

Detectron2 Analysis Report

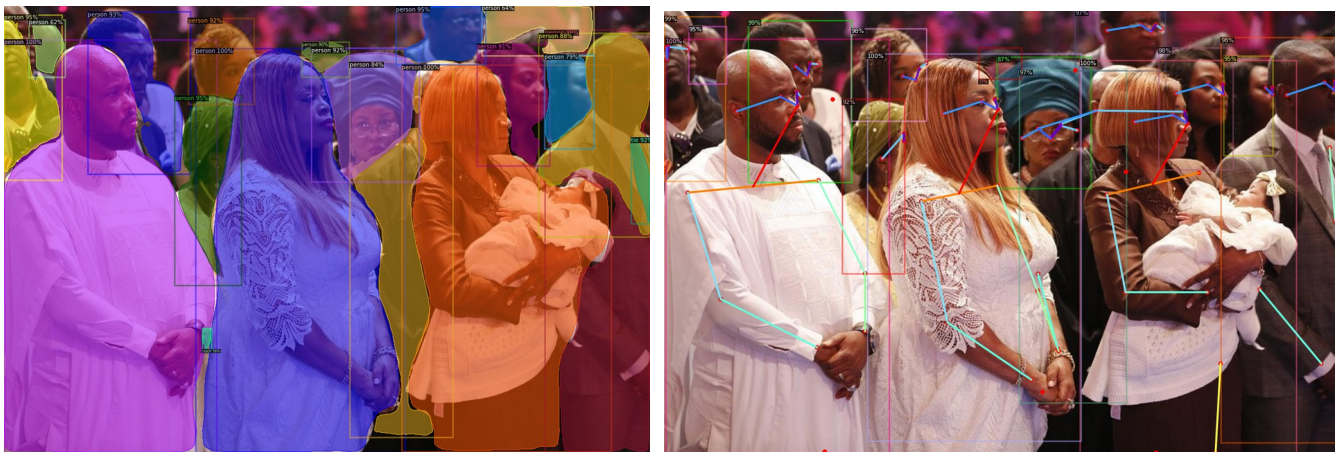
Introduction: Detectron2 is a ground-up rewrite of Detectron that started with maskrcnn-benchmark. It is Facebook's new vision library that allows us to easily use and create object detection, instance segmentation. It also provides a model zoo, a library of pre-trained models trained on a variety of datasets.

Model Architecture: To first step was to install Detectron2 following the installation tutorial, making sure that I met all of the requirement, e.g, Python ≥ 3.6 , PyTorch ≥ 1.4 et.c. once that was settled, the next step was to see the backbone architecture and the dataset it was pre-trained on, For Faster/Mask R-CNN, baselines was provided based on 3 different backbone combinations: FPN, CA and DC5, with a ResNet50 backbone pre-trained on the COCO data set which has more than 300k images annotated with bounding boxes and instance masks of 80 object categories.

Example1 and it Implementation

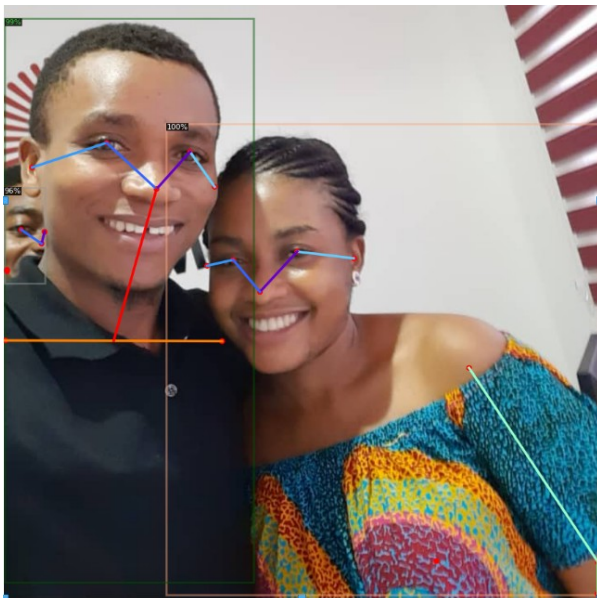
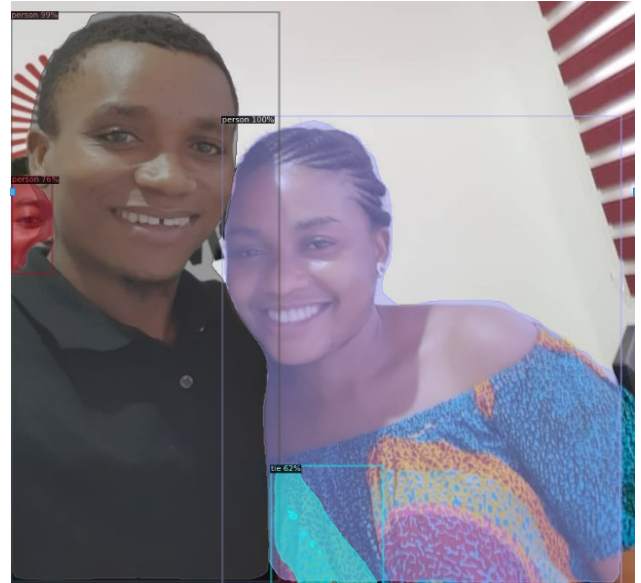


Using this random Image from the internet, I created a detectron2 config and a detectron2 Default Predictor to run inference on this image. the outputs were, (1) A prediction class and a prediction bounding box for each predicted class. Bellow is the visualization and it's respective observations.



Observation: From the above, we noticed that the instance segmentation (left) was able to segment almost all the instances in the picture, even to the point of capturing the man's wrist watch (as tiny as it might be), this is something interesting, however, the model was unable to capture the baby (maybe the coco data set was not trained on baby images). The pose estimation (right) was also good but there was no pose for the baby.

Example 2 and its Implementation



Observation: The instance segmentation for the image above is not too good, although, it was amazing how it could capture only the eyes of the third person at the back, as still predicted it to be 76% human. However, the lady (I) was not wearing a tie, but the model predicted the design on my cloth as a tie, in-fact with 62% probability. This might be as a result of the model being trained with ties of that colour, so seeing such design on my cloth, and the fact that it was close to my neck region, it

predicted a tie. As for the pose estimation, it was able to capture all the pose, including the eye of that young man at the back. This is amazing.