Our team chose to use YOLO v8 Small for the baseline model. We will discuss the performance metrics and diagnostic figures below. However, it is important to note that we improved our Data’s validation set because we added 1000 single frame positive images from: <https://datasets.simula.no/kvasir/>. The reason we decided to add this additional data is because in [the competition](https://ceur-ws.org/Vol-2886/paper4.pdf), they went ahead and did all the validation on this training set. Another reason is that our previous validation dataset mainly consisted of two sequence datasets and included very few single-image samples. We felt that this setup lacked generalizability, which is why we added new data to the validation set. Along with the additional data in the validation set, we tried two augmentation techniques. One that is very standard and does minimal augmentation and the other comes as a suggestion from the competition.

**Note on Metrics:**

The major metrics used for the baseline model coincide with best practices used on medical imaging data; specifically, polyp detection. We give brief intuitive definitions below:

* A prediction is considered a *True Positive* if:
  + In short, the predicted box overlaps a ground-truth polyp above the IoU threshold.
* A *False Positive* is when the predicted box does not match any ground truth polyp.
* A *False Negative* is when the ground-truth polyp was not detected by the model.
* Precision: “Of all *detected* polyps, how many were real?”
* Recall: “Of all *real* polyps, how many were detected?”
* mAP @ 50: Average Precision at IoU greater than or equal to 0.5
* mAP @ 50-95: Mean Average Precision averaged over IoU thresholds between 0.5-0.95

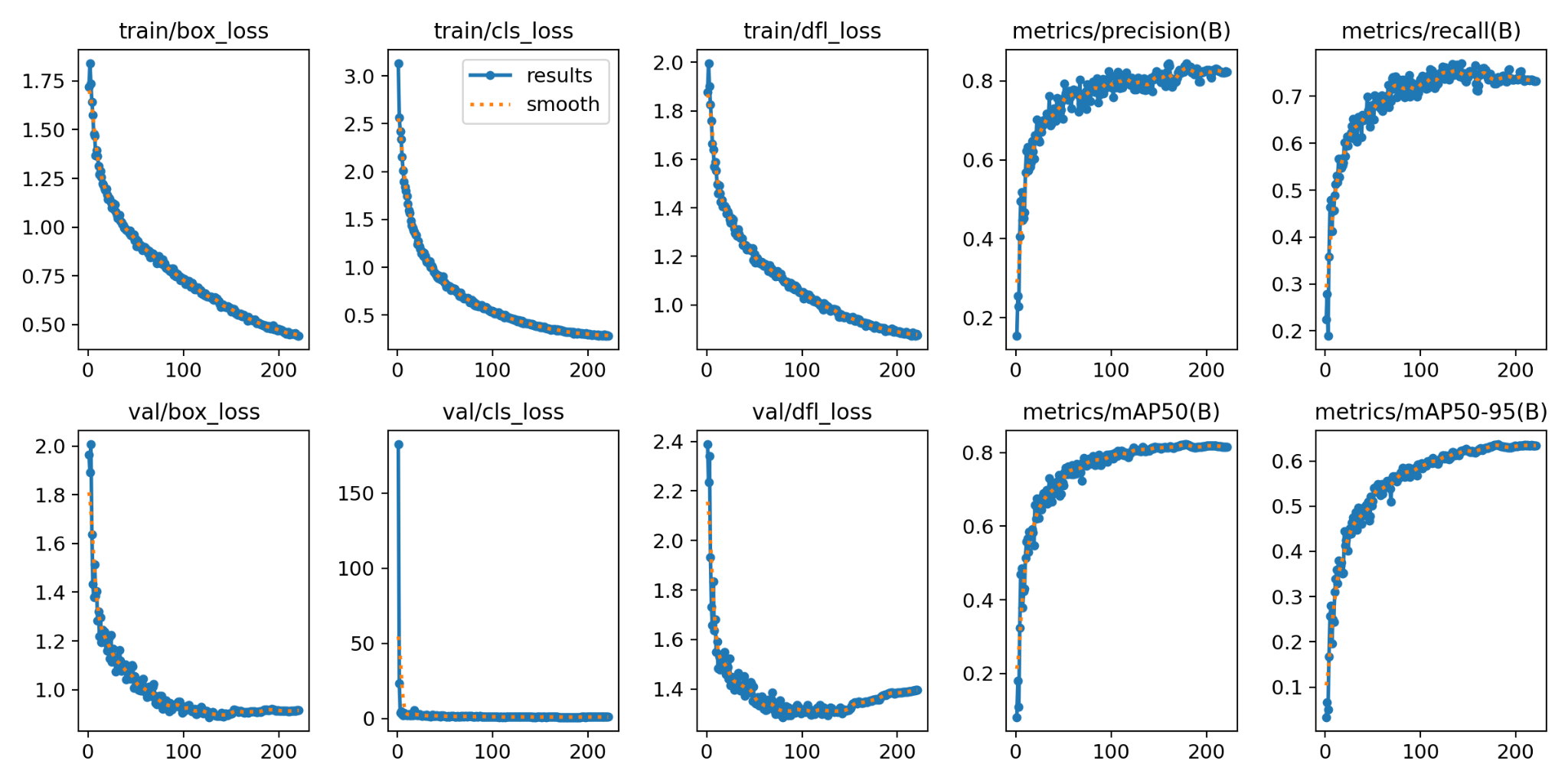
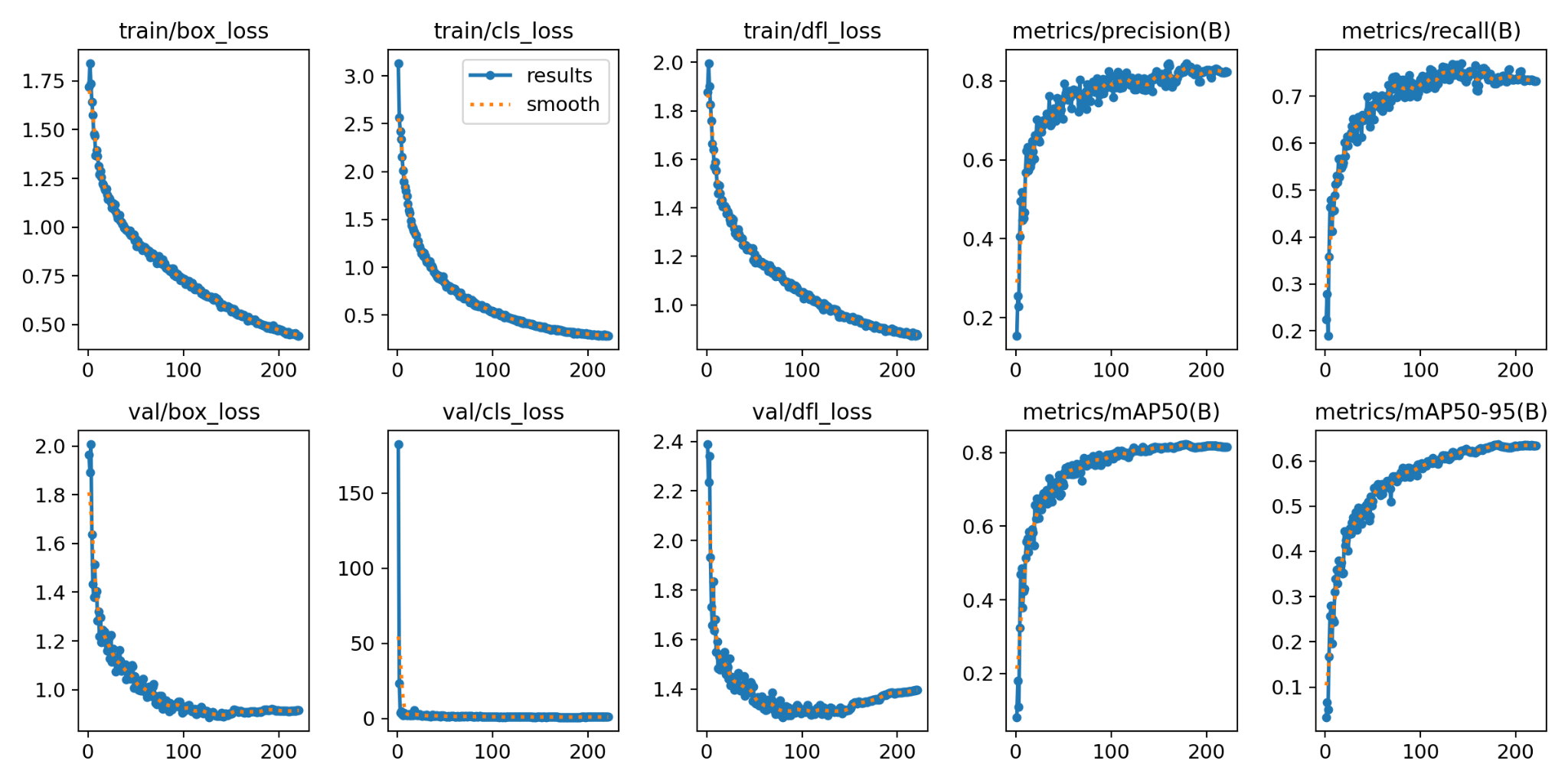
We also would like to include inference time as a separate metric, which will be measured after figuring out a common setup for the hardware.

**Note on Augmentation:**

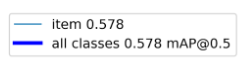
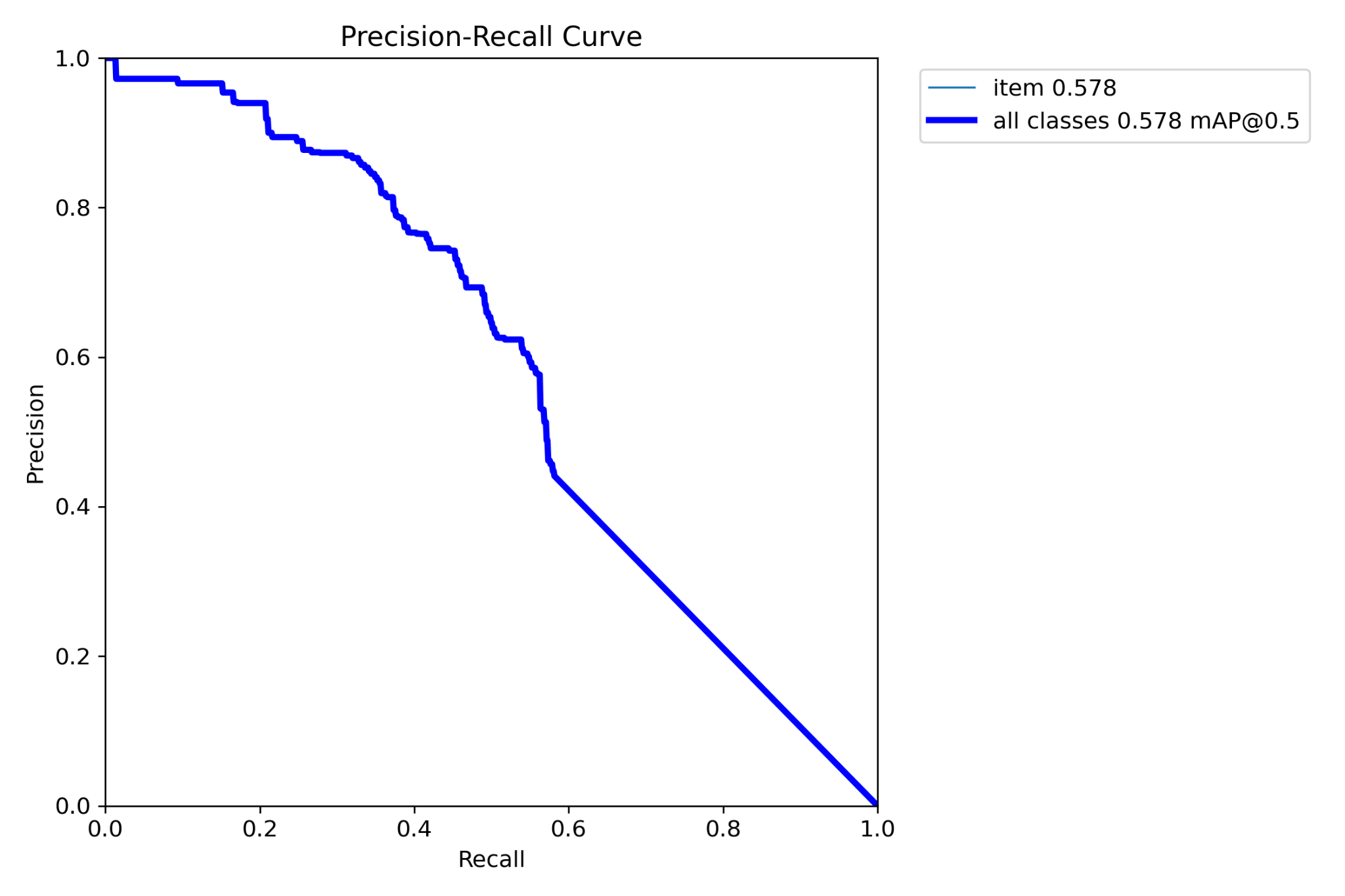
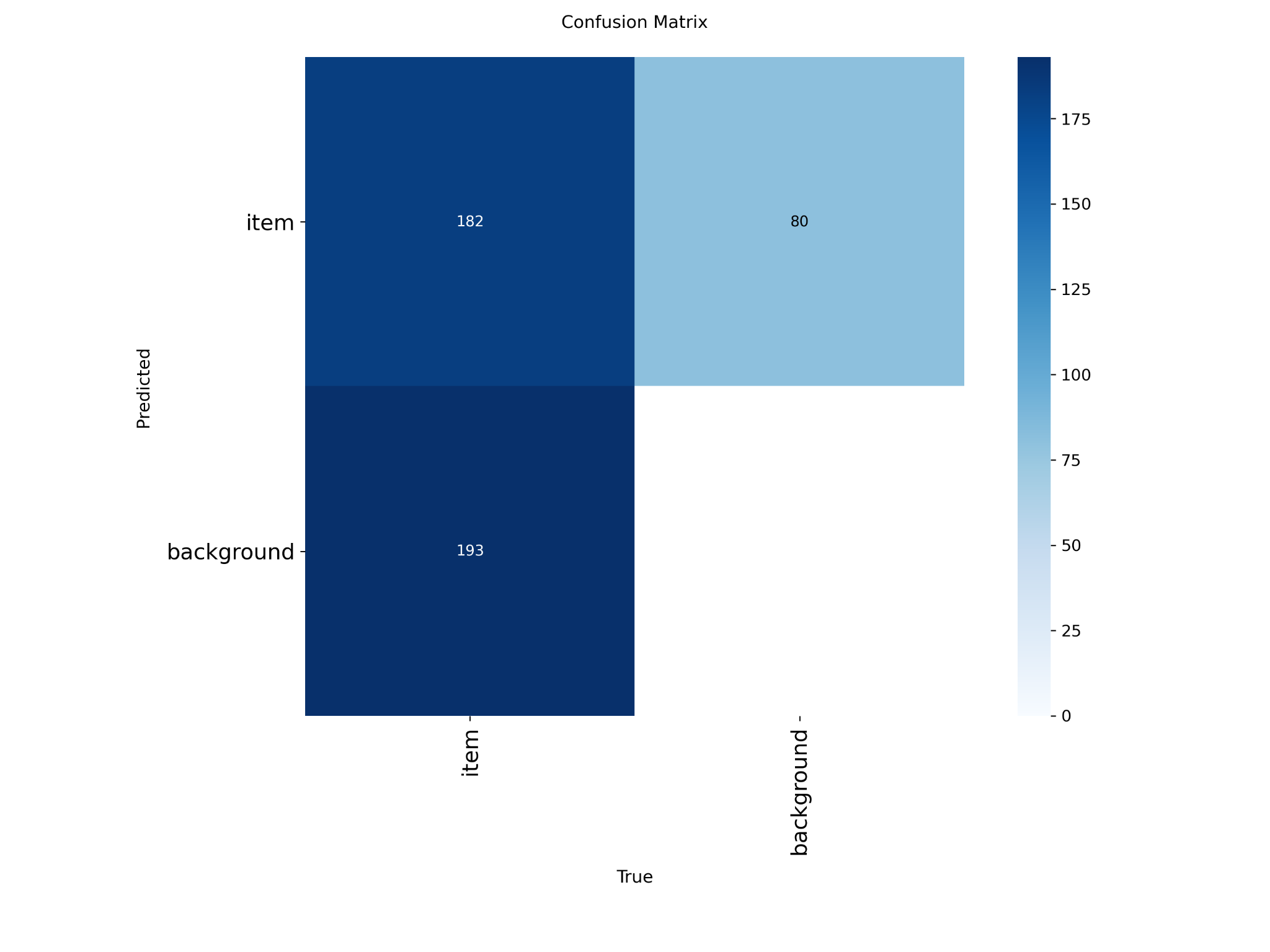
We implemented two types of augmentation:

1. Standard Augmentation (SA): few simple adjustments such as illumination, geometric shifts, flips, and mild erasing
2. Competition Augmentation (CA): more complex adjustments such as gauss noise, brightness contrast, blurring, cropping, and flipping

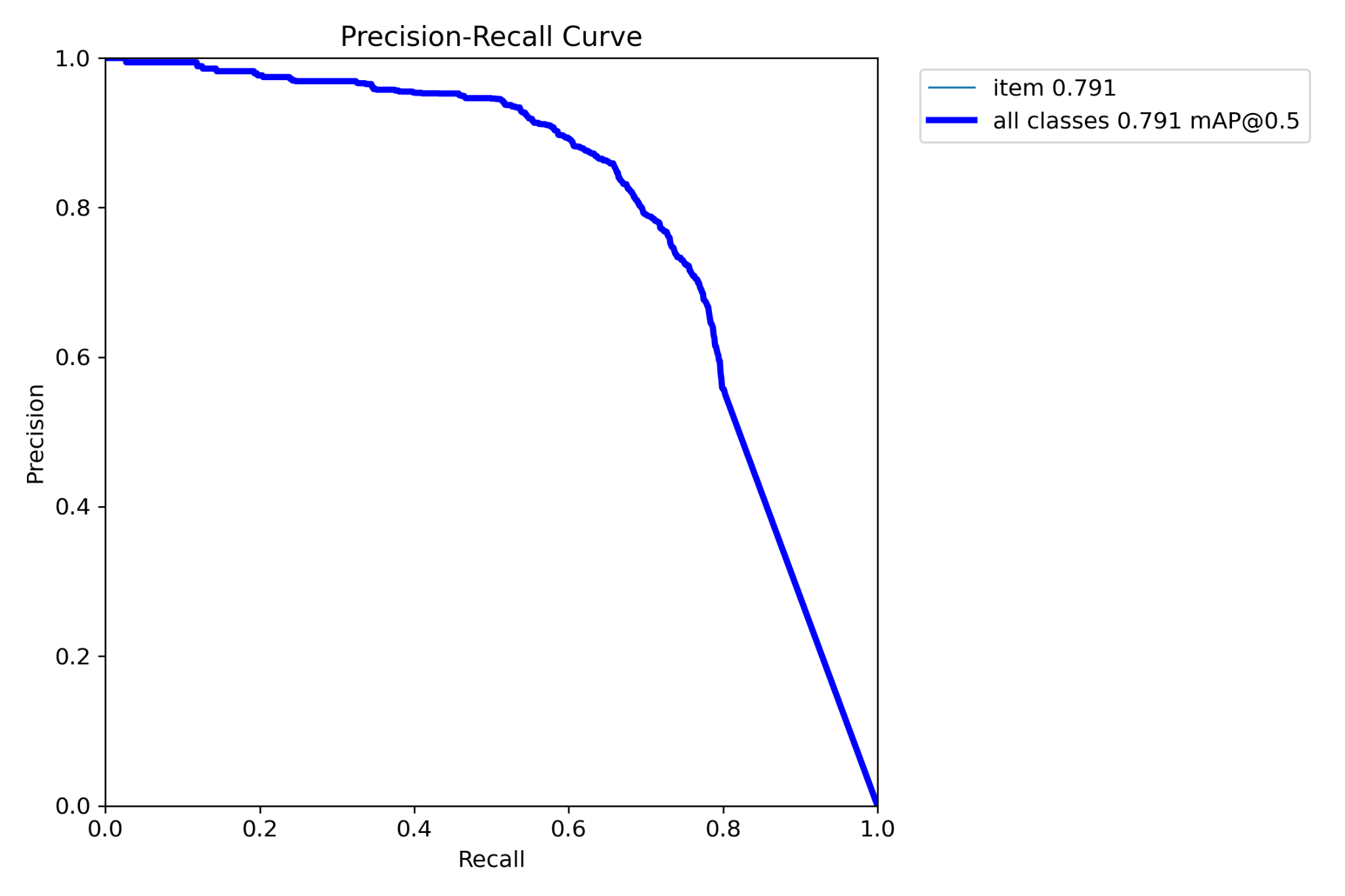
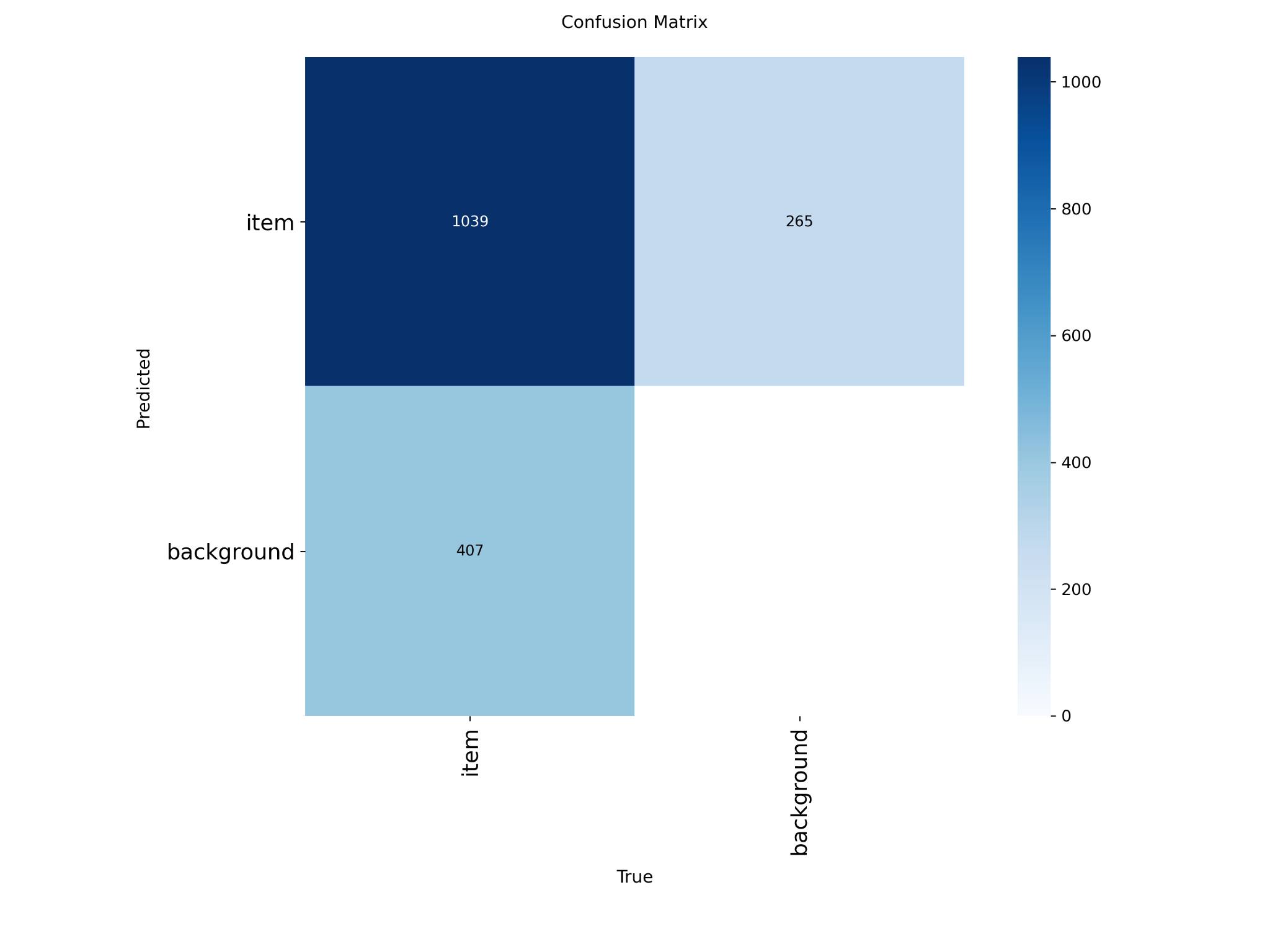
| **Data and Augmentation** | **Precision** | **Recall** | **mAP @ 50** | **mAP @ 50-95** |
| --- | --- | --- | --- | --- |
| Raw Data, SA | 0.686 | 0.488 | 0.578 | 0.439 |
| Extra Data, SA | 0.775 | 0.718 | 0.791 | 0.596 |
| Extra Data, CA | 0.828 | 0.84 | 0.821 | 0.637 |

Here and in the following page, we display selected critical statistics for the training stage and the results.

Original Data, Simple Augmentation:



Additional Val Data, Simple Augmentation:



Additional Val Data, Competition Augmentation:

