Introduction:

The Backbone architecture the model was pre-trained on is R-CNN with a ResNet50 architecture pre-trained on a coco dataset for the task of instance segmentation and object detection and runs it on an image. COCO is a dataset of more than 300k images annotated with bounding boxes and instance masks of 80 object categories. This applied for both sample prediction and pose estimation while using a keypoint detection in pose estimation.

The sample prediction: image(1):

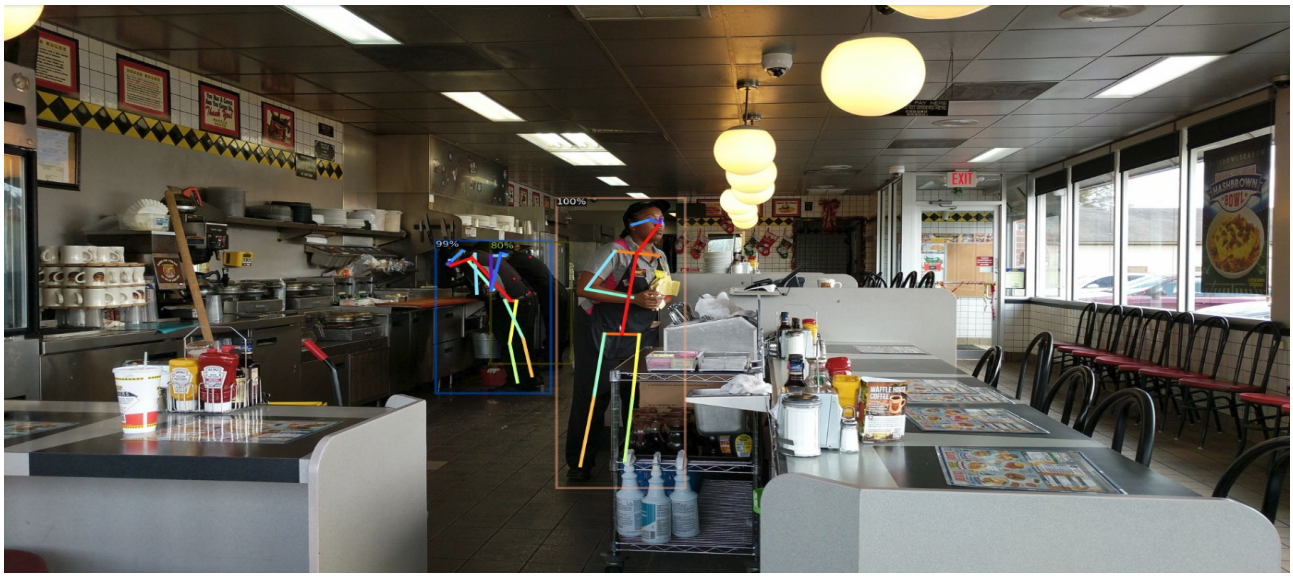
Observations:

Giving the original image and its prediction we can notice there are correct and false predictions. According to the Visualizer that has drawn the predictions on the image, i observe that the model has figured the humans' shapes correctly with accuracy in the range between (60-100%) while missed captioning the other objects within the image such as the street, grass, water and the sky. Additionally, Some part of human 2 &3 in the middle have left without classification.

Error modes:

The model error is that it does misclassify some part of the humans captioning where person figures overlapped as the person in purple has a bad prediction across the other four individuals regarding his stretched hands. The second person's two hands have been mistakenly captioned in figure one and figure 5 in purple.

Pose estimation: Image (2):



Observations:

Using Inference with a keypoint detection and panoptic segmentation model, the model predicted the humans' pose correctly while not having great pose estimation for the far individuls. The model detects the car, oven, cabinet, wall and windows properly while left with some chairs and lights unlabeled. Few objects have not been detected neither labelled.

Error modes:

Using Inference with a keypoint detection and panoptic segmentation model, the model predicted the humans' pose correctly while not having great pose estimation for the far individuals. The model detects the car, oven, cabinet, wall and windows properly while left with some chairs and lights unlabeled.