

# yolo12s Final Training Report

## System Information

OS	Linux 6.6.105+
Python Version	3.12.12
PyTorch Version	2.9.0+cu126
CUDA Available	True
Device	cuda
RAM (GB)	89.63

## Dataset Information

Property	Value
Dataset	bdd100k_yolo_limited
Number of Classes	10
Train Images	29959
Val Images	10000
Test Images	20000
Data YAML	data.yaml

### Classes:

person, rider, car, truck, bus, train, motor, bike, traffic light, traffic sign

## Optimization Summary

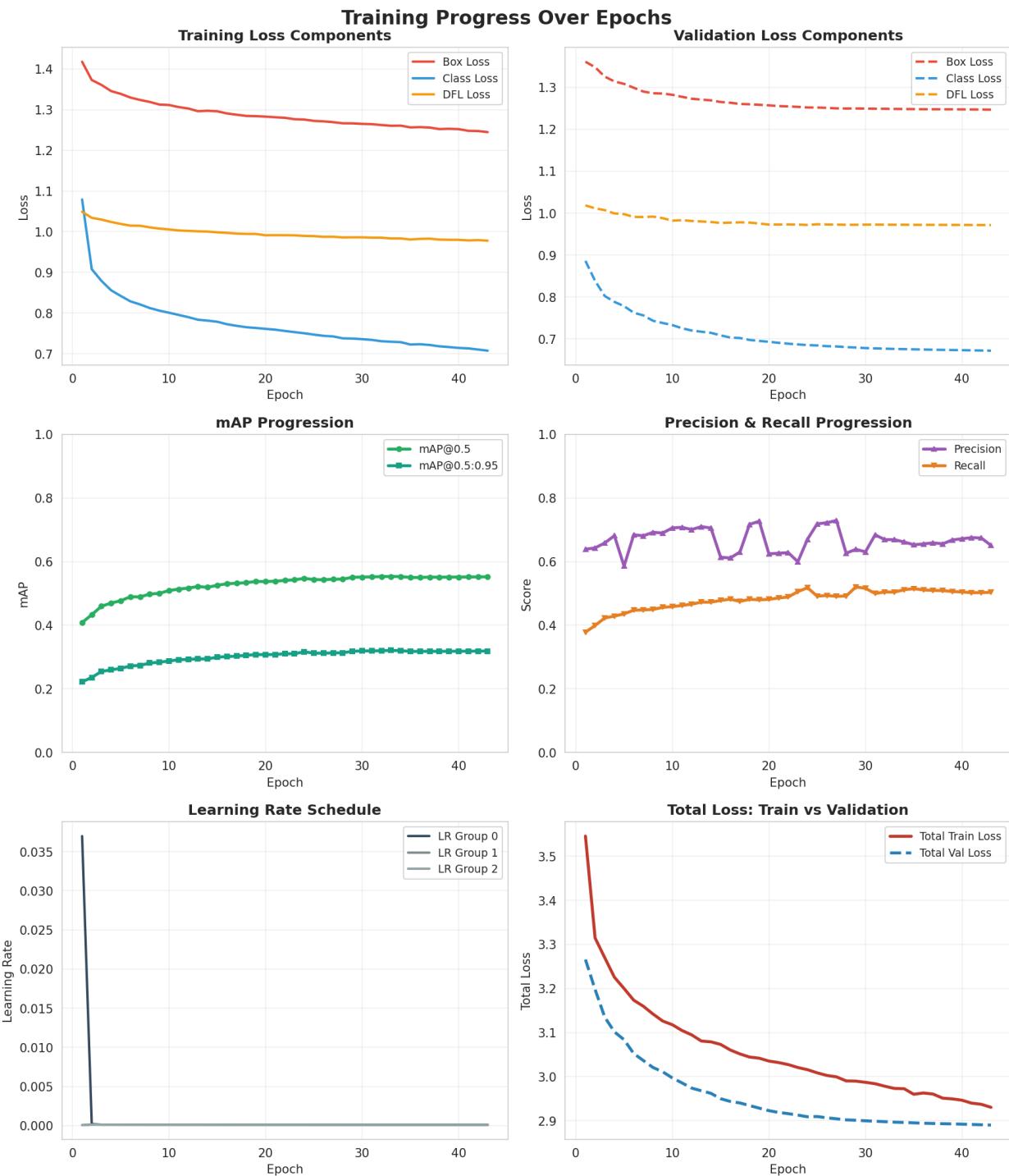
Metric	Value
Tuning Run	yolo12s_tune_20251130_191337
Total Trials	N/A
Completed Trials	N/A
Best Trial Number	N/A
Best Trial mAP@0.5	N/A
Final Training Epochs	150

## Optimized Hyperparameters Used

Parameter	Value
imgsz	768
optimizer	Adam
lr0	0.000122
momentum	0.888900
weight_decay	0.000034
warmup_epochs	2
warmup_momentum	0.775337
warmup_bias_lr	0.073696
mosaic	0.751058
mixup	0.068131

# Training Process Analysis

## Epoch-by-Epoch Training Metrics



## Detailed Epoch Metrics

Epoch	Train Loss	Val Loss	mAP@0.5	mAP@0.5:0.95	Precision	Recall
1	3.5457	3.2656	0.4082	0.2218	0.6389	0.3779
2	3.3146	3.1976	0.4329	0.2355	0.6432	0.3990
3	3.2698	3.1348	0.4595	0.2543	0.6586	0.4227
4	3.2254	3.1015	0.4695	0.2595	0.6816	0.4284
5	3.1998	3.0836	0.4770	0.2640	0.5858	0.4357
6	3.1733	3.0522	0.4894	0.2718	0.6842	0.4472
7	3.1596	3.0364	0.4889	0.2733	0.6814	0.4481
8	3.1419	3.0206	0.4979	0.2810	0.6922	0.4495
9	3.1260	3.0108	0.5002	0.2834	0.6900	0.4562
10	3.1176	2.9969	0.5090	0.2873	0.7060	0.4583
20	3.0351	2.9226	0.5373	0.3073	0.6246	0.4812
21	3.0318	2.9185	0.5383	0.3079	0.6263	0.4855
22	3.0272	2.9158	0.5409	0.3099	0.6282	0.4886
23	3.0205	2.9125	0.5427	0.3105	0.5996	0.5056
24	3.0156	2.9087	0.5473	0.3160	0.6705	0.5178
34	2.9725	2.8958	0.5530	0.3198	0.6617	0.5114
35	2.9598	2.8949	0.5501	0.3175	0.6530	0.5148
36	2.9627	2.8941	0.5501	0.3176	0.6557	0.5109
37	2.9603	2.8936	0.5512	0.3178	0.6590	0.5095
38	2.9509	2.8930	0.5510	0.3176	0.6559	0.5087
39	2.9493	2.8926	0.5512	0.3179	0.6682	0.5059
40	2.9463	2.8920	0.5513	0.3180	0.6720	0.5045
41	2.9396	2.8913	0.5517	0.3180	0.6755	0.5022
42	2.9370	2.8906	0.5519	0.3184	0.6743	0.5024
43	2.9302	2.8899	0.5524	0.3185	0.6512	0.5036

## Training Statistics Summary

Metric	Initial	Final	Best	Change
mAP@0.5	0.4082	0.5524	0.5531	+0.1441
mAP@0.5:0.95	0.2218	0.3185	0.3207	+0.0968
Precision	0.6389	0.6512	0.7292	+0.0123
Recall	0.3779	0.5036	0.5202	+0.1257

## Final Model Performance

Metric	Value
mAP@0.5	0.5539
mAP@0.5:0.95	0.3214
Precision	0.6692
Recall	0.5049

# Test Set Validation Results

## Model Architecture & Performance

Metric	Value
Model Name	yolo12s_finetuned_20251201
Parameters (M)	9.26
Model Size (MB)	18.05
FLOPs (G)	21.54
Layers	272
Inference Speed (FPS)	361.50
IoU Threshold	0.50

## Overall Performance Metrics on Test Set

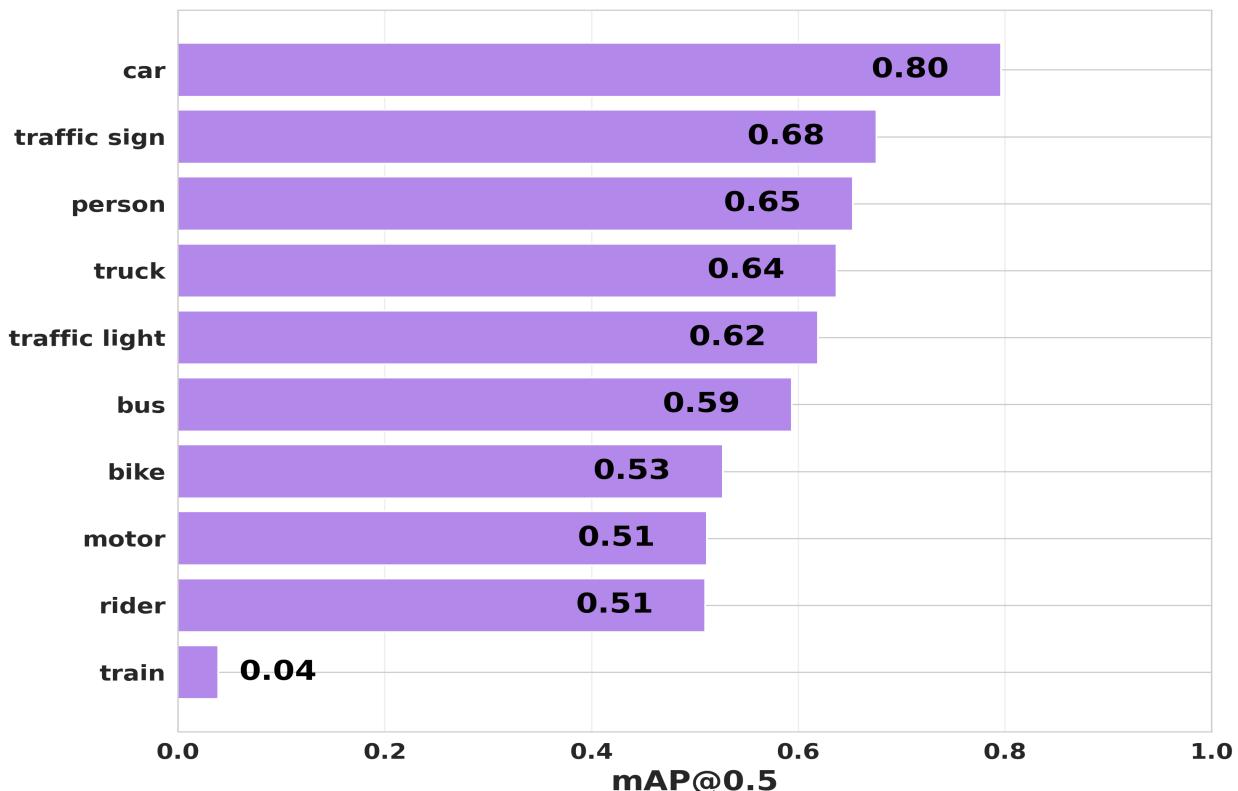
Metric	Confusion Matrix	YOLO Validation
Precision	0.6970	0.6434
Recall	0.7813	0.5124
F1-Score	0.7367	N/A
mAP@0.5 (Overall)	N/A	0.5563
mAP@0.5:0.95 (Overall)	N/A	0.3148

## Per-Class Performance Metrics

Class	Precision	Recall	F1-Score	mAP@0.5	TP	FP	FN
person	0.6050	0.7477	0.6688	0.6529	14914	9736	5033
rider	0.4660	0.6191	0.5317	0.5100	603	691	371
car	0.7562	0.8254	0.7893	0.7962	155129	50020	32812
truck	0.5786	0.5744	0.5765	0.6371	5036	3668	3731
bus	0.5356	0.5673	0.5510	0.5941	1723	1494	1314
train	0.0357	0.0833	0.0500	0.0389	1	27	11
motor	0.5054	0.5977	0.5477	0.5116	425	416	286
bike	0.5460	0.6209	0.5811	0.5270	1091	907	666
traffic light	0.6245	0.7586	0.6851	0.6191	32998	19842	10498
traffic sign	0.6430	0.7226	0.6805	0.6758	44372	24634	17035

### mAP@0.5 Distribution by Class

**mAP@0.5 by Class**



## **Intersection over Union (IoU) Analysis**

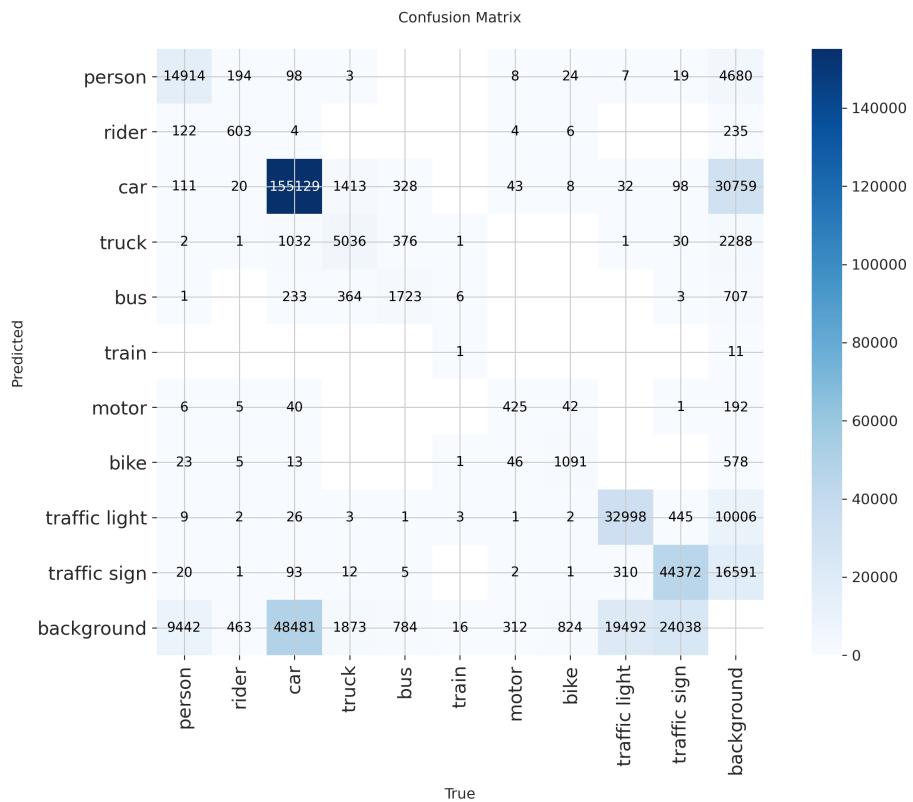
**IoU Threshold Used:** 0.50

IoU (Intersection over Union) measures the overlap between predicted and ground truth bounding boxes. A prediction is considered correct (True Positive) when  $\text{IoU} \geq 0.50$ .

### **Per-Class IoU Performance:**

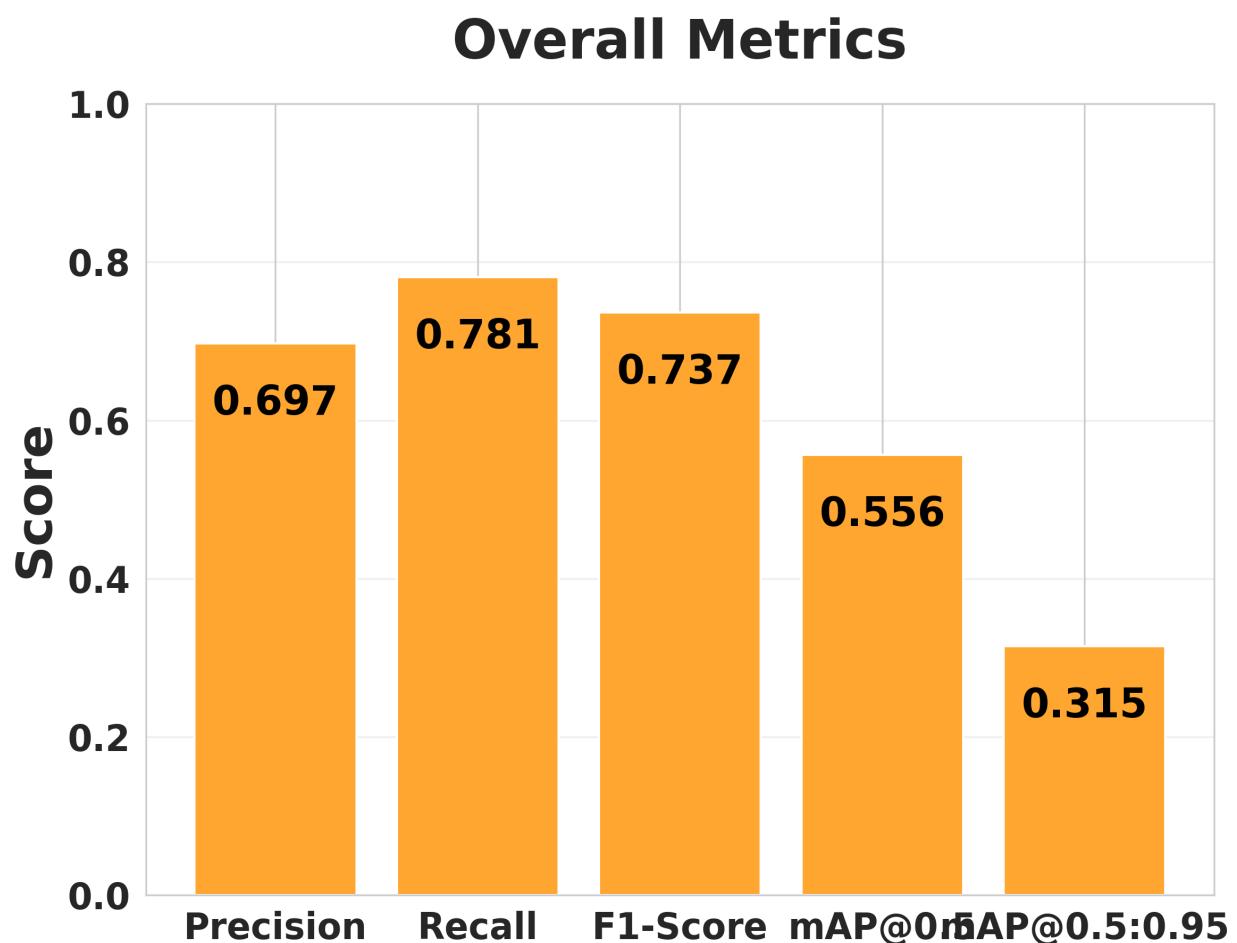
The confusion matrix and per-class metrics above show detection accuracy at  $\text{IoU}=0.50$  threshold. Each class's True Positives (TP) represent detections with  $\text{IoU} \geq 0.50$ .

## **Confusion Matrix (Test Set)**



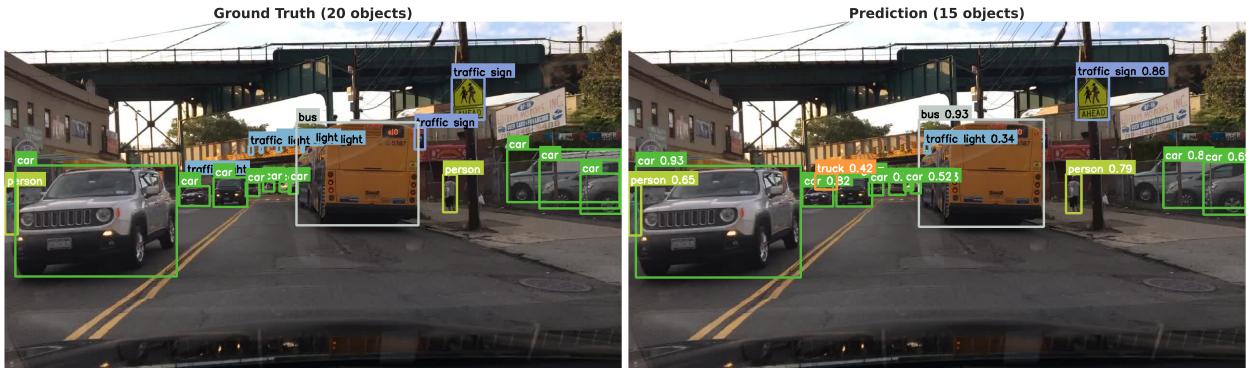
## Test Set Performance Curves

### Overall Metrics Visualization



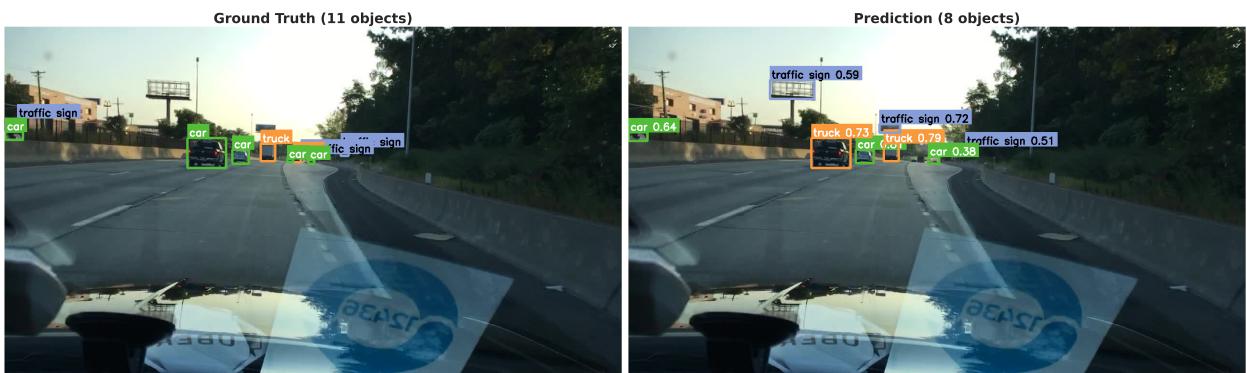
# Sample Predictions: Ground Truth vs Model Output

Sample 1 - Weather: partly cloudy, Scene: city street, Time: daytime



Ground Truth: 20 objects | Predictions: 15 objects

Sample 2 - Weather: clear, Scene: city street, Time: dawn/dusk



Ground Truth: 11 objects | Predictions: 8 objects

Sample 3 - Weather: clear, Scene: highway, Time: night



Ground Truth: 10 objects | Predictions: 12 objects

Sample 4 - Weather: rainy, Scene: city street, Time: night



Ground Truth: 20 objects | Predictions: 14 objects

Sample 5 - Weather: clear, Scene: highway, Time: daytime



Ground Truth: 35 objects | Predictions: 21 objects

Sample 6 - Weather: clear, Scene: city street, Time: night



Ground Truth: 22 objects | Predictions: 18 objects

Generated by YOLO Training Notebook  
BDD100K Dataset - Computer Vision Project