

Lab 1: Detectron2 Warmup

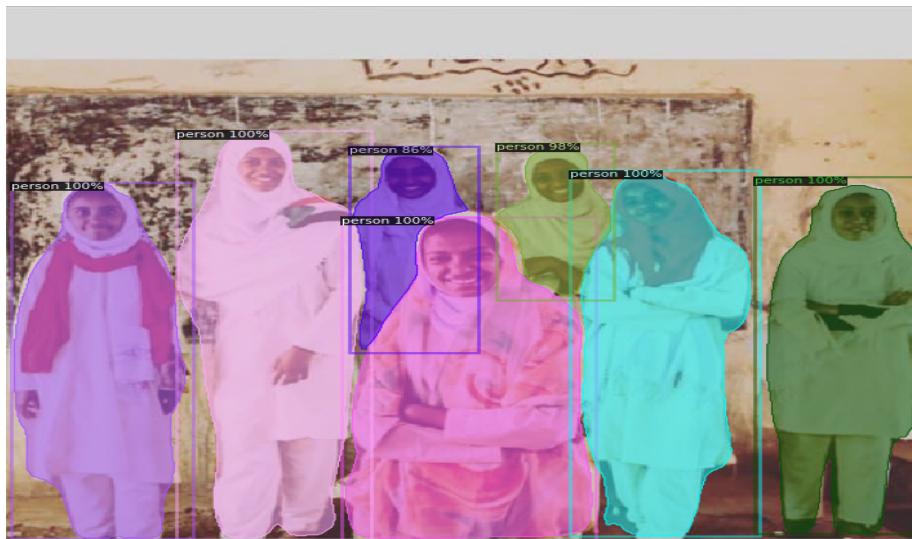
Task1:

Method: Mask R-CNN

Backbone: ResNet50+FPN with the 3x schedule (~37 COCO epochs)

Dataset on which the model is pretrained on :COCO

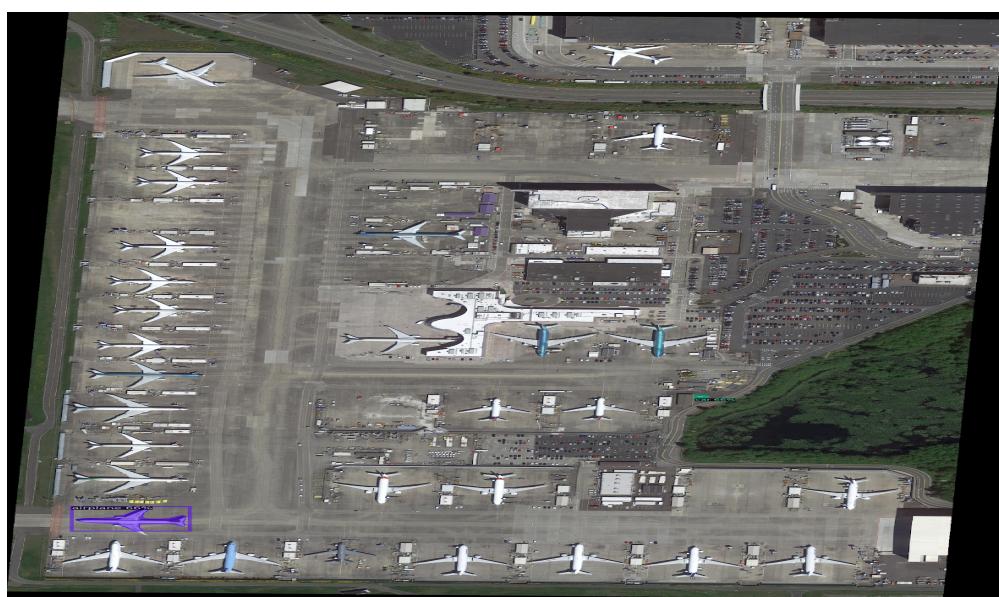
Examples:



The detection of persons when they aren't crowded is easily done , person instances are abundant in COCO



The detection of other objects that are in coco dataset is easily detected also



The detection of planes in a high resolution satellite image is not correctly , only one plane is detected ,maybe due to the high resolution and the small size of the objects which are difficult to detect if the dataset doesn't contain various sizes of the objects.



crowd labeling is not correctly done maybe the NMS algorithm doesn't take in consideration different overlapping objects

Observations:



In those first 2 images above there is more than one label for one object with different level of confidence for each , which makes me question the underlying architecture of the model of how it minimizes the loss .also in the third image it is interesting to see that the model is robust to light intensity and predicted objects correctly.

Error modes:

One of the most important errors of course is the crowd labeling which may result from the implementation of the NMS algorithm in which it doesn't take in consideration different overlapping objects , also when experimenting with different satellite images other than the one above the model cannot detect any object of them maybe due to small size of the objects & the high resolution hyperparameter that is required by the model in this case which maybe traded off for the speed .

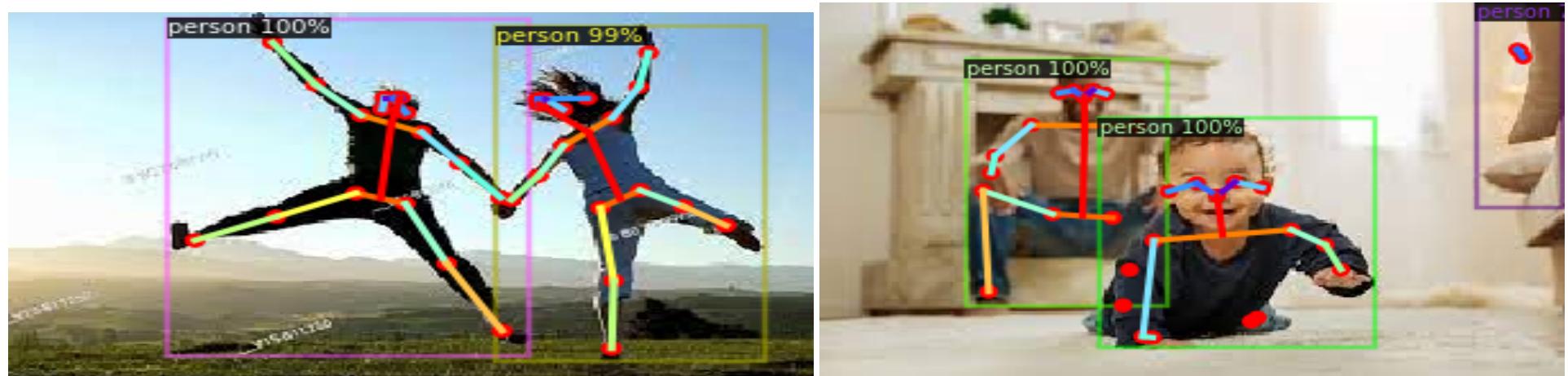
Task2:

Method: keypoint R-CNN

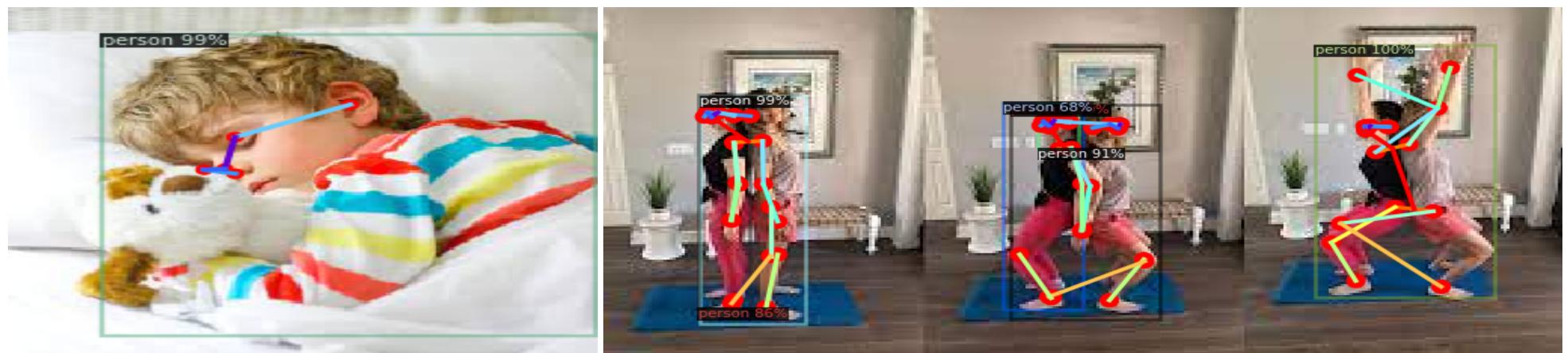
Backbone: ResNet50+FPN with the 3x schedule (~37 COCO epochs)

Dataset on which the model is pretrained on :COCO

Examples:



Key Points in both those different poses is well detected as the key points although the crawling pose is not easy to detect it was easily detected Are clearly apparent in the image objects



Sleeping pose is not detected because maybe key points Of interset Are not well defined/apparent in this case.

Multi person poses cause the model to misidentify the pose

Observation:



In the first 2 case it is interesting that the model cannot detect keypoints in fast moving or upside down pose while in the third pose although is involves many poses it was well detected because the objects are well separated from each other.

Errors modes :

It is observable that the model faces challenges with multi-person poses where occlusion & overlapping is present & it is challenging to assign homogeneous points to different persons. Also it failed in poses that involve undefined key points in the model dataset like sleeping or if the photo is flipped because this flipped ordering of pose keypointst is not well defined in the data set Also,a fast moving pose is not well detected because the input image in this case will be blurred or have low resolution