

# YOLO Model Testing Report

<b>Model:</b>	yolo11n
<b>Model Size:</b>	5.4 MB
<b>Parameters:</b>	2.6 M
<b>FLOPs (640x640):</b>	6.61 GFLOPs
<b>Run Name:</b>	yolo11n_testing_20251128_000944
<b>Timestamp:</b>	2025-11-28 00:22:45
<b>Dataset:</b>	bdd100k_yolo_limited - test split
<b>Images Processed:</b>	20000
<b>Total Execution Time</b>	100.05s
<b>IoU Threshold:</b>	0.5

## Inference Performance

Metric	Value
Average Inference Time	1.54 ms per image
FPS (Frames Per Second)	647.75

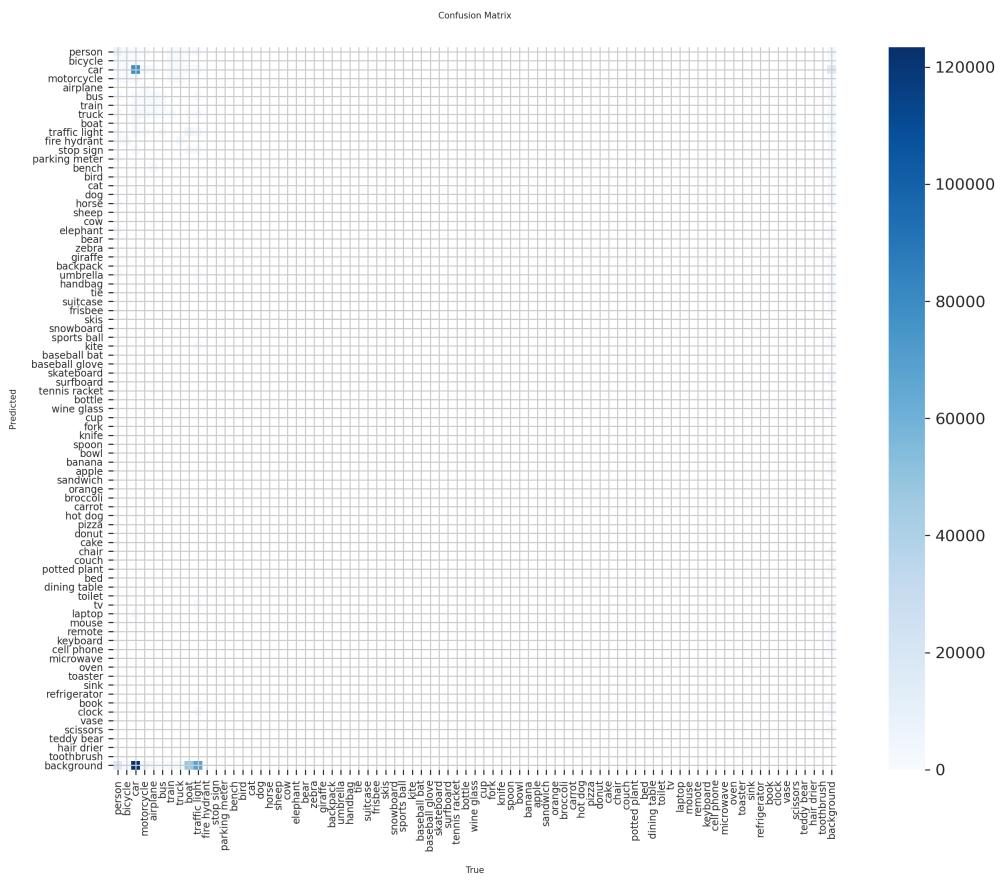
Note: Inference time includes preprocessing, model inference, and postprocessing.

## Overall Accuracy Metrics

Metric	Value
Precision	0.2362
Recall	0.6890
F1-Score	0.3518
mAP@0.5	0.0946
mAP@0.5:0.95	0.0491
True Positives	86859
False Positives	280868
False Negatives	39201

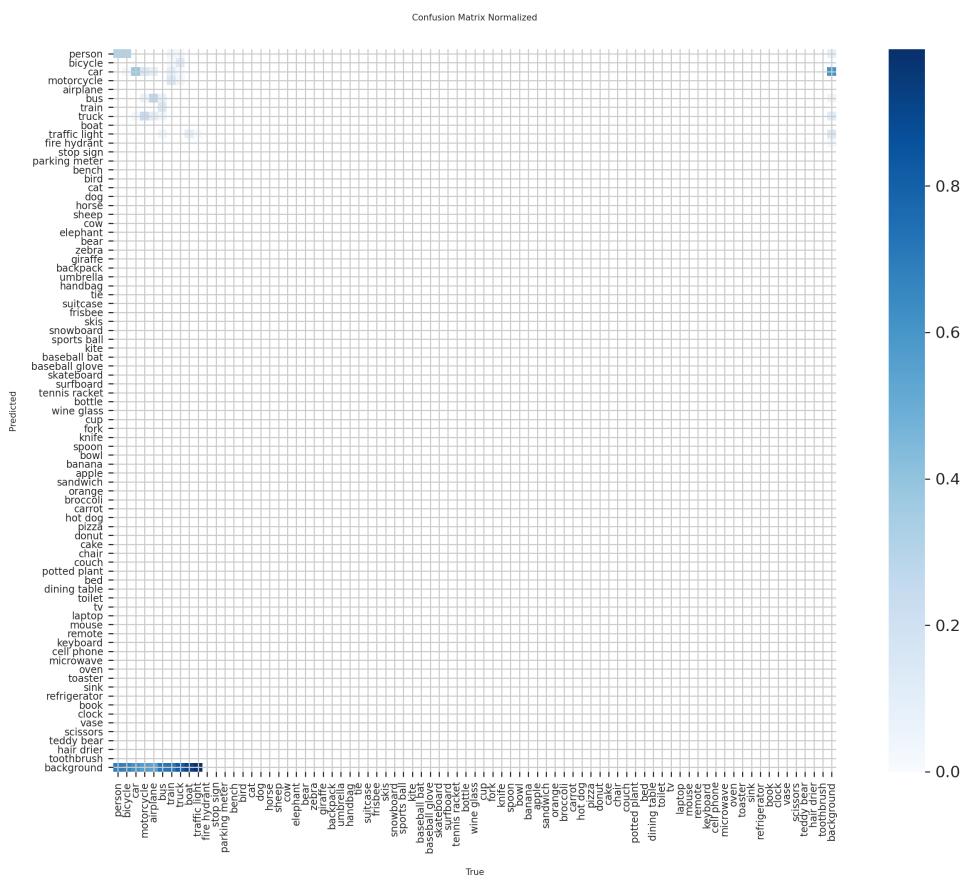
# Confusion Matrices

## Confusion Matrix (Counts)



Note: Shows raw prediction counts for each class combination.

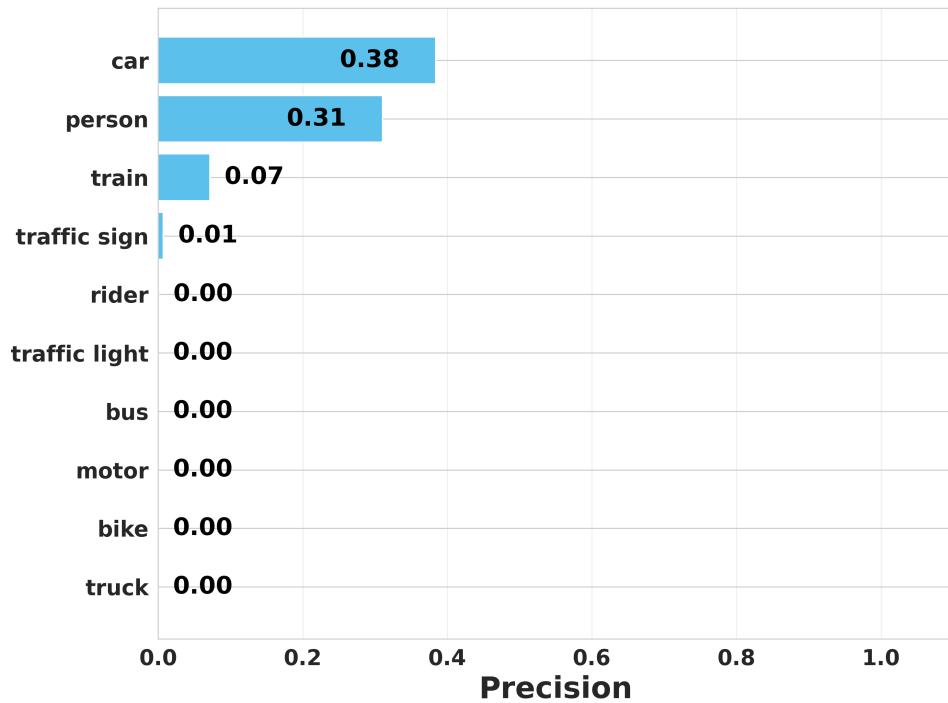
## Confusion Matrix (Normalized)



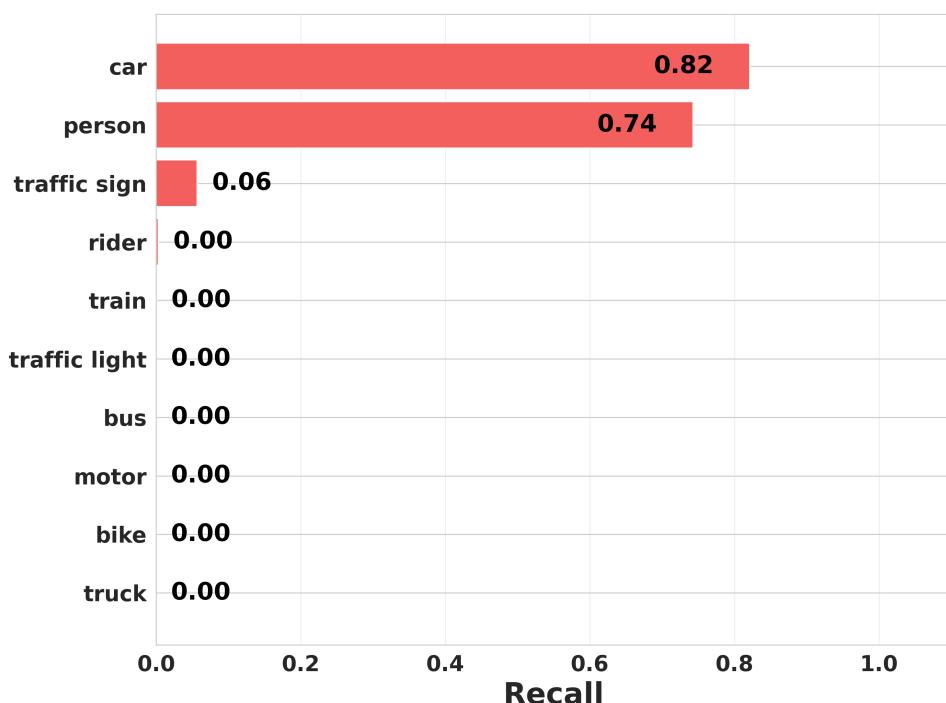
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

# Performance Visualizations

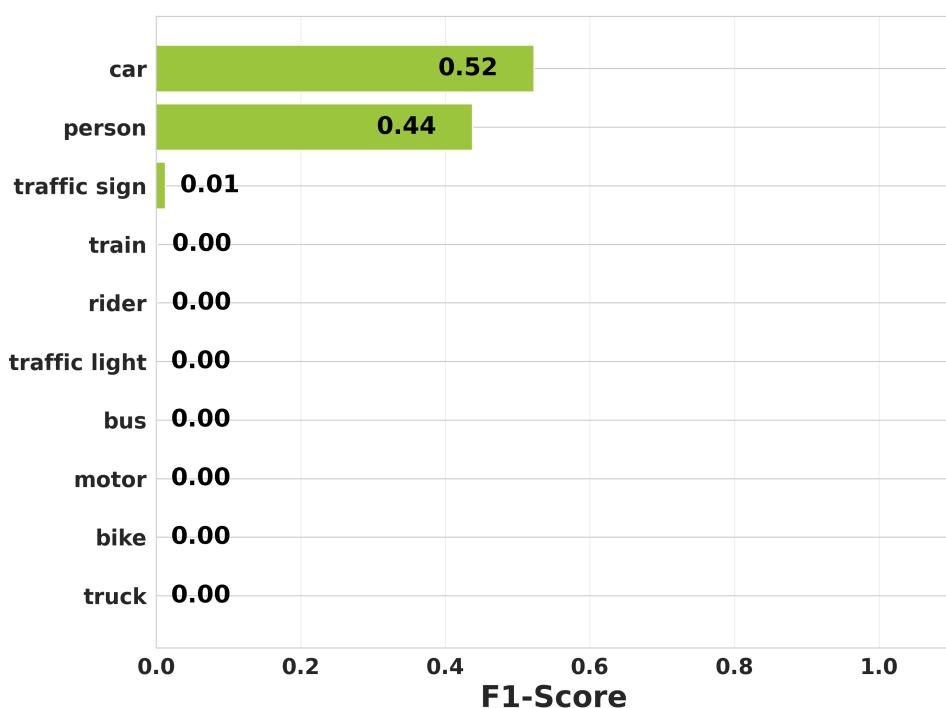
## Precision by Class



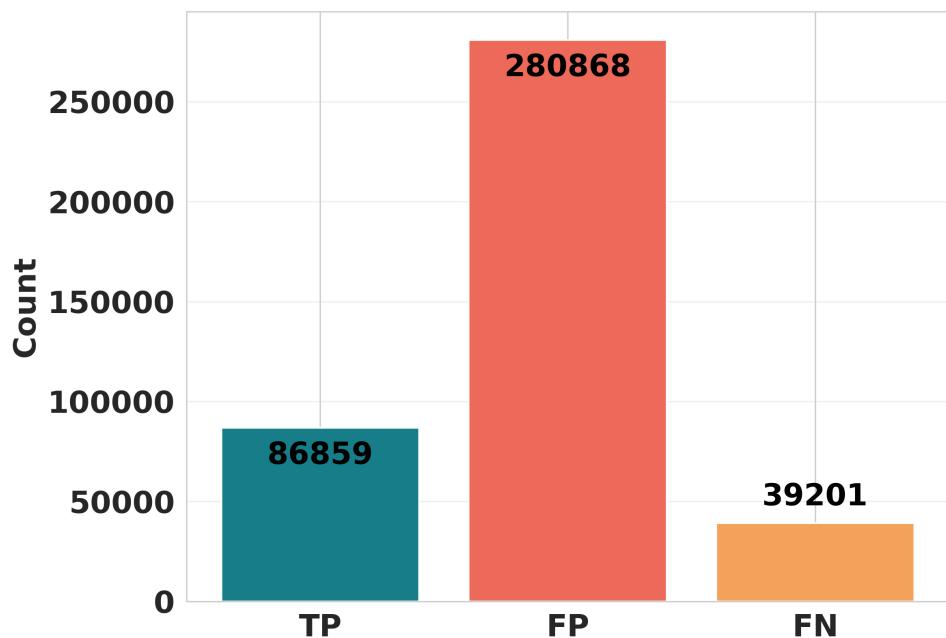
### Recall by Class



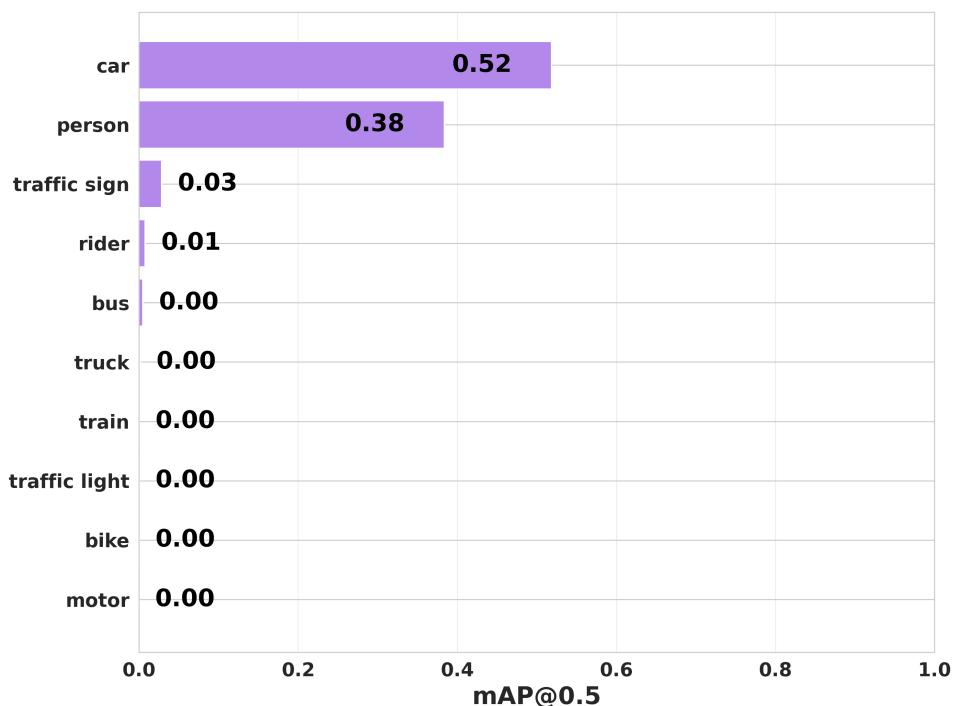
### F1-Score by Class



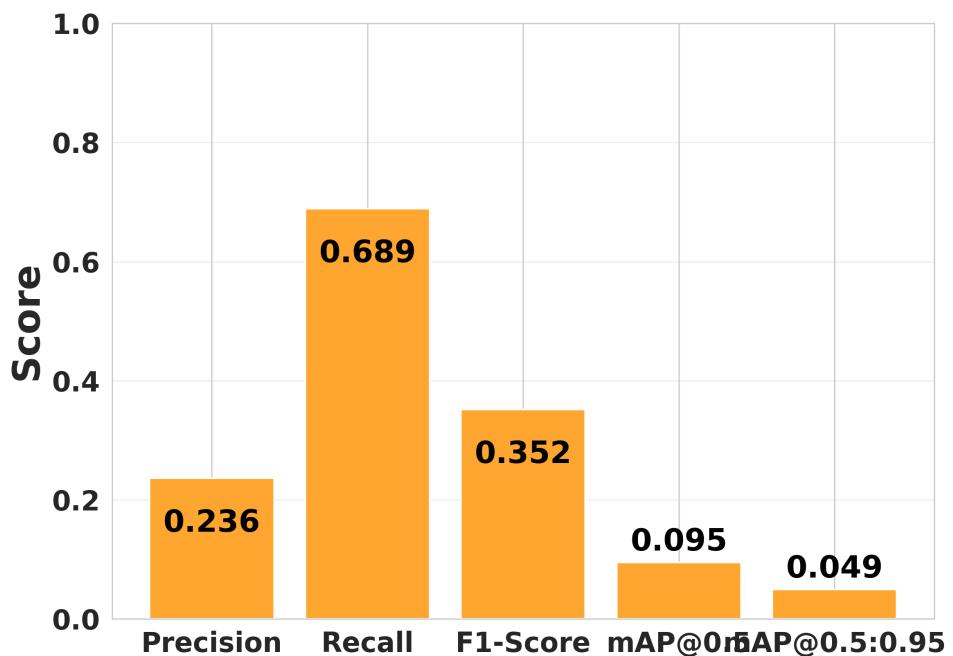
## Overall Detection Outcomes



## mAP@0.5 by Class



## Overall Metrics



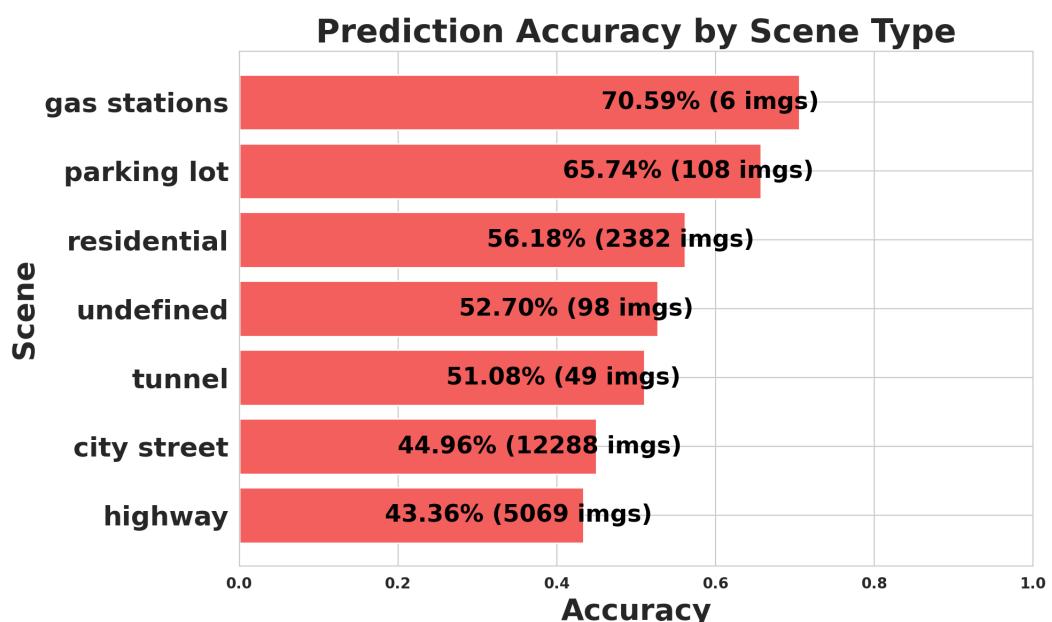
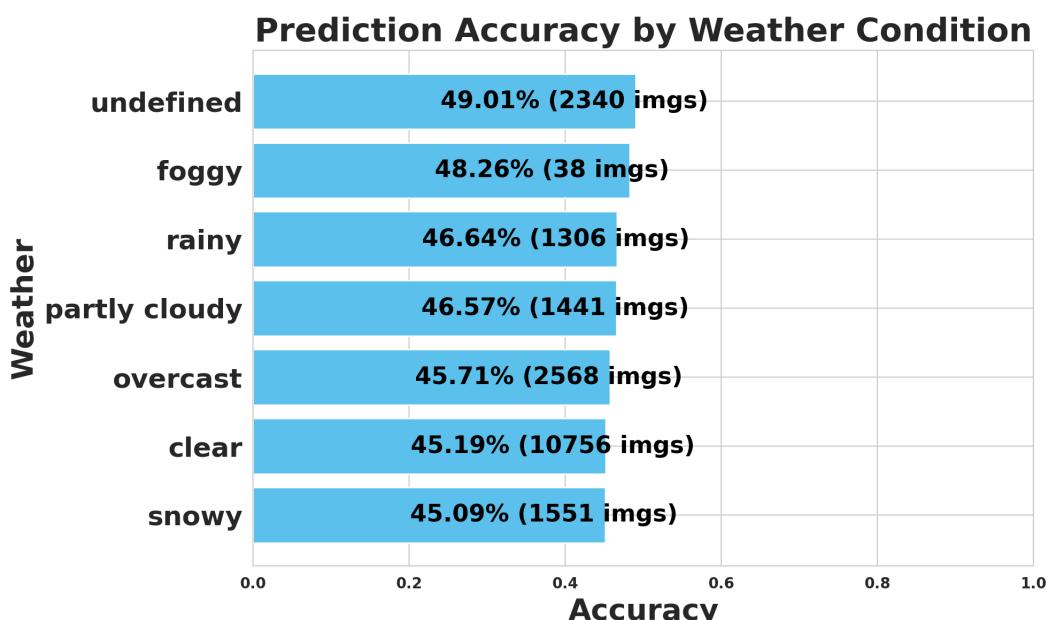
## Per-Class Performance

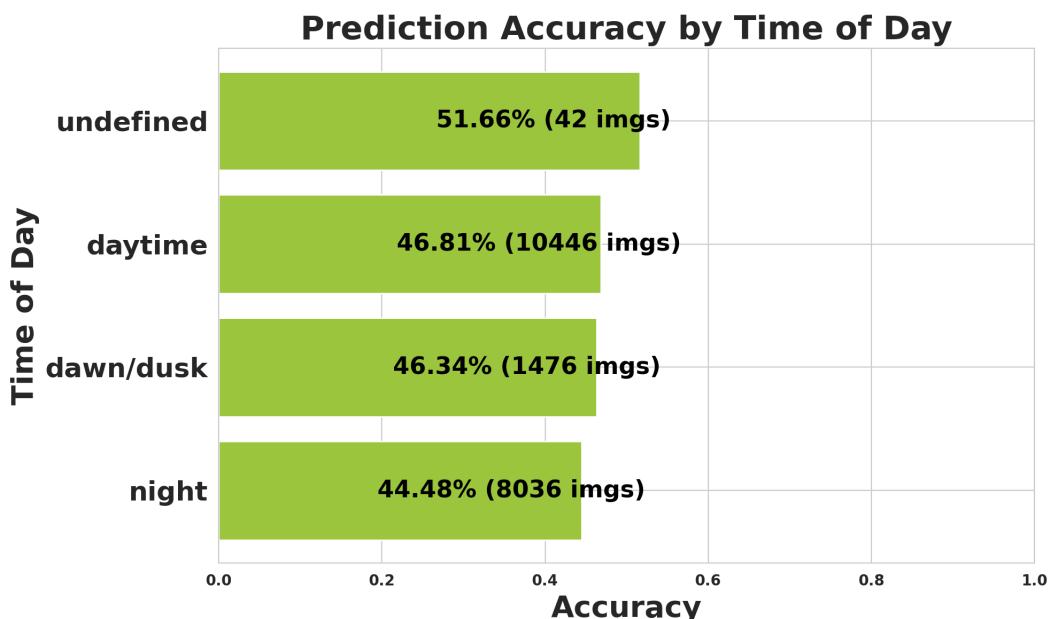
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	7652	16998	2650	0.3104	0.7428	0.4379	0.3838
rider	1	1293	324	0.0008	0.0031	0.0012	0.0078
car	78713	126436	17149	0.3837	0.8211	0.5230	0.5187
truck	0	8704	293	0.0000	0.0000	0.0000	0.0015
bus	0	3217	16	0.0000	0.0000	0.0000	0.0049
train	2	26	2851	0.0714	0.0007	0.0014	0.0003
motor	0	841	69	0.0000	0.0000	0.0000	0.0001
bike	0	1998	7701	0.0000	0.0000	0.0000	0.0002
traffic light	0	52840	7	0.0000	0.0000	0.0000	0.0002
traffic sign	491	68515	8141	0.0071	0.0569	0.0126	0.0283

# Comprehensive Failure Analysis

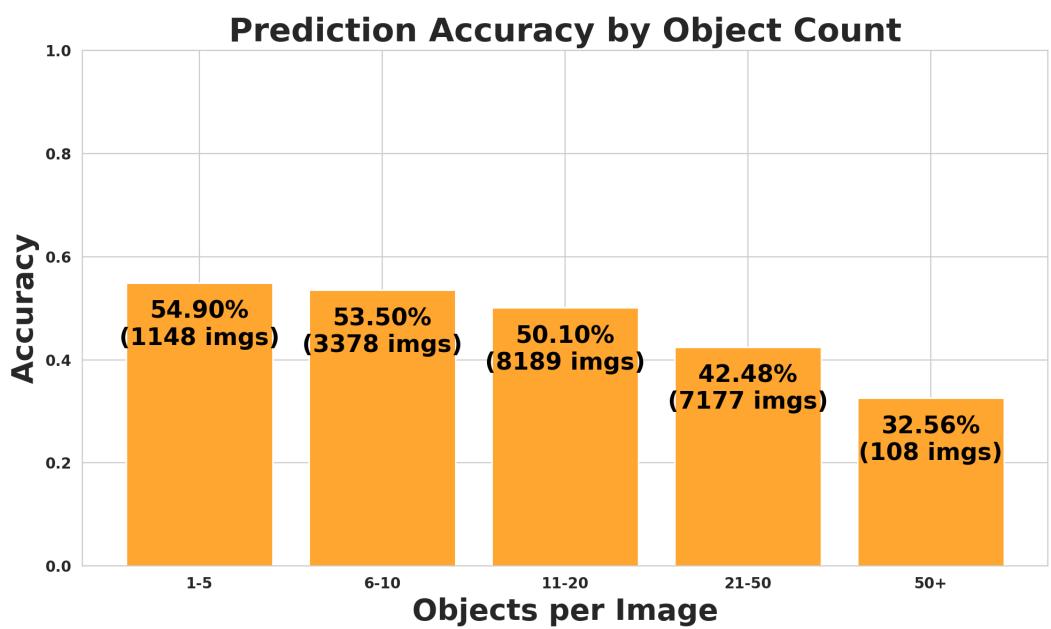
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	168953
Overall Accuracy	45.95%

## Accuracy by Environmental Attributes

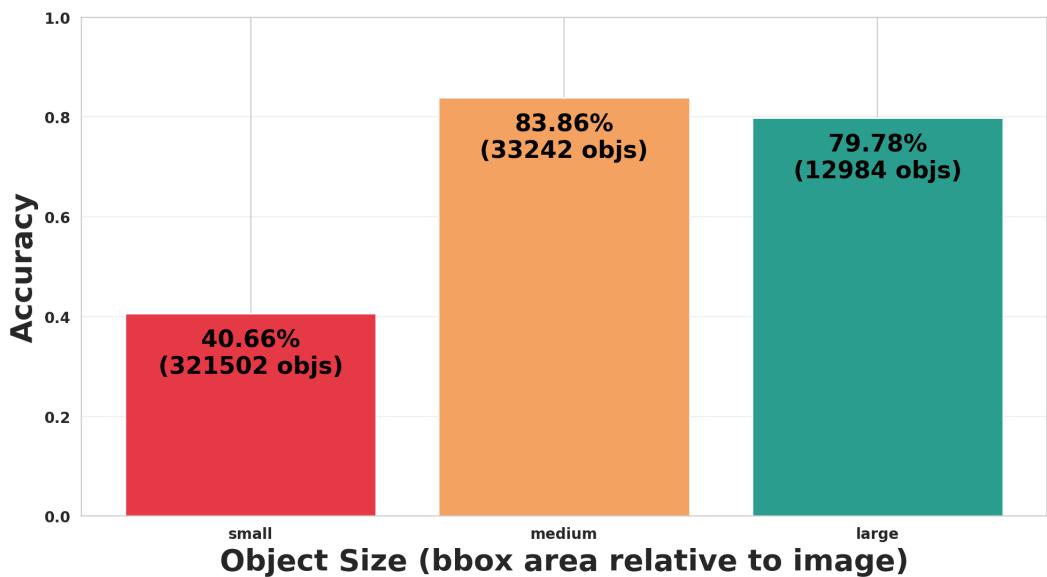




## Accuracy by Object Characteristics

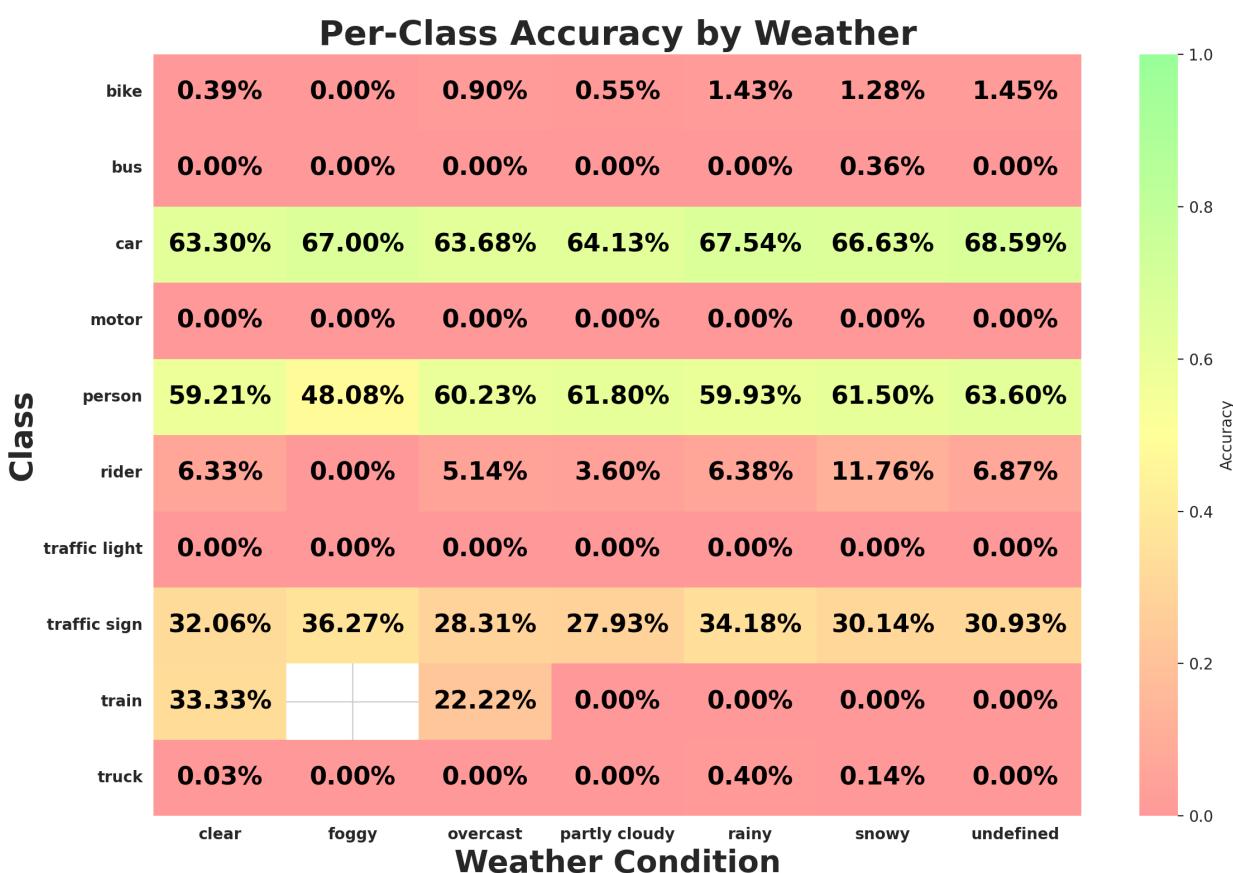
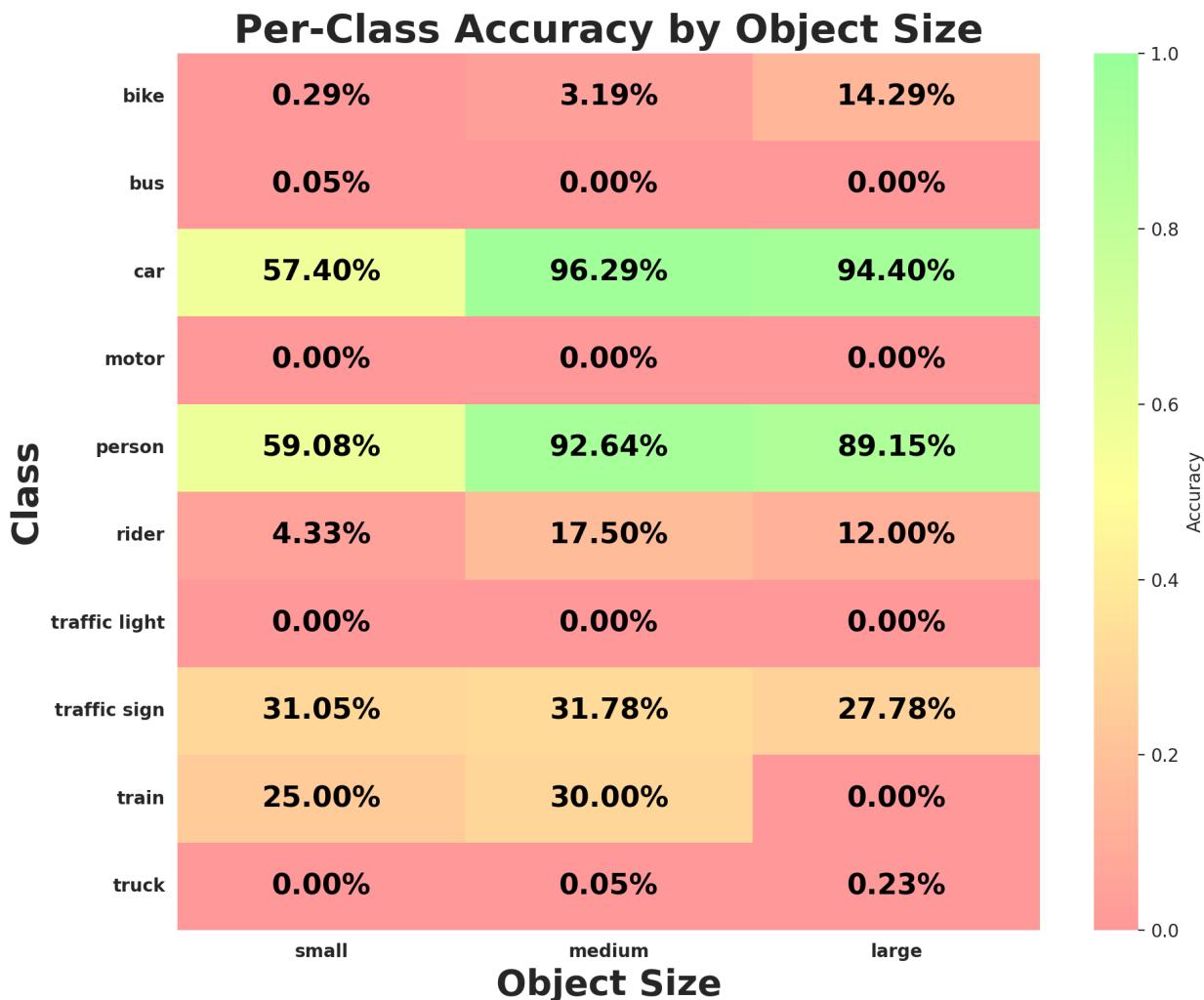


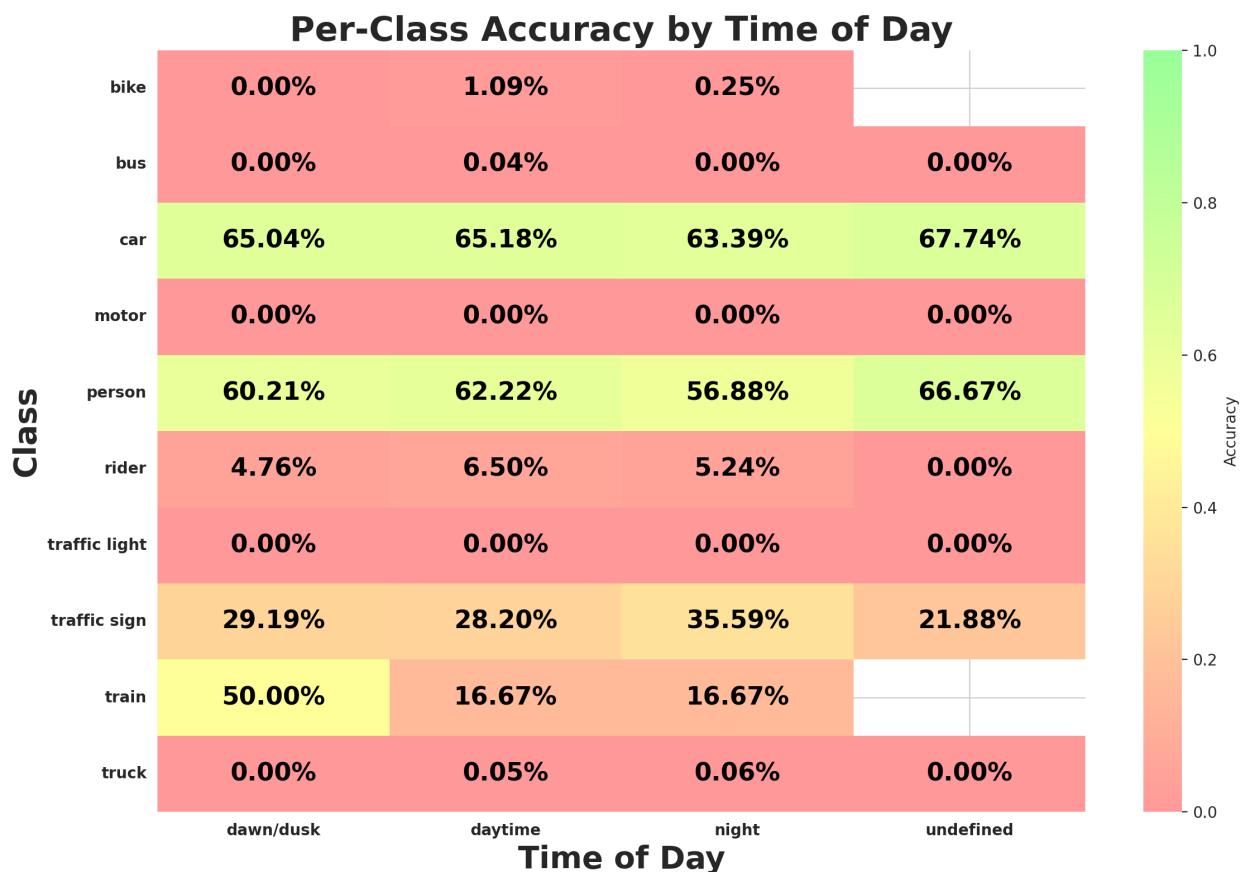
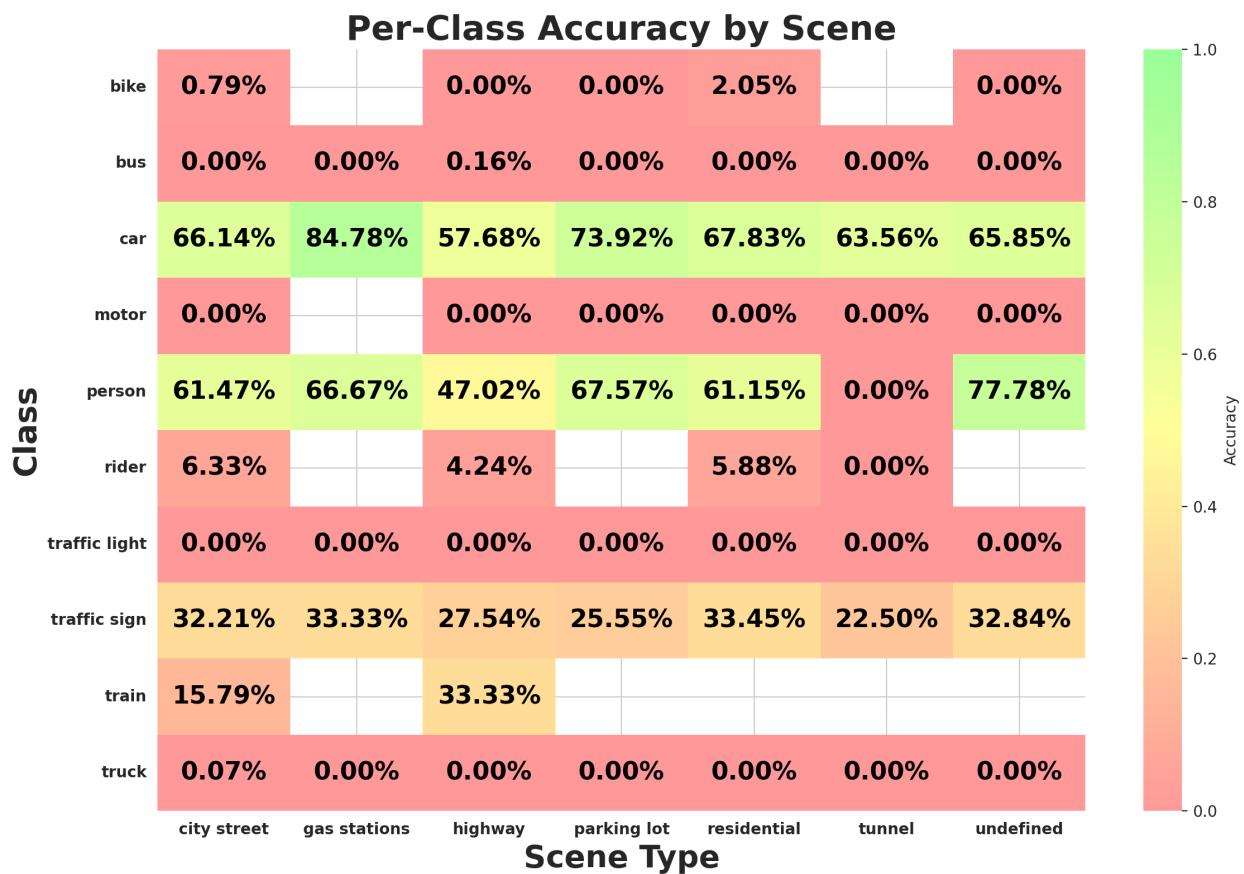
### Prediction Accuracy by Object Size



**Note:** Object size categories are defined based on bounding box area as a percentage of total image area: **Small** objects occupy less than 1% of the image (typically distant or small objects), **Medium** objects occupy between 1% and 5% of the image (moderate-sized objects at medium distances), and **Large** objects occupy more than 5% of the image (close or large objects). These thresholds help identify how object scale affects detection accuracy.

# Per-Class Performance Analysis

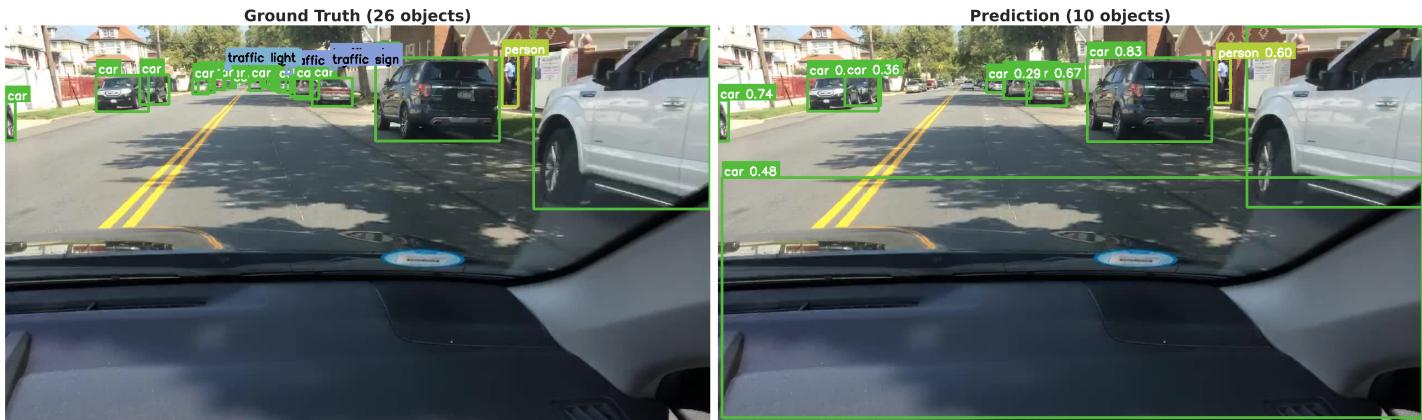




# Sample Predictions: Ground Truth vs Model

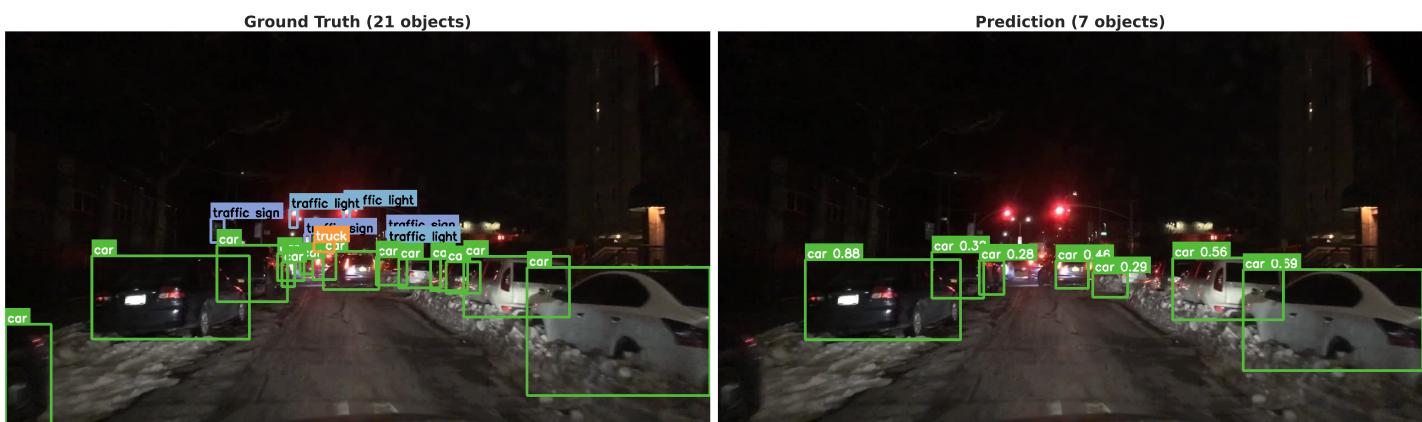
## Sample #1: f9e5e3f0-9b710929.jpg

Ground Truth Objects: 26 | Predicted Objects: 10  
Weather: clear | Scene: city street | Time of Day: daytime



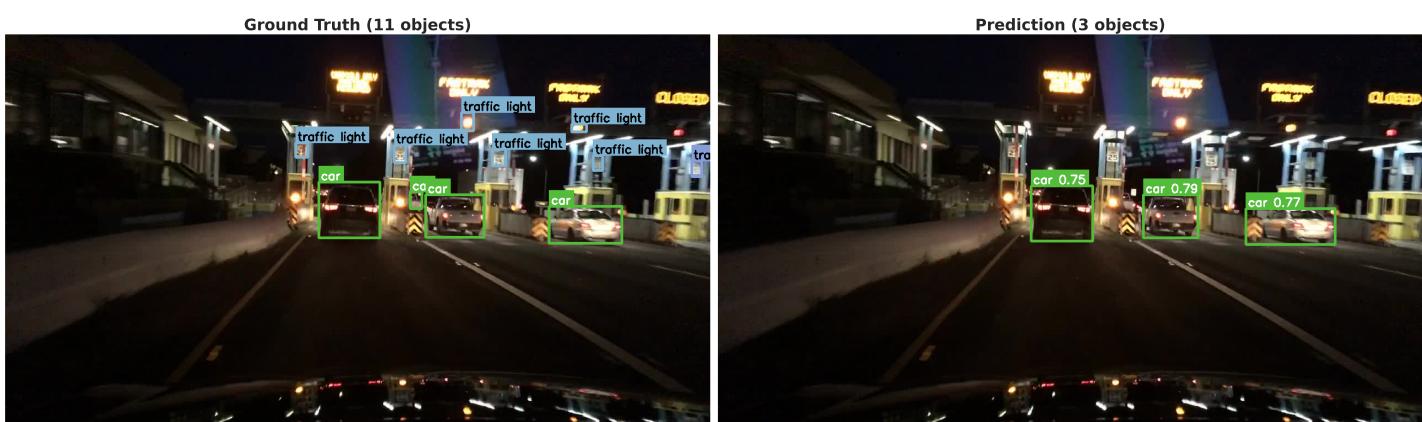
## Sample #2: df2de550-54338fca.jpg

Ground Truth Objects: 21 | Predicted Objects: 7  
Weather: snowy | Scene: residential | Time of Day: night



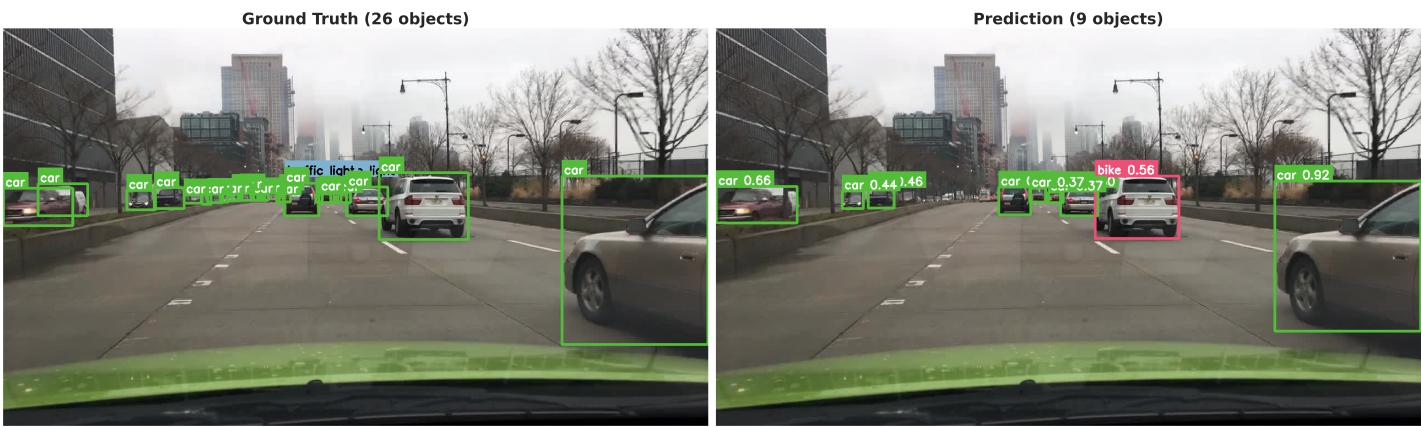
## Sample #3: eb987472-72501d44.jpg

Ground Truth Objects: 11 | Predicted Objects: 3  
Weather: clear | Scene: highway | Time of Day: night



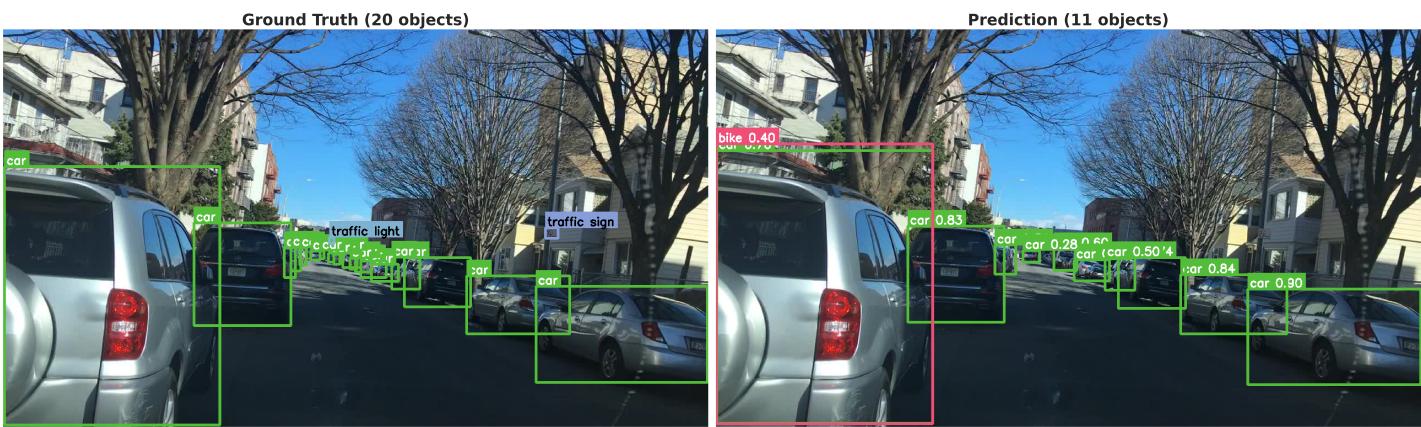
## Sample #4: e11f85ce-4a15f9ec.jpg

Ground Truth Objects: 26 | Predicted Objects: 9  
Weather: foggy | Scene: city street | Time of Day: daytime



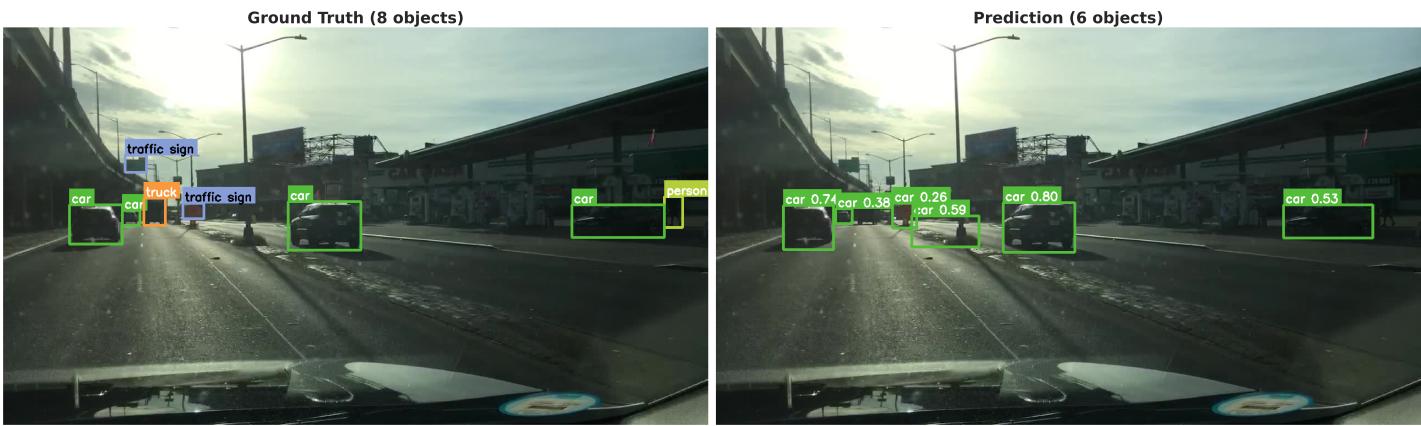
**Sample #5: f87bad26-87f4f7a0.jpg**

Ground Truth Objects: 20 | Predicted Objects: 11  
Weather: clear | Scene: city street | Time of Day: daytime



**Sample #6: f32390d9-4abfe411.jpg**

Ground Truth Objects: 8 | Predicted Objects: 6  
Weather: overcast | Scene: highway | Time of Day: dawn/dusk



Generated by YOLO Testing Pipeline