Detectron2

Reference:
Modern Computer Vision with PyTorch
https://www.packtpub.com/product/modern-computer-vision-with-pytorch/9781839213472
https://www.youtube.com/watch?v=eUSgtfK4ivk

Detectron2 is developed by Facebook AI Research and provide detection and segmentation algorithms.

Example image

Implementation

Use model *COCO-Detection/faster_rcnn_X_101_32x8d_FPN_3x.yaml* to do faster rcnn on the example image.

```
model_path = "COCO-Detection/faster_rcnn_X_101_32x8d_FPN_3x.yaml"

cfg = get_cfg()
cfg.merge_from_file(model_zoo.get_config_file(model_path))
cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST = 0.5
cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url(model_path)
predictor = DefaultPredictor(cfg)
```

Get a configure, load the model and specify the parameter.
Do the prediction and observe the model output.

```
outputs = predictor(im)

print(outputs["instances"].pred_classes[0].item())
print(outputs["instances"].pred_boxes[0])
print(MetadataCatalog.get(cfg.DATASETS.TRAIN[0]).thing_classes[outputs["instances"].pred_classes[0].item()])
```

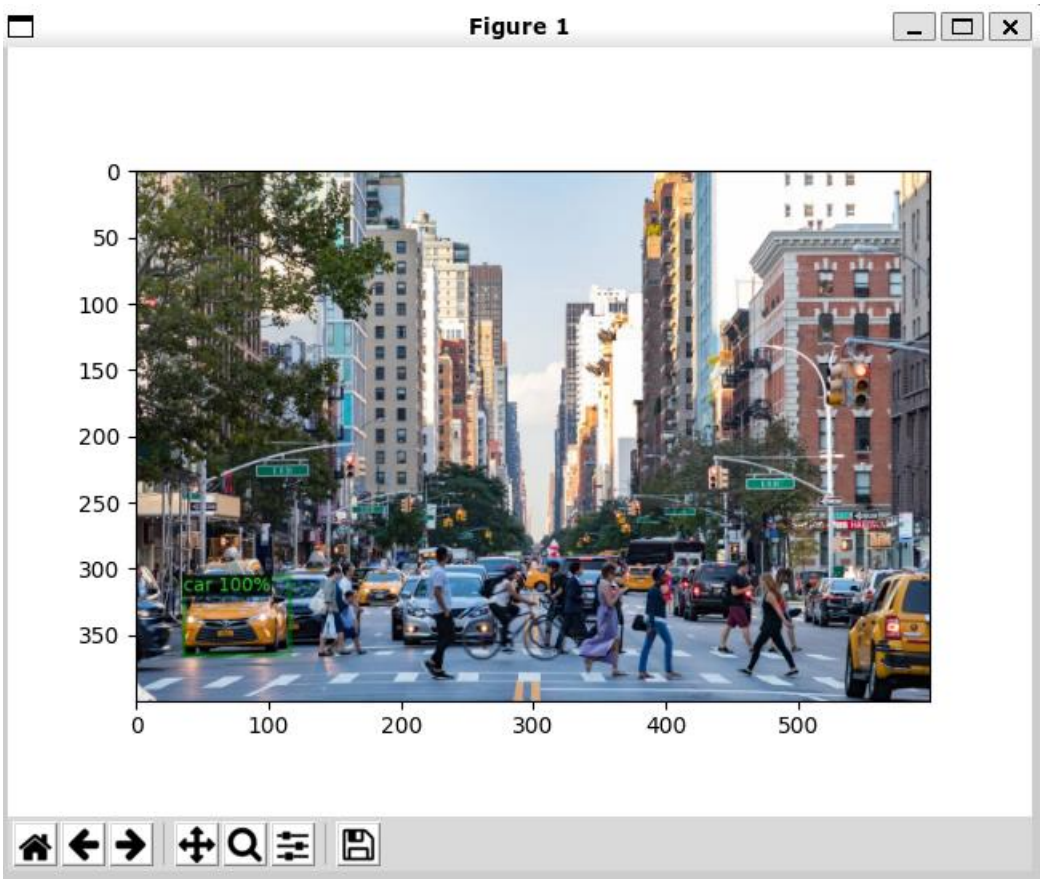
```
2
Boxes(tensor([[ 35.9478, 305.7766, 116.1322, 365.4618]], device='cuda:0'))
car
```

The first line print out the label ID of the first item that being detected; the second line print out the bounding box coordinate ([top-left.x, top-left.y, bottom-right.x, bottom-right.y]). Finally, the label name of the object.

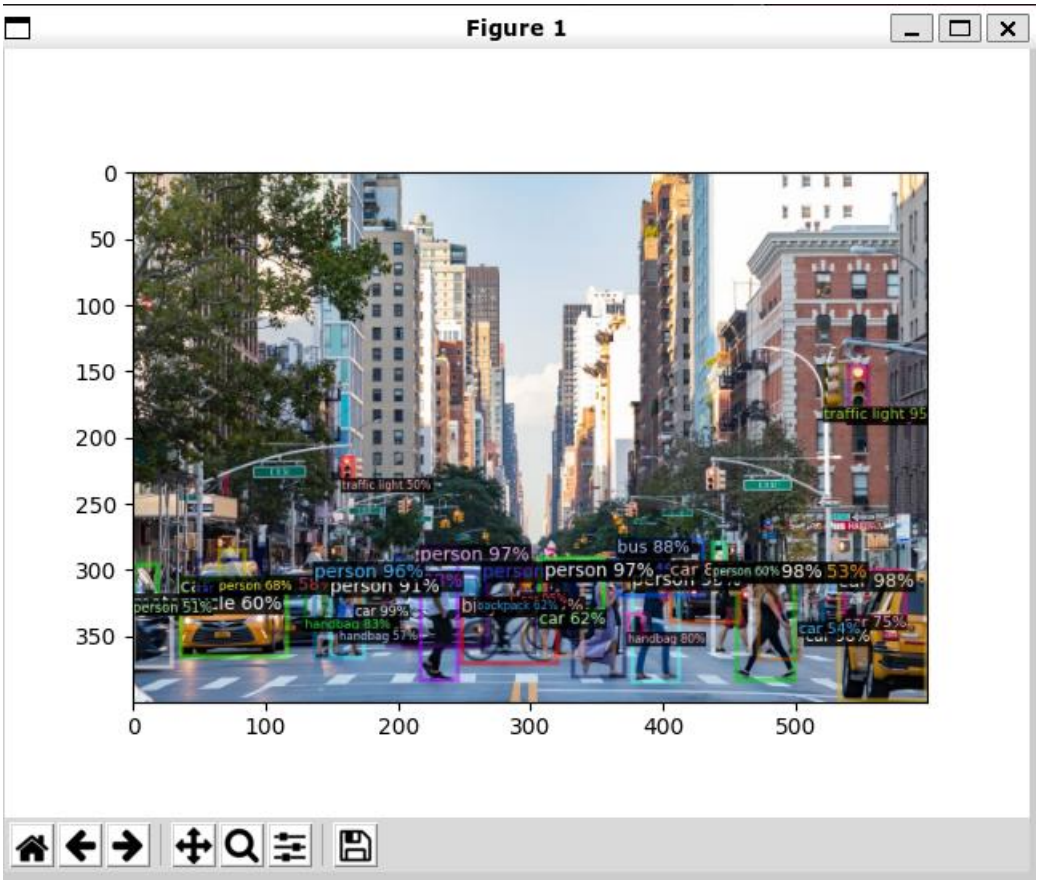
Draw the bounding box and the detected label of the first detected item on to the picture.

```
v = Visualizer(im[:, :, :],
               MetadataCatalog.get(cfg.DATASETS.TRAIN[0]),
               )

out = v.draw_instance_predictions(outputs["instances"][0].to("cpu"))
plt.imshow(out.get_image())
plt.show()
```



By pass all the output instances into to visualizer, we can print out all the detected objects and its bounding boxes.



If we want to do human pose detection rather than object detection, we can simply change the model path and the rest of the code need not be changed.

Use model *COCO-Keypoints/keypoint_rcnn_R_50_FPN_3x.yaml* to do human pose detection.