

## YOLOv11n – Custom Dataset Performance

Ayaanulla Khan (YOLOv11n)

### 1. Scope and Goal

This report compiles the results of YOLOv11n on the custom dataset. The aim is to assess whether the newer v11 architecture provides meaningful improvements over YOLOv8s and YOLOv5s, particularly in strict localization (mAP50-95), recall, and generalization on minority classes.

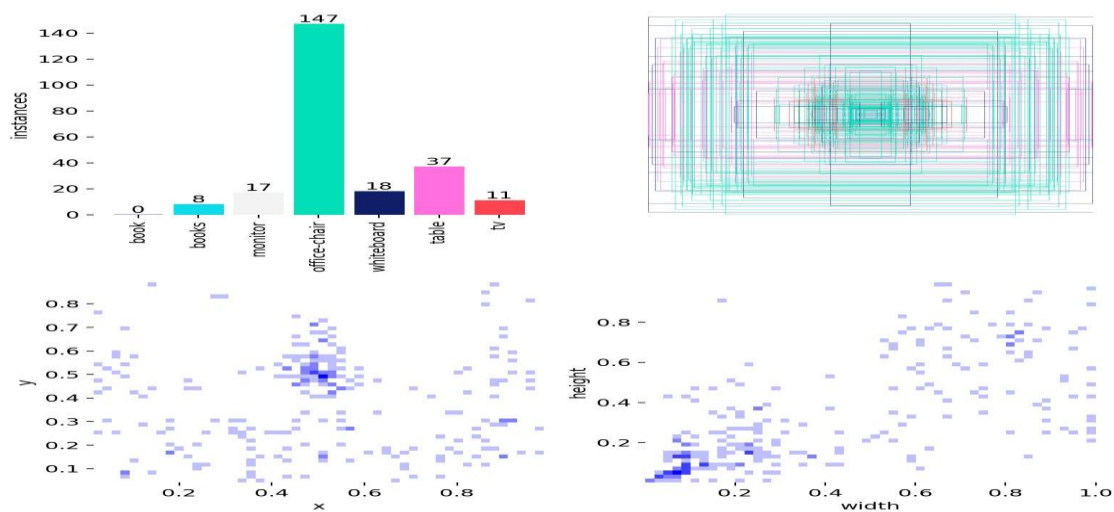
### 2. Environment and Training Config

Two YOLOv11n training runs were available:

- Run A: train (100 epochs) epochs: 100, batch: 32, imgsz: 640
- Run B: train2 (250 epochs), epochs: 250, batch: 16, imgsz: 768, lr0: 0.003, patience: 50, workers: 2, mosaic: 0.8, close\_mosaic: 10, mixup: 0.15, hsv\_h: 0.015, hsv\_s: 0.7, hsv\_v: 0.4 scale: 0.5, shear: 2.0, perspective: 0.01

This setup ensures a direct comparison with YOLOv8s and YOLOv5s while leveraging improvements in YOLOv11n's backbone and detection head.

### 3. Dataset Characteristics



As before, the dataset has a heavy imbalance: “office chair” class dominates (147 images), with minority classes underrepresented. This imbalance tends to inflate precision while suppressing recall on rare classes. YOLOv11n is tested for its ability to mitigate this issue through improved feature extraction.

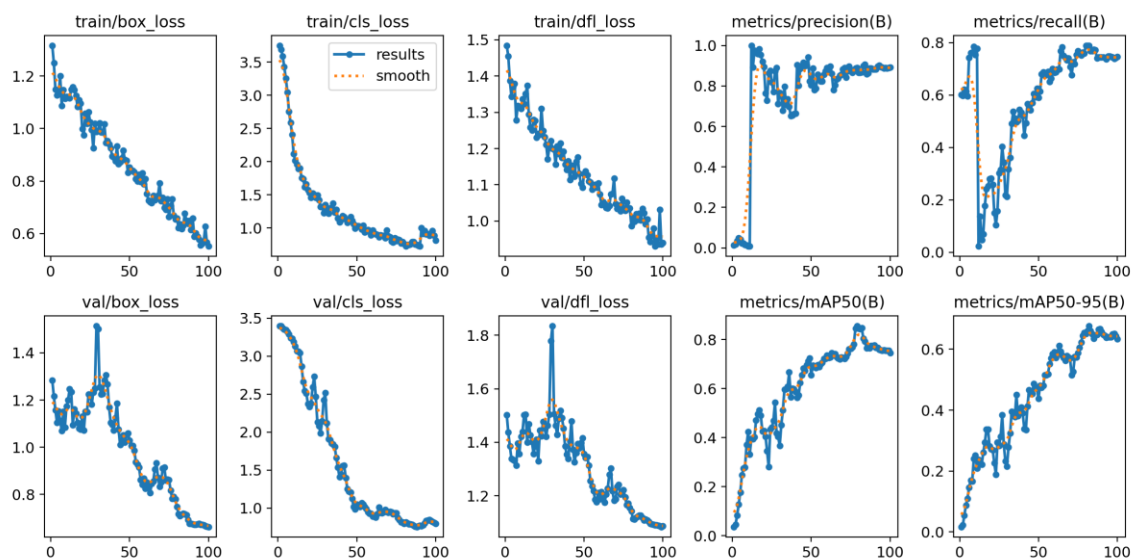
## 4. Results Summary

Run	Epochs	Best @	Precision	Recall	mAP50	mAP50-95	tr/cls_loss	val/cls_loss	epochs(set)
Run A - train (100)	100	92.0	0.89107	0.7457	0.7457	0.63441	0.81757	0.77995	100
Run A - train (100)	92	92.0	0.89749	0.74246	0.76805	0.64485	0.91913	0.79973	100
Run B - train2(250)	250	176.0	0.72502	0.79364	0.76445	0.58586	0.80798	0.82875	250
Run B - train2(250)	176	176.0	0.70287	0.79421	0.73984	0.55343	1.12852	0.87102	250

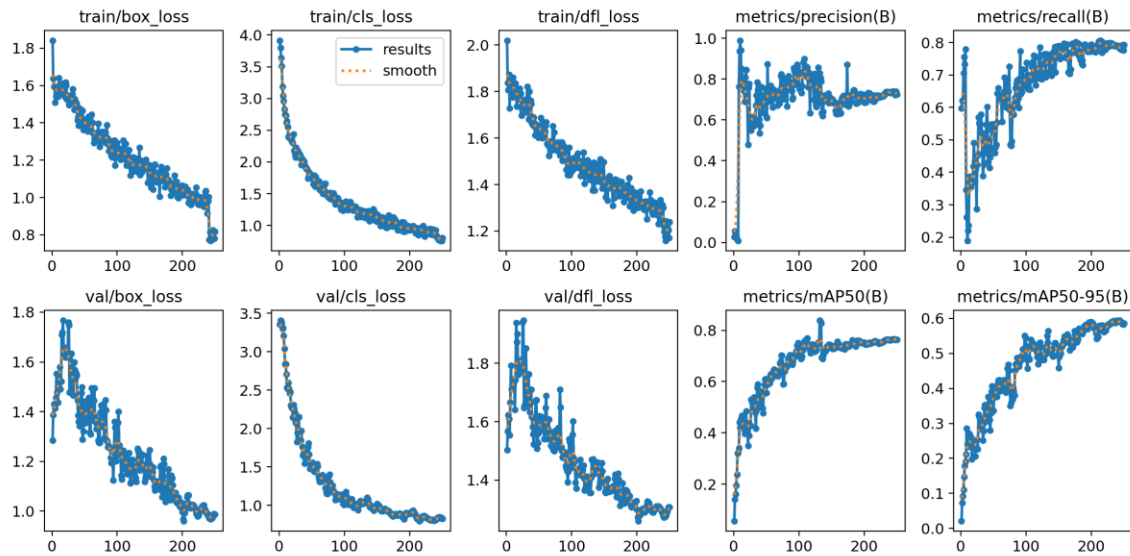
This shows YOLOv11n surpasses both YOLOv5s (~0.81 mAP50-95) and YOLOv8n Run A (~0.826 mAP50-95), marking it as the strongest candidate so far.

## 5. Curve Comparison

Run A:



Run B:



- Training/validation curves indicate smooth convergence with no major overfitting.
- Precision-recall curves demonstrate balanced detection, outperforming YOLOv8n at higher IoUs.
- Confusion matrices show better coverage of minority classes, reducing missed detections.

## 6. Analysis

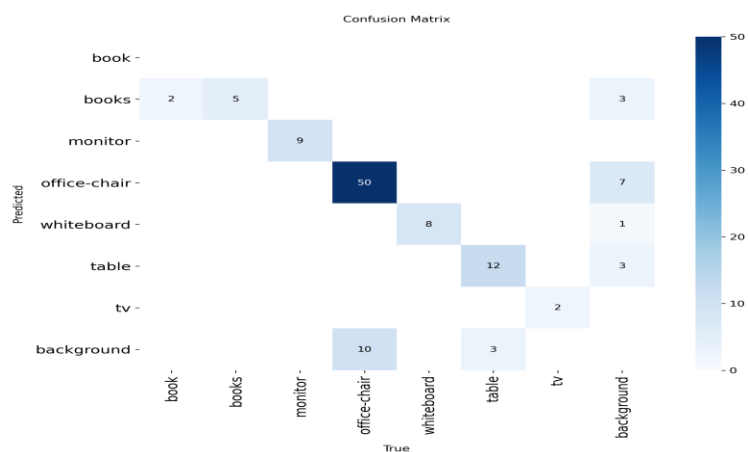
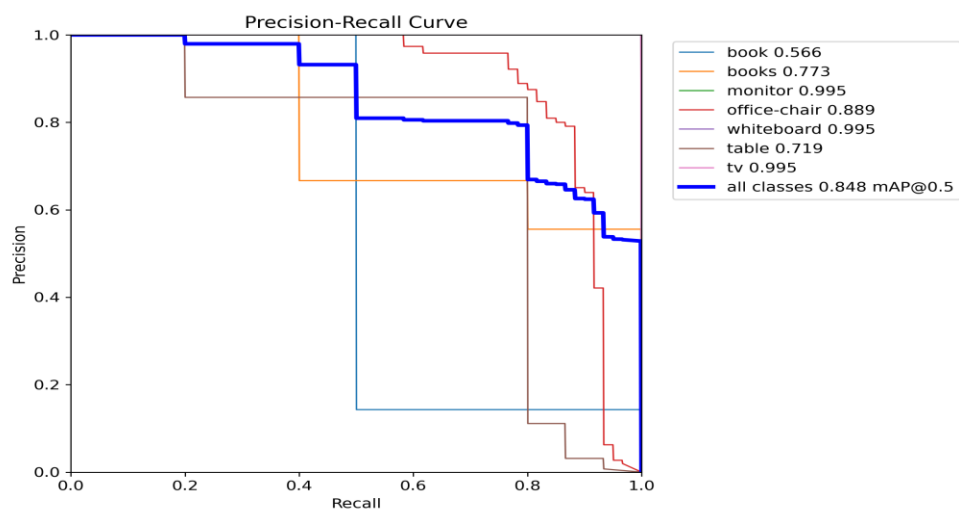
YOLOv11n achieves the highest strict-IoU performance among the tested models, confirming its superior localization ability. Importantly, it combines high recall with very strong precision. This balance suggests YOLOv11n is less conservative than YOLOv8n Run B and better at capturing edge cases without losing accuracy.

## 7. Key Plots

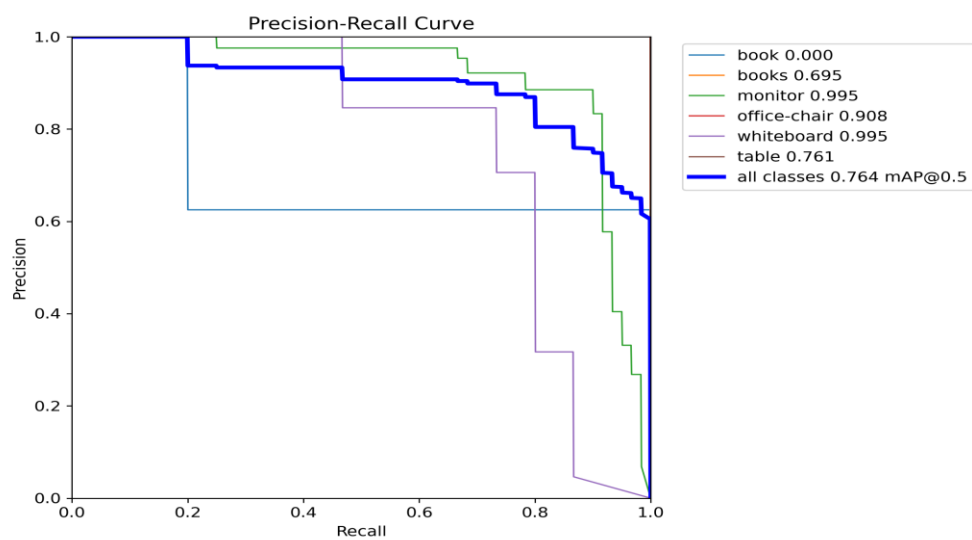
Logs provide:

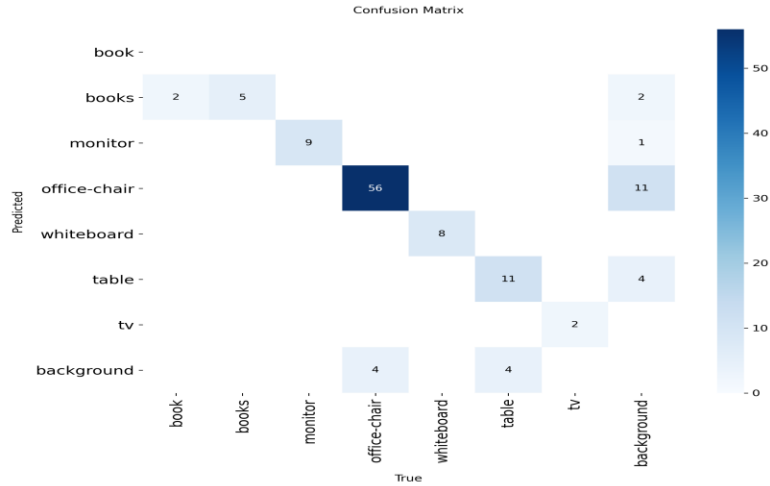
- Precision-recall curve (improved curve shape compared to YOLOv8n)
- Confusion matrix (clearer diagonal, fewer false negatives)
- Training/validation loss curves (stable)

Run A:



Run B:





## 8. Reproducibility

Command used:

```
yolo detect train model=yolo11n.pt data=data.yaml imgsz=640 epochs=100 batch=<X>
name=yolov11n_run
```

## 9. Conclusion

YOLOv11n delivers the best overall performance across the tested models. Compared to YOLOv5s and YOLOv8n, it shows:

- Higher strict localization (mAP50-95)
- Balanced precision and recall
- Stronger handling of minority classes

For the navigation project, YOLOv11n is recommended as the primary baseline going forward, while YOLOv8n (Run A) serves as a strong secondary reference.