ITI107 Assignment

Object Detection

Pamela Sin Hui (21A530M)

This was an assignment done solely.

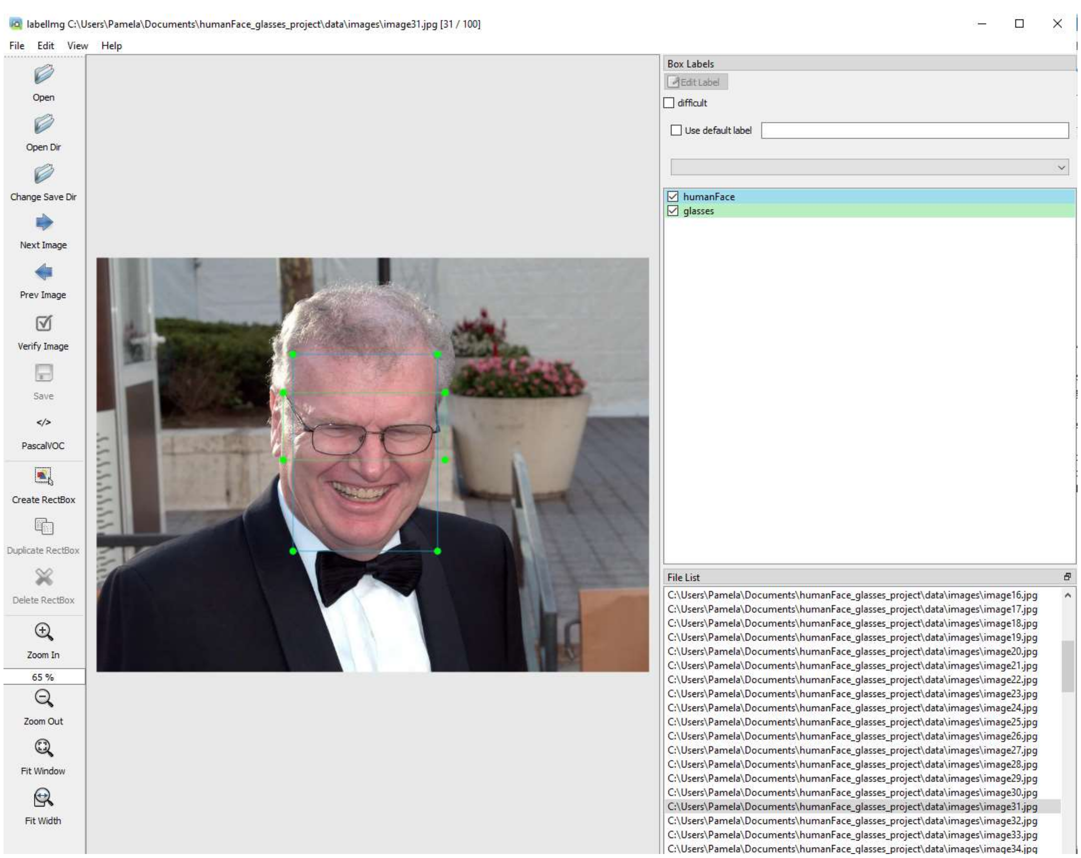
Section 1: Data Collection and Annotation

A pip install of the fiftyone library was done locally. The below code was run, where 150 images containing objects of “Glasses” and “Human face” was downloaded from the Open Images dataset (https://storage.googleapis.com/openimages/web/download.html).

```
import fiftyone as fiftyone
dataset = fiftyone.zoo.load_zoo_dataset(
              "open-images-v6",
              split="train",
              label_types=["detections"],
              classes=["Glasses", "Human face"],
              max_samples=150,
          )
```

From the 150 images, the first 100 were taken to be our final dataset, where subsequently specific images were swapped out, such as in an instance where an image showcased a large crowd of people which would have made annotation difficult. The 100 images were renamed from image1.jpg to image100.jpg for consistency. As “Glasses” and “Human face” objects tend to go hand-in-hand and be in the same image (i.e. the image contains a person wearing glasses), each of the 100 images have at least one “Glasses” object and one “Human face” object in the image.

Each of the 100 images was manually annotated using the LabelImg tool, using the PascalVOC format. For the annotation of “Glasses” objects, both clear glasses and sunglasses were annotated, and the entire object including the temples are being included in the ground truth label. For the “Human face” objects, only the human face, excluding the hair and ears are being included in the ground truth label. The images are taken from a variety of settings, be it indoors or outdoors, as it is the intention for the object detection model to be able to detect objects in different settings.



Section 2: Training Process

Training was done following step-by-step instructions from https://github.com/nyp-sit/iti107/blob/main/session-4/custom_training_with_tfod_api.md

Firstly, the images and annotation files are copied to separate folders. Next, the label map (.pbtxt file) of each numeric label to its corresponding text label was created. Subsequently, the images and annotations data was converted to TF Records binary format, with a train to validation split of 80% to 20%.

The pretrained model selected was ssd_resnet101_v1_fpn_640x640_coco17_tpu-8, given its relatively higher COCO mAP versus speed in comparison with other SSD models (https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_detection_zoo.md).

Model name	Speed (ms)	COCO mAP	Outputs
SSD MobileNet v2 320x320	19	20.2	Boxes
SSD MobileNet V1 FPN 640x640	48	29.1	Boxes
SSD MobileNet V2 FPNLite 320x320	22	22.2	Boxes
SSD MobileNet V2 FPNLite 640x640	39	28.2	Boxes
SSD ResNet50 V1 FPN 640x640 (RetinaNet50)	46	34.3	Boxes
SSD ResNet50 V1 FPN 1024x1024 (RetinaNet50)	87	38.3	Boxes
SSD ResNet101 V1 FPN 640x640 (RetinaNet101)	57	35.6	Boxes
SSD ResNet101 V1 FPN 1024x1024 (RetinaNet101)	104	39.5	Boxes
SSD ResNet152 V1 FPN 640x640 (RetinaNet152)	80	35.4	Boxes
SSD ResNet152 V1 FPN 1024x1024 (RetinaNet152)	111	39.6	Boxes

To perform hyperparameter tuning, the object detection pipeline file was configured differently for separate runs of the training. A batch size of 4 was used for all runs due to limited GPU memory.

Section 3: Experimental Results

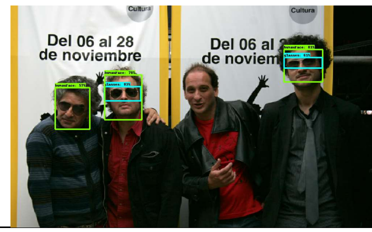
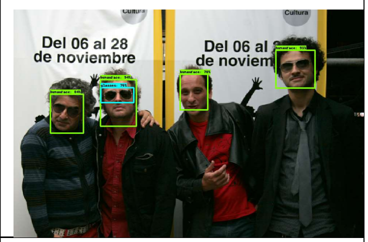
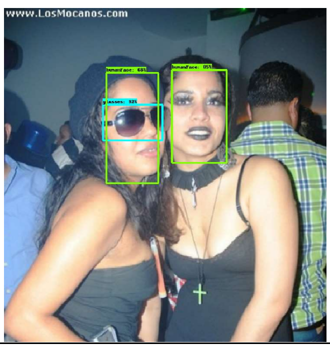
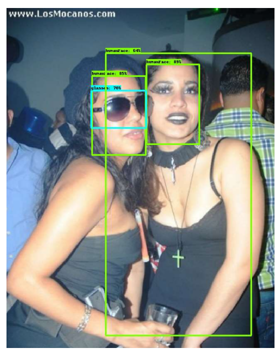
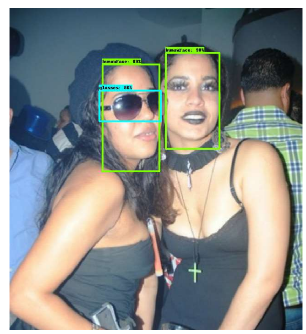
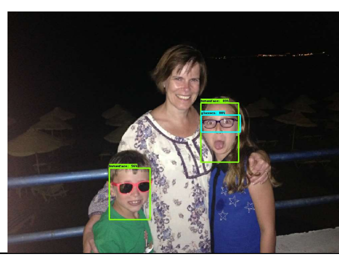
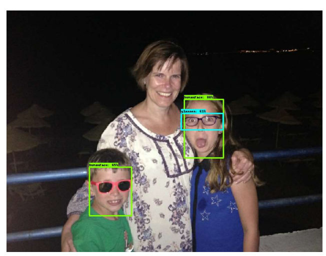
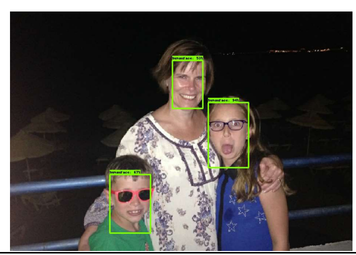
Run 2 contains the standard default hyperparameters defined in the pretrained model's configuration file.

For Run 3, the first anchor aspect ratio tweaked was from 2.0 to 1.5. This newly selected arbitrary aspect ratio seems to better capture the typical height to width ratio of the object "Human face", making this anchor less elongated. The second anchor aspect ratio tweaked was from 0.5 to 0.4, where this aspect ratio appears to better encompass the object "Glasses" sitting on a person's face is slightly wider. In comparing the mAP@IoU=0.50 between Run 2 and Run 3, Run 2 performed better at 0.643 versus Run 3 at 0.604.

In considering the experimental results of Run 2 and Run 3, for Run 4, Run 2's anchor aspect ratios was used, while the classification weight to localization weight was tweaked from 1.0:1.0 to 1.2:1.0. This was done as Run 2's Tensorboard plots of classification loss against number of batches hovers around 0.10 towards the tail end of training, while that for the localization loss is around 0.05. Hence in Run 4, we want to determine if placing greater weights on classification versus localization in the loss function will improve the performance metric. However, in comparing the mAP@IoU=0.50 between Run 2 and Run 4, Run 2 performed better at 0.643 versus Run 3 at 0.628.

We note that training for an even greater number of batches of data across the Runs could have led to different performance metric results. However, due to constraints in allocated GPU hours, we select the best model based on the experimental Runs thus far.

Section 3.1: Summary of Results

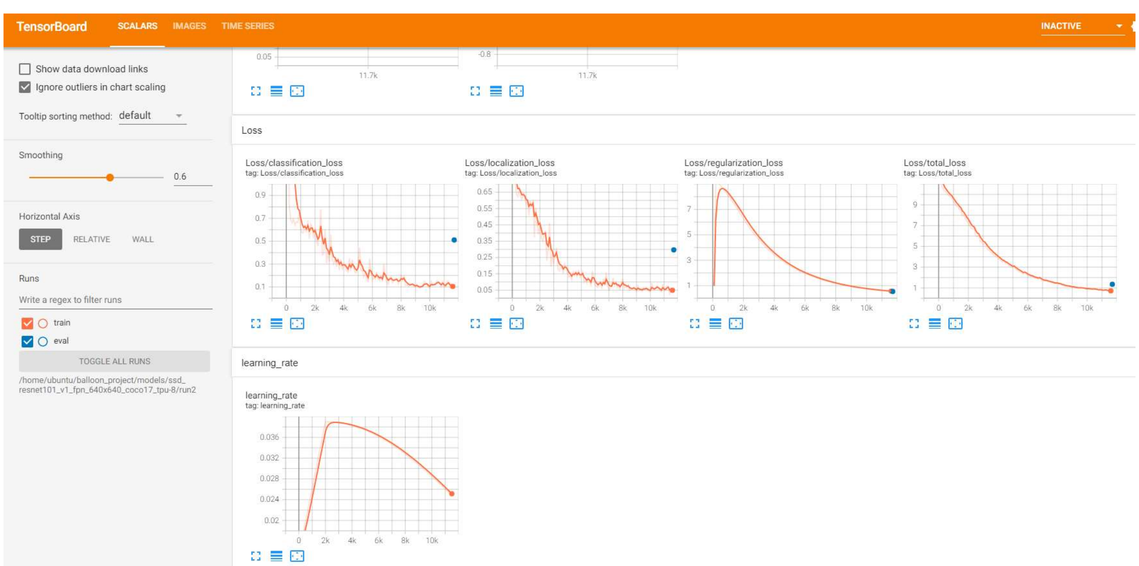
Run #	Run 2	Run 3	Run 4
Classification weight : Localization weight	1.0 : 1.0	1.0 : 1.0	1.2 : 1.0
Anchor Aspect Ratios	1.0, 2.0, 0.5	1.0, 1.5, 0.4	1.0, 2.0, 0.5
Step eval metrics are performed	11700	12900	12700
mAP@IoU=0.50	0.643	0.604	0.628
			
			
			

Section 3.2: Run 2 Detailed Results

The graphs of classification loss, localization loss, regularization loss and total loss over number of batches all have a general downward trend, where as the number of training batches increases, the losses decreases at a decreasing rate and plateaus at approximately 9000 training batches. This general trend is also seen across runs 3 and 4.

As Run 2 has the highest mAP@IoU=0.50, the side-by-side comparison of a few test images against the ground truth label is additional provided in this section.

```
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.314
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.643
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.183
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.200
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.342
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.200
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.396
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.438
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.331
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.463
INFO:tensorflow:Eval metrics at step 11700
I0109 22:59:20.827951 140110343554496 model_lib_v2.py:1007] Eval metrics at step 11700
INFO:tensorflow:        + DetectionBoxes_Precision/mAP: 0.313534
I0109 22:59:20.835347 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP: 0.313534
INFO:tensorflow:        + DetectionBoxes_Precision/mAP@.50IOU: 0.642573
I0109 22:59:20.836146 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP@.50IOU: 0.642573
INFO:tensorflow:        + DetectionBoxes_Precision/mAP@.75IOU: 0.182703
I0109 22:59:20.836694 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP@.75IOU: 0.182703
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (small): 0.000000
I0109 22:59:20.837198 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (small): 0.000000
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (medium): 0.200446
I0109 22:59:20.837694 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (medium): 0.200446
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (large): 0.342099
I0109 22:59:20.838170 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (large): 0.342099
INFO:tensorflow:        + DetectionBoxes_Recall/AR@1: 0.200263
I0109 22:59:20.838645 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@1: 0.200263
INFO:tensorflow:        + DetectionBoxes_Recall/AR@10: 0.396053
I0109 22:59:20.839138 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@10: 0.396053
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100: 0.437807
I0109 22:59:20.839626 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100: 0.437807
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (small): 0.000000
I0109 22:59:20.840081 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (small): 0.000000
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (medium): 0.331250
I0109 22:59:20.840557 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (medium): 0.331250
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (large): 0.462894
I0109 22:59:20.841067 140110343554496 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (large): 0.462894
INFO:tensorflow:        + Loss/localization_loss: 0.296767
I0109 22:59:20.841483 140110343554496 model_lib_v2.py:1010]     + Loss/localization_loss: 0.296767
INFO:tensorflow:        + Loss/classification_loss: 0.510632
I0109 22:59:20.841907 140110343554496 model_lib_v2.py:1010]     + Loss/classification_loss: 0.510632
INFO:tensorflow:        + Loss/regularization_loss: 0.542540
I0109 22:59:20.842327 140110343554496 model_lib_v2.py:1010]     + Loss/regularization_loss: 0.542540
INFO:tensorflow:        + Loss/total_loss: 1.349939
I0109 22:59:20.842776 140110343554496 model_lib_v2.py:1010]     + Loss/total_loss: 1.349939
```



eval_side_by_side_0_0
tag: eval_side_by_side_0_0
step **11,700**

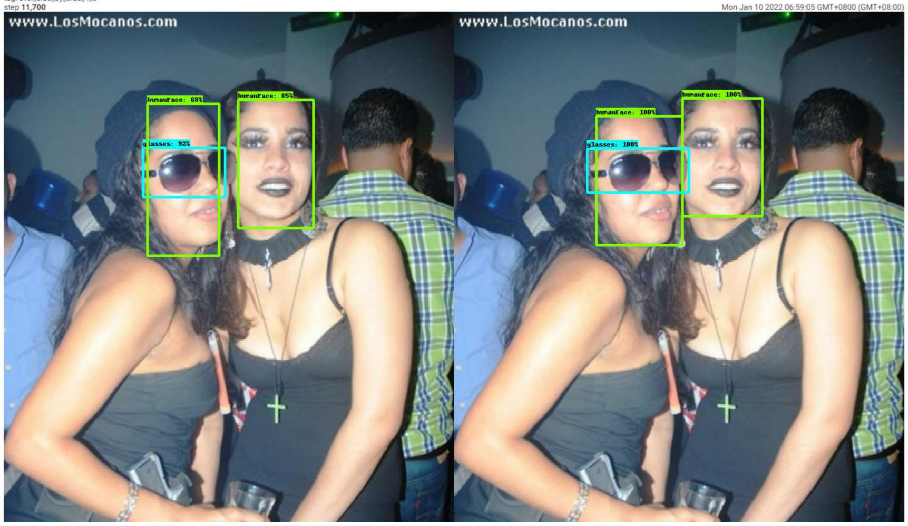
eval

Mon Jan 10 2022 06:59:00 GMT+0800 (GMT+08:00)

eval_side_by_side_4_0
tag: eval_side_by_side_4_0
step **11,700**

eval

Mon Jan 10 2022 06:59:05 GMT+0800 (GMT+08:00)



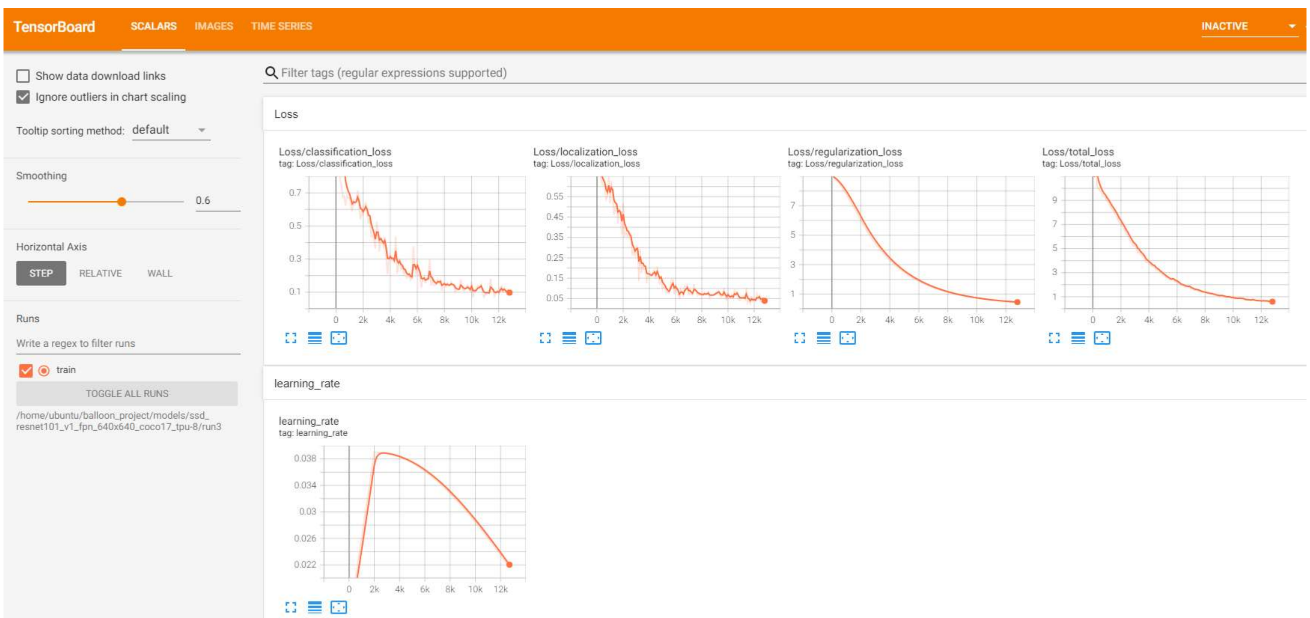
eval_side_by_side_6_0
tag: eval_side_by_side_6_0
step **11,700**

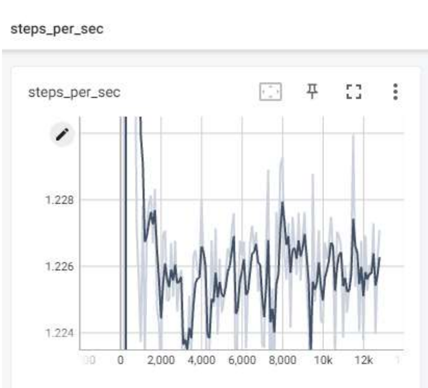
eval

Mon Jan 10 2022 06:59:07 GMT+0800 (GMT+08:00)

Section 3.3: Run 3 Detailed Results

```
I0109 21:35:09.657657 140498521413056 model_lib_v2.py:701] {'Loss/classification_loss': 0.10052558,
 'Loss/localization_loss': 0.030555733,
 'Loss/regularization_loss': 0.4614604,
 'Loss/total_loss': 0.5925417,
 'learning_rate': 0.022155251}
INFO:tensorflow:Step 12600 per-step time 0.817s
I0109 21:36:31.358970 140498521413056 model_lib_v2.py:698] Step 12600 per-step time 0.817s
INFO:tensorflow:{'Loss/classification_loss': 0.08817355,
 'Loss/localization_loss': 0.051073972,
 'Loss/regularization_loss': 0.454397,
 'Loss/total_loss': 0.5936445,
 'learning_rate': 0.02189114}
I0109 21:36:31.359282 140498521413056 model_lib_v2.py:701] {'Loss/classification_loss': 0.08817355,
 'Loss/localization_loss': 0.051073972,
 'Loss/regularization_loss': 0.454397,
 'Loss/total_loss': 0.5936445,
 'learning_rate': 0.02189114}
INFO:tensorflow:Step 12700 per-step time 0.816s
I0109 21:37:52.910320 140498521413056 model_lib_v2.py:698] Step 12700 per-step time 0.816s
INFO:tensorflow:{'Loss/classification_loss': 0.12050601,
 'Loss/localization_loss': 0.051203266,
 'Loss/regularization_loss': 0.44749594,
 'Loss/total_loss': 0.61920524,
 'learning_rate': 0.021626579}
I0109 21:37:52.910622 140498521413056 model_lib_v2.py:701] {'Loss/classification_loss': 0.12050601,
 'Loss/localization_loss': 0.051203266,
 'Loss/regularization_loss': 0.44749594,
 'Loss/total_loss': 0.61920524,
 'learning_rate': 0.021626579}
INFO:tensorflow:Step 12800 per-step time 0.815s
I0109 21:39:14.404221 140498521413056 model_lib_v2.py:698] Step 12800 per-step time 0.815s
INFO:tensorflow:{'Loss/classification_loss': 0.079613246,
 'Loss/localization_loss': 0.02292065,
 'Loss/regularization_loss': 0.4408505,
 'Loss/total_loss': 0.5433844,
 'learning_rate': 0.021361625}
I0109 21:39:14.404499 140498521413056 model_lib_v2.py:701] {'Loss/classification_loss': 0.079613246,
 'Loss/localization_loss': 0.02292065,
 'Loss/regularization_loss': 0.4408505,
 'Loss/total_loss': 0.5433844,
 'learning_rate': 0.021361625}
INFO:tensorflow:Step 12900 per-step time 0.815s
I0109 21:40:35.866988 140498521413056 model_lib_v2.py:698] Step 12900 per-step time 0.815s
INFO:tensorflow:{'Loss/classification_loss': 0.08342803,
 'Loss/localization_loss': 0.03190211,
 'Loss/regularization_loss': 0.43443272,
 'Loss/total_loss': 0.54976285,
 'learning_rate': 0.021096328}
I0109 21:40:35.867253 140498521413056 model_lib_v2.py:701] {'Loss/classification_loss': 0.08342803,
 'Loss/localization_loss': 0.03190211,
 'Loss/regularization_loss': 0.43443272,
 'Loss/total_loss': 0.54976285,
 'learning_rate': 0.021096328}
```

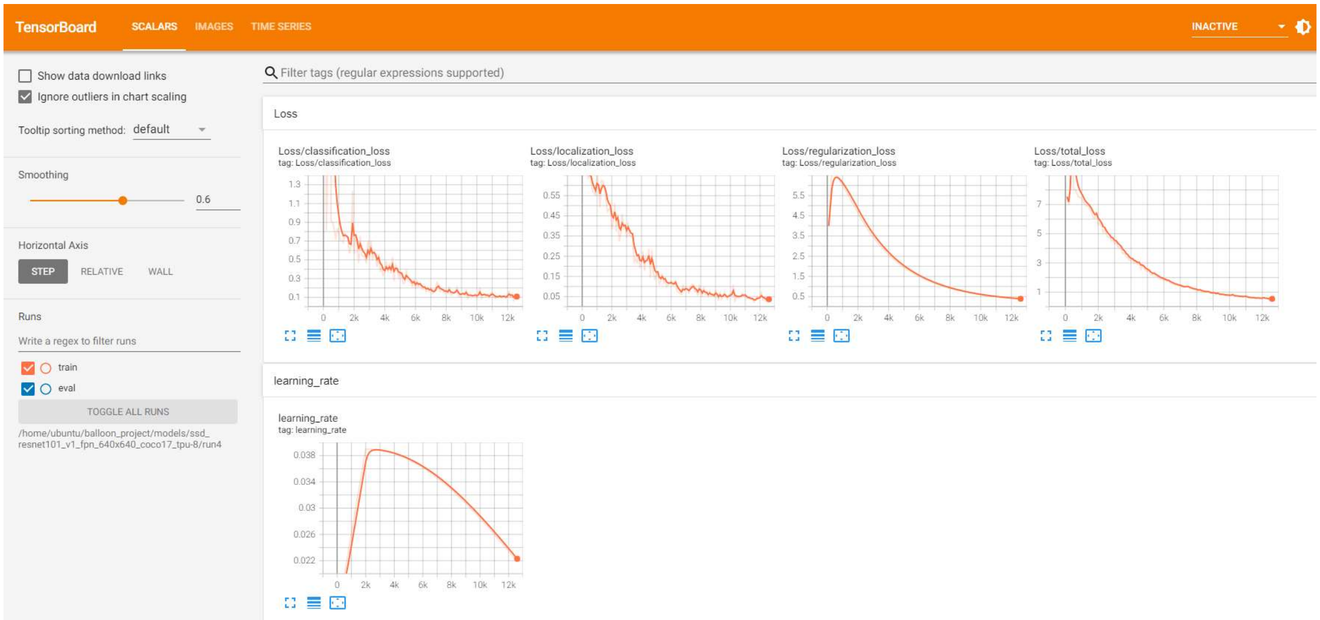


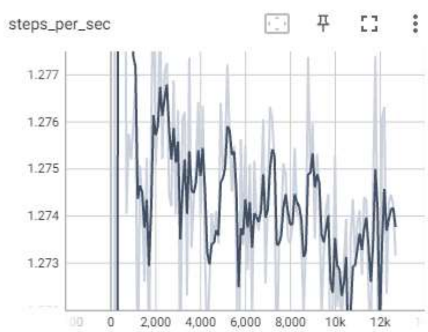


```
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.309
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.604
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.313
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.189
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.323
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.210
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.398
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.443
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.327
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.459
INFO:tensorflow:Eval metrics at step 12900
I0109 22:55:50.112673 139659940021696 model_lib_v2.py:1007] Eval metrics at step 12900
INFO:tensorflow:        + DetectionBoxes_Precision/mAP: 0.308588
I0109 22:55:50.119797 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP: 0.308588
INFO:tensorflow:        + DetectionBoxes_Precision/mAP@.50IOU: 0.604386
I0109 22:55:50.120655 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP@.50IOU: 0.604386
INFO:tensorflow:        + DetectionBoxes_Precision/mAP@.75IOU: 0.312803
I0109 22:55:50.121204 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP@.75IOU: 0.312803
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (small): 0.000000
I0109 22:55:50.121691 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (small): 0.000000
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (medium): 0.188822
I0109 22:55:50.122173 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (medium): 0.188822
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (large): 0.323309
I0109 22:55:50.122655 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (large): 0.323309
INFO:tensorflow:        + DetectionBoxes_Recall/AR@1: 0.210088
I0109 22:55:50.123156 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@1: 0.210088
INFO:tensorflow:        + DetectionBoxes_Recall/AR@10: 0.398421
I0109 22:55:50.123669 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@10: 0.398421
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100: 0.442719
I0109 22:55:50.124159 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100: 0.442719
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (small): 0.000000
I0109 22:55:50.124633 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (small): 0.000000
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (medium): 0.327206
I0109 22:55:50.125130 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (medium): 0.327206
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (large): 0.459369
I0109 22:55:50.125670 139659940021696 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (large): 0.459369
INFO:tensorflow:        + Loss/localization_loss: 0.308534
I0109 22:55:50.126093 139659940021696 model_lib_v2.py:1010]     + Loss/localization_loss: 0.308534
INFO:tensorflow:        + Loss/classification_loss: 0.566454
I0109 22:55:50.126524 139659940021696 model_lib_v2.py:1010]     + Loss/classification_loss: 0.566454
INFO:tensorflow:        + Loss/regularization_loss: 0.434367
I0109 22:55:50.126949 139659940021696 model_lib_v2.py:1010]     + Loss/regularization_loss: 0.434367
INFO:tensorflow:        + Loss/total_loss: 1.309355
I0109 22:55:50.127390 139659940021696 model_lib_v2.py:1010]     + Loss/total_loss: 1.309355
```

Section 3.4: Run 4 Detailed Results

```
I0110 02:06:22.563429 140644024722880 model_lib_v2.py:701] {'Loss/classification_loss': 0.12946908,
 'Loss/localization_loss': 0.040291637,
 'Loss/regularization_loss': 0.401158,
 'Loss/total_loss': 0.57091874,
 'learning_rate': 0.022681937}
INFO:tensorflow:Step 12400 per-step time 0.785s
I0110 02:07:41.039009 140644024722880 model_lib_v2.py:698] Step 12400 per-step time 0.785s
INFO:tensorflow:{'Loss/classification_loss': 0.06427159,
 'Loss/localization_loss': 0.028724974,
 'Loss/regularization_loss': 0.39531687,
 'Loss/total_loss': 0.48831344,
 'learning_rate': 0.022418866}
I0110 02:07:41.039325 140644024722880 model_lib_v2.py:701] {'Loss/classification_loss': 0.06427159,
 'Loss/localization_loss': 0.028724974,
 'Loss/regularization_loss': 0.39531687,
 'Loss/total_loss': 0.48831344,
 'learning_rate': 0.022418866}
INFO:tensorflow:Step 12500 per-step time 0.785s
I0110 02:08:59.502703 140644024722880 model_lib_v2.py:698] Step 12500 per-step time 0.785s
INFO:tensorflow:{'Loss/classification_loss': 0.099045835,
 'Loss/localization_loss': 0.029932687,
 'Loss/regularization_loss': 0.38945168,
 'Loss/total_loss': 0.51843023,
 'learning_rate': 0.022155251}
I0110 02:08:59.502973 140644024722880 model_lib_v2.py:701] {'Loss/classification_loss': 0.099045835,
 'Loss/localization_loss': 0.029932687,
 'Loss/regularization_loss': 0.38945168,
 'Loss/total_loss': 0.51843023,
 'learning_rate': 0.022155251}
INFO:tensorflow:Step 12600 per-step time 0.785s
I0110 02:10:17.982205 140644024722880 model_lib_v2.py:698] Step 12600 per-step time 0.785s
INFO:tensorflow:{'Loss/classification_loss': 0.120541014,
 'Loss/localization_loss': 0.0406868,
 'Loss/regularization_loss': 0.3838333,
 'Loss/total_loss': 0.5450611,
 'learning_rate': 0.02189114}
I0110 02:10:17.982475 140644024722880 model_lib_v2.py:701] {'Loss/classification_loss': 0.120541014,
 'Loss/localization_loss': 0.0406868,
 'Loss/regularization_loss': 0.3838333,
 'Loss/total_loss': 0.5450611,
 'learning_rate': 0.02189114}
INFO:tensorflow:Step 12700 per-step time 0.785s
I0110 02:11:36.527723 140644024722880 model_lib_v2.py:698] Step 12700 per-step time 0.785s
INFO:tensorflow:{'Loss/classification_loss': 0.08524129,
 'Loss/localization_loss': 0.041681,
 'Loss/regularization_loss': 0.37828946,
 'Loss/total_loss': 0.5052117,
 'learning_rate': 0.021626579}
I0110 02:11:36.527987 140644024722880 model_lib_v2.py:701] {'Loss/classification_loss': 0.08524129,
 'Loss/localization_loss': 0.041681,
 'Loss/regularization_loss': 0.37828946,
 'Loss/total_loss': 0.5052117,
 'learning_rate': 0.021626579}
```



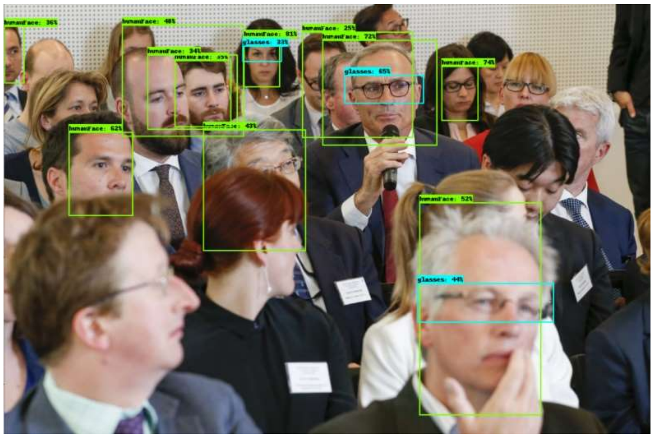


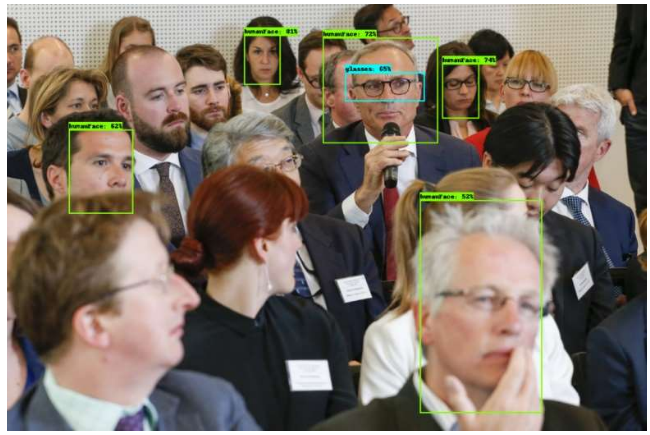

```
DONE (t=0.02s).
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.351
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.628
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.354
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.199
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.410
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.232
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.433
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.478
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = 0.383
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.514
INFO:tensorflow:Eval metrics at step 12700
I0110 15:49:34.923910 139862456309184 model_lib_v2.py:1007] Eval metrics at step 12700
INFO:tensorflow:        + DetectionBoxes_Precision/mAP: 0.351333
I0110 15:49:34.930751 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP: 0.351333
INFO:tensorflow:        + DetectionBoxes_Precision/mAP@.50IOU: 0.627843
I0110 15:49:34.931349 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP@.50IOU: 0.627843
INFO:tensorflow:        + DetectionBoxes_Precision/mAP@.75IOU: 0.354309
I0110 15:49:34.931893 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP@.75IOU: 0.354309
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (small): 0.000000
I0110 15:49:34.932399 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (small): 0.000000
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (medium): 0.198564
I0110 15:49:34.932919 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (medium): 0.198564
INFO:tensorflow:        + DetectionBoxes_Precision/mAP (large): 0.410086
I0110 15:49:34.933418 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Precision/mAP (large): 0.410086
INFO:tensorflow:        + DetectionBoxes_Recall/AR@1: 0.232368
I0110 15:49:34.933918 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@1: 0.232368
INFO:tensorflow:        + DetectionBoxes_Recall/AR@10: 0.432544
I0110 15:49:34.934418 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@10: 0.432544
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100: 0.478246
I0110 15:49:34.934917 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100: 0.478246
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (small): 0.000000
I0110 15:49:34.935395 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (small): 0.000000
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (medium): 0.382721
I0110 15:49:34.935893 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (medium): 0.382721
INFO:tensorflow:        + DetectionBoxes_Recall/AR@100 (large): 0.514471
I0110 15:49:34.936424 139862456309184 model_lib_v2.py:1010]     + DetectionBoxes_Recall/AR@100 (large): 0.514471
INFO:tensorflow:        + Loss/localization_loss: 0.254354
I0110 15:49:34.936853 139862456309184 model_lib_v2.py:1010]     + Loss/localization_loss: 0.254354
INFO:tensorflow:        + Loss/classification_loss: 0.640374
I0110 15:49:34.937306 139862456309184 model_lib_v2.py:1010]     + Loss/classification_loss: 0.640374
INFO:tensorflow:        + Loss/regularization_loss: 0.378237
I0110 15:49:34.937743 139862456309184 model_lib_v2.py:1010]     + Loss/regularization_loss: 0.378237
INFO:tensorflow:        + Loss/total_loss: 1.272964
I0110 15:49:34.938177 139862456309184 model_lib_v2.py:1010]     + Loss/total_loss: 1.272964
```

Section 4: Test Image and Test Video

Using the trained Run 2 model that had the best performance metric, object detection for the two classes was done on an out-of-sample test image and test video.

Below shows the results from the object detection model on the test image, comparing a minimum threshold of 0.25 versus 0.50.





Below shows the a few frames of results from the object detection model on the test video. The predicted bounding boxes do not appear consistently throughout the video, where at quite a few intervals there is no output detection. We also note that the pretrained model was trained on the COCO dataset, where one of the objects is “Person”. Our “Human face” object in could be said to have an overlap with the COCO “Person” object the model was pretrained on. In contrast, “Glasses” is a completely new type of object. This could explain why the object detection model is better able to detect “Human face” over “Glasses”.

In addition, we note that 80 train images and 20 validation images was used was for the experimental runs. Lifting time constraints and increasing the number of annotated train and validation images will allow the model to “see” and be trained on many more examples of our selected two object types, and make the model more robust, possibly leading to an increase in the mAP@IoU=0.50 performance metric