Relevant 2-Step Pretraining Improves Semantic Segmentation on Limited Data

An exploration into how pretraining can be used to assist semantic segmentation of COVID-19 chest x-rays even though the dataset is tiny dataset

Tau Merand
Supervisor: Dr. Richard Klein
School of Computer Science and Applied Mathematics, University of the Witwatersrand
tau.taylor.merand@gmail.com
github.com/TauMerand

Semantic Segmentation

Semantic segmentation involves labeling or classifying each pixel that belongs to a particular object or class in an image. Segmentation is useful for accurate accurate object localisation, particularly when the objects are non-regularly shaped. This has applications in vision problems where it is required to now the exact boundaries of an object such as self driving cars and medical diagnostics.

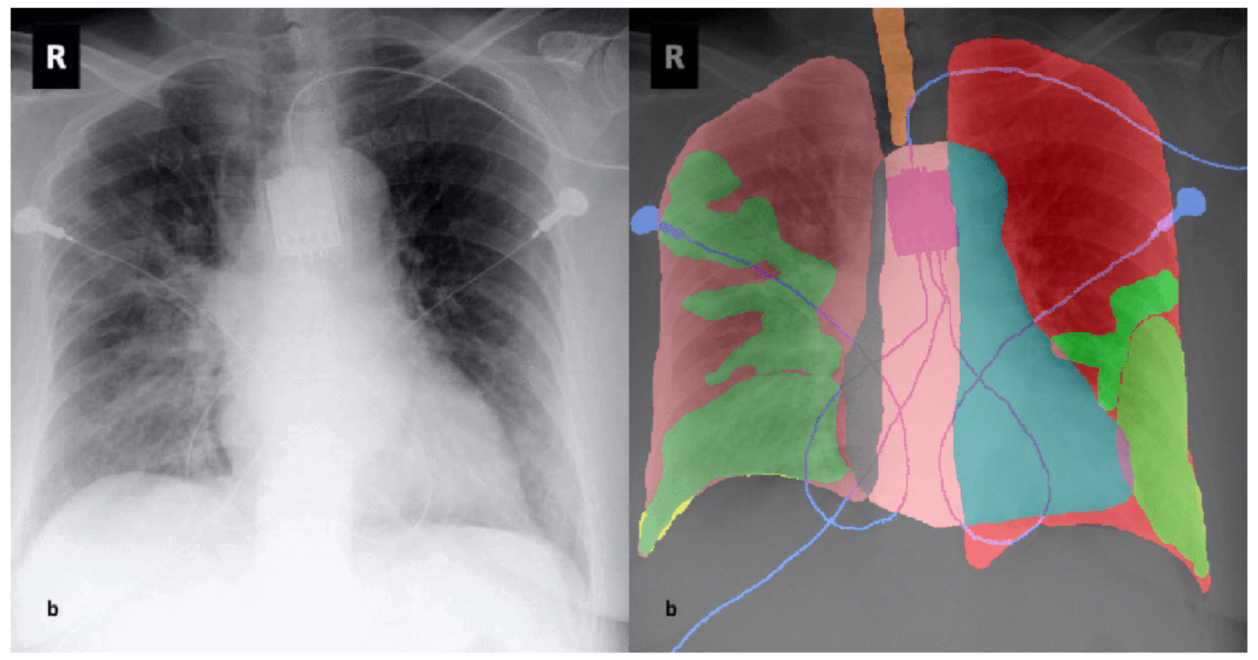


Figure 1: The left image is an ordinary x-ray while the right has its segmentation mask overlaid and coloured. It is possible to see how each pixel denoting the heart, lungs, medical tubing etc. are labelled. Observe that a bounding box would include a large amount of normal lung to enclose the entire ground glass opacity indicated by the green mask. The segmentation mask also has the benefit of emphasising parts of the x-ray that may have been indistinct to the naked eye.

Data, Data, Data

- **COVID-19 Chest X-ray Segmentation (CCXS)** dataset is a small, 100 image dataset of x-rays of COVID-19 positive adults which have been semantically segmented. CCXS has 13 segmentation categories. Figure 1 is an example from CCXS.
- **SIIM-FISABIO-RSNA COVID-19 (SFRC)** dataset consists of 7597 images that have been annotated by radiologists with bounding boxes around COVID-19 pneumonia related abnormalities.
- **CheXpert** dataset consists of 224,316 chest x-ray. It is considered to be the most accurate, large chest x-ray dataset available[2]. Unlike SFRC or CCXS CheXpert is an image classification dataset and has no object localisation only image level labels.

Pretraining

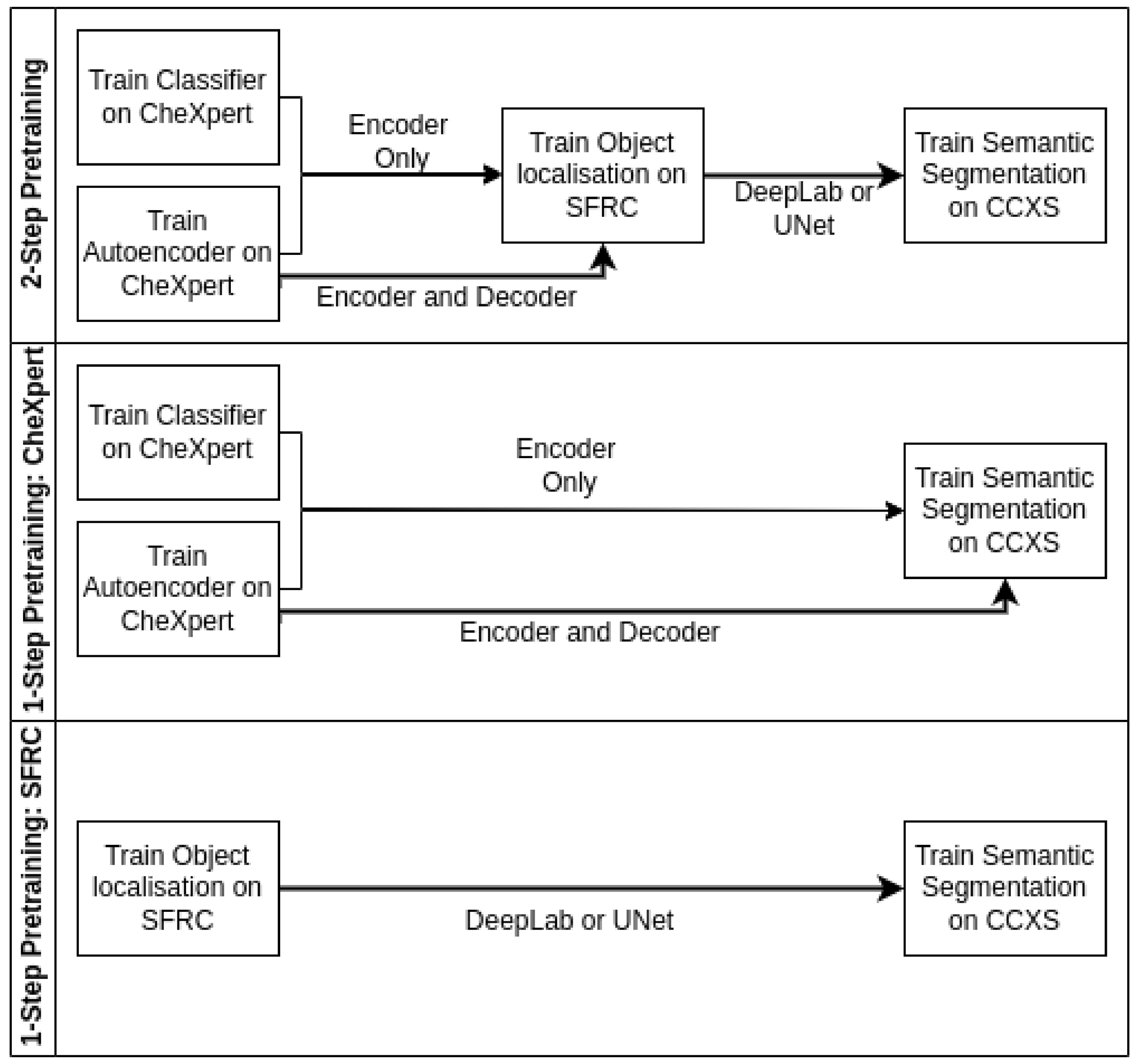
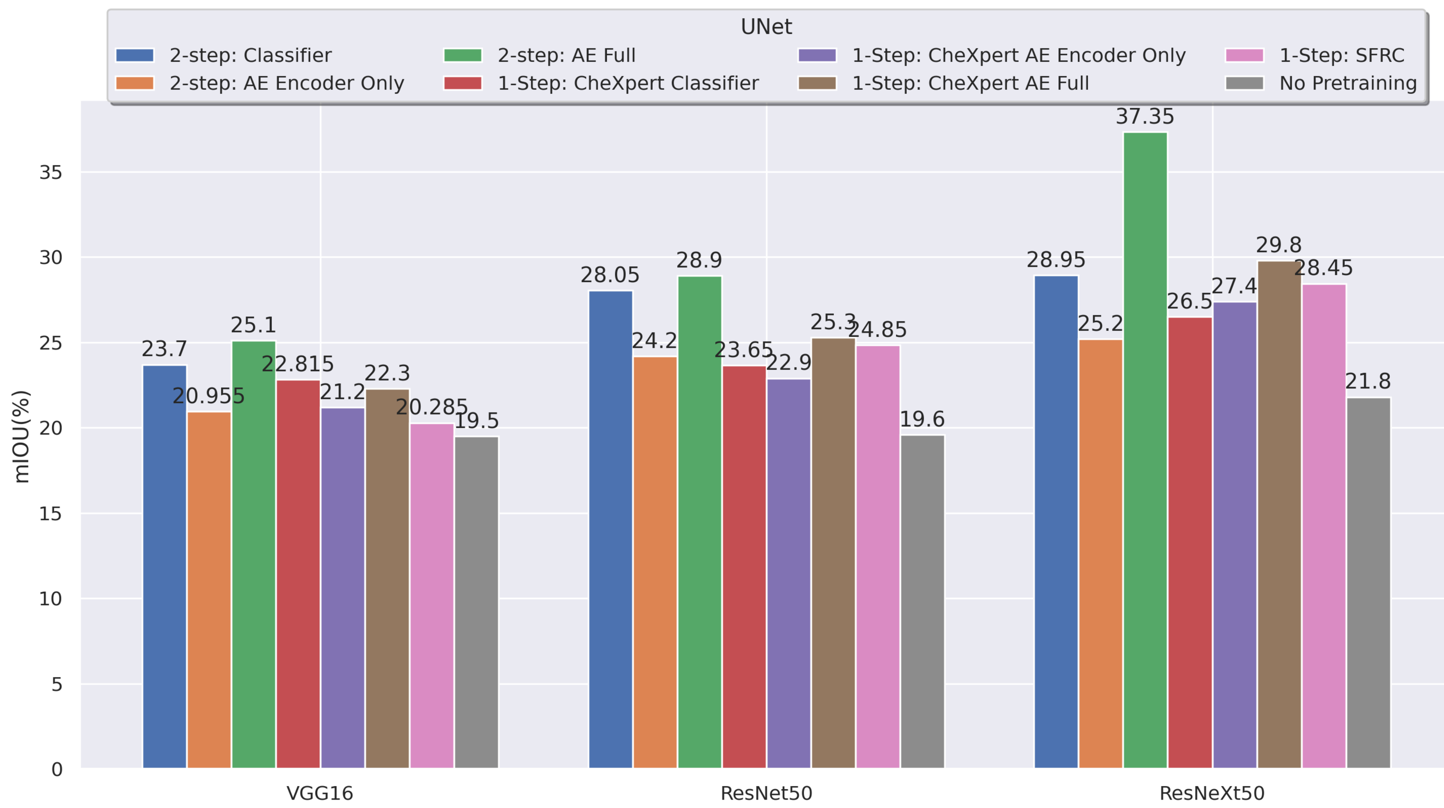


Figure 2: 2-Step pretraining involves training models first on CheXpert, then using those trained weights to train DeepLabv3 or UNet on SFRC then finally tuning the segmentation models on CCXS.

Results



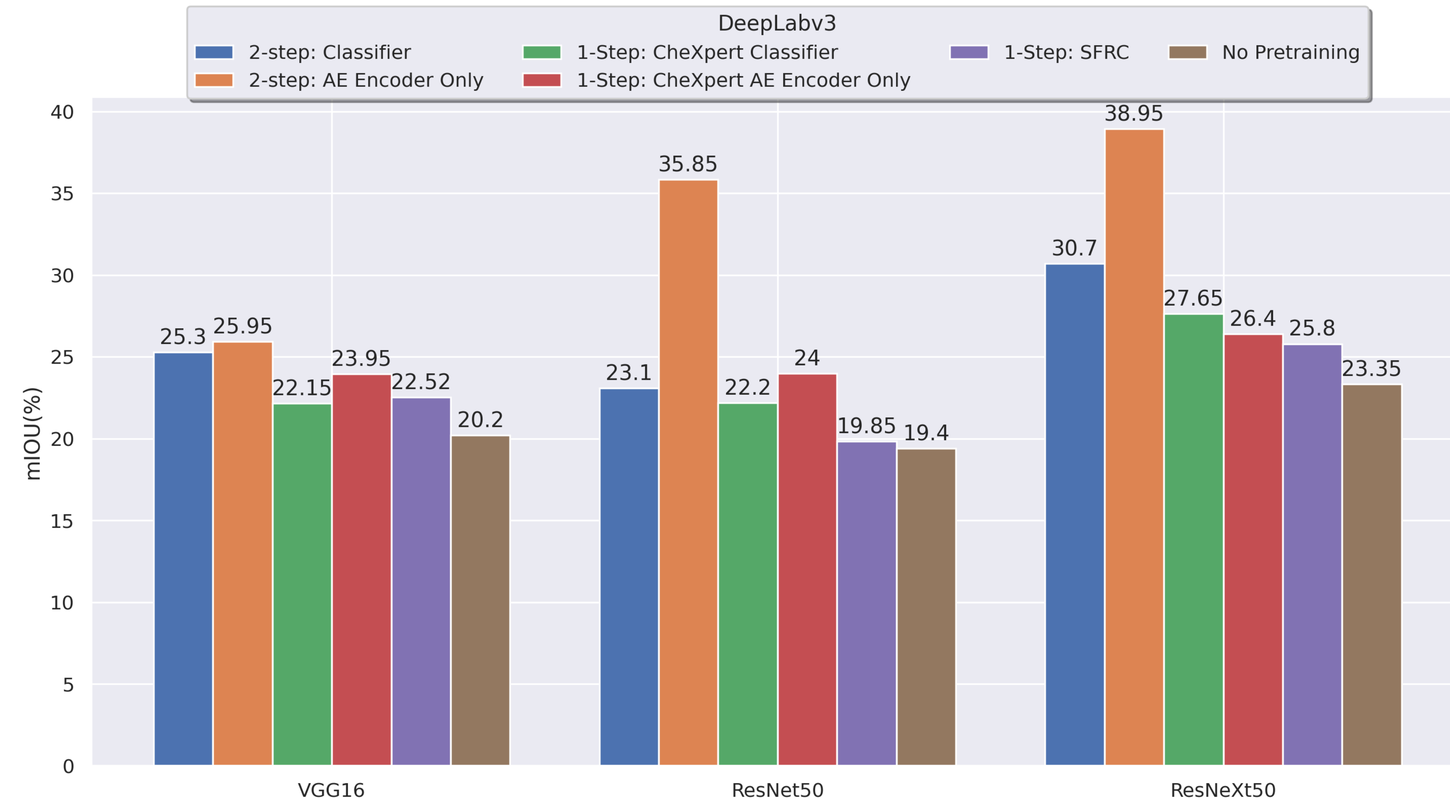


Figure 3: Comparison of average mIOU(%) achieved by UNet and DeepLab w.r.t 1- and 2-step pretraining for VGG16, ResNet50 and ResNeXt50 backbones. Note that UNet can use both the encoder and the decoder from a pretrained autoencoder whereas DeepLab can only use an encoder.

Conclusion

2-step pretraining is an effective way of improving the accuracy of a semantic segmentation model when faced with limited labelled segmentation data. DeepLabv3 with a ResNeXt50 backbone 2-step pretrained starting as an autoencoder on CheXpert is overall the best of the tested models for this approach. Overall all models that were pretrained did better than no-pretraining and all 6 of the top models were in some way 2-step pretrained.

References

[1] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.

[2] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silviana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 590–597, 2019.

[3] Paras Lakhani, John Mongan, Chinmay Singhal, Quan Zhou, Katherine P Andriole, William F Auffermann, Prasanth Prasanna, Tessie Pham, Michael Peterson, Peter J Bergquist, and et al. The 2021 siim-fisabio-rsna machine learning covid-19 challenge: Annotation and standard exam classification of covid-19 chest radiographs., Oct 2021.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.