MSc in Computational Software Techniques in Engineering

Evaluation of 2D Semantic Segmentation Methods for Scene Understanding for Autonomous Driving Applications

Objectives

1. Establish baseline performance to compare against the literature.
2. Evaluate performance and robustness of DeepLabv3 when evaluated with ApolloScape data.
3. Analyse the combination of the factors that led to the performance observed.

This helped determine the appropriate backbone and whether to attempt training the network on weak hardware. ResNet is slower than MobileNet, but more accurate. Autonomous driving motivation pushed more towards MobileNet.

Preliminary results
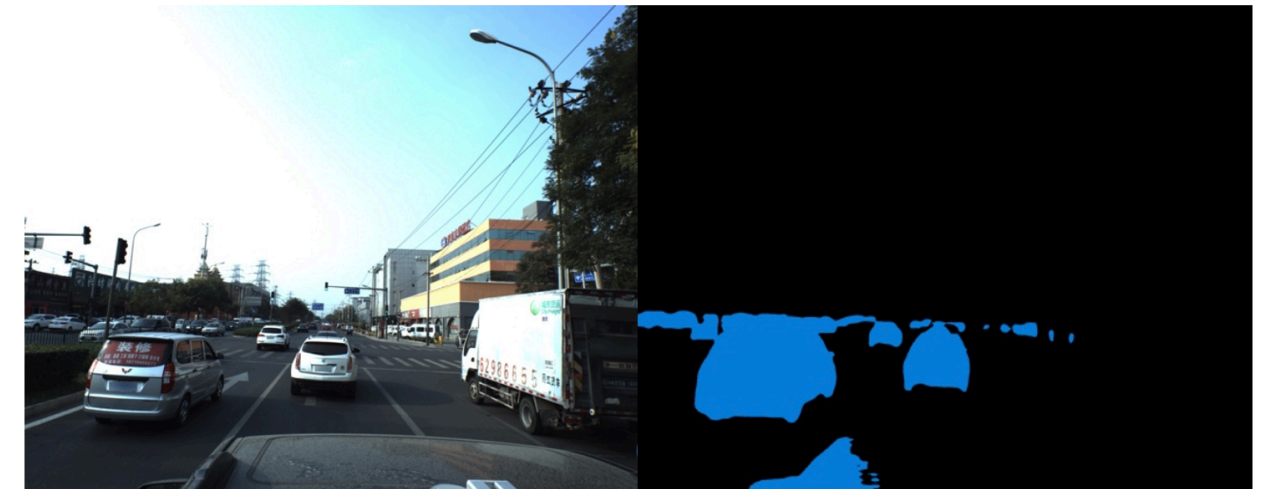
Methodology
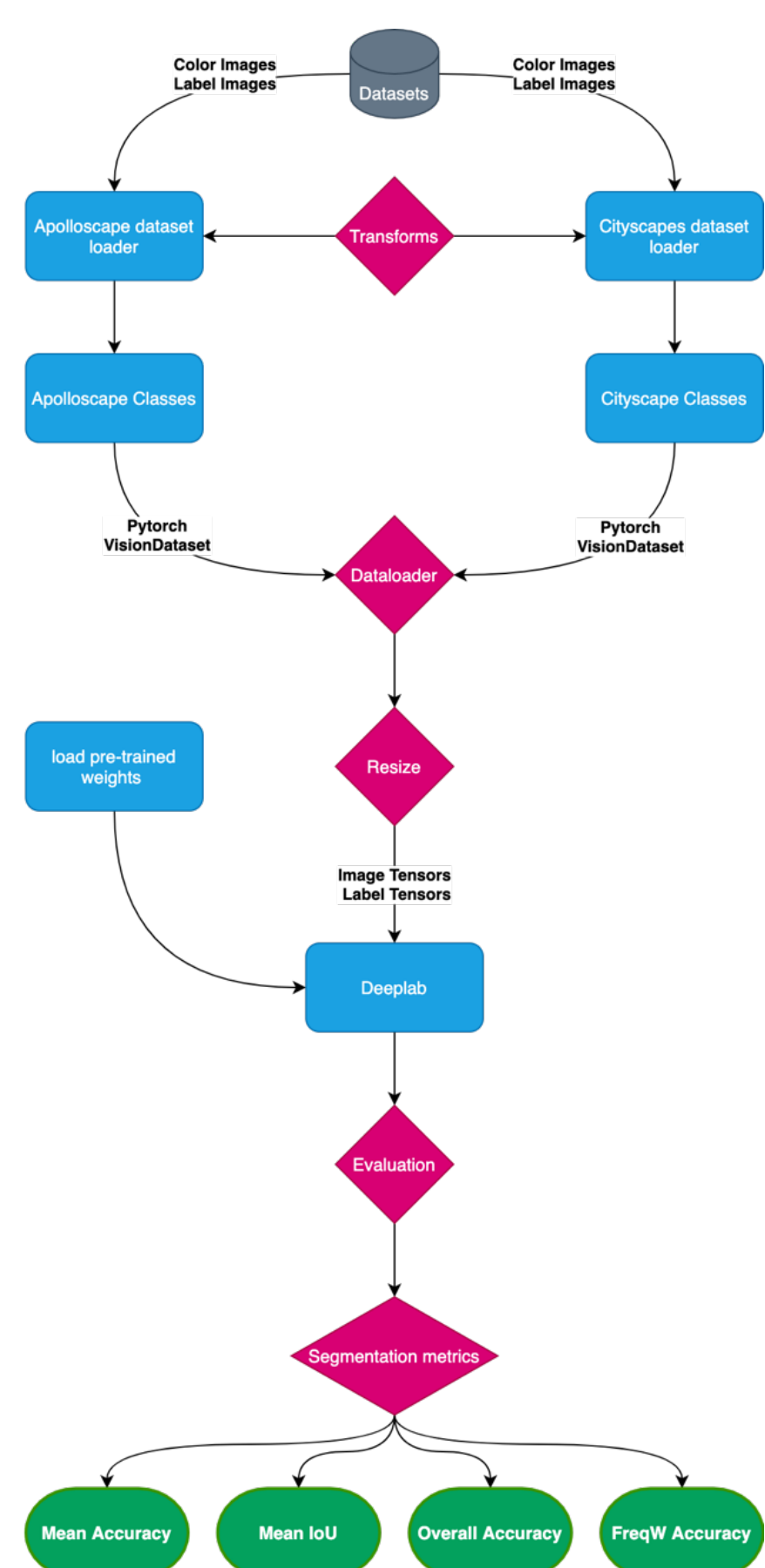
Key insights

- MobileNet is not noticeably faster.
- The largest differences came from image size.
- Iterating with different image sizes showcased that DeepLabv3Plus_MobileNet is pretty poor at smaller image sizes.
- Mapping between Cityscapes and ApolloScape was very doable
- ApolloScape is not labelled in the same way as Cityscapes so there is a lot of unclear labels where it would fall between class IDs.
- ApolloScape and Cityscapes have similarities but are quite different from their conception to execution.
- This also explains the class mismatch when DeepLabv3Plus_MobileNet detects certain classes that do not show up in the label image.
- This is a clear pattern in the results.

Conclusion

The purpose of this project was to evaluate the 2D semantic segmentation approach DeepLabv3 that was pre-trained on the dataset Cityscapes with the ApolloScape dataset. One of the key objectives was to determine the robustness of the model and the also the problem areas caused by the mismatch of classes between the datasets.

The findings demonstrate that the difference in datasets, especially with regards to the format of the Label files, is sufficient enough to derail quite a number of classes, and there by the mIoU results. The findings also suggest the model that was trained on Cityscapes and evaluated with the ApolloScape dataset struggled with a lot of common classes like Sky and Road, and therefore heavily impacted the mIOU of the approaches.

The evaluation of DeepLabv3Plus_MobileNet on Cityscapes achieved a mIoU of 0.626 when input images were resized to 512px. The evaluation with the ApolloScape dataset achieved a mIoU of 0.193 when input images were resized to 512px and 0.237 when 1024px. The poor performance of the latter is largely caused by label mismatches and the fact that the datasets were not designed to be perfectly compatible with each other.

Results
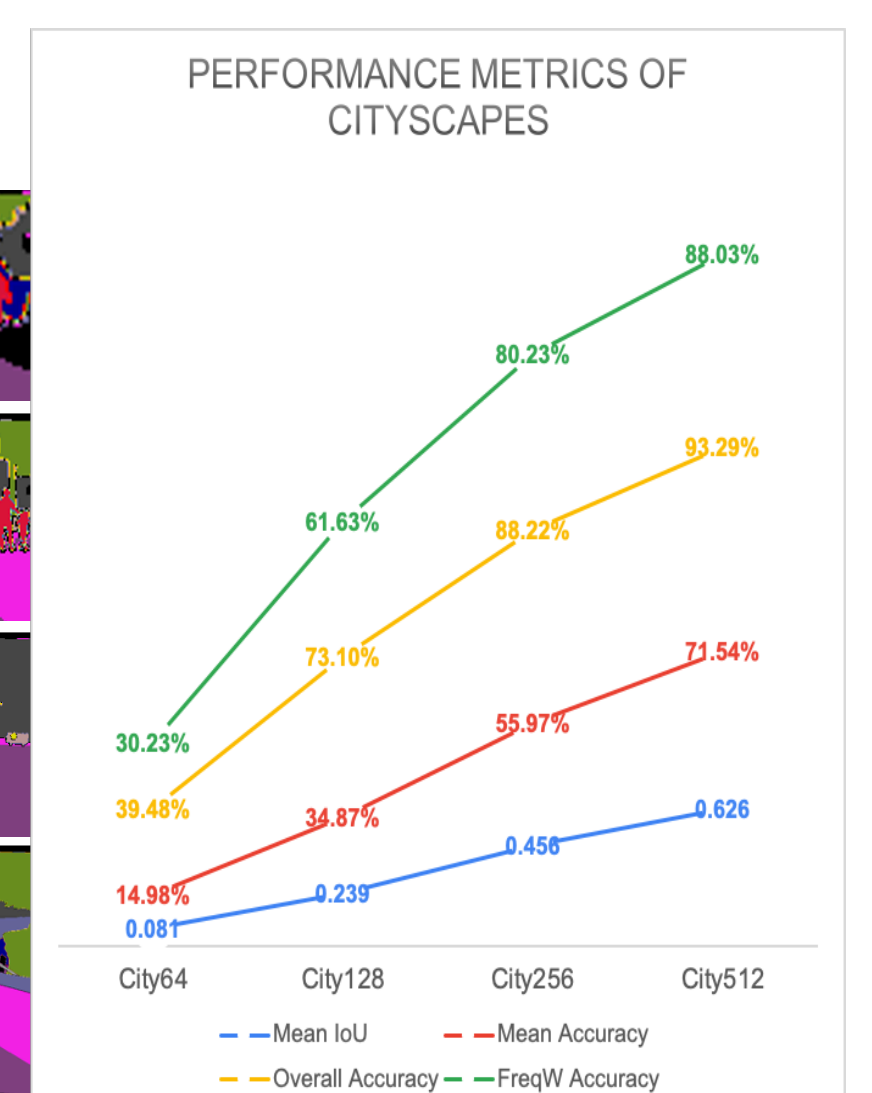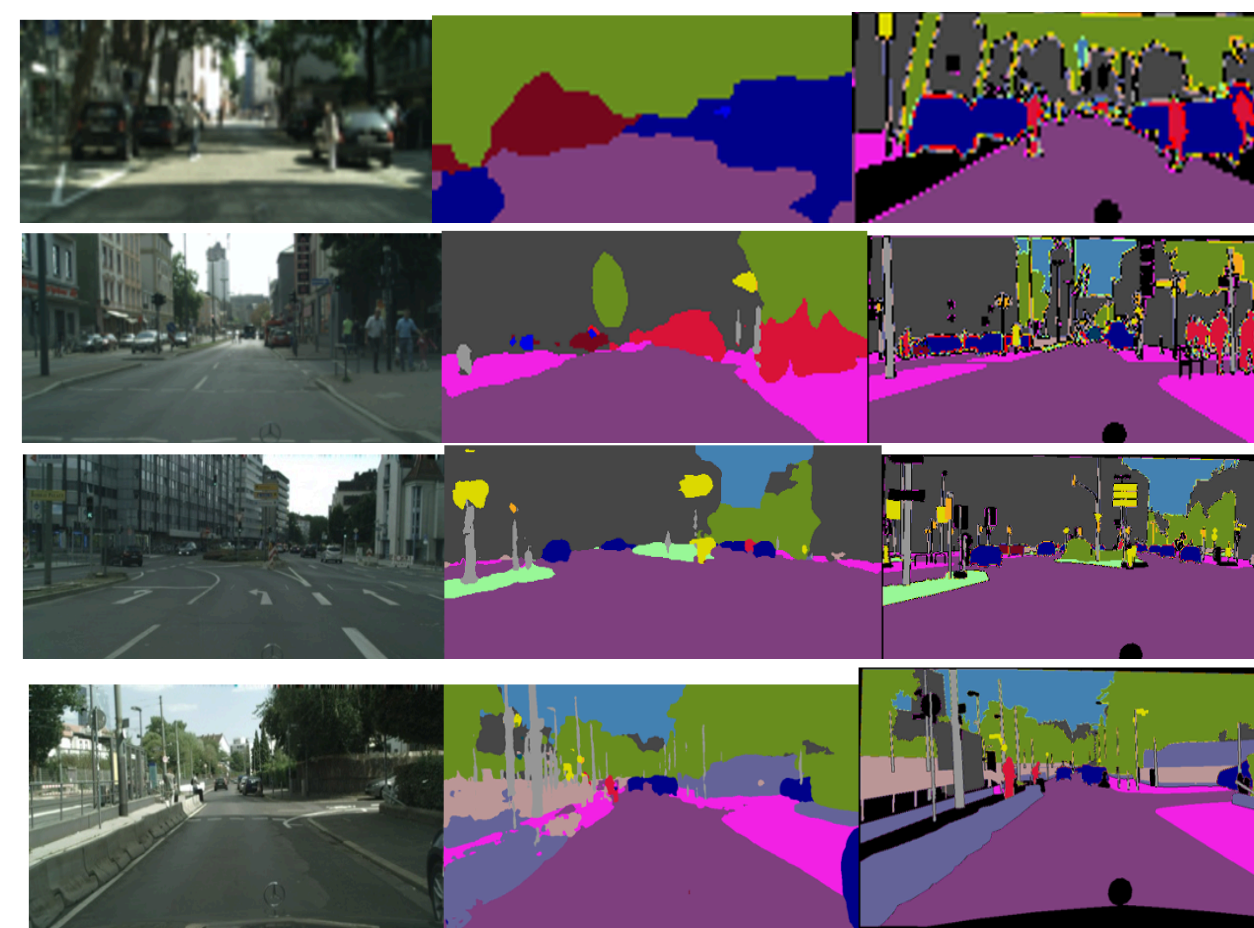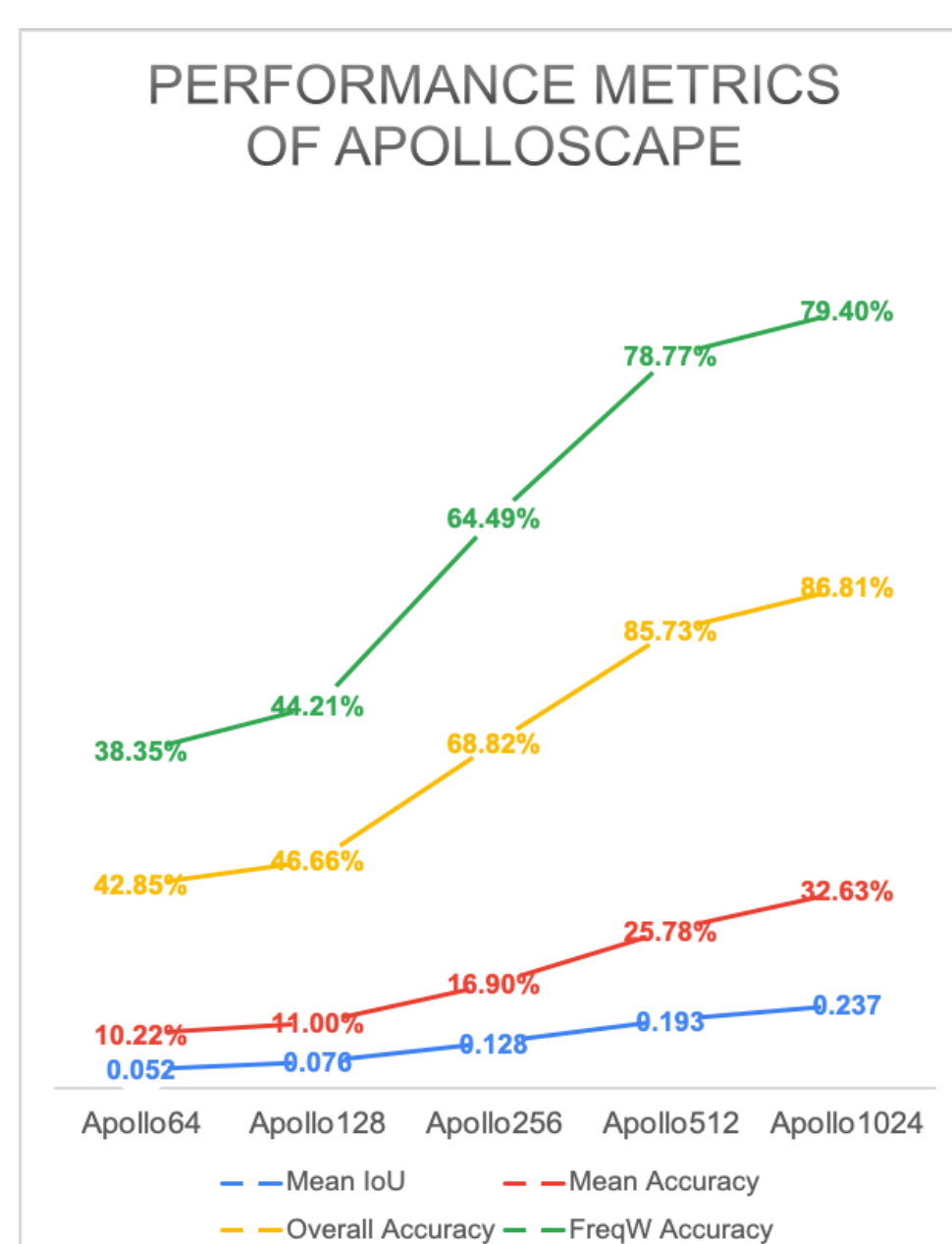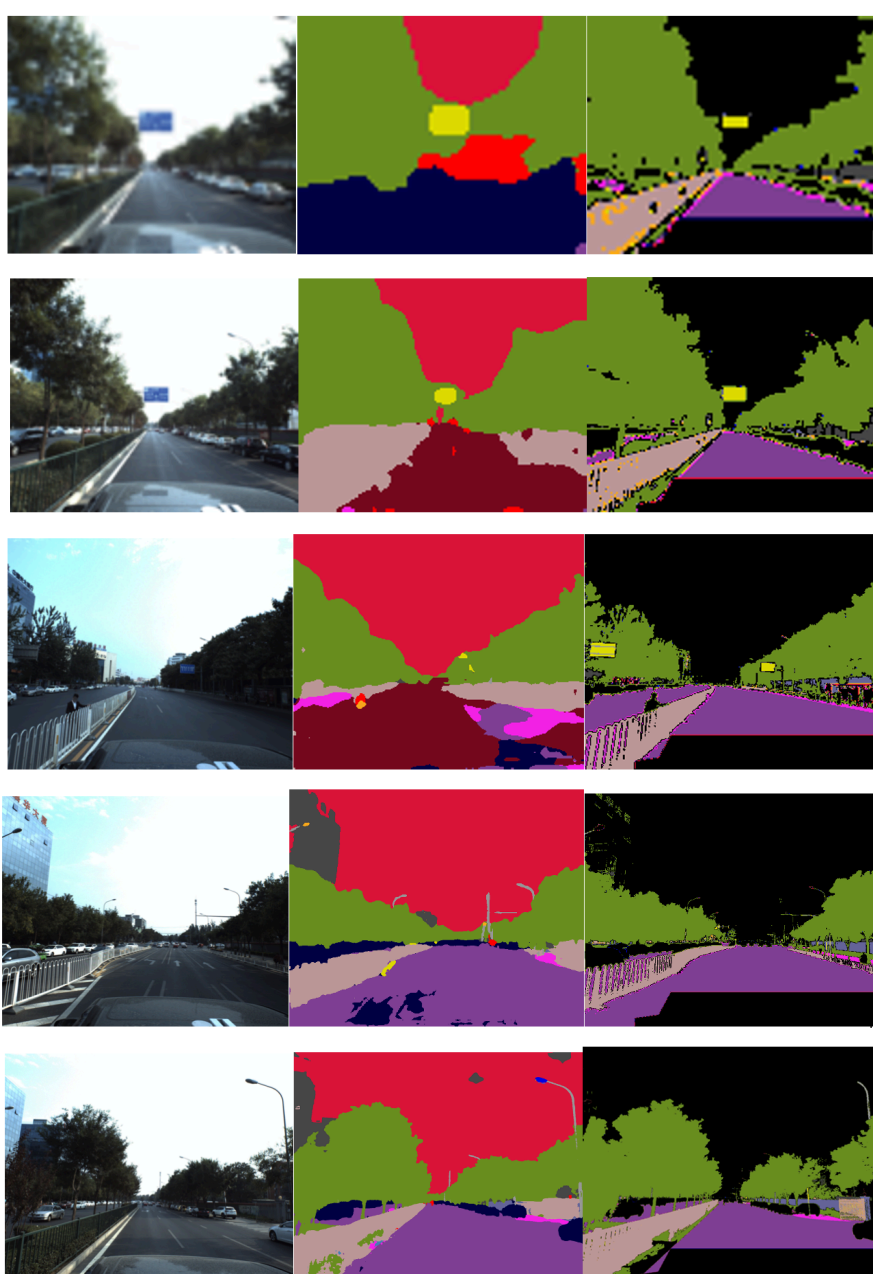
Muhammad Hasan Ali

Supervisor: Dr. Yifan Zhao

www.cranfield.ac.uk

Individual Research Project 2019/2020