

▼ Detectron2 Beginner's Tutorial



Welcome to detectron2! This is the official colab tutorial of detectron2. Here, we will go through some basics usage of detectron2, including the following:

- Run inference on images or videos, with an existing detectron2 model
- Train a detectron2 model on a new dataset.

```

1 !python -m pip install pyyaml
2 import sys, os, distutils.core
3 # Note: This is a faster way to install detectron2 in Colab, but it does not include all functionalities (e.g. compiled operators).
4 # See https://detectron2.readthedocs.io/tutorials/install.html for full installation instructions
5 !git clone 'https://github.com/facebookresearch/detectron2'
6 dist = distutils.core.run_setup("./detectron2/setup.py")
7 !python -m pip install {''.join(['f"'{}"' for x in dist.install_requires])}
8 sys.path.insert(0, os.path.abspath('./detectron2'))

Requirement already satisfied: pyyaml in /usr/local/lib/python3.10/dist-packages (6.0.1)
fatal: destination path 'detectron2' already exists and is not an empty directory.
Requirement already satisfied: Pillow>=7.1 in /usr/local/lib/python3.10/dist-packages (9.4.0)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (3.7.1)
Requirement already satisfied: pycocotools>=2.0.2 in /usr/local/lib/python3.10/dist-packages (2.0.7)
Requirement already satisfied: termcolor>=1.1 in /usr/local/lib/python3.10/dist-packages (2.3.0)
Requirement already satisfied: yacs>=0.1.8 in /usr/local/lib/python3.10/dist-packages (0.1.8)
Requirement already satisfied: tabulate in /usr/local/lib/python3.10/dist-packages (0.9.0)
Requirement already satisfied: cloudpickle in /usr/local/lib/python3.10/dist-packages (2.2.1)
Requirement already satisfied: tqdm>4.29.0 in /usr/local/lib/python3.10/dist-packages (4.66.1)
Requirement already satisfied: tensorboard in /usr/local/lib/python3.10/dist-packages (2.14.1)
Requirement already satisfied: fvcore<0.1.6,>=0.1.5 in /usr/local/lib/python3.10/dist-packages (0.1.5.post20221221)
Requirement already satisfied: iopath<0.1.10,>=0.1.7 in /usr/local/lib/python3.10/dist-packages (0.1.9)
Requirement already satisfied: omegaconf<2.4,>=2.1 in /usr/local/lib/python3.10/dist-packages (2.3.0)
Requirement already satisfied: hydra-core>=1.1 in /usr/local/lib/python3.10/dist-packages (1.3.2)
Requirement already satisfied: black in /usr/local/lib/python3.10/dist-packages (23.11.0)
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (23.2)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.2.0)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (4.44.0)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.4.5)
Requirement already satisfied: numpy>=1.20 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (1.23.5)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (3.1.1)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib) (2.8.2)
Requirement already satisfied: PyYAML in /usr/local/lib/python3.10/dist-packages (from yacs>=0.1.8) (6.0.1)
Requirement already satisfied: absl-py>=0.4 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (1.4.0)
Requirement already satisfied: grpcio>=1.48.2 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (1.59.2)
Requirement already satisfied: google-auth<3,>=1.6.3 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (2.17.3)
Requirement already satisfied: google-auth-oauthlib<1.1,>=0.5 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (1.0.0)
Requirement already satisfied: markdown>=2.6.8 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (3.5.1)
Requirement already satisfied: protobuf>=3.19.6 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (3.20.3)
Requirement already satisfied: requests<3,>=2.21.0 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (2.31.0)
Requirement already satisfied: setuptools>=41.0.0 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (67.7.2)
Requirement already satisfied: six>1.9 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (1.16.0)
Requirement already satisfied: tensorboard-data-server<0.8.0,>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (0.7.)
Requirement already satisfied: werkzeug>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from tensorboard) (3.0.1)
Requirement already satisfied: portalocker in /usr/local/lib/python3.10/dist-packages (from iopath<0.1.10,>=0.1.7) (2.8.2)
Requirement already satisfied: antlr4-python3-runtime==4.9.* in /usr/local/lib/python3.10/dist-packages (from omegaconf<2.4,>=2.1) (4.9.)
Requirement already satisfied: click>=8.0.0 in /usr/local/lib/python3.10/dist-packages (from black) (8.1.7)
Requirement already satisfied: mpy-py-extensions>=0.4.3 in /usr/local/lib/python3.10/dist-packages (from black) (1.0.0)
Requirement already satisfied: pathspec>=0.9.0 in /usr/local/lib/python3.10/dist-packages (from black) (0.11.2)
Requirement already satisfied: platformdirs>=2 in /usr/local/lib/python3.10/dist-packages (from black) (3.11.0)
Requirement already satisfied: tomli>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from black) (2.0.1)
Requirement already satisfied: typing-extensions>=4.0.1 in /usr/local/lib/python3.10/dist-packages (from black) (4.5.0)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3->tensorboar)
Requirement already satisfied: pyasn1-modules>=0.2.1 in /usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3->tensorboard)
Requirement already satisfied: rsa<5,>=3.1.4 in /usr/local/lib/python3.10/dist-packages (from google-auth<3,>=1.6.3->tensorboard) (4.9)
Requirement already satisfied: requests-oauthlib>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from google-auth-oauthlib<1.1,>=0.5-
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0->tensorboar
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0->tensorboard) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0->tensorboard) (2.
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests<3,>=2.21.0->tensorboard) (28
Requirement already satisfied: MarkupSafe>=2.1.1 in /usr/local/lib/python3.10/dist-packages (from werkzeug>=1.0.1->tensorboard) (2.1.3)
Requirement already satisfied: pyasn1<6.0.0,>=0.4.6 in /usr/local/lib/python3.10/dist-packages (from pyasn1-modules>=0.2.1->google-auth<
Requirement already satisfied: oauthlib>=3.0.0 in /usr/local/lib/python3.10/dist-packages (from requests-oauthlib>=0.7.0->google-auth-oa

```

```

1 import torch, detectron2
2 !nvcc --version
3 TORCH_VERSION = ".".join(torch.__version__.split(".")[:2])
4 CUDA_VERSION = torch.__version__.split("+")[-1]
5 print("torch: ", TORCH_VERSION, "; cuda: ", CUDA_VERSION)
6 print("detectron2:", detectron2.__version__)

nvcc: NVIDIA (R) Cuda compiler driver
Copyright (c) 2005-2022 NVIDIA Corporation
Built on Wed_Sep_21_10:33:58_PDT_2022
Cuda compilation tools, release 11.8, V11.8.89
Build cuda_11.8.r11.8/compiler.31833905_0
torch: 2.1 ; cuda: cu118
detectron2: 0.6

1 # Some basic setup:
2 # Setup detectron2 logger
3 import detectron2
4 from detectron2.utils.logger import setup_logger
5 setup_logger()
6
7 # import some common libraries
8 import numpy as np
9 import os, json, cv2, random
10 from google.colab.patches import cv2_imshow
11
12 # import some common detectron2 utilities
13 from detectron2 import model_zoo
14 from detectron2.engine import DefaultPredictor
15 from detectron2.config import get_cfg
16 from detectron2.utils.visualizer import Visualizer
17 from detectron2.data import MetadataCatalog, DatasetCatalog

```

▼ Run a pretrained Detectron2 model

We first download some image from the given URLs:

```

1 !wget http://images.cocodataset.org/val2017/00000007574.jpg -q -O input.jpg
2 im_input = cv2.imread("./input.jpg")
3 cv2_imshow(im_input)

```



```

1 !wget http://images.cocodataset.org/val2017/000000013923.jpg -q -O test1.jpg
2 im_test1 = cv2.imread("./test1.jpg")
3 cv2_imshow(im_test1)

```



```
1 !wget http://images.cocodataset.org/val2017/00000018380.jpg -q -O test2.jpg
2 im_test2 = cv2.imread("./test2.jpg")
3 cv2_imshow(im_test2)
```



We can see there are multiple objects in these images: bottles, tables, chairs, people, etc. Let us see if we can detect them all by using a pre-trained model given by Detectron2.

Let's take a look at the model output.

In inference mode, the builtin model outputs a `list[dict]`, one dict for each image. For the object detection task, the dict contain the following fields:

- "instances": Instances object with the following fields:
 - "pred_boxes": Storing N boxes, one for each detected instance.
 - "scores": a vector of N scores.
 - "pred_classes": a vector of N labels in range [0, num_categories].

For more details, please see <https://detectron2.readthedocs.io/tutorials/models.html#model-output-format> for specification

```
1 cfg = get_cfg()
2 cfg.merge_from_file(model_zoo.get_config_file("COCO-Detection/faster_rcnn_R_50_FPN_1x.yaml"))
3 cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST= 0.5 # set threshold for this model
4 cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url("COCO-Detection/faster_rcnn_R_50_FPN_1x.yaml")
5 predictor = DefaultPredictor(cfg)
6 outputs = predictor(im_input)
```

```
[11/11 05:49:59 d2.checkpoint.detection_checkpoint]: [DetectionCheckpointer] Loading from https://dl.fbaipublicfiles.com/detectron2/COCO/
```

```
1 print(outputs)

{'instances': Instances(num_instances=17, image_height=480, image_width=640, fields=[pred_boxes: Boxes(tensor([[ 0.6764, 158.7422, 105.480.6438, 298.5876, 509.7858, 378.3471], [293.3246, 287.8804, 344.7386, 388.2297], [501.2999, 306.3395, 533.4362, 387.4037], [300.8816, 148.8771, 366.4546, 227.1448], [104.8806, 73.4484, 148.7521, 88.7021], [205.1707, 55.5022, 231.9424, 93.1075], [357.5101, 394.1846, 393.9808, 425.1449], [38.0390, 355.4640, 315.9145, 442.3582], [357.1301, 286.4151, 387.7237, 365.4046], [197.1212, 267.4070, 208.2965, 292.6432], [241.4115, 57.8996, 267.1059, 93.9473], [454.1900, 310.0651, 484.8028, 360.4231], [243.9083, 382.6956, 275.9801, 407.7662], [185.0177, 355.2176, 304.0785, 409.8493], [441.3940, 307.2240, 470.4242, 347.5514], [330.0779, 267.5388, 355.4027, 303.1052]], device='cuda:0')), scores: tensor([0.9938, 0.9790, 0.9769, 0.9570, 0.9114, 0.8425, 0.6727, 0.6304, 0.6266, 0.6026, 0.6025, 0.5396, 0.5248, 0.5182], device='cuda:0'), pred_classes: tensor([72, 39, 39, 39, 68, 45, 75, 55, 71, 41, 39, 75, 41, 45, 71, 41, 39], device='cuda:0'))}
```

```
1 print(outputs["instances"].pred_classes)
2 print(outputs["instances"].pred_boxes)

tensor([72, 39, 39, 39, 68, 45, 75, 55, 71, 41, 39, 75, 41, 45, 71, 41, 39], device='cuda:0')
Boxes(tensor([[ 0.6764, 158.7422, 105.0817, 397.2280], [480.6438, 298.5876, 509.7858, 378.3471], [293.3246, 287.8804, 344.7386, 388.2297], [501.2999, 306.3395, 533.4362, 387.4037], [300.8816, 148.8771, 366.4546, 227.1448], [104.8806, 73.4484, 148.7521, 88.7021], [205.1707, 55.5022, 231.9424, 93.1075], [357.5101, 394.1846, 393.9808, 425.1449], [38.0390, 355.4640, 315.9145, 442.3582], [357.1301, 286.4151, 387.7237, 365.4046], [197.1212, 267.4070, 208.2965, 292.6432], [241.4115, 57.8996, 267.1059, 93.9473], [454.1900, 310.0651, 484.8028, 360.4231], [243.9083, 382.6956, 275.9801, 407.7662], [185.0177, 355.2176, 304.0785, 409.8493], [441.3940, 307.2240, 470.4242, 347.5514], [330.0779, 267.5388, 355.4027, 303.1052]], device='cuda:0'))
```

```
1 outputs_q1q2 = {'q1': [], 'q2': []}
2 outputs_q1q2['q1'].append(outputs["instances"])

1 # We can use "Visualizer" to draw the predictions on the image
2 v = Visualizer(im_input[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
3 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
4 cv2_imshow(out.get_image()[:, :, ::-1])
```



AWESOME!!! Great progress so far! We are able to detect sink, microwave, bottle and even refrigerator! At this point, we have used the pre-trained model to do the inference on the given image. There are in total 17 objects are being detected. The image is adopted from the [MS-COCO](#) dataset and there are 81 classes including person, bicycle, car, etc. You may find the id-category mapping [here](#).

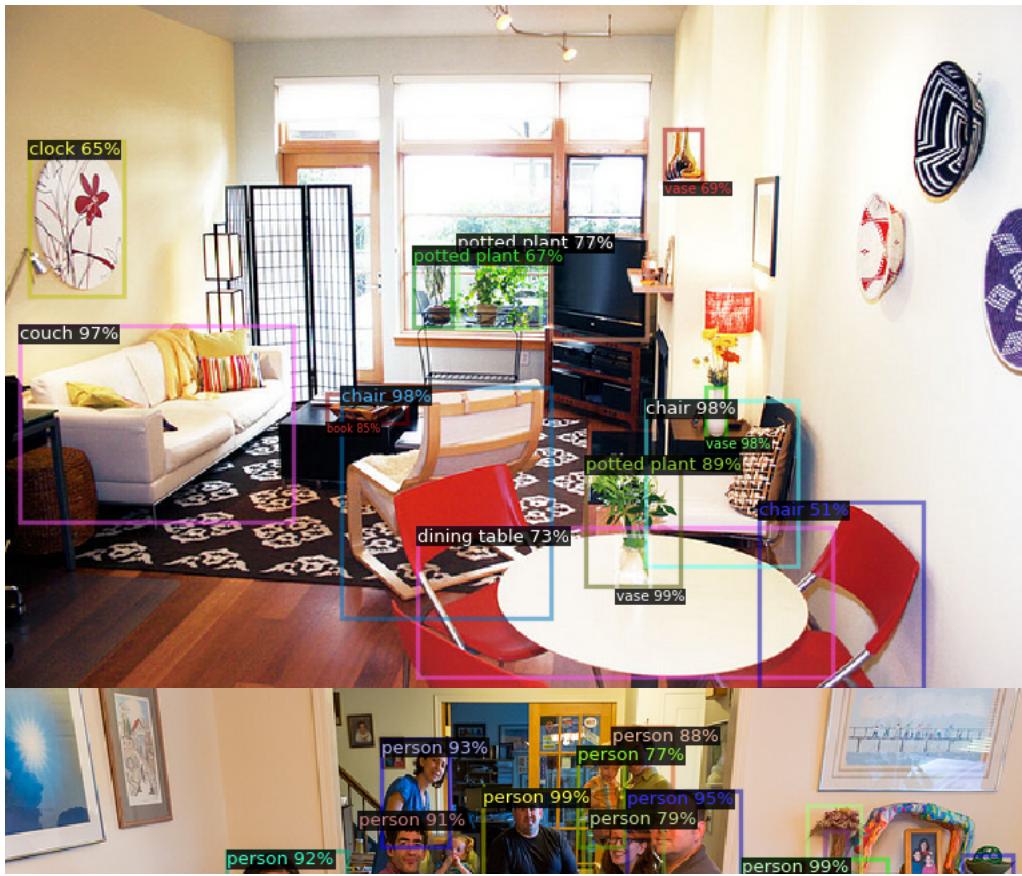
The model we just used is `COCO-Detection/faster_rcnn_R_50_FPN_1x.yaml`. Actually, the Detectron2 provides us more than that, you may find great amounts of models for different tasks in the given [MODEL ZOO](#). What about we try a different model to see what its output will look like?

- Q1 (5%): Object Detection. Use the same configuration `COCO-Detection/faster_rcnn_R_50_FPN_1x.yaml`, with IoU threshold of 0.5 (`SCORE_THRESH_TEST=0.5`), to also run inference on the rest two images (`test1.jpg & test2.jpg`) and view the outputs with bounding boxes.
- Q2: Object Detection. Use the `COCO-Detection/faster_rcnn_R_101_FPN_3X.yaml`, which has a ResNet-101 as the backbone, with IoU threshold of 0.5 and view the outputs of all three images with bounding boxes. By looking at the outputs, can you find the difference with the one `COCO-Detection/faster_rcnn_R_50_FPN_1x.yaml` we used in Q1? (e.g., numbers of objects, confidence scores, ...)
- Q3: Object Detection. Use the `COCO-Detection/faster_rcnn_R_101_FPN_3X.yaml` with an IoU threshold of 0.9 and view the outputs of all three images with bounding boxes.
- Q4 (5%): Instance Segmentation. The models we have tried in Q1-Q3 are the Faster R-CNN models for object detection. Here, let's try a Mask R-CNN model `COCO-InstanceSegmentation/mask_rcnn_R_101_FPN_3X.yaml`, with IoU threshold of 0.5, to perform the instance segmentation and view the outputs of all three images with segmentation masks. Compare the difference of outputs between an object detection model with an instance segmentation model.

```

1 ## Q1
2
3 outputs = predictor(im_test1)
4 outputs_q1q2 = {'q1': [], 'q2': []}
5 outputs_q1q2['q1'].append(outputs["instances"])
6 # We can use "Visualizer" to draw the predictions on the image
7 v = Visualizer(im_test1[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
8 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
9 cv2_imshow(out.get_image()[:, :, ::-1])
10 outputs = predictor(im_test2)
11 outputs_q1q2 = {'q1': [], 'q2': []}
12 outputs_q1q2['q1'].append(outputs["instances"])
13 # We can use "Visualizer" to draw the predictions on the image
14 v = Visualizer(im_test2[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
15 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
16 cv2_imshow(out.get_image()[:, :, ::-1])

```



```

1 # todo: Q2
2 cfg = get_cfg()
3 cfg.merge_from_file(model_zoo.get_config_file("COCO-Detection/faster_rcnn_R_101_FPN_3x.yaml"))
4 cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST= 0.5 # set threshold for this model
5 cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url("COCO-Detection/faster_rcnn_R_101_FPN_3x.yaml")
6 predictor = DefaultPredictor(cfg)
7 outputs = predictor(im_test2)
8 outputs_q1q2 = {'q1': [], 'q2': []}
9 outputs_q1q2['q1'].append(outputs["instances"])
10 # We can use "Visualizer" to draw the predictions on the image
11 v = Visualizer(im_test2[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
12 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
13 cv2_imshow(out.get_image()[:, :, ::-1])
14 outputs = predictor(im_test1)
15 outputs_q1q2 = {'q1': [], 'q2': []}
16 outputs_q1q2['q1'].append(outputs["instances"])
17 # We can use "Visualizer" to draw the predictions on the image
18 v = Visualizer(im_test1[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
19 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
20 cv2_imshow(out.get_image()[:, :, ::-1])
21 outputs = predictor(im_input)
22 outputs_q1q2 = {'q1': [], 'q2': []}
23 outputs_q1q2['q1'].append(outputs["instances"])
24 # We can use "Visualizer" to draw the predictions on the image
25 v = Visualizer(im_input[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
26 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
27 cv2_imshow(out.get_image()[:, :, ::-1])

```

[11/11 05:50:08 d2.checkpoint.detection_checkpoint]: [DetectionCheckpoint] Loading from <https://dl.fbaipublicfiles.com/detectron2/COCC>

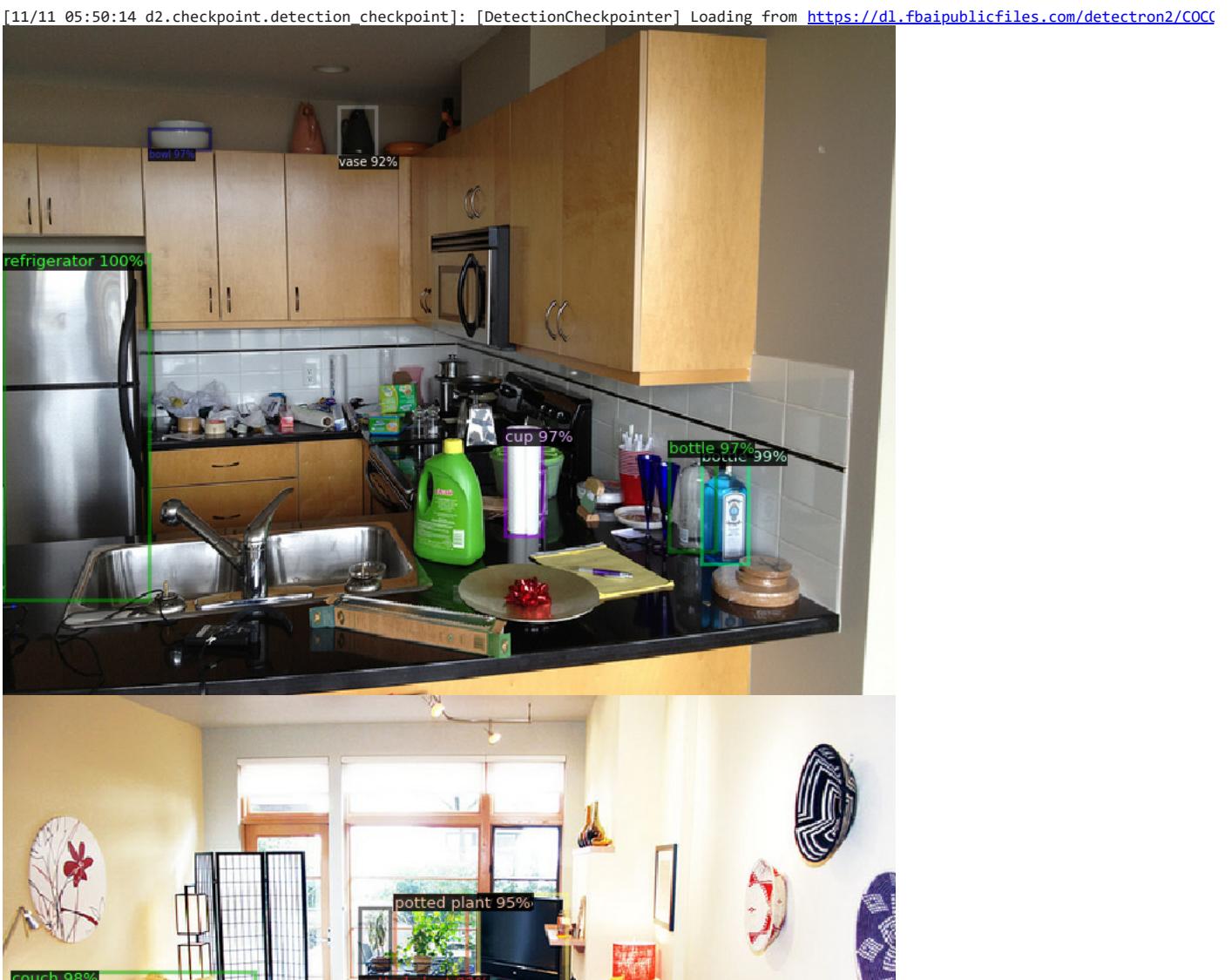


In this new 101 iteration of CNN, the bounding boxes are more "tightly packed". In addition, some of the mistakes previously made aren't. For example, if you look at the im_input image, the ribbon was marked cake, but is not anymore. In the test 2 image, a fork is wrongly detected as a

pizza, and is correctly determined as a fork. Such inaccuracies are reduced in this 101 iteration after the 50 iteration.

The confidence scores are all generally higher as well, and the number of boxes, though not by much, have reduced. For example, in the Test 1 image, the value reduced from 13 to 12.

```
1 # todo: Q3
2 cfg = get_cfg()
3 cfg.merge_from_file(model_zoo.get_config_file("COCO-Detection/faster_rcnn_R_101_FPN_3x.yaml"))
4 cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST= 0.9 # set threshold for this model
5 cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url("COCO-Detection/faster_rcnn_R_101_FPN_3x.yaml")
6 predictor = DefaultPredictor(cfg)
7 outputs = predictor(im_input)
8 outputs_q1q2 = {'q1': [], 'q2': []}
9 outputs_q1q2['q1'].append(outputs["instances"])
10 # We can use "Visualizer" to draw the predictions on the image
11 v = Visualizer(im_input[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
12 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
13 cv2_imshow(out.get_image()[:, :, ::-1])
14
15 outputs = predictor(im_test1)
16 outputs_q1q2 = {'q1': [], 'q2': []}
17 outputs_q1q2['q1'].append(outputs["instances"])
18 # We can use "Visualizer" to draw the predictions on the image
19 v = Visualizer(im_test1[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
20 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
21 cv2_imshow(out.get_image()[:, :, ::-1])
22 outputs = predictor(im_test2)
23 outputs_q1q2 = {'q1': [], 'q2': []}
24 outputs_q1q2['q1'].append(outputs["instances"])
25 # We can use "Visualizer" to draw the predictions on the image
26 v = Visualizer(im_test2[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
27 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
28 cv2_imshow(out.get_image()[:, :, ::-1])
```



```

1 # todo: Q4
2 cfg = get_cfg()
3 cfg.merge_from_file(model_zoo.get_config_file("COCO-InstanceSegmentation/mask_rcnn_R_101_FPN_3x.yaml"))
4 cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST= 0.9 # set threshold for this model
5 cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url("COCO-InstanceSegmentation/mask_rcnn_R_101_FPN_3x.yaml")
6 predictor = DefaultPredictor(cfg)
7 outputs = predictor(im_input)
8 outputs_q1q2 = {'q1': [], 'q2': []}
9 outputs_q1q2['q1'].append(outputs["instances"])
10 # We can use "Visualizer" to draw the predictions on the image
11 v = Visualizer(im_input[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
12 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
13 cv2_imshow(out.get_image()[:, :, ::-1])
14
15 outputs = predictor(im_test1)
16 outputs_q1q2 = {'q1': [], 'q2': []}
17 outputs_q1q2['q1'].append(outputs["instances"])
18 # We can use "Visualizer" to draw the predictions on the image
19 v = Visualizer(im_test1[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
20 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
21 cv2_imshow(out.get_image()[:, :, ::-1])
22 outputs = predictor(im_test2)
23 outputs_q1q2 = {'q1': [], 'q2': []}
24 outputs_q1q2['q1'].append(outputs["instances"])
25 # We can use "Visualizer" to draw the predictions on the image
26 v = Visualizer(im_test2[:, :, ::-1], MetadataCatalog.get(cfg.DATASETS.TRAIN[0]), scale=1.2)
27 out = v.draw_instance_predictions(outputs["instances"].to("cpu"))
28 cv2_imshow(out.get_image()[:, :, ::-1])

```



This is different in that there is quite literally a mask over whatever is being detected. In addition, the accuracies are quite high as well, all usually above the 90% mark. All in all, it seems to be a better alternative.

▼ Now let's train on the provided sportsmot dataset

please upload the provided sportsmot dataset first

```
1 from google.colab import drive
2 drive.mount('/content/drive')

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

1 !unzip "/content/drive/MyDrive/sportsmot.zip"
   inflating: sportsmot/JPEGImages/train_000056.jpg
   inflating: sportsmot/JPEGImages/train_000057.jpg
   inflating: sportsmot/JPEGImages/train_000058.jpg
   inflating: sportsmot/JPEGImages/train_000059.jpg
   inflating: sportsmot/JPEGImages/train_000060.jpg
   inflating: sportsmot/JPEGImages/train_000062.jpg
```

```

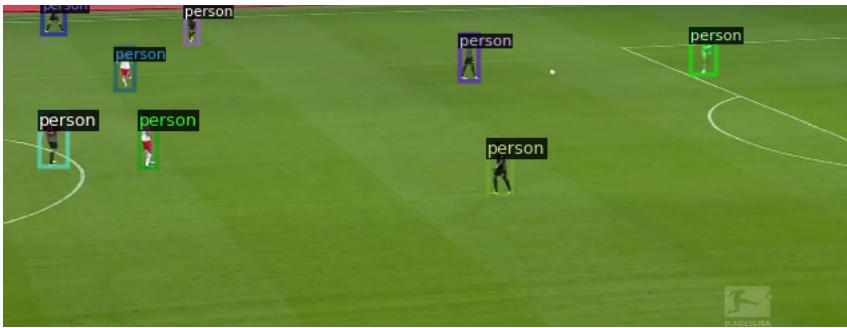
1 from detectron2.structures import BoxMode
2
3 def get_dataset_dicts(data_root, txt_file):
4     dataset_dicts = []
5     filenames = []
6     csv_path = os.path.join(data_root, txt_file)
7     with open(csv_path, "r") as f:
8         for line in f:
9             filenames.append(line.rstrip())
10
11    for idx, filename in enumerate(filenames):
12        record = {}
13
14        image_path = os.path.join(data_root, filename)
15
16        im = cv2.imread(image_path)
17        if im is None:
18            continue
19        height, width = im.shape[:2]
20
21        record['file_name'] = image_path
22        record['image_id'] = idx
23        record['height'] = height
24        record['width'] = width
25
26        image_filename = os.path.basename(filename)
27        image_name = os.path.splitext(image_filename)[0]
28        annotation_path = os.path.join(data_root, 'labels', '{}.txt'.format(image_name))
29        annotation_rows = []
30
31        with open(annotation_path, "r") as f:
32            for line in f:
33                temp = line.rstrip().split(" ")
34                annotation_rows.append(temp)
35
36        objs = []
37        for row in annotation_rows:
38            xcentre = int(float(row[1])*width)
39            ycentre = int(float(row[2])*height)
40            bwidth = int(float(row[3])*width)
41            bheight = int(float(row[4])*height)
42
43            xmin = int(xcentre - bwidth/2)
44            ymin = int(ycentre - bheight/2)
45            xmax = xmin + bwidth
46            ymax = ymin + bheight
47
48            obj= {
49                'bbox': [xmin, ymin, xmax, ymax],
50                'bbox_mode': BoxMode.XXY_ABS,
51                # alternatively, we can use bbox_mode = BoxMode.XYWH_ABS
52                # 'bbox': [xmin, ymin, bwidth, bheight],
53                # 'bbox_mode': BoxMode.XYWH_ABS,
54                'category_id': int(row[0]),
55                'iscrowd': 0
56            }
57
58            objs.append(obj)
59        record['annotations'] = objs
60        dataset_dicts.append(record)
61    return dataset_dicts
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
709
710
711
712
713
714
715
716
717
718
719
719
720
721
722
723
724
725
726
727
728
729
729
730
731
732
733
734
735
736
737
738
739
739
740
741
742
743
744
745
746
747
748
749
749
750
751
752
753
754
755
756
757
758
759
759
760
761
762
763
764
765
766
767
768
769
769
770
771
772
773
774
775
776
777
778
779
779
780
781
782
783
784
785
786
787
787
788
789
789
790
791
792
793
794
795
796
797
797
798
799
799
800
801
802
803
804
805
806
807
808
809
809
810
811
812
813
814
815
816
817
817
818
819
819
820
821
822
823
824
825
826
827
827
828
829
829
830
831
832
833
834
835
836
837
837
838
839
839
840
841
842
843
844
845
846
846
847
848
848
849
849
850
851
852
853
854
855
856
856
857
858
858
859
859
860
861
862
863
864
865
865
866
867
867
868
868
869
869
870
871
872
873
874
875
875
876
877
877
878
878
879
879
880
881
882
883
884
885
885
886
887
887
888
888
889
889
890
891
892
893
893
894
894
895
895
896
896
897
897
898
898
899
899
900
901
901
902
902
903
903
904
904
905
905
906
906
907
907
908
908
909
909
910
910
911
911
912
912
913
913
914
914
915
915
916
916
917
917
918
918
919
919
920
920
921
921
922
922
923
923
924
924
925
925
926
926
927
927
928
928
929
929
930
930
931
931
932
932
933
933
934
934
935
935
936
936
937
937
938
938
939
939
940
940
941
941
942
942
943
943
944
944
945
945
946
946
947
947
948
948
949
949
950
950
951
951
952
952
953
953
954
954
955
955
956
956
957
957
958
958
959
959
960
960
961
961
962
962
963
963
964
964
965
965
966
966
967
967
968
968
969
969
970
970
971
971
972
972
973
973
974
974
975
975
976
976
977
977
978
978
979
979
980
980
981
981
982
982
983
983
984
984
985
985
986
986
987
987
988
988
989
989
990
990
991
991
992
992
993
993
994
994
995
995
996
996
997
997
998
998
999
999
1000
1000
1001
1001
1002
1002
1003
1003
1004
1004
1005
1005
1006
1006
1007
1007
1008
1008
1009
1009
1010
1010
1011
1011
1012
1012
1013
1013
1014
1014
1015
1015
1016
1016
1017
1017
1018
1018
1019
1019
1020
1020
1021
1021
1022
1022
1023
1023
1024
1024
1025
1025
1026
1026
1027
1027
1028
1028
1029
1029
1030
1030
1031
1031
1032
1032
1033
1033
1034
1034
1035
1035
1036
1036
1037
1037
1038
1038
1039
1039
1040
1040
1041
1041
1042
1042
1043
1043
1044
1044
1045
1045
1046
1046
1047
1047
1048
1048
1049
1049
1050
1050
1051
1051
1052
1052
1053
1053
1054
1054
1055
1055
1056
1056
1057
1057
1058
1058
1059
1059
1060
1060
1061
1061
1062
1062
1063
1063
1064
1064
1065
1065
1066
1066
1067
1067
1068
1068
1069
1069
1070
1070
1071
1071
1072
1072
1073
1073
1074
1074
1075
1075
1076
1076
1077
1077
1078
1078
1079
1079
1080
1080
1081
1081
1082
1082
1083
1083
1084
1084
1085
1085
1086
1086
1087
1087
1088
1088
1089
1089
1090
1090
1091
1091
1092
1092
1093
1093
1094
1094
1095
1095
1096
1096
1097
1097
1098
1098
1099
1099
1100
1100
1101
1101
1102
1102
1103
1103
1104
1104
1105
1105
1106
1106
1107
1107
1108
1108
1109
1109
1110
1110
1111
1111
1112
1112
1113
1113
1114
1114
1115
1115
1116
1116
1117
1117
1118
1118
1119
1119
1120
1120
1121
1121
1122
1122
1123
1123
1124
1124
1125
1125
1126
1126
1127
1127
1128
1128
1129
1129
1130
1130
1131
1131
1132
1132
1133
1133
1134
1134
1135
1135
1136
1136
1137
1137
1138
1138
1139
1139
1140
1140
1141
1141
1142
1142
1143
1143
1144
1144
1145
1145
1146
1146
1147
1147
1148
1148
1149
1149
1150
1150
1151
1151
1152
1152
1153
1153
1154
1154
1155
1155
1156
1156
1157
1157
1158
1158
1159
1159
1160
1160
1161
1161
1162
1162
1163
1163
1164
1164
1165
1165
1166
1166
1167
1167
1168
1168
1169
1169
1170
1170
1171
1171
1172
1172
1173
1173
1174
1174
1175
1175
1176
1176
1177
1177
1178
1178
1179
1179
1180
1180
1181
1181
1182
1182
1183
1183
1184
1184
1185
1185
1186
1186
1187
1187
1188
1188
1189
1189
1190
1190
1191
1191
1192
1192
1193
1193
1194
1194
1195
1195
1196
1196
1197
1197
1198
1198
1199
1199
1200
1200
1201
1201
1202
1202
1203
1203
1204
1204
1205
1205
1206
1206
1207
1207
1208
1208
1209
1209
1210
1210
1211
1211
1212
1212
1213
1213
1214
1214
1215
1215
1216
1216
1217
1217
1218
1218
1219
1219
1220
1220
1221
1221
1222
1222
1223
1223
1224
1224
1225
1225
1226
1226
1227
1227
1228
1228
1229
1229
1230
1230
1231
1231
1232
1232
1233
1233
1234
1234
1235
1235
1236
1236
1237
1237
1238
1238
1239
1239
1240
1240
1241
1241
1242
1242
1243
1243
1244
1244
1245
1245
1246
1246
1247
1247
1248
1248
1249
1249
1250
1250
1251
1251
1252
1252
1253
1253
1254
1254
1255
1255
1256
1256
1257
1257
1258
1258
1259
1259
1260
1260
1261
1261
1262
1262
1263
1263
1264
1264
1265
1265
1266
1266
1267
1267
1268
1268
1269
1269
1270
1270
1271
1271
1272
1272
1273
1273
1274
1274
1275
1275
1276
1276
1277
1277
1278
1278
1279
1279
1280
1280
1281
1281
1282
1282
1283
1283
1284
1284
1285
1285
1286
1286
1287
1287
1288
1288
1289
1289
1290
1290
1291
1291
1292
1292
1293
1293
1294
1294
1295
1295
1296
1296
1297
1297
1298
1298
1299
1299
1300
1300
1301
1301
1302
1302
1303
1303
1304
1304
1305
1305
1306
1306
1307
1307
1308
1308
1309
1309
1310
1310
1311
1311
1312
1312
1313
1313
1314
1314
1315
1315
1316
1316
1317
1317
1318
1318
1319
1319
1320
1320
1321
1321
1322
1322
1323
1323
1324
1324
1325
1325
1326
1326
1327
1327
1328
1328
1329
1329
1330
1330
1331
1331
1332
1332
1333
1333
1334
1334
1335
1335
1336
1336
1337
1337
1338
1338
1339
1339
1340
1340
1341
1341
1342
1342
1343
1343
1344
1344
1345
1345
1346
1346
1347
1347
1348
1348
1349
1349
1350
1350
1351
1351
1352
1352
1353
1353
1354
1354
1355
1355
1356
1356
1357
1357
1358
1358
1359
1359
1360
1360
1361
1361
1362
1362
1363
1363
1364
1364
1365
1365
1366
1366
1367
1367
1368
1368
1369
1369
1370
1370
1371
1371
1372
1372
1373
1373
1374
1374
1375
1375
1376
1376
1377
1377
1378
1378
1379
1379
1380
1380
1381
1381
1382
1382
1383
1383
1384
1384
1385
1385
1386
1386
1387
1387
1388
1388
1389
1389
1390
1390
1391
1391
1392
1392
1393
1393
1394
1394
1395
1395
1396
1396
1397
1397
1398
1398
1399
1399
1400
1400
1401
1401
1402
1402
1403
1403
1404
1404
1405
1405
1406
1406
1407
1407
1408
1408
1409
1409
1410
1410
1411
1411
1412
1412
1413
1413
1414
1414
1415
1415
1416
1416
1417
1417
1418
1418
1419
1419
1420
1420
1421
1421
1422
1422
1423
1423
1424
1424
1425
1425
1426
1426
1427
1427
1428
1428
1429
1429
1430
1430
1431
1431
1432
1432
1433
1433
1434
1434
1435
1435
1436
1436
1437
1437
1438
1438
1439
1439
1440
1440
1441
1441
1442
1442
1443
1443
1444
1444
1445
1445
1446
1446
1447
1447
1448
1448
1449
1449
1450
1450
1451
1451
1452
1452
1453
1453
1454
1454
1455
1455
1456
1456
1457
1457
1458
1458
1459
1459
1460
1460
1461
1461
1462
1462
1463
1463
1464
1464
1465
1465
1466
1466
1467
1467
1468
1468
1469
1469
1470
1470
1471
1471
1472
1472
1473
1473
1474
1474
1475
1475
1476
1476
1477
1477
1478
1478
1479
1479
1480
1480
1481
1481
1482
1482
1483
1483
1484
1484
1485
1485
1486
1486
1487
1487
1488
1488
1489
1489
1490
1490
1491
1491
1492
1492
1493
1493
1494
1494
1495
1495
1496
1496
1497
1497
1498
1498
1499
1499
1500
1500
1501
1501
1502
1502
1503
1503
1504
1504
1505
1505
1506
1506
1507
1507
1508
1508
1509
1509
1510
1510
1511
1511
1512
1512
1513
1513
1514
1514
1515
1515
1516
1516
1517
1517
1518
1518
1519
1519
1520
1520
1521
1521
1522
1522
1523
1523
1524
1524
1525
1525
1526
1526
1527
1527
1528
1528
1529
1529
1530
1530
1531
1531
1532
1532
1533
1533
1534
1534
1535
1535
1536
1536
1537
1537
1538
1538
1539
1539
1540
1540
1541
1541
1542
1542
1543
1543
1544
1544
1545
1545
1546
1546
1547
1547
1548
1548
1549
1549
1550
1550
1551
1551
1552
1552
1553
1553
1554
1554
1555
1555
1556
1556
1557
1557
1558
1558
1559
1559
1560
1560
1561
1561
1562
1562
1563
1563
1564
1564
1565
1565
1566
1566
1567
1567
1568
1568
1569
1569
1570
1570
1571
1571
1572
1572
1573
1573
1574
1574
1575
1575
1576
1576
1577
1577
1578
1578
1579
1579
1580
1580
1581
1581
1582
1582
1583
1583
1584
1584
1585
1585
1586
1586
1587
1587
1588
1588
1589
1589
1590
1590
1591
1591
1592
1592
1593
1593
1594
1594
1595
1595
1596
1596
1597
1597
1598
1598
1599
1599
1600
1600
1601
1601
1602
1602
1603
1603
1604
1604
1605
1605
1606
1606
1607
1607
1608
1608
1609
1609
1610
1610
1611
1611
1612
1612
1613
1613
1614
1614
1615
1615
1616
1616
1617
1617
1618
1618
1619
1619
1620
1620
1621
1621
1622
1622
1623
1623
1624
1624
1625
1625
1626
1626
1627
1627
1628
1628
1629
1629
1630
1630
1631
1631
1632
1632
1633
1633
1634
1634
1635
1635
1636
1636
1637
1637
1638
1638
1639
1639
1640
1640
16
```

```
13 output_dir = "./outputs"
14
15 def count_lines(fname):
16     with open(fname) as f:
17         for i, l in enumerate(f):
18             pass
19     return i + 1
20
21 train_img_count = count_lines(os.path.join(data_root, train_txt))
22 print("There are {} samples in training data".format(train_img_count))
```

There are 725 samples in training data

```
1 # Register the traffic_sign_train datasets
2 DatasetCatalog.register(name=train_data_name,
3                         func=lambda: get_dataset_dicts(data_root, train_txt))
4 train_metadata = MetadataCatalog.get(train_data_name).set(thing_classes=thing_classes)
5
6 # Register the traffic_sign_test datasets
7 DatasetCatalog.register(name=test_data_name,
8                         func=lambda: get_dataset_dicts(data_root, test_txt))
9 test_metadata = MetadataCatalog.get(test_data_name).set(thing_classes=thing_classes)

1 train_data_dict = get_dataset_dicts(data_root, train_txt)
2
3 for d in random.sample(train_data_dict, 3):
4     img = cv2.imread(d["file_name"])
5     visualizer = Visualizer(img[:, :, ::-1], metadata=train_metadata, scale=0.5)
6     out = visualizer.draw_dataset_dict(d)
7     cv2.imshow(out.get_image()[:, :, ::-1])
```



```

1 from detectron2.engine import DefaultTrainer
2
3 cfg = get_cfg()
4 cfg.merge_from_file(model_zoo.get_config_file("COCO-Detection/faster_rcnn_R_50_FPN_1x.yaml"))
5 cfg.DATASETS.TRAIN = (train_data_name,)
6 cfg.DATASETS.TEST = ()
7 cfg.DATALOADER.NUM_WORKERS = 2
8 cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url("COCO-Detection/faster_rcnn_R_50_FPN_1x.yaml") # let's trainining initialize from model z
9 cfg.SOLVER.IMS_PER_BATCH = 3
10 cfg.SOLVER.BASE_LR = 0.001 # pick a good LR
11 cfg.SOLVER.MAX_ITER = 300 # 300 iterations seems good enough for this toy dataset; you will need to train longer for a practical dataset
12 cfg.MODEL.ROI_HEADS.NUM_CLASSES = len(thing_classes) # only has one class (traffic-sign)
13 cfg.MODEL.ROI_HEADS.BATCH_SIZE_PER_IMAGE = 256 # faster, and good enough for this toy dataset (default: 512)
14 cfg.OUTPUT_DIR = output_dir

```

```

1 os.makedirs(cfg.OUTPUT_DIR, exist_ok=True)
2 trainer = DefaultTrainer(cfg)
3 trainer.resume_or_load(resume=False)
4 trainer.train()

```

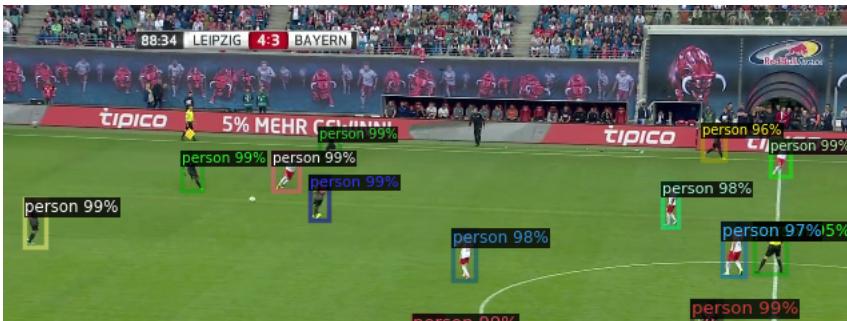
```

[11/11 06:20:50 d2.engine.defaults]: Model:
GeneralizedRCNN(
  (backbone): FPN(
    (fpn_lateral12): Conv2d(256, 256, kernel_size=(1, 1), stride=(1, 1))
    (fpn_output2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (fpn_lateral13): Conv2d(512, 256, kernel_size=(1, 1), stride=(1, 1))
    (fpn_output3): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (fpn_lateral14): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1))
    (fpn_output4): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (fpn_lateral15): Conv2d(2048, 256, kernel_size=(1, 1), stride=(1, 1))
    (fpn_output5): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (top_block): LastLevelMaxPool()
  (bottom_up): ResNet(
    (stem): BasicStem(
      (conv1): Conv2d(
        3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False
        (norm): FrozenBatchNorm2d(num_features=64, eps=1e-05)
      )
    )
    (res2): Sequential(
      (0): BottleneckBlock(
        (shortcut): Conv2d(
          64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False
          (norm): FrozenBatchNorm2d(num_features=256, eps=1e-05)
        )
        (conv1): Conv2d(
          64, 64, kernel_size=(1, 1), stride=(1, 1), bias=False
          (norm): FrozenBatchNorm2d(num_features=64, eps=1e-05)
        )
        (conv2): Conv2d(
          64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
          (norm): FrozenBatchNorm2d(num_features=64, eps=1e-05)
        )
        (conv3): Conv2d(
          64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False
          (norm): FrozenBatchNorm2d(num_features=256, eps=1e-05)
        )
      )
      (1): BottleneckBlock(
        (conv1): Conv2d(
          256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False
          (norm): FrozenBatchNorm2d(num_features=64, eps=1e-05)
        )
        (conv2): Conv2d(
          64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
          (norm): FrozenBatchNorm2d(num_features=64, eps=1e-05)
        )
      )
    )
  )
)
```

```
(conv3): Conv2d(  
    64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False  
    (norm): FrozenBatchNorm2d(num_features=256, eps=1e-05)  
)  
(2): BottleneckBlock(  
    (conv1): Conv2d(  
        256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False  
        (norm): FrozenBatchNorm2d(num_features=64, eps=1e-05)  
)  
  
1 # cfg already contains everything we've set previously. Now we changed it a little bit for inference:  
2 cfg.MODEL.WEIGHTS = os.path.join(cfg.OUTPUT_DIR, "model_final.pth") # path to the model we just trained  
3 cfg.MODEL.ROI_HEADS.SCORE_THRESH_TEST = 0.65  
4 predictor = DefaultPredictor(cfg)
```

```
[11/11 06:25:05 d2.checkpoint.detection_checkpoint]: [DetectionCheckpointer] Loading from ./outputs/model_final.pth ...
```

```
1 from detectron2.utils.visualizer import ColorMode  
2  
3 test_data_dict = get_dataset_dicts(data_root, test_txt)  
4  
5 for d in random.sample(test_data_dict, 3):  
6     im = cv2.imread(d["file_name"])  
7     outputs = predictor(im)  
8     v = Visualizer(im[:, :, ::-1],  
9                     metadata=test_metadata,  
10                     scale=0.5,  
11                     )  
12     out = v.draw_instance_predictions(outputs["instances"].to("cpu"))  
13     cv2.imshow(out.get_image()[:, :, ::-1])
```



▼ Problem 2

Multi-Object Tracking

After training the detector, now we want to implement tracking on the testing video.

```
1 # Let's start with a detector class
2 class detector:
3     def __init__(self,predictor):
4         self.model = predictor
5
6     def predict(self,img):
7         pred = self.model(img)
8         pred = [pred['instances'][i].pred_boxes.tensor.tolist()[0] for i in range(len(pred['instances']))]
9         return pred
10
11 # TODO
12 # Initiate a detector and inference on the first test image('sportsmot/JPEGImages/test_000001.jpg') and print the bounding box prediction.
13 # The output format should be x1,y1,x2,y2
```

```
1 test_data_dict = get_dataset_dicts(data_root, test_txt)

1 im = cv2.imread(test_data_dict[0]['file_name'])
2
3 outputs = predictor(im)

1 thisDetector = detector(predictor)
1 thisDetector.predict(im)

[[695.1017456054688, 329.04803466796875, 727.0886840820312, 395.6009216308594],
 [198.265625, 240.1858673095703, 231.60711669921875, 291.2448425292969],
 [608.6708984375, 380.8650817871094, 638.8543090820312, 440.002685546875],
 [767.8047485351562, 455.3278503417969, 804.1311645507812, 538.384033203125],
 [1243.25048828125, 387.00177001953125, 1274.087158203125, 466.32476806640625],
 [868.5335083007812,
 257.01336669921875,
 898.5045776367188,
 311.70037841796875],
 [545.5725708007812, 209.90061950683594, 572.3731079101562, 260.3439636230469],
 [1062.9161376953125,
 292.83416748046875,
 1092.3922119140625,
 357.8119201660156],
 [348.142822265625, 453.1558532714844, 381.67755126953125, 525.9806518554688],
 [1110.0975341796875,
 293.57086181640625,
 1133.2042236328125,
 359.6961975097656],
 [1171.104736328125, 245.98287963867188, 1189.6796875, 302.8831481933594],
 [487.81304931640625,
 327.46539306640625,
 510.3650207519531,
 391.7694891796875],
 [746.527587890625, 254.35745239257812, 772.0448608398438, 310.05157470703125],
 [65.74864959716797, 274.23773193359375, 93.96112060546875, 335.4070129394531],
 [33.90310287475586, 308.1076965332031, 59.97958755493164, 378.60284423828125],
 [89.94819641113281,
 281.0174560546875,
 115.04015350341797,
 336.06024169921875],
 [754.9312133789062, 326.55084228515625, 775.4976196289062, 397.2494812011719],
```

```
[885.1402587890625, 402.6461181640625, 910.5313110351562, 479.73583984375],
[210.8836212158203, 286.97467041015625, 233.4933624267578, 349.0710754394531],
[493.33306884765625,
208.96238708496094,
510.21954345703125,
260.8963317871094],
[662.82080078125, 306.70709228515625, 686.7090454101562, 372.1379089355469]]
```

Now you will implement your own tracker!

Let's start with the IoU function and tracklet class.

```
1 # calculate the overlap ratio of two bounding boxes
2 def calculate_iou(bbox1, bbox2):
3
4     x1_1, y1_1, x2_1, y2_1 = bbox1
5     x1_2, y1_2, x2_2, y2_2 = bbox2
6     x_left = max(x1_1, x1_2)
7     y_top = max(y1_1, y1_2)
8     x_right = min(x2_1, x2_2)
9     y_bottom = min(y2_1, y2_2)
10
11    if x_right < x_left or y_bottom < y_top:
12        return 0.0
13
14    area_bbox1 = (x2_1 - x1_1 + 1) * (y2_1 - y1_1 + 1)
15    area_bbox2 = (x2_2 - x1_2 + 1) * (y2_2 - y1_2 + 1)
16    intersection_area = (x_right - x_left + 1) * (y_bottom - y_top + 1)
17
18    iou = intersection_area / float(area_bbox1 + area_bbox2 - intersection_area)
19
20    return iou

1 # base class for tracklet, you can of course add more features and try to improve the performance!
2 class tracklet:
3     def __init__(self, tracking_ID, box):
4         self.ID = tracking_ID
5         self.cur_box = box
6         self.alive = True
7
8     def update(self, box):
9         self.cur_box = box
10
11    def close(self):
12        self.alive = False

1 from scipy.optimize import linear_sum_assignment
2
3 class IoU_Tracker:
4     def __init__(self):
5         self.all_tracklets = [] # this saves all the tracklets so that we can know how many tracklets we have
6         self.cur_tracklets = [] # this saves tracklets from the last frame for current frame's association
7         self.online_tracklets = [] # this saves the tracklets after association, so we can pass the tracking result to output
8
9     def update(self, frame_id, detection):
10
11        if frame_id%100 == 0:
12            print(f'Running tracking || current frame {frame_id}')
13
14        if len(self.cur_tracklets) == 0:
15            for det in detection:
16                new_tracklet = tracklet(len(self.all_tracklets)+1, det)
17                self.cur_tracklets.append(new_tracklet)
18                self.all_tracklets.append(new_tracklet)
19        else:
20            cost_matrix = np.zeros((len(self.cur_tracklets), len(detection)))
21
22            # build up cost matrix, each element in cost matrix should be 1-IoU between tracklet and detection
23            for row in range(len(self.cur_tracklets)):
24                for col in range(len(detection)):
25                    cost_matrix[row][col] = 1 - calculate_iou(self.cur_tracklets[row].cur_box, detection[col])
26
27            row_inds, col_inds = linear_sum_assignment(cost_matrix)
28
29            matches = min(len(row_inds), len(col_inds))
```

```

30
31     for idx,trk in enumerate(self.cur_tracklets):
32         if idx not in row_inds: # if it is not matched in the above Hungarian algorithm stage
33             # TODO
34             # use tracklet's close function to kill those unmatched tracklets
35             self.cur_tracklets[idx].close()
36             self.all_tracklets[idx].close()
37
38     for idx,det in enumerate(detection):
39         if idx not in col_inds: # if it is not matched in the above Hungarian algorithm stage
40             # TODO
41             # initiate unmatched detections as new tracklets
42             new_tracklet = tracklet(idx,det)
43             self.all_tracklets.append(new_tracklet)
44             self.cur_tracklets.append(new_tracklet)
45
46     for idx in range(matches):
47         row,col = row_inds[idx],col_inds[idx]
48         if cost_matrix[row][col] == 1:
49             # TODO 1. Kill the tracklet using tracklet's close function
50             self.cur_tracklets[row].close()
51             self.all_tracklets[row].close()
52             # TODO 2. Initiate a new tracklet for the new detection
53             new_tracklet = tracklet(len(self.all_tracklets),det)
54             # TODO 3. Append new tracklet to the current tracklets and all tracklets
55             self.all_tracklets.append(new_tracklet)
56             self.cur_tracklets.append(new_tracklet)
57         else:
58             self.cur_tracklets[row].update(detection[col])
59
60     self.cur_tracklets = [trk for trk in self.cur_tracklets if trk.alive]
61
62     return self.cur_tracklets

```

Now it's time to run tracking!

```

1 import glob
2
3 # TODO
4 # initialize your detector (the one you just trained)
5 thisDetector = detector(predictor)
6
7 # TODO
8 # initialize tracker
9 thisIOUTracker = IoU_Tracker()
10
11 # TODO run tracking
12 images = np.loadtxt('sportsmot/sportsmot_test.txt',dtype=str)
13 results = []
14
15 print(f'length of sequence is {len(images)}')
16
17 for frame_id, img_path in enumerate(images,1):
18     img = cv2.imread('sportsmot/'+img_path)
19     # TODO get detection with your model
20     detection = thisDetector.predict(img)
21
22     # TODO update tracker with detection
23     result = thisIOUTracker.update(frame_id,detection)
24
25     for track in result:
26         x1,y1,x2,y2 = track.cur_box
27         track_id = track.ID
28         results.append(f'{int(frame_id)},{int(track_id)},{int(x1)},{int(y1)},{int(x2-x1)},{int(y2-y1)},{1},{1},{1}') # format: frame_id, t
29
30 with open('results.txt','w') as f:
31     for line in results:
32         f.writelines(line)
33         f.writelines('\n')
34
35 length of sequence is 500
36 Running tracking || current frame 100
37 Running tracking || current frame 200
38 Running tracking || current frame 300
39 Running tracking || current frame 400
40 Running tracking || current frame 500

```