MID EXAM: Detectron2 using Mask-RCNN

GitHub Repository: https://github.com/pshruthi5/FS2022.git

GitHub URL: https://github.com/pshruthi5/FS2022/tree/main/Applied-Artificial-Intelligence

This is a code implementing a detectron using a mask RCNN network. I modified the detectron tutorial uploaded by facebook in their github repository. From the model zoo, I picked the mask RCNN. There are three types of outputs in the code, one is a video of my choice from youtube and and put its segmented output. The second is a an image using my local webcam, and the segmented image as the output. Third is a real-time(?) feed from my webcam and its output.

The first part of the code is downloading the dependencies and running it exactly as the code from the detectron tutorial.

We then check the version of pytorch, CUDA and the detectron we have downloaded (just to verify for our information)

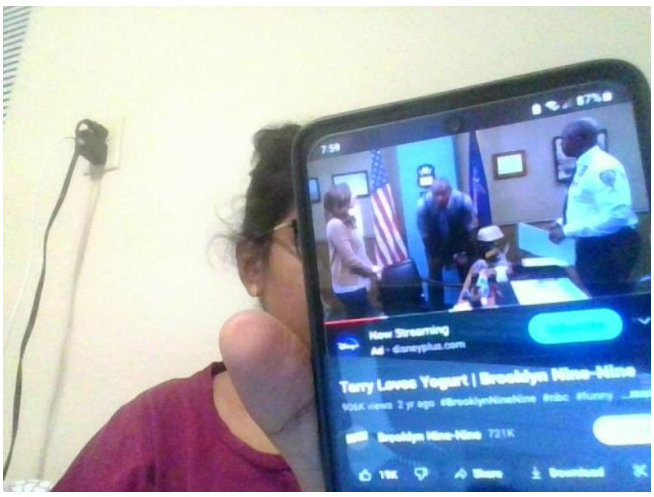
We then download all the required all the required dependencies and libraries from the detectron repository

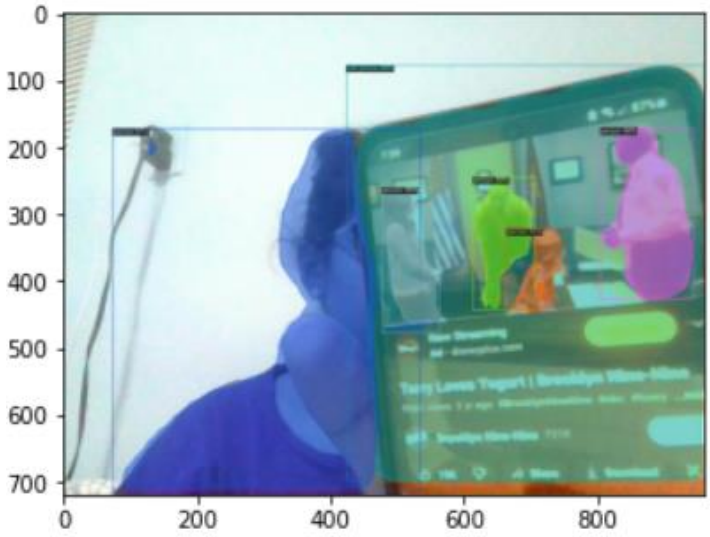
```
# Some basic setup:
# Setup detectron2 logger
import detectron2
from detectron2.utils.logger import setup_logger
setup_logger()

# import some common libraries
import numpy as np
import os, json, cv2, random
from google.colab.patches import cv2_imshow

# import some common detectron2 utilities
from detectron2 import model_zoo
from detectron2.engine import DefaultPredictor
from detectron2.config import get_cfg
from detectron2.utils.visualizer import Visualizer
from detectron2.data import MetadataCatalog, DatasetCatalog
```

We then run it with the example that is already present in the code for a trial run. On the right side is the segmented output:

Now to create the first type of output, using a youtube video and getting its segmented output. I got a downloaded video of the segmented output.

The next type of output was a picture from the webcam and its segmented output (using matplotlib),





The third type of output was using a real-time video and putting out the segmented output. However, the output was very very slow… This is a frame from webcam which is segmented.