

Computer Vision, Object Detection Lab 1

Rami Ahmed

1. Model Architecture

Method	Mask R-CNN R 50 FPN 3x
Backbone	ResNet50
Pre-Train Data	COCO Dataset

2. Examples

Part B: Instance Segmentation



Figure 1.a. Correct Prediction.

The predictions are detailed and perfect.



Figure 1.b. Correct Prediction.

The predictions are almost perfect, despite that it's not a realistic picture [Slam Dunk Manga].



Figure 2.a. Incorrect Prediction.

Some of the predictions are incorrect.
The model predicts an empty space with [mouse 63%], highlighters with [book 57%], and a bottle brush with [toothbrush 65%] and mouse 71%].



Figure 2.b. Incorrect Prediction.

Part C: Pose Estimation



Figure 3.a. Correct Prediction.

The predictions follow the two players perfectly. Some of the people in the background do have partial predictions; because they're blurry.



Figure 3.b. Correct Prediction.

The predictions follow the character almost perfectly. However, some parts are missing; because some are occluded, and probably because they're animated characters [ATLA].



Figure 4.a. Incorrect Prediction.

The shape of the prediction doesn't follow the body of the man, because of the strange posture.

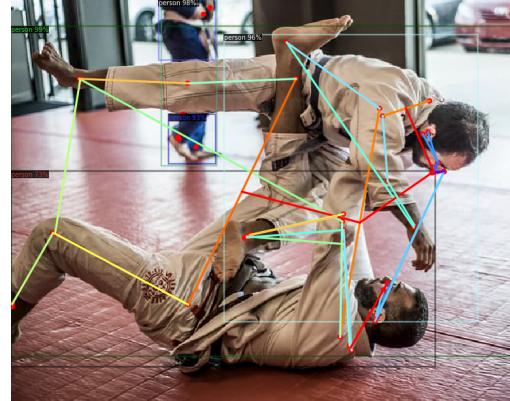


Figure 4.b. Incorrect Prediction.

The shape of the predictions don't follow the two bodies; because of the complex interaction.

3. Observations

In instance segmentation, we noticed that the model performs well when the objects are clearly visible, even when the objects are not realistic it gives a good performance. However, when the objects interfere with each other, the model fails to predict some of them, or gives wrong predictions. In **pose estimation**, we noticed that the model performs well when the objects are clear, even when the objects are not realistic but they've a human shape, it gives a good performance. However, when the pose is strange, or they're interference it gives messy or missing predictions.

4. Error Models

In **instance segmentation**, the model gives wrong predictions in some cases, such as: [When multiple objects interfere with each others, The posture of the object is different from what the model trained on, standing/laying]

In **pose estimation**, the model gives wrong predictions in some cases, such as: [When multiple objects interfere with each others, The posture of the object is strange, Animation characters may have missing prediction points]