AudioScript 2

Before discussing the results of both the baseline and final models, we first explain how the mean average precision , that was used for evaluating the performance is calculated.

Mean average precision (mAP) is a common and accepted way of evaluating models for object detection.

The first step in calculating mAP for a set of images is to calculate the precision and recall for each individual object in each image.

For each detection made for the image we calculate the intersection over union between it and the ground truth for that prediction. If the intersection is above than threshold, usually 0.5, then the prediction is classed as a true positive (TP). Whereas, if the intersection is less than 0.5 or there is duplicated bounding box prediction, then the detection is classed as false positive (FP). When the prediction is wrong classification or there is no prediction at all, then the detection is classed as false negative (FN).

With precision and recall, the precision-recall (PR) curve is then plotted, and the AP is calculated by taking the area under the curve. The higher the mAP is, the better the model prediction is.

We trained this model and the baseline model over 12 epochs, calculating the Mean average Precision at each epoch using a minimum intersection over union threshold of 0.5.

The Figure shows the results of mean average precision at each epoch for both the Baseline and final model on the train and test set respectively.

From the graph we can see that the final model outperforms the baseline at almost every epoch on both data sets, only falling lower at epoch 9.

The highest mAP score for the final model on the test set was 0.665, which was an increase of just over 2%, over the highest mAP for the baseline model, which was 0.651.

——————————————

Change Slide

——————————————

The model performed very well in locating people. The numerical results of the models show that the model has indeed been able to learn to detect and classify objects. Upon visual inspection of the predictions made by the model on images in the test set, we can see the model to be performing excellently in locating objects in the image.

These images below shows some of the predictions made by our model on test images.

————————————

Change Slide

————————————

….