COVIDNet-CT: Detection of COVID-19 from Chest CT Images using a Tailored Deep Convolutional Neural Network Architecture

Hayden Gunraj — University of Waterloo
Linda Wang — University of Waterloo, DarwinAI Corp.
Alexander Wong — University of Waterloo, Waterloo AI Institute, DarwinAI Corp.
Email: {hayden.gunraj, linda.wang, a28wong}@uwaterloo.ca

Abstract

The COVID-19 pandemic continues to have a tremendous impact on patients and healthcare systems around the world. To combat this disease, there is a need for effective screening tools to identify patients infected with COVID-19, and to this end CT imaging has been proposed as a key screening method to complement RT-PCR testing. Early studies have reported abnormalities in chest CT images which are characteristic of COVID-19 infection, but these abnormalities may be difficult to distinguish from abnormalities caused by other lung conditions. Motivated by this, we introduce COVIDNet-CT, a deep convolutional neural network architecture tailored for detection of COVID-19 cases from chest CT images. We also introduce COVIDx-CT, a CT image dataset comprising 104,009 images across 1,489 patient cases. Finally, we leverage explainability to investigate the decision-making behaviour of COVIDNet-CT and ensure that COVIDNet-CT makes predictions based on relevant indicators in CT images.

1 Introduction

Currently, real-time reverse transcription polymerase chain reaction (RT-PCR) testing is the primary means of screening for COVID-19, as it can detect SARS-CoV-2 ribonucleic acid (RNA) in sputum samples collected from the upper respiratory tract [1]. While RT-PCR testing for COVID-19 is highly specific, its sensitivity is variable depending on sampling method and time since onset of symptoms [2–4], and some studies have reported relatively low COVID-19 sensitivity [3, 5].

Chest computed tomography (CT) imaging has been proposed as an effective screening tool for COVID-19 infection due to its high sensitivity, particularly when used as a complement to RT-PCR testing [4–6]. In early studies, it was found that certain abnormalities in chest CT images are indicative of COVID-19 infection [4–10]. However, these imaging abnormalities may not be specific to COVID-19 infection, and the performance of radiologists in distinguishing COVID-19-related abnormalities from abnormalities of other etiology may vary considerably [11, 12].

In this study, we introduce COVIDNet-CT, a deep convolutional neural network architecture tailored specifically for detection of COVID-19 cases from chest CT images via a machine-driven design exploration approach. We also introduce COVIDx-CT, a benchmark CT image dataset derived from CT imaging data collected by the China National Center for Bioinformation (CNCB) [13] comprising 104,009 images across 1,489 patient cases. Additionally, to investigate the decision-making behaviour of COVIDNet-CT, we perform an explainability-driven performance validation and analysis of its predictions, allowing us to explore the critical factors associated with COVID-19 infection while also auditing COVIDNet-CT to ensure that its decisions are based on relevant CT image features.

2 Methods

2.1 COVIDx-CT dataset

To build the proposed COVIDNet-CT, we constructed a dataset of 104,009 chest CT images across 1,489 patient cases, which we refer to as COVIDx-CT. This dataset is derived from CT imaging data collected by the CNCB [13], which is comprised of chest CT volumes across three different infection types: novel coronavirus pneumonia due to SARS-CoV-2 viral infection (NCP), common pneumonia (CP), and normal controls. For NCP and CP CT volumes, slices marked as containing lung abnormalities were leveraged. We split the COVIDx-CT dataset into training, validation, and test sets, using an approximate 60%-20%-20% split for training, validation, and test, respectively.

2.2 Machine-driven design exploration

Inspired by [14], in this study we leveraged generative synthesis [15] as our machine-driven design exploration strategy, where the problem of identifying a tailored deep neural network architecture for the task and data at hand is formulated as a constrained optimization problem based on a universal performance function $\mathcal{U}$ (e.g., [16]) and a set of quantitative constraints. For the initial network design prototype, we leveraged residual architecture design principles [17, 18], as they have been shown to enable reliable deep architectures which are easier to train to high performance.

2.3 Network architecture

The proposed COVIDNet-CT architecture is shown in Figure 1, and is made publicly available[1]. As can be seen, the machine-driven design exploration strategy made heavy use of unstrided and strided projection-replication-projection-expansion design patterns (which we denote as PRPE and PRPE-S for unstrided and strided patterns respectively) consisting of a projection to lower channel dimensionality via pointwise convolutions, a replication of the projections to increase channel dimensionality efficiently, an efficient spatial feature representation via depthwise convolutions, a projection to lower channel dimensionality via pointwise convolutions, and finally an expansion of channel dimensionality conducted by pointwise convolutions. Selective long-range connectivity can also be observed in the proposed COVIDNet-CT architecture, which enables greater representational capabilities in a more efficient manner than densely-connected deep neural network architectures.

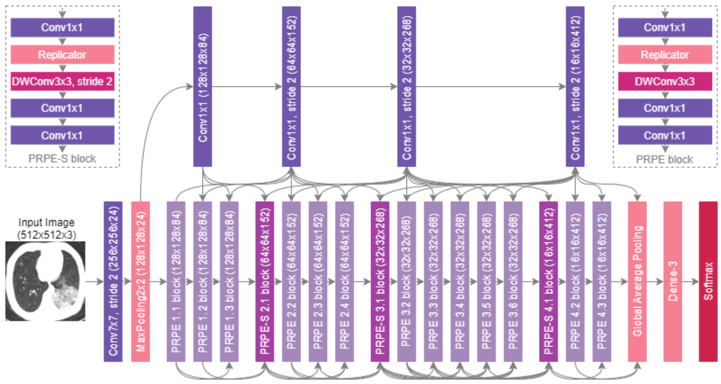

Fig. 1: The proposed COVIDNet-CT architecture designed via machine-driven design exploration. Notable characteristics include high architectural diversity, selective long-range connectivity, and lightweight design patterns (e.g., PRPE and PRPE-S patterns).

2.4 Explainability-driven performance validation of COVIDNet-CT

In clinical applications, the ability to understand how a deep neural network makes decisions is critical, as these decisions may ultimately affect the health of patients. Motivated by this, we audit COVIDNet-CT via an explainability-driven performance analysis strategy in order to better understand which CT imaging features are critical to its detection decisions. In this study, we leverage GSInquire [19] as the explainability method of choice for explainability-driven performance validation to visualize critical factors in CT images. GSInquire leverages the generative synthesis strategy [15] that was employed for machine-driven design exploration, and was previously shown quantitatively to provide explanations that better reflect the decision-making process of deep neu-

[1] https://github.com/haydengunraj/COVIDNet-CT

ral networks when compared to other state-of-the-art explainability methods [19].

3 Results and discussion

3.1 Quantitative results

We quantitatively evaluate the performance of the proposed COVIDNet-CT on the COVIDx-CT dataset. The test accuracy, architectural complexity, and computational complexity of COVIDNet-CT are shown in Table 1. As shown, COVIDNet-CT achieves a relatively high test accuracy of 99.1%, which is 0.4% higher than that achieved with the ResNet-50 architecture [18] and 0.8% higher than that achieved with the EfficientNet-B0 architecture [20]. Moreover, COVIDNet-CT has significantly lower architectural and computational complexity than ResNet-50 (94.1% fewer parameters and 90.2% fewer FLOPs), and compared to EfficientNet-B0 it exhibits reduced architectural complexity and similar computational complexity. From Table 2, we observe that COVIDNet-CT achieves good COVID-19 sensitivity (97.3%), which ensures a low proportion of false-negatives for COVID-19 cases. Moreover, given that RT-PCR testing is highly specific, COVIDNet-CT's high sensitivity would allow it to effectively complement RT-PCR testing. Next, from Table 3, we observe that COVIDNet-CT also achieves a high COVID-19 PPV, thereby ensuring a low proportion of false-positives which could cause an unnecessary burden on the healthcare system in the form of isolation, testing, and treatment. Based on these results, it is shown that COVIDNet-CT could be used as an effective standalone screening tool for COVID-19 patients, and could also be used in conjunction with RT-PCR testing.

Table 1: Comparison of parameters, FLOPs, and accuracy (image-level) for tested network architectures on the COVIDx-CT dataset. Best results highlighted in **bold**.

Architecture	Param. (M)	FLOPs (G)	Accuracy (%)
ResNet-50 [18]	23.55	42.72	98.7
EfficientNet-B0 [20]	4.05	**4.07**	98.3
COVIDNet-CT	**1.40**	4.18	**99.1**

Table 2: Sensitivity for each infection type at the image level on the COVIDx-CT dataset. Best results highlighted in **bold**.

	Sensitivity (%)		
Architecture	Normal	Non-COVID-19	COVID-19
ResNet-50 [18]	99.9	98.7	96.2
EfficientNet-B0 [20]	99.8	97.8	95.8
COVIDNet-CT	**100.0**	**99.0**	**97.3**

Table 3: Positive predictive value (PPV) for each infection type at the image level on the COVIDx-CT dataset. Best results highlighted in **bold**.

	PPV (%)		
Architecture	Normal	Non-COVID-19	COVID-19
ResNet-50 [18]	99.3	97.8	99.1
EfficientNet-B0 [20]	98.7	97.6	98.6
COVIDNet-CT	**99.4**	**98.4**	**99.7**

3.2 Qualitative results

In this study, we leveraged GSInquire [19] to perform explainability-driven performance validation of COVIDNet-CT in order to better understand its decision-making behaviour, and to ensure that its decisions are based on diagnostically-relevant imaging features rather than irrelevant visual indicators. As previously mentioned, GSInquire leverages the generative synthesis strategy in order to identify the critical factors leveraged by COVIDNet-CT to make its detection decisions. Figure 2 shows the critical factors identified by GSInquire for three chest CT images of patients with COVID-19 pneumonia. Examining these visual interpretations, we observe that COVIDNet-CT primarily leverages abnormalities within the lungs in the chest CT images to identify COVID-19 cases, as well as to differentiate these cases from non-COVID-19 pneumonia cases. As previously mentioned, our initial experiments yielded deep neural networks that were found via explainability-driven performance validation to be basing their detection decisions on irrelevant indicators such as patient tables and imaging artifacts, which highlights the importance of leveraging explainability methods when building and evaluating deep neural networks for clinical applications.

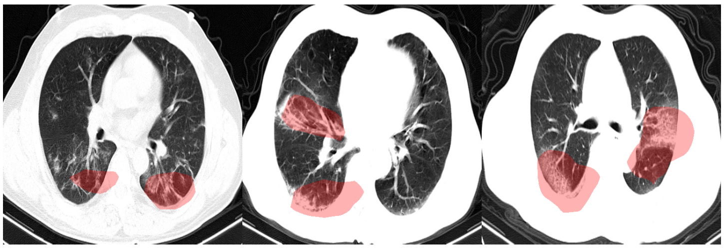

Fig. 2: Example chest CT images of COVID-19 cases and their associated critical factors (highlighted in red) as identified by GSInquire [19].

Based on both quantitative and qualitative results, it can be seen that not only does COVIDNet-CT achieve high performance, but it is leveraging relevant abnormalities in the lungs in its decision-making process rather than erroneous visual cues.

4 Conclusion

In this study, we introduced COVIDNet-CT, a deep convolutional neural network architecture tailored for detection of COVID-19 cases from chest CT images via machine-driven design exploration. Additionally, we introduced COVIDx-CT, a benchmark CT image dataset consisting of 104,009 chest CT images across 1,489 patients. We quantitatively evaluated COVIDNet-CT using the COVIDx-CT test dataset in terms of accuracy, sensitivity, and PPV. Furthermore, we analysed the predictions of COVIDNet-CT via explainability-driven performance validation to ensure that its predictions are based on relevant image features and to better understand the CT image features associated with COVID-19 infection. In our analyses, we observed that COVIDNet-CT is highly performant when tested on the COVIDx-CT test dataset, and that abnormalities in the lungs are leveraged by COVIDNet-CT in its decision-making process.

Acknowledgments

We would like to thank Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Research Chairs program, CIFAR, DarwinAI Corp., NVIDIA Corp., and Hewlett Packard Enterprise Co.

References

[1] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, and W. Tan, "Detection of SARS-CoV-2 in Different Types of Clinical Specimens," *JAMA*, vol. 323, pp. 1843–1844, 05 2020.

[2] Y. Yang, M. Yang, C. Shen, F. Wang, J. Yuan, J. Li, M. Zhang, Z. Wang, L. Xing, J. Wei, L. Peng, G. Wong, H. Zheng, M. Liao, K. Feng, J. Li, Q. Yang, J. Zhao, Z. Zhang, L. Liu, and Y. Liu, "Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections," *medRxiv*, 2020.

[3] Y. Li, L. Yao, J. Li, L. Chen, Y. Song, Z. Cai, and C. Yang, "Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19," *Journal of Medical Virology*, vol. 92, no. 7, pp. 903–908, 2020.

[4] 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 for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, 2020. PMID: 32101510.

[5] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, and W. Ji, "Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR," *Radiology*, vol. 296, no. 2, pp. E115–E117, 2020. PMID: 32073353.

[6] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, and J. Liu, "Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: Relationship to Negative RT-PCR Testing," *Radiology*, vol. 296, no. 2, pp. E41–E45, 2020. PMID: 32049601.

[7] 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, B. Du, L.-j. Li, G. Zeng, K.-Y. Yuen, R.-c. Chen, C.-l. Tang, T. Wang, P.-y. Chen, J. Xiang, S.-y. Li, J.-l. Wang, Z.-j. Liang, Y.-x. Peng, L. Wei, Y. Liu, Y.-h. Hu, P. Peng, J.-m. Wang, J.-y. Liu, Z. Chen, G. Li, Z.-j. Zheng, S.-q. Qiu, J. Luo, C.-j. Ye, S.-y. Zhu, and N.-s. Zhong, "Clinical Characteristics of Coronavirus Disease 2019 in China," *New England Journal of Medicine*, vol. 382, no. 18, pp. 1708–1720, 2020.

[8] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, and Z. Peng, "Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China," *JAMA*, vol. 323, pp. 1061–1069, 03 2020.

[9] M. Chung, A. Bernheim, X. Mei, N. Zhang, M. Huang, X. Zeng, J. Cui, W. Xu, Y. Yang, Z. A. Fayad, A. Jacobi, K. Li, S. Li, and H. Shan, "CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. 202–207, 2020. PMID: 32017661.

[10] F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Li, D. Zheng, J. Wang, R. L. Hesketh, L. Yang, and C. Zheng, "Time Course of Lung Changes at Chest CT during Recovery from Coronavirus Disease 2019 (COVID-19)," *Radiology*, vol. 295, no. 3, pp. 715–721, 2020. PMID: 32053470.

[11] H. X. Bai, B. Hsieh, Z. Xiong, K. Halsey, J. W. Choi, T. M. L. Tran, I. Pan, L.-B. Shi, D.-C. Wang, J. Mei, X.-L. Jiang, Q.-H. Zeng, T. K. Egglin, P.-F. Hu, S. Agarwal, F.-F. Xie, S. Li, T. Healey, M. K. Atalay, and W.-H. Liao, "Performance of Radiologists in Differentiating COVID-19 from Non-COVID-19 Viral Pneumonia at Chest CT," *Radiology*, vol. 296, no. 2, pp. E46–E54, 2020. PMID: 32155105.

[12] X. Mei, H.-C. Lee, K.-y. 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, and Y. Yang, "Artificial intelligence–enabled rapid diagnosis of patients with COVID-19," *Nature Medicine*, pp. 1–5, 05 2020.

[13] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Zha, W. Liang, C. Wang, K. Wang, L. Ye, M. Gao, Z. Zhou, L. Li, J. Wang, Z. Yang, H. Cai, J. Xu, L. Yang, W. Cai, W. Xu, S. Wu, W. Zhang, S. Jiang, L. Zheng, X. Zhang, L. Wang, L. Lu, J. Li, H. Yin, W. Wang, O. Li, C. Zhang, L. Liang, T. Wu, R. Deng, K. Wei, Y. Zhou, T. Chen, J. Y.-N. Lau, M. Fok, J. He, T. Lin, W. Li, and G. Wang, "Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography," *Cell*, vol. 18, no. 6, pp. 1423–1433, 2020.

[14] 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," 2020.

[15] A. Wong, M. J. Shafiee, B. Chwyl, and F. Li, "FermiNets: Learning generative machines to generate efficient neural networks via generative synthesis," 2018.

[16] A. Wong, "NetScore: Towards Universal Metrics for Large-scale Performance Analysis of Deep Neural Networks for Practical Usage," *CoRR*, vol. abs/1806.05512, 2018.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Identity Mappings in Deep Residual Networks," in *Computer Vision - ECCV 2016* (B. Leibe, J. Matas, N. Sebe, and M. Welling, eds.), (Cham), pp. 630–645, Springer International Publishing, 2016.

[19] Z. Q. Lin, M. J. Shafiee, S. Bochkarev, M. S. Jules, X. Y. Wang, and A. Wong, "Do Explanations Reflect Decisions? A Machine-centric Strategy to Quantify the Performance of Explainability Algorithms," 2019.

[20] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *2019 International Conference on Machine Learning (ICML)*, 2019.