Segmenting Lung Lesions And Accurate COVID-19 Detection Through Chest CT Screening

Abhishek Nagrecha
Department of Computer Science
Lakehead University
Thunder Bay, Canada
nagrechaa@lakeheadu.ca

Rajesh Aytha
Department of Computer Science
Lakehead University
Thunder Bay, Canada
raytha@lakeheadu.ca

***Abstract*—**
The communicable disease caused by the extreme acute respiratory syndrome is Coronavirus. More than thirty four million people around the globe have tested positive for the disease and in access to a million people have lost their lives. One major hurdle in controlling the spread of this virus has been the inefficiency and shortage of medical tests.There is an immediate stipulation for fast and reliable screening methods to classify patients afflicted with COVID-19 in the battle against this deadly disease, as we are gradually discovering that a majority of individuals who test positive for COVID-19 are immunocompromised. To this end, CT imaging has been suggested as one of the primary screening methods which can be adopted as a substitute to the testing of RT-PCR.Compelled by this, we contemplate to precisely segment lung lesions with computing-time accuracy and reliability from healthy lung fields using CT scans and implement a neural network architecture during this analysis that is designed to detect novel cases of coronavirus pneumonia (NCP) from chest CT scans.In addition, a benchmark CT image dataset extracted from CT imaging data collected by the China National Center for Bioinformation (CNCB) containing more than 100,000 scans is proposed as the COVID19 dataset. The data consists of chest CT volumes across three different infection types: novel coronavirus pneumonia (NCP), common pneumonia (CP), and normal control (NC).We want our research to demonstrate the effects of CT scans of NCP patients and to examine the distinction between NCP and other viral and bacterial pneumonia. In contrast with other state-of-the-art standards, we firmly presume that our research studies would reveal the optimal effectiveness of the theorized architecture.

I. INTRODUCTION

The coronavirus infection (COVID-19) is a global pandemic that first emerged in December 2019, in Wuhan, China, the capital city of mainland China's Hubei province [1]. There is no authorised human vaccine at present to prevent it. When individuals are in close vicinity, the transmission of COVID-19 is quicker and the disease has been declared a pandemic by the World Health Organisation (WHO). Studies have shown that more than 60 % of patients lose their lives as soon as they advance to the stage of serious or critical illness [2] Thus, the WHO has strictly indicated that travel precautions will control the progression of the virus, and it has been reiterated to keep our hands sanitized regularly for avoiding possible virus infections.Most of the federal governments around the globe, including Canada have imposed travel restrictions and sealed their borders to stop any further spread.To add to that, a set of rules have been established for travellers entering a new territory under the Quarantine Act. The most prominent illness signs, are fever and cough. There may be other effects, including chest pain, growth of sputum, and a stuffy nose that appears to have a significant effect on patients and health services around the world. There is indeed an urgent need for quick and reliable screening methods to classify patients afflicted with COVID-19 in order to ensure prompt isolation and care in the battle against this unusual disease. Digital technology applications, such as contact tracing apps, have been used since February 2020 to curb the possible risk of infection. When anyone is affected with the virus, the smartphone apps warn users to self-quarantine and notify the health authority concerned. They also track infected individuals and the hindmost individuals with whom they were in close contact [3].However,the fast spread of COVID-19 is one of the major threats, with an average of 1-3 individuals getting infected by the disease upon contact with an infected person [4]. This advocates that they are more likely to infect 15-35 other individuals if 10 individuals are COVID-19 positive.However,WHO is still reviewing ongoing studies on the forms in which COVID-19 is being transmitted and therefore will continue to post certified data for more information.Furthermore, unless intervention steps are enforced, COVID-19 could infect a very large number of ethnic groups in the coming days since transmission from individual to individual has been clearly identified [5].

Besides, a new gene close to the virus causing SARS and MERS is caused by COVID-19. While the spread of COVID-19 has a lower mortality rate than the SARS and MERS outbreaks [6],but it spreads more quickly and has infected larger inhabitants than the SARS and MERS outbreaks [6]. The reverse transcription polymerase chain reaction (RT-PCR) research method [1] is currently the main diagnostic technique for the primary means of screening for COVID-19, a laboratory protocol that interacts with other ribonucleic (RNA) and deoxyribonucleic acids (DNA) to determine the volume of specific ribonucleic acids using fluorescence.While RT-PCR is capable of identifying the strain of severe acute respiratory syndrome SARS-CoV-2 that causes COVID-19, in some instances it gave negative testing results even though sufferers showed improvement on follow-up chest computed tomography (CT) scans [6] and its susceptibility is also varied, depending on the method of sampling and the period after symptom onset [7] [8].Several studies including [7] have advocated the use of chest computed tomography (CT) scans rather than RT-PCR , because its available in all major public health facilities, including hospital emergency rooms (ERs), and even at rural clinics, and its sensitivity is also high ,more importantly considering the limitations of RT-PCR research and the time required to obtain data. The method of CT scanning has a quicker processing time than the RT-PCR test and can offer more detailed pathology-related details and explain the extent of lung involvement [9]. Additionally, the exposure of COVID-19 symptoms in the lower parts of the lungs has a higher accuracy when using CT scans than that when using RT-PCR [10].Likewise, radiological imaging is also considered an important screening method for COVID-19 diagnosis [9]. Ai et al. [11] demonstrated the consistency of the radiological history of COVID-19-associated pneumonia with the clinical nature of the condition. When examined by CT scans, almost all COVID-19 patients have exhibited similar features including ground glass opacities in the early stages and pulmonary consolidation in the latter stages. In fact, the morphology and peripheral lung distribution can be rounded.

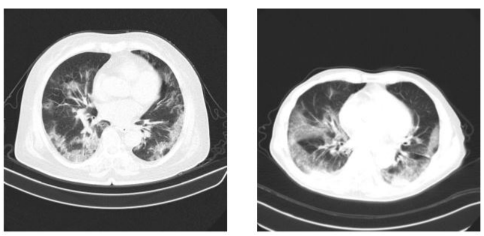

Fig. 1. Left-to-Right, CT of patients with NCP vs CP

Also, deep learning models can be employed to initially evaluate a COVID-19 patient as an alternative solution to traditional approaches that are time-consuming and labour-intensive. However, they cannot exclusively address the whole problem due to the high volume of re-examinations of infected people who wish to know the progression of their illness. Over and above,in this study, we intend to use a deep convolutional neural network architecture specifically for detection of COVID-19 cases from chest CT images via a machine-driven design exploration approach. The dataset we are going to use would be a benchmark COVID19 image dataset derived from CT imaging data collected by CNCB [12] [13] comprising more than 100,000 images.Besides, this approach involves validating and testing its forecasts, enabling us to investigate the essential visual variables associated with infection with COVID-19.In an attempt to promote ongoing research and growth, we will ensure that our COVID19 dataset, is available for public viewing in an open source and open access way.Our primary goal will be to do a research work that can assist the radiologists in distinguishing NCP from the other bacterial and viral pneumonia using the chest CT scans,which would allow them to concentrate all their energies in right direction and would also be useful for optimum resource allocation.

II. Literature Review

The very first paper which caught our attention was [12]. We found this paper from the CNCB website from where the data was garnered. They devised a new system that was able to offer precise clinical prognosis that could help clinicians identify suitable early clinical intervention and properly distribute resources, they made this AI system accessible internationally to help clinicians tackle COVID-19. Several research papers, including [14], have recapitulated chest computed tomography (CT) as a useful tool in the assessment of suspected SARS-CoV-2 infection patients. Furthermore they included the number, height, and density of lesions, as well as the total lung parenchyma, lung injury, and residual lung reserve as lesion characteristics, and also tested the hypothesis as to whether an AI method can be established using both clinical data and CT parameters to produce an reliable clinical prognostic model that can succor clinicians to prepare for early surveillance.A broad CT dataset on NCP and other typical pneumonia and standard controls had therefore been developed and an AI diagnostic method has been developed to assist with correct diagnosis for implementation in an epidemic area and two non-epidemic areas in China [12]. But, one thing which may be disputable is their interpretation of data.The data which they have put on public domain is very difficult to understand and not enough instructions or documentation of any sorts are provided for using it efficiently.Likewise,the data which they have used for their 3 pilot studies are not made public.Moreover, they trained their model using 80 percent of the total data and tested only on 10 percent without cross validating it.

We also studied another similar research to [12] which used the same data and discovered some discrepancies.we felt that there was a need for doing additional analysis of the explainability results for coming to an acceptable outcome.Identification of key patterns in the CT images was also missing which would have aided clinicians in manual screening. In addition, by using a mixture of CT and clinical criteria, they imparted prognostic signs for patients with NCP, with the goal of offering another instrument to assist physicians,According to [15], the lack of polished radiologists risks the availability and adequacy of COVID19 screening facilities in affected countries.Suspicious patients anywhere, especially in developed countries, have fair access to the correct diagnosis, appropriate treatment, and isolation through the implementation of AI diagnostic algorithms. With 10,250 CT scans from three centres in China and three publicly accessible databases, they developed clinically representative large-scale datasets.They developed both CT-based and CXR-based diagnostic systems and tested them using paired data to explain the relative efficiency of CT and CXR for the identification of COVID-19 [16]. DeepPneumonia was developed to help doctors diagnose the pneumonia-causing COVID-19 and localise the key lesions. Three key measures have been developed for their fully automatic lung CT diagnostic method.

Next, the key regions of the lungs were removed and the lung itself filled the blank of lung segmentation to prevent noises caused by different lung contours.Although, after going through their research we found their pre-processing steps highly debatable, specifically their selection of libraries.Also the data which was used for this experiment was relatively small, having only 88 COVID-19 patients.In order to extract the top-K information in the CT images and corral the image-level predictions, an Information Relation Extraction neural network (DRE-Net) was developed. Both two-dimensional local and 3D global model characteristics were being collected. As the backbone, the COVNet structure consisted of RestNet50, which took as input a set of CT slices and created functionality for the corresponding slices [17].Nevertheless, there were few inconsistencies in this research, such as the random choice of CAP from Aug'16 to Feb'20, and the analysis relied only on whether one sample was of COVID-19 or not, but the classification of the virus into various levels of severity was not discussed. Moving on.by using a max-pooling process, characteristics derived from both slices were then integrated. To fabricate a probability score for each form, the final feature map was fed to a totally connected layer and the softmax activation function. Also,we have seen a variety of papers which have put-to-use well-known convolutional neural networks to separate COVID-19 infection from non-COVID-19 classes, including [18], [19], and [20].

We checked a number of papers for some detailed studies on image segmentation and found the methodology of [21], [22] and [23] useful. [21] proposes a recent and realistic deep, completely convolutional neural network architecture named SegNet for semantic pixel-wise segmentation. This central trainable segmentation system consists of a network of encoders, a corresponding network of decoders preceded by a classification stage pixel-wise.SegNet 's innovation lies in the fact in which the decoder up-samples its function map with lower resolution information. And from both road scenes and SUN RGB-D indoor scene segmentation activities, they also conducted a managed benchmark of SegNet and other frameworks. In addition , the submitted cadre was contrasted with the commonly accepted FCN [24] and also with the well-known architectures of DeepLab-LargeFOV [25], DeconvNet [26].After detailed evaluation we contrived that in [24] they trained and validated on the PASCAL VOC 2011 segmentation challenge which is relatively ancient. Apart from that, we agree that both accurate diagnosis and predictive management can be supported by quantitative instruments by enhancing disease identification and quantification and-the reproducibility of disease intensity evaluation. However, some difficulties limiting the production and realistic use of current quantitative analysis methods include the consistency of training records, and lack of real-world assessments of the impact on outcomes. [27]

They had described a fully convolutional network (FCN) when it comes to [23], trained end-to-end, pixels-to-pixels on semantic segmentation. Since their redefinition of classification networks,generates output maps for inputs of any size, the output parameters are generally decreased by sub-sampling.Sub-sampled classification networks to retain low philtres and fair computing requirements. The U-Net architecture is split into two directions in [22], contracting symmetric path resp for image segmentation. The Overlap Tile solution was chosen because for large images it is necessary to extend to the network. A 1x1 convolution was then used in the final layer to assign each 64-component function vector to the desired number of levels. The network has a total of 23 convolutional layers.Now, when it comes to standard criterions in this domain, we want to validate our architecture's output with [28] as well as [21] as we described above, which is a deep learning based approach that captures local and contextual information to segment the ONH's individual neural and connective tissues. Their long-term aim was to provide a basis for the segmentation and automated collection of structural parameters from OCT volumes in 3D that could be generalised. It consists of a down sampling an up-sampling tower, linked to each other by skip connexions with each tower consisting of one regular block and two residual blocks. Both the normal and the residual blocks were built using two dilated convolution layers, each with 16 philtres (size 3-3). We found that most of the researchers, including [29], used a dice similarity coefficient, a location-based calculation used to determine the overlap of the location.Also a boundary-based test, Normalized Surface Dice, was used to decide how similar the segmentation and ground truth thresholds at a given tolerance are to each other.

Now, since we have scrutinized a lot about the segmentation frameworks, let's move towards the classification aspect for the CT scans.For this we came across limited papers since the disease is very nascent. We found [30], [31], [32] useful for our project framework and hence we studied them in detail. [30] recommended the use of an attentive fully convolutional network that could concentrate on contaminated areas of the chest, allowing a more detailed forecast to be made.They trained the model on a relatively small dataset comprising of roughly 2000 CT images and also presented a diagram of their prototype attention maps and demonstrated how close they were to the board-certified radiologists' manually identified infected regions as we would be actually working extensively with images, it was important for us to have one survey paper in our literature, so we could navigate to the availability easily.for this purpose we chose [31] which discusses the approach and performance of the computer vision CT-based pathological condition diagnosis, having selected some recent representative works to provide an overview of their effectiveness.Based on their role in contagion management, they grouped the approaches listed into 3 categories: computed tomography (CT) scans, X-ray imaging, and control and prevention. Last but not least, [32] implements a computer-aided diagnosis (CAD) web site for online identification of COVID19.In this research, a public chest CT scan database was utilized, including 746 participants. In order to identify the most effective model for the hybrid system, a range of very well-known deep neural network architectures consisting of ResNet, Inception and MobileNet were inspected. To differentiate between COVID-19 and safe controls, a variation of the Tightly connected convolutional network (DenseNet) was chosen to minimise image dimensions.Nevertheless, the limited data set size they have chosen will have to be a big drawback, because the training phase for deep neural networks demands a significant number of samples and such limited samples trigger overfitting.

III. Discussion/ Conclusion

In the sections above,we have introduced and investigated the existing literature on this research area in order to discover relevant loopholes,in doing so we managed to get detailed insights which would be extremely useful for us while going deeper into our approach. To conclude, in this research, we want to use a deep convolutional neural network for detection of COVID-19 using CT-scan images. Furthermore, we want to achieve a high accuracy rate when compared to already existing works and try to obtain more features from the dataset not manually but by automating it with end-to-end structure.In addition, for the segmentation part, we intend to apply one of the methods which we have described in our literature review in order for our network to work optimally.The key motivation for reviewing all these journals, aside from obtaining successful results, was to find a most appropriate method that could be adhered to our input data.

References

[1] W. H. Organization *et al.*, "Laboratory testing for coronavirus disease 2019 (covid-19) in suspected human cases: interim guidance, 2 march 2020," World Health Organization, Tech. Rep., 2020.

[2] W.-j. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu, H. Shan, C.-l. Lei, D. S. Hui *et al.*, "Clinical characteristics of coronavirus disease 2019 in china," *New England journal of medicine*, vol. 382, no. 18, pp. 1708–1720, 2020.

[3] A. Chen, "China's coronavirus app could have unintended consequences," 2020.

[4] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: summary of a report of 72 314 cases from the chinese center for disease control and prevention," *Jama*, vol. 323, no. 13, pp. 1239–1242, 2020.

[5] J. F.-W. Chan, S. Yuan, K.-H. Kok, K. K.-W. To, H. Chu, J. Yang, F. Xing, J. Liu, C. C.-Y. Yip, R. W.-S. Poon *et al.*, "A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster," *The Lancet*, vol. 395, no. 10223, pp. 514–523, 2020.

[6] J. Chan, "W, yuan s, kok k-h, et al," *A familial cluster of pneumonia associated with the*, pp. 30 154–9, 2019.

[7] D. S. W. Ting, L. Carin, V. Dzau, and T. Y. Wong, "Digital technology and covid-19," *Nature medicine*, vol. 26, no. 4, pp. 459–461, 2020.

[8] M. L. Kyoung and J. Lee, "Drive-through trend sweeps across multiple sectors, korea. net."

[9] H. Shi, X. Han, N. Jiang, Y. Cao, O. Alwalid, J. Gu, Y. Fan, and C. Zheng, "Radiological findings from 81 patients with covid-19 pneumonia in wuhan, china: a descriptive study," *The Lancet Infectious Diseases*, 2020.

[10] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849*, 2020.

[11] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, "Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: a report of 1014 cases," *Radiology*, p. 200642, 2020.

[12] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Zha, W. Liang, C. Wang, K. Wang *et al.*, "Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography," *Cell*, 2020.

[13] L. Wang and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *arXiv preprint arXiv:2003.09871*, 2020.

[14] X. Mei, H. Lee, K. Diao, M. Huang, B. Lin, C. Liu, Z. Xie, Y. Ma, P. Robson, M. Chung, A. Bernheim, V. Mani, C. Calcagno, K. Li, S. Li, H. Shan, J. Lv, T. Zhao, J. Xia, Q. Long, S. Steinberger, A. Jacobi, T. Deyer, M. Luksza, F. Liu, B. Little, Z. Fayad, and Y. Yang, "Artificial intelligenceenabled rapid diagnosis of patients with covid-19," *Nature Medicine*, vol. 26, no. 8, Aug. 2020.

[15] C. Jin, W. Chen, Y. Cao, Z. Xu, X. Zhang, L. Deng, C. Zheng, J. Zhou, H. Shi, and J. Feng, "Development and evaluation of an ai system for covid-19 diagnosis," *medRxiv*, 2020.

[16] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, H. Zhao, Y. Jie, R. Wang *et al.*, "Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images," *medRxiv*, 2020.

[17] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song *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, 2020.

[18] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "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*, p. 103795, 2020.

[19] V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, and N. Mehendale, "Diagnosis of covid-19 using ct scan images and deep learning techniques," *medRxiv*, 2020.

[20] J. Chen, L. Wu, J. Zhang, L. Zhang, D. Gong, Y. Zhao, S. Hu, Y. Wang, X. Hu, B. Zheng *et al.*, "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study," *MedRxiv*, 2020.

[21] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[22] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[23] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 4, pp. 640–651, 2017.

[24] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[25] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.

[26] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1520–1528.

[27] A. Chen, R. A. Karwoski, D. S. Gierada, B. J. Bartholmai, and C. W. Koo, "Quantitative ct analysis of diffuse lung disease," *Radiographics*, vol. 40, no. 1, pp. 28–43, 2020.

[28] S. K. Devalla, T. H. Pham, S. K. Panda, L. Zhang, G. Subramanian, A. Swaminathan, C. Z. Yun, M. Rajan, S. Mohan, R. Krishnadas *et al.*, "Towards label-free 3d segmentation of optical coherence tomography images of the optic nerve head using deep learning," *arXiv preprint arXiv:2002.09635*, 2020.

[29] J. Ma, Y. Wang, X. An, C. Ge, Z. Yu, J. Chen, Q. Zhu, G. Dong, J. He, Z. He *et al.*, "Towards efficient covid-19 ct annotation: A benchmark for lung and infection segmentation," *arXiv preprint arXiv:2004.12537*, 2020.

[30] S. Yazdani, S. Minaee, R. Kafieh, N. Saeedizadeh, and M. Sonka, "Covid ct-net: Predicting covid-19 from chest ct images using attentional convolutional network," *arXiv preprint arXiv:2009.05096*, 2020.

[31] A. Ulhaq, A. Khan, D. Gomes, and M. Pau, "Computer vision for covid-19 control: A survey," *arXiv preprint arXiv:2004.09420*, 2020.

[32] A. Saeedi, M. Saeedi, and A. Maghsoudi, "A novel and reliable deep learning web-based tool to detect covid-19 infection form chest ct-scan," *arXiv preprint arXiv:2006.14419*, 2020.