Lab 2 : Training a D2 Model on a Custom Dataset

By:

Henock Makumbu Mboko

Lecturer:

Natalia Neverova

9 May 2024

Plan

Introduction

As in the first lab, we began by installing the lab tools; in this case, we used the following:

I. Data

In this lab, we show how to train an existing detectron2 model on a custom dataset in a new format.

Data Preparation and Registration

Before training our model, we need to prepare and register the data. The training data will be registered as nuts_train and the val data as nuts_val. The metadata for both train and val will be in accordance with the following:

NUTS_CATEGORIES ["color": [0, 125, 92], "isthing": 1, "id": 1, "name": "date", "color": [119, 11, 32], "isthing": 1, "id": 2, "name": "fig", "color": [0, 0, 142], "isthing": 1, "id": 3, "name": "hazelnut",]

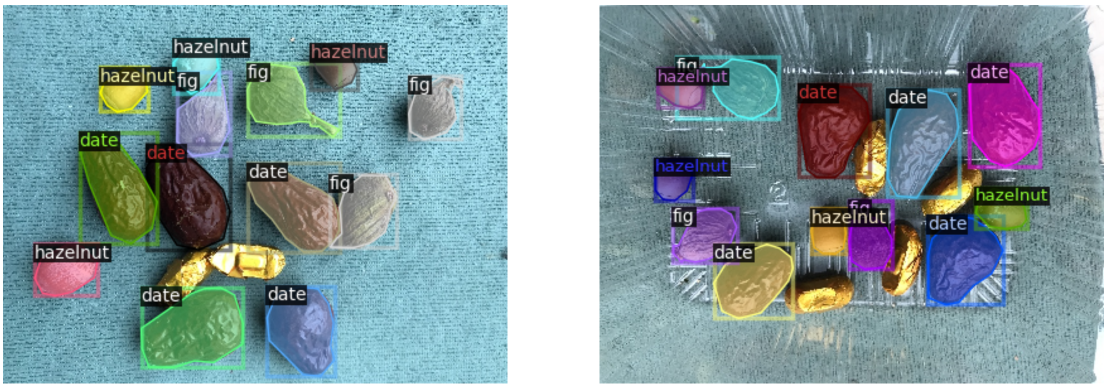


Figure 1: CoCo Dataset

II. Model and Training Schedule

Model initialization and Training process

Let's break down some key points:

1. Model Architecture: The model architecture is a Generalized Region Convolutional Neural Network (GeneralizedRCNN). It consists of a Feature Pyramid Network (FPN) backbone, a Region Proposal Network (RPN), and ROI (Region of Interest) Heads for bounding box detection and instance segmentation.

2. Data Preparation: 13 images in COCO format are loaded from a specified location. Data augmentation techniques such as resizing and random flipping are applied to the training dataset.

3. Training Initialization : The training process starts from iteration 0.

4. Training Process: The training progresses iteratively, with updates on the loss and other metrics reported for every few iterations. The reported metrics include total loss, classification loss (loss_cls), bounding box regression loss (loss_box_reg), mask loss (loss_mask), RPN classification loss (loss_rpn_cls), and RPN localization loss (loss_rpn_loc)

From COCOinit Model

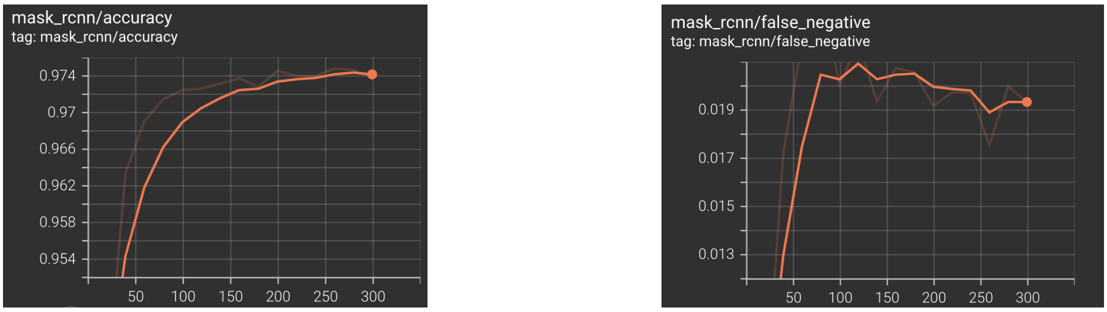


Figure 2: Accuracy-Mask-RCNN

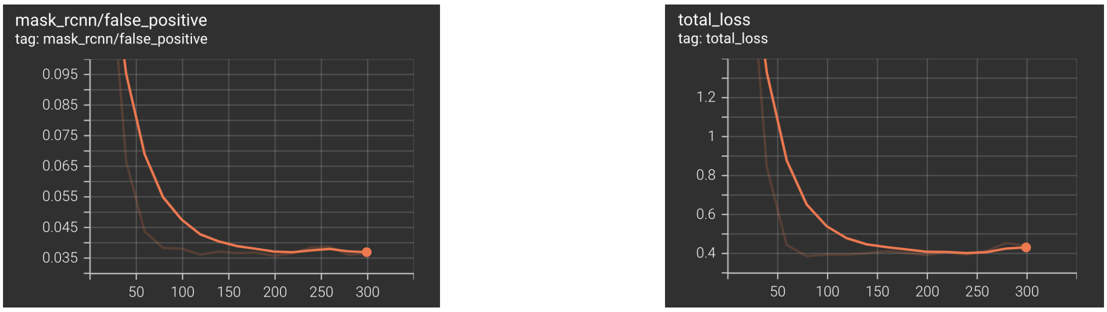


Figure 3: loss-Mask-RCNN & total loss

From INinit Mddel

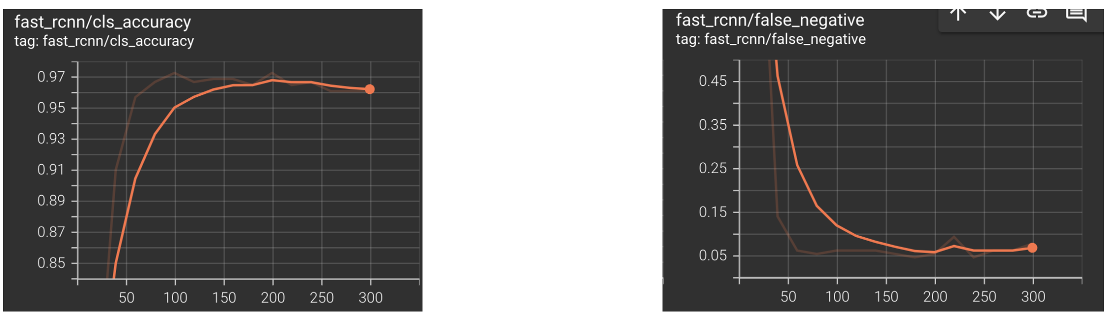


Figure 4: Accuracy-Mask-RCNN

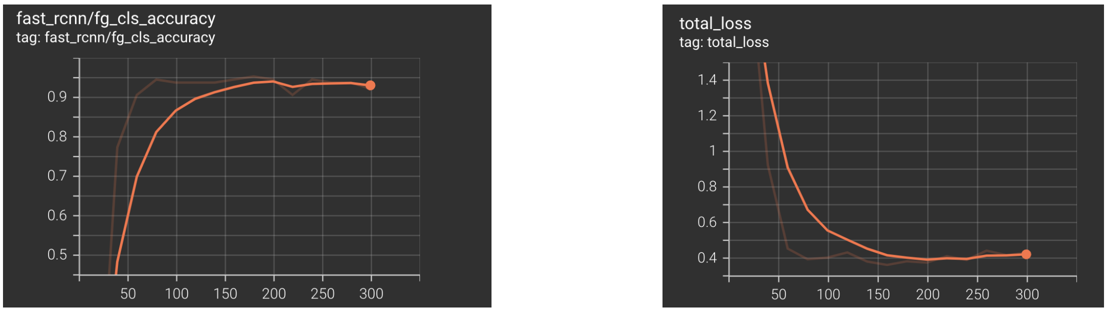


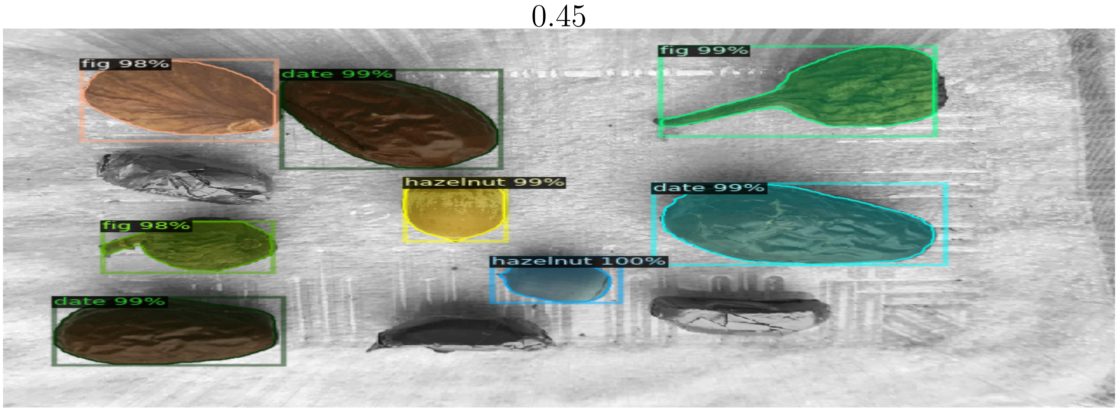
Figure 5: loss-Mask-RCNN & total loss

III. Inference using trained model

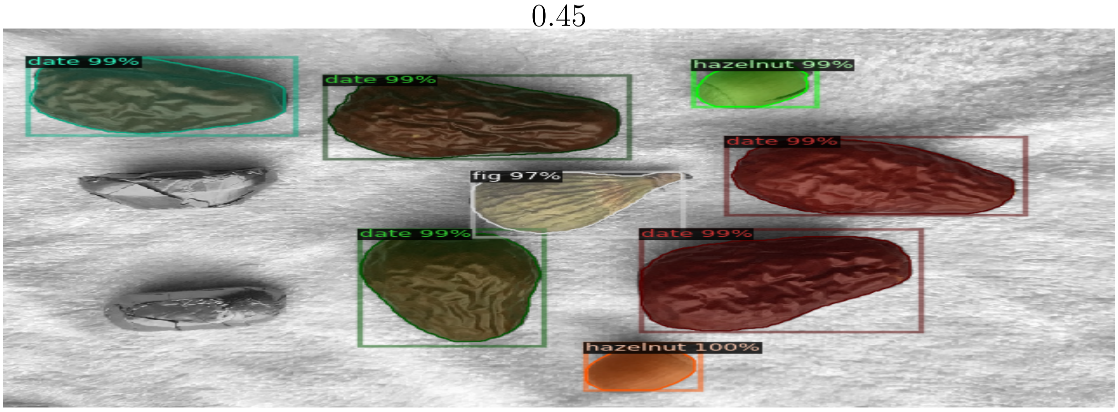
Visualize predictions of both trained models, on the images of the val set :

From COCOinit Model

0.45

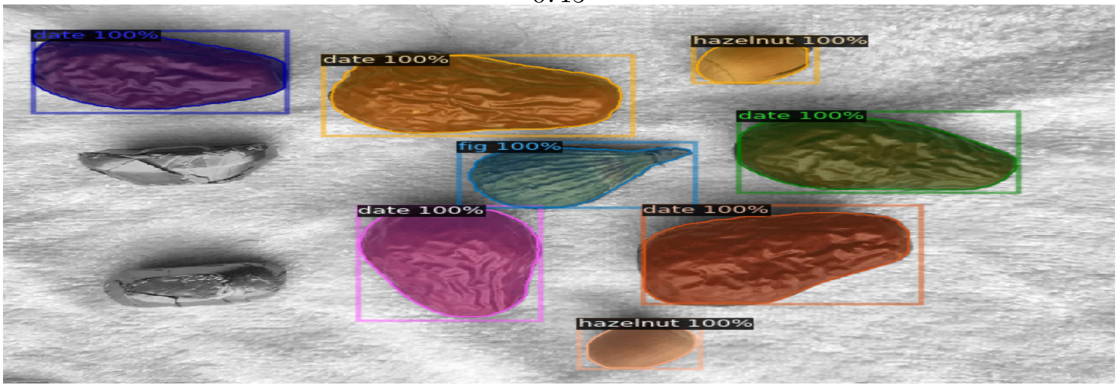


0.45

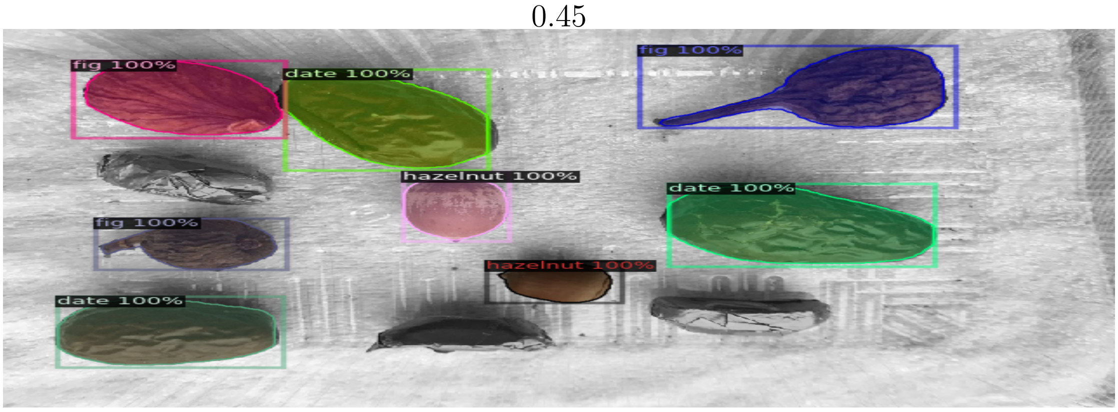


From INinit Mddel

0.45



0.45



IV. Evaluation of trained model

The results exhibit an outstanding performance, especially in segmentation where the model achieved a remarkable AP of 92.15%. This indicates high accuracy in delineating object boundaries, crucial for tasks like image understanding and analysis. The perfect AP50 score for both bounding box and segmentation highlights flawless detection at a certain IoU threshold, suggesting the model's precision in identifying objects within images.

Table 1: Evaluation Metrics with COCOinit

Metric	Bounding Box (bbox)	Segmentation (segm)
Average Precision (AP)	75.24%	92.15%
AP50	100.00%	100.00%
AP75	95.05%	100.00%
AP Small	NaN	NaN
AP Medium	69.29%	90.51%
AP Large	78.62%	92.12%

Table 2: Evaluation Results (bbox)

Category	AP (bbox)	AP50 (bbox)
Date	84.58	100.00
Fig	83.68	-
Hazelnut	74.41	-
Overall	**80.89**	**-**

Bounding Box Evaluation (Table 1):

- The model achieves an Average Precision (AP) of 80.89% for bounding box detection, indicating a high level of accuracy in localizing objects within images.
- At a 50% Intersection over Union (IoU) threshold (AP50), the model achieves perfect precision, correctly identifying all objects with at least 50% overlap with ground truth boxes.
- The model also demonstrates perfect precision at a 75% IoU threshold (AP75), emphasizing its ability to precisely localize objects with stricter criteria.
- The category-wise breakdown reveals varying levels of performance across different object types, with dates being detected most accurately (AP = 84.58%) and hazelnuts showing slightly lower accuracy (AP = 74.41%).

Table 3: Evaluation Results (segm)

Category	AP (segm)	AP75 (segm)	APm (segm)	APl (segm)
Date	97.38	100.00	-	-
Fig	89.55	-	-	-
Hazelnut	87.40	-	-	-
Overall	**91.44**	**100.00**	**87.89**	**86.46**

Segmentation Evaluation (Table 2):

- For segmentation, the model achieves an overall Average Precision (AP) of 91.44%, indicating highly accurate delineation of object boundaries.
- The model maintains perfect precision at a 75% IoU threshold (AP75), emphasizing its ability to precisely delineate object boundaries, especially crucial for tasks requiring fine-grained segmentation.
- Performance across different object sizes varies, with the model demonstrating slightly lower accuracy for small objects (APm = 87.89%) compared to medium-sized (APm $\bar{8}$7.40%) and large objects (APl = 95.37%).
- Similar to bounding box detection, the category-wise breakdown reveals varying levels of segmentation accuracy across different object types, with dates showing the highest segmentation precision (AP = 97.38

Click here to visit Notebook Lab .

Contents