. Oncoradiomics SA, Liège, Belgium
3. Department of Radiology, University Hospital of Liège, Liège, Belgium
4. The D-Lab, Department of Precision Medicine, Department of Nuclear Medicine, GROW–School for Oncology, Maastricht University, Maastricht, The Netherlands
5. Department of Infectious Diseases, University Hospital of Liège, Liège, Belgium
6. Department of Medico-economic Information, University Hospital of Liège, Liège, Belgium
7. Department of Computer Applications, University Hospital of Liège, Liège, Belgium
8. Department of Nuclear Medicine and Oncological Imaging, University Hospital of Liège, Liège, Belgium

Corresponding Author:

Dr. Julien Guiot
Pneumology department
C.H.U. - SART TILMAN - B35
4000 Liège
BELGIUM
J.Guiot@chuliege.be

Word count: 2448

Key words: COVID-19, pandemic, coronavirus, radiomics, chest CT, AI, deep learning, machine learning, diagnosis, lung disease

Abstract

Background The coronavirus disease 2019 (COVID-19) outbreak has reached pandemic status. Drastic measures of social distancing are enforced in society and healthcare systems are being pushed to and over their limits.

Objectives To develop a fully automatic framework to detect COVID-19 by applying AI to chest CT and evaluate validation performance.

Methods In this retrospective multi-site study, a fully automated AI framework was developed to extract radiomics features from volumetric chest CT exams to learn the detection pattern of COVID-19 patients. We analysed the data from 181 RT-PCR confirmed COVID-19 patients as well as 1200 other non-COVID-19 control patients to build and assess the performance of the model. The datasets were collected from 2 different hospital sites of the CHU Liège, Belgium. Diagnostic performance was assessed by the area under the receiver operating characteristic curve (AUC), sensitivity and specificity.

Results 1381 patients were included in this study. The average age was 64.4 ± 15.8 and 63.8 ± 14.4 years with a gender balance of 56% and 52% male in the COVID-19 and control group, respectively. The final curated dataset used for model construction and validation consisted of chest CT scans of 892 patients. The model sensitivity and specificity for detecting COVID-19 in the test set (training 80% and test 20% of patients) were 78.94% and 91.09%, respectively, with an AUC of 0.9398 (95% CI: 0.875-1). The negative predictive value of the algorithm was found to be larger than 97%.

Conclusions Benchmarked against RT-PCR confirmed cases of COVID-19, our AI framework can accurately differentiate COVID-19 from routine clinical conditions in a fully automated fashion. Thus, providing rapid accurate diagnosis in patients suspected of COVID-19 infection, facilitating the timely implementation of isolation procedures and early intervention.

Introduction

The rapid outbreak of coronavirus disease 2019 (COVID-19), originating from severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) infection, has plainly become a public health emergency of international concern [1]. COVID-19 has created worldwide an immense adverse impact. Globally there has been 1,131,713 confirmed cases and 60,115 deaths as of April 4th 2020 [2].

The disease is currently confirmed by reverse-transcription polymerase chain reaction (RT-PCR) [3]. There is, however, evidence that the sensitivity of RT-PCR may not be optimal for the objective of very early detection and early intervention of COVID-19 patients [4]. Due to the limited supply of RT-PCR kit, the lengthy turnaround times, and the emergence of false-negative nucleic acid cases, some experts propose to diagnose suspected cases using the widely available time-saving and non-invasive imaging approach of chest computed tomography (CT) rather than RT-PCR [5,6] as CT can capture imaging features from the lung associated with COVID-19 [7] early in the course of the disease [8]. CT could thus serve as an efficient and effective way to screen, diagnosis, and possibly triage COVID-19 patients. Despite these apparent strengths, there are notable calls of caution with respect to using CT for these purposes [9,10], due to increased radiation exposure of the population and the risk of cross-infection if disinfection is not properly implemented.

Nevertheless, if chest CT is to be utilized in this pandemic it needs to be applied optimally. Artificial intelligence (AI) using machine learning technology in the medical imaging domain has accomplished impressive results due to the intrinsic properties of machine vision [11–14] and can be leveraged in this crisis.

Radiomics is the high-throughput mining of quantitative image features from standard-of-care medical imaging that enables data to be extracted and applied within clinical decision support systems to improve diagnostic, prognostic, and/or predictive accuracy [15]. Conceptually, radiomics is a bridge between imaging and precision medicine [16].

In this study, we hypothesize that a radiomics analysis can identify a diagnostic signature for COVID-19 infection, based on standard of care chest CT imaging. As a result, we present a fully automated AI framework to detect COVID-19 using chest CT, referred to as COVIA ('coronavirus intelligence artificielle'). Other non-COVID-19 patients were included as controls to test the robustness of the model performance.

Materials and methods

Ethics

The study has been approved by the local ethics committee of the CHU-Liège (EC number 116/2020). The institutional review board waived the requirement to obtain written informed consent for this retrospective case series, since all analyses were performed on de-identified (i.e., anonymized) data and there was no potential risk to patients.

Subjects

Two cohorts of patients were included retrospectively in this study. Cohorts came from two sites (CHU Sart-Tilman and CHU Notre Dame des Bruyères) in Liège, Belgium. The first cohort (label: COVID) consists of all patients with COVID-19 infection confirmed by RT-PCR that underwent chest CT imaging before March 28th, 2020. The second cohort (label: Control) consisted of consecutive patients that underwent chest CT imaging between October 1st, 2019 and October 24th, 2019, which ensures that none of these patients were infected by COVID-19. No other inclusion or exclusion criteria were considered while collecting the data. This resulted in a set of CT images from either COVID-19 infected patients or non-infected patients (normal and a variety of diseases) representing real life conditions.

Radiomics

Radiomics is an in-depth automatic analysis that aims to quantify all the information contained in medical images. Some of this information might go undetected for by the naked eye but is present in the image and, when detected by radiomics, opens a pathway towards precision imaging medicine. The schematic representation in

Figure 1 depicts the radiomics workflow applied in this study. The following sections will detail each step in the workflow.

Imaging

All CT images used in the study were acquired on one of five multidetector CT scanners (Siemens Edge Plus (2), GE Revolution CT (1), GE Brightspeed (2)) available at the sites. Since CT images were collected retrospectively, no standardized scan protocol was available over the complete dataset. In order to prevent excess variability in the imaging used for model generation the following criteria for radiomic analysis was used:

- Lungs completely visible in the scan
- Slice increment less than 1.5 mm
- No missing slices
- For GE scans: STANDARD reconstruction kernel
- For Siemens scans: B30-range reconstruction intervals

Lung segmentation

The lungs were segmented as a single structure using RadiomiX (OncoRadiomics SA, Liège, Belgium) based on convolutional neural networks by combining 2D and 3D architectures. The predicted segmentations of each architecture are ensembled and the intersection constitutes the final lung segmentation which is used for extraction of radiomics features. Figure 2 shows example segmentations for two four patients from both the COVID and Control groups. Complete details on the segmentation methods can be found in the online supplement.

Feature Extraction

For each patient, 166 image features were extracted from the lung segmentation using RadiomiX (OncoRadiomics SA, Liège, Belgium) based upon quantitative image analysis technology. The extracted features comprised first order statistics, texture, and shape. A bin width of 25 Hounsfield unit was used for image intensity discretization. Mathematical descriptions of all features are described in [16].

Modelling

A statistical model that can provide a binary classification of patients (COVID-19 infected or not) has been trained and validated using a regularized logistic regression model [17].

Stratified partitioning was performed to randomly split the data into 80% for model training and 20% for validation. Stratification was performed for age, gender, diagnosis of COPD and COVID-19 status, ensuring a balanced distribution of these characteristics in both data partitions.

Multivariable logistic regression with Elastic Net regularization was performed in the training data set. Highly correlated features, features with near zero variance and linear combinations between features were first eliminated from further analysis. For each highly correlated feature pair (Pearson correlation coefficient $\rho > 0.9$), the variable with the largest mean absolute correlation with all remaining features was removed. Model training was performed using 100 times repeated 10-fold cross-validation to select the optimal model hyperparameters, optimizing for the area under the receiver operator curve (AUC). All features were standardized before modelling. Class weighting was applied to address for class imbalance in the training data. To further reduce the chance of overfitting to the training data, we selected the simplest

candidate model (i.e., the model with fewest non-zero coefficients) within one standard error of the best performing model.

Model performance was validated in the test data partition. Here, the AUC was used to assess model performance in discriminating between COVID-19 positive and COVID-19 negative patients. Classification performance was assessed with sensitivity and specificity analysis. All statistical analysis was performed in R (version 3.6.2).

Results

Study population

Table 1 lists the study population characteristics for the COVID and Control cohorts, as well as the main CT findings as scored by radiologists. Both cohorts have a similar mean age and male/female distribution. For the COVID-19 infected patients 69% needed O₂ at admission, resulting in 37% of patients ending up in the ICU. 17% of COVID-19 patients needed mechanical ventilation and 4% died.

Data curation

After an automated quality check on the inclusion criteria, CT images and lung segmentations for a total number of 892 patients were included for further processing. A flow chart describing the overall workflow from data collection to model training and testing is shown in Figure 3.

COVID-19 infection status prediction

The final prediction model included 30 radiomics features with a non-zero regression coefficient. In the test data partition, the sensitivity and specificity for COVID-19 classification are 78.9% and 91.1%, respectively. The overall accuracy is 89.7% (95% CI: 84 – 93.9). The positive predictive value and negative predictive value are 53.6% and 97.1%, respectively. The ROC curve is shown in Figure 4. The corresponding AUC value for discriminating between COVID-19 positive and negative cases is 0.939 (95% CI: 0.875-1).

Discussion

COVID-19 has spread rapidly across the globe and rate of infection is accelerating. Therefore, rapid early diagnosis of the disease is essential for early intervention and swift isolation of patients in order to prevent the spread of the virus. RT-PCR is considered as the 'gold standard', however, there are reports of false-negatives occurring which are eventually confirmed as true-positive by repeated swab tests [18]. False negatives can be a significant problem in high-throughput settings operating under severe pressure [19]. The correct operation of the test is crucial and there is ambiguity with respect to the kinetics of SARS-CoV-2 viral shedding, thus the timing of the test may very well dictate the result. Furthermore, it is also unclear what kind of clinical sample is most appropriate as nasopharyngeal swabs may offer greater consistency than sputum samples. When considering the limited supply of RT-PCR, the growing backlog and the likely increasing pressure and turnaround times in laboratories, the issues pertaining to false-negatives, some experts have concluded that to diagnose suspected cases using the widely available time-saving and non-invasive imaging approach of chest CT is justified as this approach can sensitively and specifically identify COVID-19 patients. The high NPV (>97%) of our model provides further justification of using CT imaging-based diagnosis a primary tool for COVID-19 patient management.

Whereas similar studies in COVID-19 focus mainly on the detection of various diseased regions (including ground-glass opacification, consolidation, bilateral involvement, peripheral and diffuse distribution amongst others) in the lung [20,21] [22], our approach performs an easy segmentation of the lungs as one single structure, which is by far an easier task to automate with AI. Features for our quantitative image analysis are extracted from this whole lung structure and

subsequently used for prediction model application and COVID-19 infection status classification. In the end this constitutes a fully automated clinical decision support tool for the diagnosis of COVID-19, which is able to provide an objective, robust (i.e., no user variability) and easy to interpret classification (yes-no) of COVID-19 infection status. The complete workflow takes between 40-60s and is thus, providing a rapid and accurate diagnosis in patients with suspected of COVID-19 infection, facilitating the timely implementation of isolation procedures and early intervention. For 160 CT scans of the patients in the COVID group a radiologist interpretation of COVID status was also available. Out of these patients, 123 were diagnosed as COVID-19 infected patients based on the image analysis from the radiologists (sensitivity=0.77). It is important to mention that our AI algorithm has similar sensitivity as compared to a trained radiologist, but our results were obtained in a large dataset with a COVID-19 prevalence of only 11% while the radiologist scoring happened in the midst of the current pandemic where COVID-19 prevalence was greatly elevated amongst the patients undergoing a chest CT exam. We therefore expect that our algorithm will outperform radiologists once the pandemic phase of COVID-19 is over. This hypothesis will be tested in a virtual clinical trial setting in the coming weeks. This aspect is vital in the context of incidental findings, which is of increasing relevance [23]. An automated AI solution is helpful in assisting the accurate identification of potentially COVID-19 positive patients, alerting the radiology department that a 'clean machine' now requires decontamination.

One argument against this image-based screening approach is the increased radiation exposure of the population. It is certainly important to keep the radiation dose to the public arising from medical examinations as low as reasonably achievable [24]. However, when we consider the exemplars of breast and lung

cancer (two indications with established or soon to be established ionizing radiation medical imaging screening programs [25,26]) the expected mortality rate linked to the disease in the US this year is approximately 42,000 and 135,000 respectively [27,28]. A recent estimate of the mortality rate in the US this year due to COVID-19 was between 1.1-2.2 million [29]. These estimates would seem to support the justification to use CT as a screening tool for COVID-19. Moreover, a recent study reported that no effect on human DNA was detected when low dose CT procedures are used [30].

Another objection is the risk of cross-infection if disinfection of imaging equipment is not properly implemented. This is of great important because one of the great benefits of early diagnosis is the ability to swiftly isolate the patient and thus limit the spread of the virus. Thus, if the CT scanner were to become a vector for infection this would completely negate any benefit accrued. It is therefore vital that radiology departments are adequately prepared for managing (suspected) COVID-19 patients in a safe, efficient, and effective fashion [31].

A general objection of AI methods is the lack of transparency and interpretability. This is not the case with our approach, as 'handcrafted' radiomics features are explicitly defined and linked to clearly specified regions of interest within the images, driving the decision of the algorithm. Thus, clearly and intelligibly quantifying the image phenotype, which has also been shown to provide a means of connecting to the underlying biology [32].

We have encountered some limitations in this study. Firstly, COVID-19 is caused by a SARS-CoV-2 and may have similar imaging characteristics as pneumonia caused by other types of viruses. However, due to the lack of laboratory confirmation of the

aetiology for each of these cases, we were not able to select other viral pneumonias for comparison in this study. Although our Control group of non-COVID-19 patients contains several patients (CAP in Table 1, 12.5%) with pneumonia (either viral, bacterial or organizing pneumonia from any cause), it would be desirable to test the performance of our algorithm in distinguishing COVID-19 from other viral pneumonias that have real time polymerase chain reaction confirmation of the viral agent in a future study.

Furthermore, there is a large commonality in how the lung responds to various pathologies and there is a considerable amount of similarity in the presentation of many diseases in the lung that depend on a plethora of factors (e.g. age, drug reactivity, immune status, underlying co-morbidities). It is likely impossible to differentiate all lung diseases by way of imaging and AI alone, and as such, drastic improvements in positive predictive value are likely difficult to achieve. A multi-omics approach for this is most likely optimal.

Future work is planned to collect additional chest CTs from multiple centres to externally validate the performance of our algorithm. Ultimately, this study focuses on diagnosis whereas prognosis on the future disease trajectory is an even more urgent unmet clinical need that would enable improved resource management (including management decisions regarding the allocation of ventilators). This is the next step for our collaborative research.

Conclusion

Benchmarked against RT-PCR confirmed cases of COVID-19, our AI framework can accurately detect COVID-19. Thus, providing rapid accurate diagnosis in patients with suspected of COVID-19 infection, facilitating the timely implementation of isolation procedures and early intervention.

Acknowledgements

The authors acknowledge financial support from Interreg V-A Euregio Meuse-Rhine (Euradiomics project), an ERC advanced grant (ERC-ADG-2015, n° 694812 - Hypoximmuno) and a European Marie Curie grant (PREDICT - ITN - n° 766276) for the execution of this work.

References

- [1] WHO, Landing page, (2020). <https://www.who.int/> (accessed April 2, 2020).
- [2] Johns Hopkins University & Medicine, Coronavirus Resource Center, (2020). <https://coronavirus.jhu.edu/map.html> (accessed April 2, 2020).
- [3] J.T. Wu, K. Leung, G.M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, *Lancet*. 395 (2020) 689–697. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9).
- [4] S. Yang, Y. Shi, H. Lu, J. Xu, F. Li, Z. Qian, X. Hua, X. Ding, F. Song, J. Shen, Y. Lu, F. Shan, Z. Zhang, Clinical and CT features of early-stage patients with COVID-19: a retrospective analysis of imported cases in Shanghai, China, *Eur. Respir. J.* (2020) 2000407. <https://doi.org/10.1183/13993003.00407-2020>.
- [5] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, L. Xia, Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases, *Radiology*. (2020) 200642. <https://doi.org/10.1148/radiol.2020200642>.
- [6] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, W. Ji, Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR, *Radiology*. (2020) 200432. <https://doi.org/10.1148/radiol.2020200432>.
- [7] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in

Wuhan, China, *Lancet*. 395 (2020) 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).

- [8] H.L. Ruili Li, Guangxue Liu, Xiaochun Zhang, Letter to the Editor: Chest CT and RT-PCR: radiologists' experience in the diagnosis of COVID-19 in China, *Eur. Radiol.* (2020). <https://www.european-radiology.org/opinions/chest-ct-and-rt-pcr-radiologists-experience-in-the-diagnosis-of-covid-19-in-china/>.
- [9] M. Hope, A role for CT in COVID-19? What data really tell us so far, *Lancet*. 0 (2020). [https://doi.org/10.1016/S0140-6736\(20\)30728-5](https://doi.org/10.1016/S0140-6736(20)30728-5).
- [10] Y. Huang, W. Cheng, N. Zhao, H. Qu, J. Tian, Correspondence CT screening for early diagnosis of SARS-CoV-2, *Lancet Infect. Dis.* 51 (2020) 30241. [https://doi.org/10.1016/S1473-3099\(20\)30241-3](https://doi.org/10.1016/S1473-3099(20)30241-3).
- [11] H.J.W.L. Aerts, E.R. Velazquez, R.T.H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M.M. Rietbergen, C.R. Leemans, A. Dekker, J. Quackenbush, R.J. Gillies, P. Lambin, Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach., *Nat. Commun.* 5 (2014) 4006. <https://doi.org/10.1038/ncomms5006>.
- [12] S.L.F. Walsh, L. Calandriello, M. Silva, N. Sverzellati, Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study, *Lancet Respir. Med.* 6 (2018) 837–845. [https://doi.org/10.1016/S2213-2600\(18\)30286-8](https://doi.org/10.1016/S2213-2600(18)30286-8).
- [13] S.M. McKinney, M. Sieniek, V. Godbole, J. Godwin, N. Antropova, H. Ashrafian, T. Back, M. Chesus, G.C. Corrado, A. Darzi, M. Etemadi, F. Garcia-

- Vicente, F.J. Gilbert, M. Halling-Brown, D. Hassabis, S. Jansen, A. Karthikesalingam, C.J. Kelly, D. King, J.R. Ledsam, D. Melnick, H. Mostofi, L. Peng, J.J. Reicher, B. Romera-Paredes, R. Sidebottom, M. Suleyman, D. Tse, K.C. Young, J. De Fauw, S. Shetty, International evaluation of an AI system for breast cancer screening, *Nature*. 577 (2020) 89–94.
<https://doi.org/10.1038/s41586-019-1799-6>.
- [14] D. Ardila, A.P. Kiraly, S. Bharadwaj, B. Choi, J.J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D.P. Naidich, S. Shetty, End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography, *Nat. Med.* 25 (2019) 954–961. <https://doi.org/10.1038/s41591-019-0447-x>.
- [15] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G.P.M. Van Stiphout, P. Granton, C.M.L. Zegers, R. Gillies, R. Boellard, A. Dekker, H.J.W.L. Aerts, Radiomics: Extracting more information from medical images using advanced feature analysis, *Eur. J. Cancer*. 48 (2012) 441–446.
<https://doi.org/10.1016/j.ejca.2011.11.036>.
- [16] P. Lambin, R.T.H. Leijenaar, T.M. Deist, J. Peerlings, E.E.C. De Jong, J. Van Timmeren, S. Sanduleanu, R.T.H.M. Larue, A.J.G. Even, A. Jochems, Y. Van Wijk, H. Woodruff, J. Van Soest, T. Lustberg, E. Roelofs, W. Van Elmpt, A. Dekker, F.M. Mottaghy, J.E. Wildberger, S. Walsh, Radiomics: The bridge between medical imaging and personalized medicine, *Nat. Rev. Clin. Oncol.* 14 (2017). <https://doi.org/10.1038/nrclinonc.2017.141>.
- [17] T.M. Deist, F.J.W.M. Dankers, G. Valdes, R. Wijsman, I.C. Hsu, C. Oberije, T. Lustberg, J. van Soest, F. Hoebbers, A. Jochems, I. El Naqa, L. Wee, O. Morin,

- D.R. Raleigh, W. Bots, J.H. Kaanders, J. Belderbos, M. Kwint, T. Solberg, R. Monshouwer, J. Bussink, A. Dekker, P. Lambin, Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers, *Med. Phys.* 45 (2018) 3449–3459.
<https://doi.org/10.1002/mp.12967>.
- [18] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, J. Liu, Chest CT for Typical 2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing, *Radiology*. (2020) 200343. <https://doi.org/10.1148/radiol.2020200343>.
- [19] C. Sheridan, Coronavirus and the race to distribute reliable diagnostics, *Nat. Biotechnol.* (2020). <https://doi.org/10.1038/d41587-020-00002-2>.
- [20] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, H. Zhao, Y. Jie, R. Wang, Y. Chong, J. Shen, Y. Zha, Y. Yang, Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images, *MedRxiv*. (2020) 2020.02.23.20026930.
<https://doi.org/10.1101/2020.02.23.20026930>.
- [21] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, B. Xu, A deep learning algorithm using CT images to screen for corona virus disease (COVID-19), *MedRxiv*. (2020) 2020.02.14.20023028.
<https://doi.org/10.1101/2020.02.14.20023028>.
- [22] J. Chen, L. Wu, J. Zhang, L. Zhang, D. Gong, Y. Zhao, S. Hu, Y. Wang, X. Hu, B. Zheng, K. Zhang, H. Wu, Z. Dong, Y. Xu, Y. Zhu, X. Chen, L. Yu, H. Yu, Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study, *MedRxiv*.

(2020) 2020.02.25.20021568. <https://doi.org/10.1101/2020.02.25.20021568>.

- [23] D. Albano, F. Bertagna, M. Bertolia, G. Bosio, S. Lucchini, F. Motta, M.B. Panarotto, A. Peli, L. Camoni, F.M. Bengel, R. Giubbini, INCIDENTAL FINDINGS SUGGESTIVE OF COVID-19 IN ASYMPTOMATIC PATIENTS UNDERGOING NUCLEAR MEDICINE PROCEDURES IN A HIGH PREVALENCE REGION., *J. Nucl. Med.* (2020).
<https://doi.org/10.2967/jnumed.120.246256>.
- [24] I.B. Atci, H. Yilmaz, V. Antar, N.G. Ozdemir, O. Baran, S. SadilliOglu, M. Ozel, O. Turk, M. Yaman, H. Topacoglu, What do we know about ALARA? Is our knowledge sufficient about radiation safety?, *J. Neurosurg. Sci.* 61 (2017) 597–602. <https://doi.org/10.23736/S0390-5616.16.03418-4>.
- [25] M. Marmot, D.G. Altman, D.A. Cameron, J.A. Dewar, S.G. Thompson, M. Wilcox, The benefits and harms of breast cancer screening: An independent review, *Lancet.* 380 (2012) 1778–1786. [https://doi.org/10.1016/S0140-6736\(12\)61611-0](https://doi.org/10.1016/S0140-6736(12)61611-0).
- [26] H.J. De Koning, C.M. Van Der Aalst, P.A. De Jong, E.T. Scholten, K. Nackaerts, M.A. Heuvelmans, J.W.J. Lammers, C. Weenink, U. Yousaf-Khan, N. Horeweg, S. Van'T Westeinde, M. Prokop, W.P. Mali, F.A.A. Mohamed Hoesein, P.M.A. Van Ooijen, J.G.J.V. Aerts, M.A. Den Bakker, E. Thunnissen, J. Verschakelen, R. Vliegenthart, J.E. Walter, K. Ten Haaf, H.J.M. Groen, M. Oudkerk, Reduced lung-cancer mortality with volume CT screening in a randomized trial, *N. Engl. J. Med.* 382 (2020) 503–513.
<https://doi.org/10.1056/NEJMoa1911793>.

- [27] ASCO, Breast cancer mortality, (2020). <https://www.cancer.net/cancer-types/breast-cancer/statistics> (accessed April 3, 2020).
- [28] American cancer society, Lung cancer mortality, (2020). <https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html>.
- [29] N.M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L.C. Okell, S. Van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Gt Walker, C. Walters, P. Winskill, C. Whittaker, C.A. Donnelly, S. Riley, A.C. Ghani, Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, (n.d.). <https://doi.org/10.25561/77482>.
- [30] H. Sakane, M. Ishida, L. Shi, W. Fukumoto, C. Sakai, Y. Miyata, T. Ishida, T. Akita, M. Okada, K. Awai, S. Tashiro, Biological Effects of Low-Dose Chest CT on Chromosomal DNA, *Radiology*. (2020) 190389. <https://doi.org/10.1148/radiol.2020190389>.
- [31] M. Mossa-Basha, C.C. Meltzer, D.C. Kim, M.J. Tuite, K.P. Kolli, B.S. Tan, Radiology Department Preparedness for COVID-19: Radiology Scientific Expert Panel, *Radiology*. (2020) 200988. <https://doi.org/10.1148/radiol.2020200988>.
- [32] R.T.H. Leijenaar, M. Bogowicz, A. Jochems, F.J.P. Hoebbers, F.W.R. Wesseling, S.H. Huang, B. Chan, J.N. Waldron, B. O'Sullivan, D. Rietveld, C.R. Leemans, R.H. Brakenhoff, O. Riesterer, S. Tanadini-Lang, M. Guckenberger, K. Ikenberg, P. Lambin, Development and validation of a

radiomic signature to predict HPV (p16) status from standard CT imaging: A

multicenter study, Br. J. Radiol. 91 (2018) 20170498.

<https://doi.org/10.1259/bjr.20170498>.

Tables

Table 1: Summary of patient characteristics (age, gender and CT findings scored by radiologist) per cohort

	CONTROL (n=1200)	COVID (n=181)
Age (years)	63.8 ± 14.4	64.4±15.8
Gender (% Male)	52%	56%
Normal (%)	33%	4.41%
Neoplasia (%)	8.73%	0%
CAP (%)	12.50%	8.10%
COPD (%)	26%	19.33%
Isolated pleurisy (%)	6.2%	1.10%
Pulmonary embolism (%)	0.77%	1.10%
Nodule (%)	19%	6.62%
Chronic inflammation* (%)	8.48%	5.52%
Pneumothorax (%)	0.68%	0%
Isolated atelectasis (%)	3.68%	3.31%

Figures

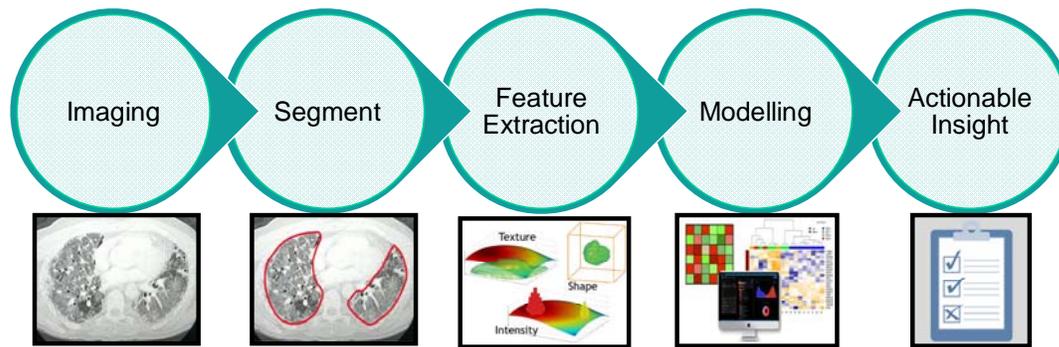


Figure 1 The radiomics workflow applied in this study: Schematic representation of the steps performed to produce a signature to support clinical decision making.

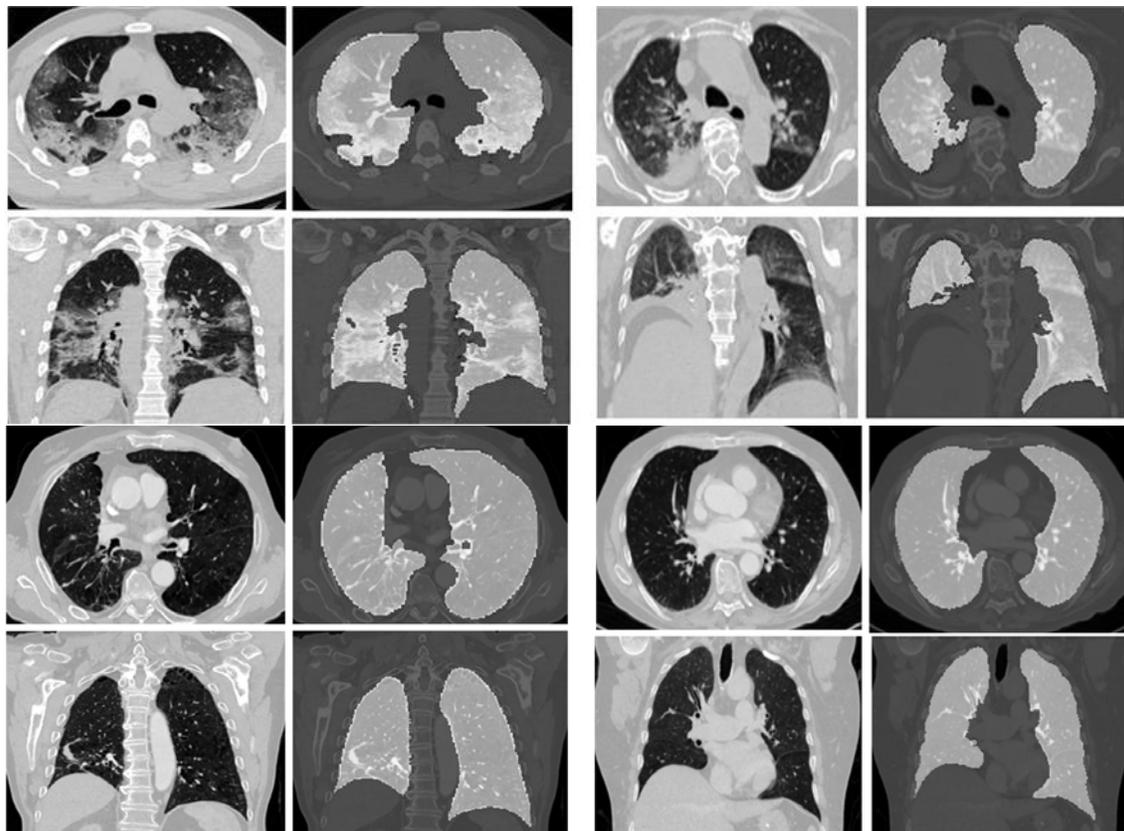


Figure 2 Axial and coronal slices with accompanying segmentation masks. Top left quadrant - Typical aspect of COVID-19 pneumonia characterized by bilateral multilobe ground-glass opacities of peripheral/subpleural distribution, with intralesional reticulations, presenting of a “crazy paving” aspect. Also found are subpleural condensations and retraction bronchiectasis, typical of organizing pneumonia; Top right quadrant - Atypical aspect of COVID-19 pneumonia, with posterior right lower lobe condensation and retraction of the ipsilateral diaphragm. Central and peripheral ground-glass opacities in right lower lobe, right upper lobe and left upper lobe; Bottom left quadrant - Typical chronic obstructive pulmonary disease COPD chest CT characterized by severe centrilobular and para-septal emphysema, associated with cylindrical bronchiectasis and bronchial walls thickening. Right peripheral upper lobe tree in bud pattern seen in bronchiolitis.

Middle lobe crescent-shaped atelectasis condensation; Bottom right quadrant -
Normal chest CT.

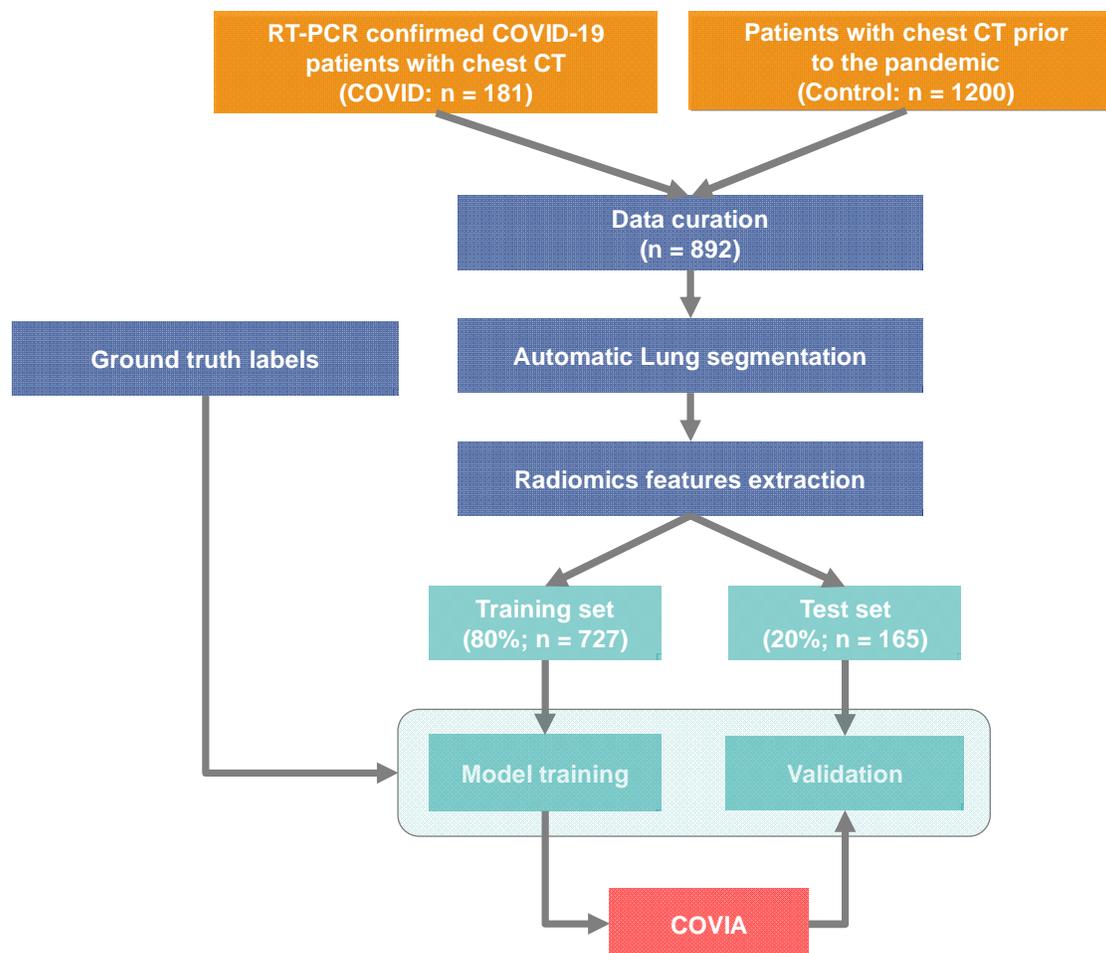


Figure 3 Flow diagram: Data was collected and combined from the COVID and Control cohorts. Lungs were segmented from the combined dataset and radiomics features were extracted. The Radiomics data was split into a training set for COVIA model development and a test set to validate the performance of COVIA with unseen patient CTs.

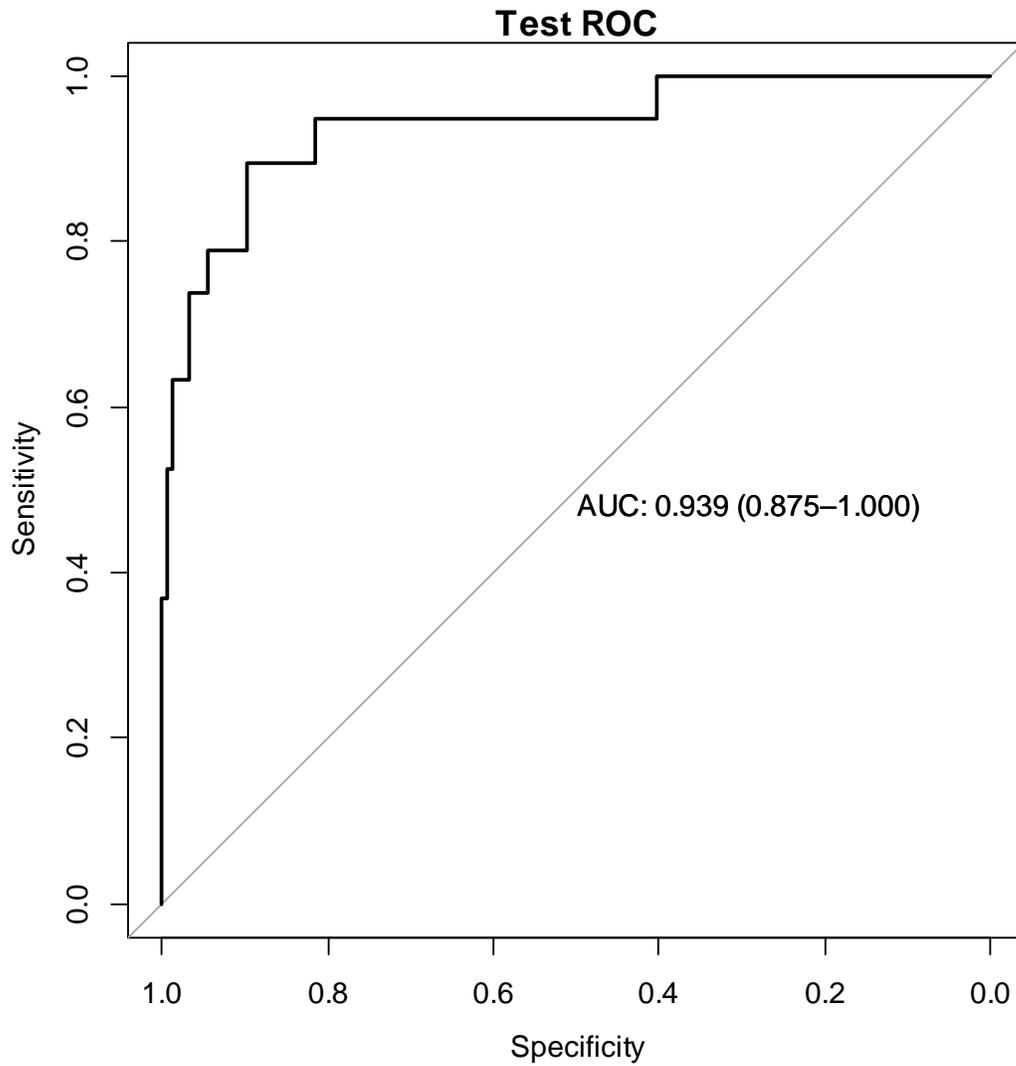


Figure 4 ROC curve. The ROC plot illustrates the performance (black curve) of the AI framework to discriminate between COVID-19 positive and negative cases in the test set (20% data partition) with an AUC of 0.939 (95% CI: 0.875-1).

Online supplement

The lungs were segmented as a single structure using RadiomiX (Oncoradiomics SA, Liège, Belgium) based on convolutional neural networks by combining 3D and 2D architectures. Details on both these architectures are given below.

3D lung segmentation

This model architecture consists of a 3D U-Net [1] with residual blocks [2] in the encoder part of the network. Publicly available data from the cancer imaging archive [3] was used to train and validate the model. The specific dataset [4] contains CT scans of 422 confirmed non-small cell lung cancer cases, along with manual segmentations of the left and right lungs. The segmentations were performed by an experienced radiologist and these segmentations were used as a reference standard. The data was randomly partitioned into a training set ($n = 322$), a tuning set ($n = 50$), and a test set ($n = 50$). In order to generate homogeneous CT volumes as input for the model, the following pre-processing steps were performed. All the volumes were resized to $160 \times 160 \times 448$ along the x, y and z axis .and image intensities were clipped at a window width of 1500 HU and a window level of -600 HU (i.e., a standard lung CT window level settings).

The following data augmentations were performed to avoid overfitting [5] on the training dataset:

- Flipping in different directions: up and down, left and right
- Randomly resampling volumes to varying voxel sizes
- Rotating (10-30 degrees) onto the left or right direction
- Reversing the sequence of axial slices

The model was trained with the pre-processed volumes and their corresponding reference labels, using Jaccard loss [6] as an objective function. Here, the loss is calculated in a mini batch of two images per iteration. The network was trained for 10 epochs and at the end of each epoch the Jaccard loss was calculated on the model's predictions to ensure validation loss convergence.

2D lung lobe segmentation

The full lung lobe segmentation model consists of two sub-networks with a similar architecture, but independently trained for the left and right lung (i.e., one model is trained on 2D axial slices for the right upper/middle/lower lobes and the other model is trained on 2D axial slices for left upper/lower lobes). The architecture of the networks is based on a 2D Feature Pyramid Network [7] adapted with ResNext blocks [8] in the encoder. The model was trained and validated on the publicly available subset of the LUNA16 challenge. The subset contains 50 CT volumes manually annotated for lung lobes [9]. Since this is only a small training dataset, the following data augmentation strategies were adopted to increase the training set:

- Resampling to varied voxel sizes and resizing to the common size (160,160,448)
- Cropping and zooming in through left or right lung and resizing the cropped volume to size (160,160,448)
- Zooming out at different ranges
- Rotating left or right at varied degrees

The networks were trained with the 2D axial slices clipped at a window width of 1500 HU and a window level of -600 HU and with their corresponding reference labels. The networks' weights were updated by using the Adam optimizer at an initial

learning rate of $1e-5$ [10]. The model was trained using customized focal Tversky loss [11] (adapted for calculating loss over multiple classes) as an objective function where the loss is calculated in a mini batch of four images per iteration. The network was trained for 5 epochs and at the end of each epoch, the customized focal Tversky loss was calculated on the model's predictions to ensure validation loss convergence.

The predicted segmentations of each architecture (i.e., the segmentation output from both the 3D and the 2D segmentation models) are ensembled and the intersection constitutes the final total lung segmentation which is used for extraction of radiomics features. The deep learning-based lung segmentation achieved a mean Dice similarity coefficient score of 0.86 across the publicly available datasets which indicates adequate precision (i.e. no significant over or under segmentation).

References

- [1] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-net: Learning dense volumetric segmentation from sparse annotation, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2016. https://doi.org/10.1007/978-3-319-46723-8_49.
- [2] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., IEEE Computer Society, 2016: pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
- [3] F. Prior, K. Smith, A. Sharma, J. Kirby, L. Tarbox, K. Clark, W. Bennett, T. Nolan, J. Freymann, The public cancer radiology imaging collections of The

- Cancer Imaging Archive, Sci. Data. 4 (2017) 170124.
<https://doi.org/10.1038/sdata.2017.124>.
- [4] NSCLC-Radiomics - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki, (n.d.).
<https://wiki.cancerimagingarchive.net/display/Public/NSCLC-Radiomics>
(accessed April 4, 2020).
- [5] R. Caruana, S. Lawrence, L. Giles, Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping, in: Adv. Neural Inf. Process. Syst., 2001.
- [6] J. Bertels, T. Eelbode, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, M. Blaschko, Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory & Practice, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 11765 LNCS (2019) 92–100. https://doi.org/10.1007/978-3-030-32245-8_11.
- [7] T.Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, Feature pyramid networks for object detection, in: Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, Institute of Electrical and Electronics Engineers Inc., 2017: pp. 936–944. <https://doi.org/10.1109/CVPR.2017.106>.
- [8] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated residual transformations for deep neural networks, in: Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, Institute of Electrical and Electronics Engineers Inc., 2017: pp. 5987–5995. <https://doi.org/10.1109/CVPR.2017.634>.
- [9] H. Tang, C. Zhang, X. Xie, Automatic pulmonary lobe segmentation using

deep learning, in: Proc. - Int. Symp. Biomed. Imaging, IEEE Computer Society, 2019: pp. 1225–1228. <https://doi.org/10.1109/ISBI.2019.8759468>.

- [10] D.P. Kingma, J.L. Ba, Adam: A method for stochastic optimization, in: 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., International Conference on Learning Representations, ICLR, 2015.
- [11] N. Abraham, N.M. Khan, A novel focal tversky loss function with improved attention u-net for lesion segmentation, in: Proc. - Int. Symp. Biomed. Imaging, 2019. <https://doi.org/10.1109/ISBI.2019.8759329>.