7 Conclusion and future work

Overall, in this project, we mostly have created and trained a Mask R-CNN model with the specific VOC2012 dataset. We have also done some experiment mainly to demonstrate the efficiency of trained model. And before start training, we have carefully analyzing the dataset and choosing a variation of model settings with the aim to improve 1) the time efficiency of training and testing this model and 2) the precision of the model when performing object detection

The results prove the conclusion that our Mask R-CNN model performance are very well and visual results seemed to be very promising. And we also find the most appropriate IoU threshold to make sure correct objects are being marked. However, the mean average precision obtained is not as high as expect from an object detection model, and it also contain a few errors that the model makes on images in the test set. But after some careful analysis we reveal that the errors for these types of images were minimal, relative to the number of objects that were correctly located and classed.

Future work will address the short-comings of our approach – it’s possible that the model may not perform very well on some images. We think this problem could be solved by adjusting the training data to bring some improvement of accuracy.