

# Leveraging state-of-the-art architectures by enriching training information - a case study

## Contribution to the MICCAI COVID-19 Lung CT Lesion Segmentation Challenge

LCSBmedAI team: Jan Sölter<sup>1</sup>, Daniele Proverbio<sup>2</sup>, Mehri Baniasadi<sup>2,1</sup>, Matias Nicolas Bossa<sup>1</sup>, Vanja Vlasov<sup>1</sup>, Beatriz Garcia Santa Cruz<sup>3,1</sup>, and Andreas Husch<sup>1,2</sup>

<sup>1</sup> Uni. of Luxembourg, Lux. Ctr. for Systems Biomedicine, Interventional Neuroscience Group

<sup>2</sup> Uni. of Luxembourg, Lux. Ctr. for Systems Biomedicine, Systems Control Group

<sup>3</sup> Centre Hospitalier de Luxembourg, National Dept. of Neurosurgery  
medAI@uni.lu

### Abstract

Our working hypothesis is that key factors in COVID-19 imaging are the available imaging data and their label noise and confounders, rather than network architectures per se. Thus, we applied existing state-of-the-art convolution neural network frameworks based on the U-Net architecture, namely nnU-Net [3], and focused on leveraging the available training data. We did not apply any pre-training nor modified the network architecture.

First, we enriched training information by generating two additional labels for lung and body area. Lung labels were created with a public available lung segmentation network and weak body labels were generated by thresholding. Subsequently, we trained three different multi-class networks: 2-label (original background and lesion labels), 3-label (additional lung label) and 4-label (additional lung and body label). The 3-label obtained the best single network performance in internal cross-validation (Dice-Score 0.756) and on the leaderboard (Dice-Score 0.755, Hausdorff95-Score 57.5).

To improve robustness, we created a weighted ensemble of all three models, with calibrated weights to optimise the ranking in Dice-Score. This ensemble achieved a slight performance gain in internal cross-validation (Dice-Score 0.760). On the validation set leaderboard, it improved our Dice-Score to 0.768 and Hausdorff95-Score to 54.8. It ranked 3rd in phase I according to mean Dice-Score.

Adding unlabelled data from the public TCIA dataset in a student-teacher manner significantly improved our internal validation score (Dice-Score of 0.770). However, we noticed partial overlap between our additional training data (although not human-labelled) and final test data and therefore submitted the ensemble without additional data, to yield realistic assessments.

## Acknowledgments

The experiments presented in this paper were carried out using the HPC facilities of the University of Luxembourg [4] – see <http://hpc.uni.lu>. The work of JS and MNB is supported by the Luxembourg National Research Fund (FNR) Grant COVID-19/2020-1/14702831/AICovIX/Husch. BGSC is supported by the FNR within the PARK-QC DTU (PRIDE17/12244779/PARK-QC). DP is supported by the FNR PRIDE DTU CriTiCS, ref 10907093. MB is supported by the FNR, AFR grant ref. 12548237.

## References

1. An, P., Xu, S., Harmon, S., Turkbey, E., Sanford, T., Amalou, A., Kassin, M., Varble, N., Blain, M., Anderson, V., Patella, F., Carrafiello, G., Turkbey, B., Wood, B.: CT Images in Covid-19 [Data set]. In: The Cancer Imaging Archive (2020). <https://doi.org/10.7937/tcia.2020.gqry-nc81>
2. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., Prior, F.: The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging* **26**(6), 1045–1057 (2013). <https://doi.org/10.1007/s10278-013-9622-7>
3. Isensee, F., Jaeger, P., Kohl, S., et al.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods* (2020). <https://doi.org/10.1038/s41592-020-01008-z>
4. Varrette, S., Bouvry, P., Cartiaux, H., Georgatos, F.: Management of an academic hpc cluster: The ul experience. In: Proc. of the 2014 Intl. Conf. on High Performance Computing & Simulation (HPCS 2014). pp. 959–967. IEEE, Bologna, Italy (July 2014)