

# YOLO Model Testing Report

<b>Model:</b>	yolo11m
<b>Model Size:</b>	38.8 MB
<b>Parameters:</b>	20.1 M
<b>FLOPs (640x640):</b>	68.53 GFLOPs
<b>Run Name:</b>	yolo11m_testing_20251128_003534
<b>Timestamp:</b>	2025-11-28 00:48:20
<b>Dataset:</b>	bdd100k_yolo_limited - test split
<b>Images Processed:</b>	20000
<b>Total Execution Time</b>	118.68s
<b>IoU Threshold:</b>	0.5

## Inference Performance

Metric	Value
Average Inference Time	2.53 ms per image
FPS (Frames Per Second)	395.36

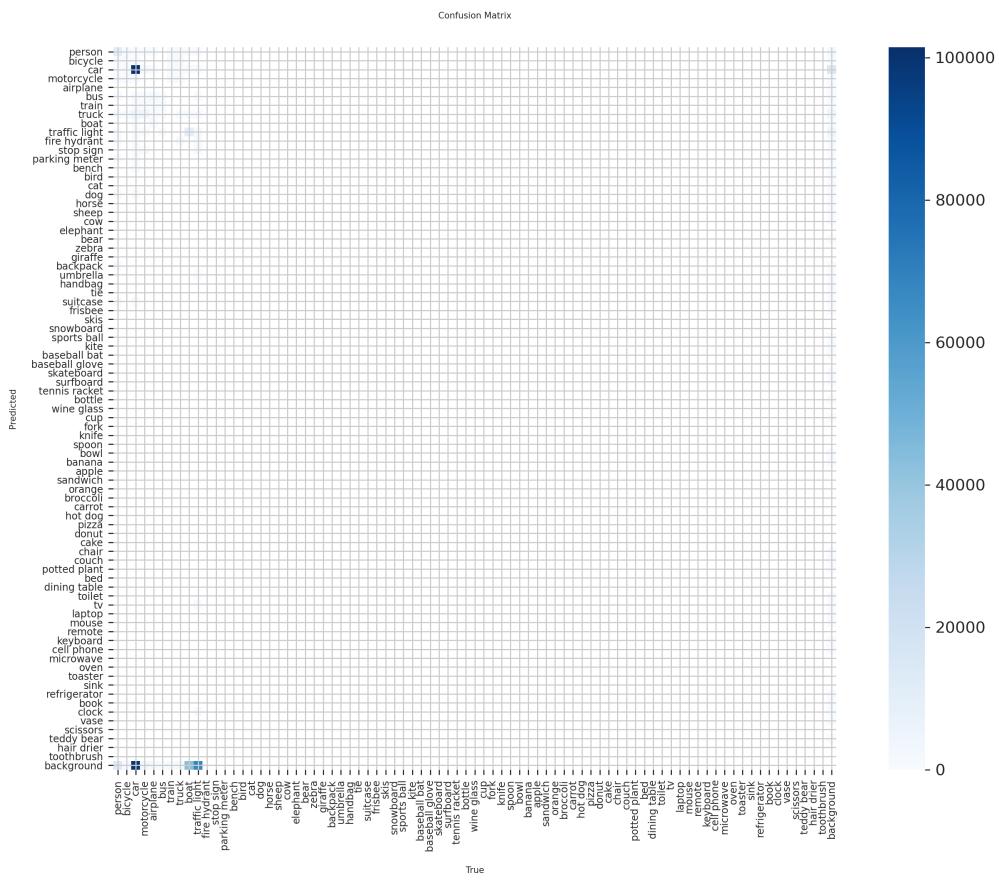
Note: Inference time includes preprocessing, model inference, and postprocessing.

## Overall Accuracy Metrics

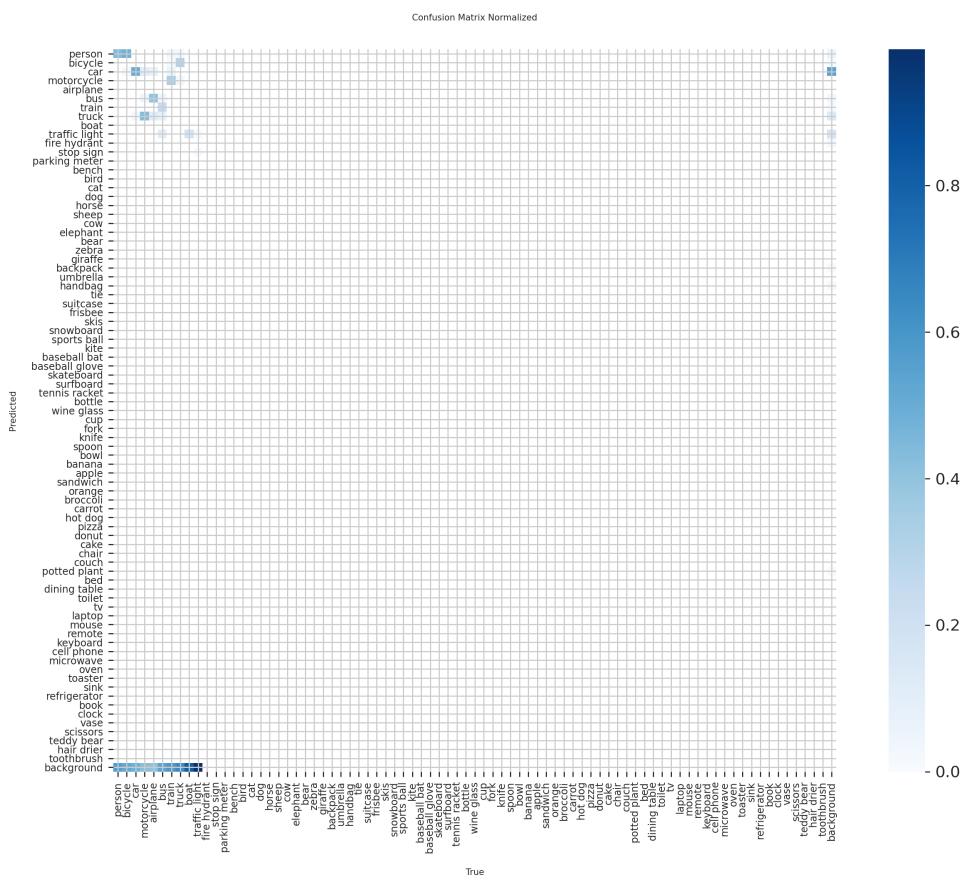
Metric	Value
Precision	0.3062
Recall	0.7008
F1-Score	0.4262
mAP@0.5	0.1243
mAP@0.5:0.95	0.0681
True Positives	112608
False Positives	255119
False Negatives	48078

# Confusion Matrices

## Confusion Matrix (Counts)



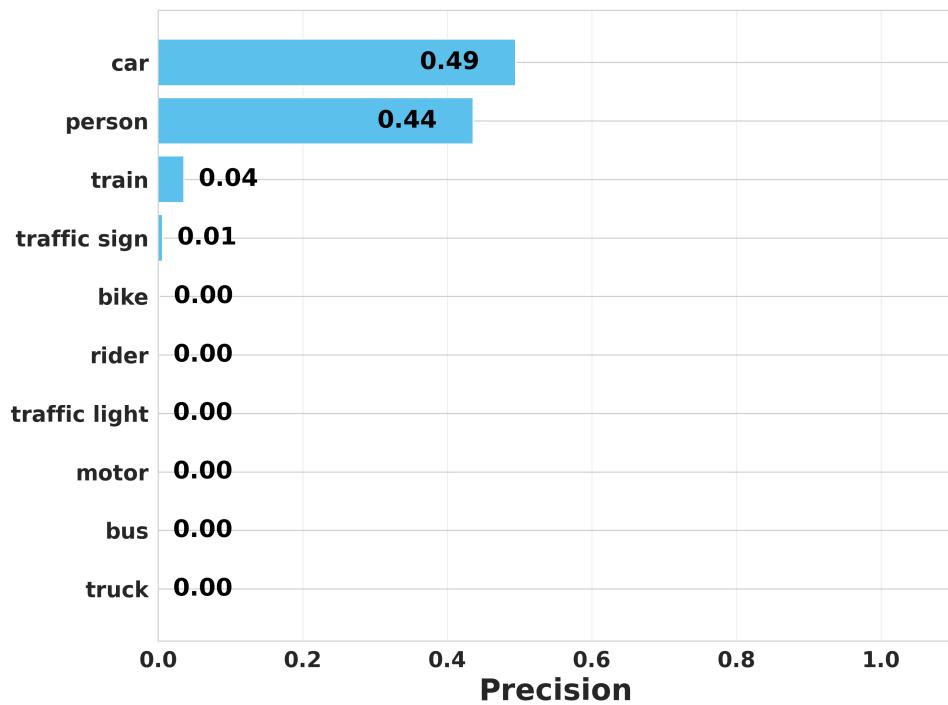
Note: Shows raw prediction counts for each class combination.



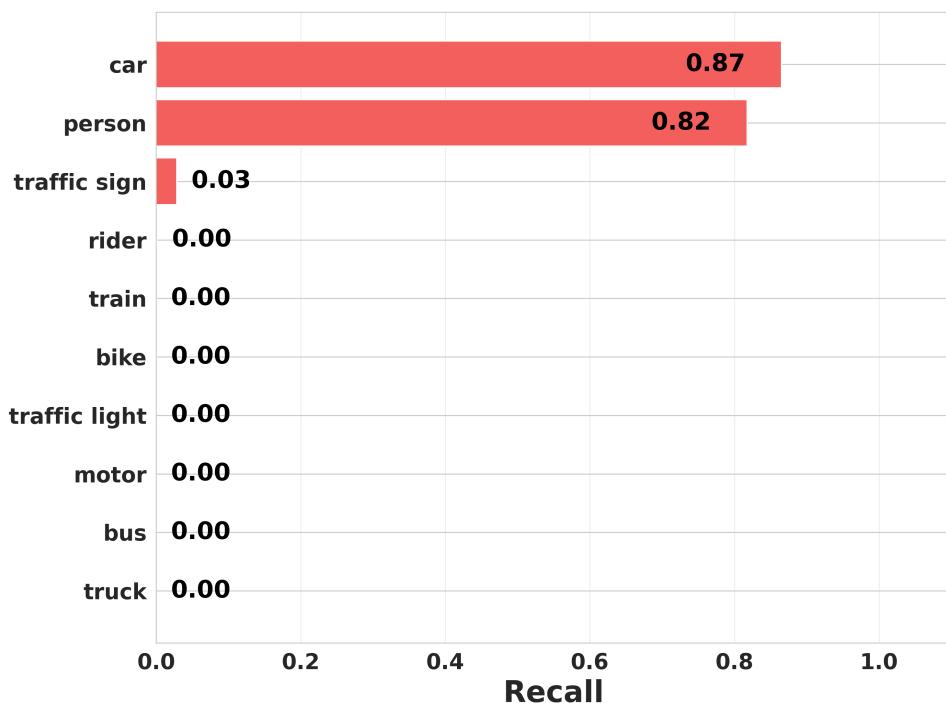
Note: Values are normalized by the number of ground truth instances per class. Shows the proportion of correct vs incorrect classifications.

# Performance Visualizations

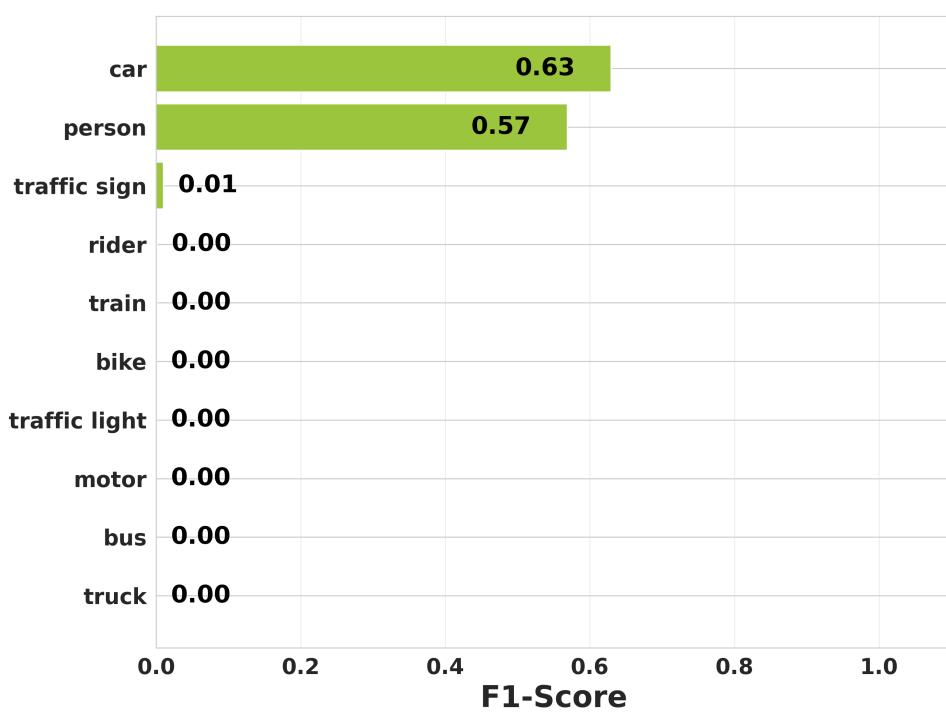
## Precision by Class



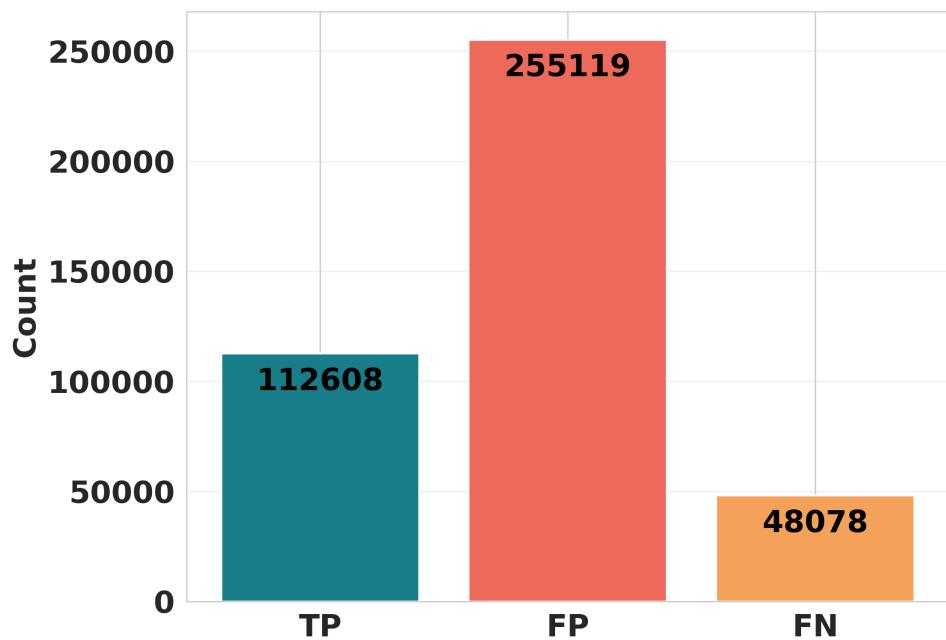
### Recall by Class



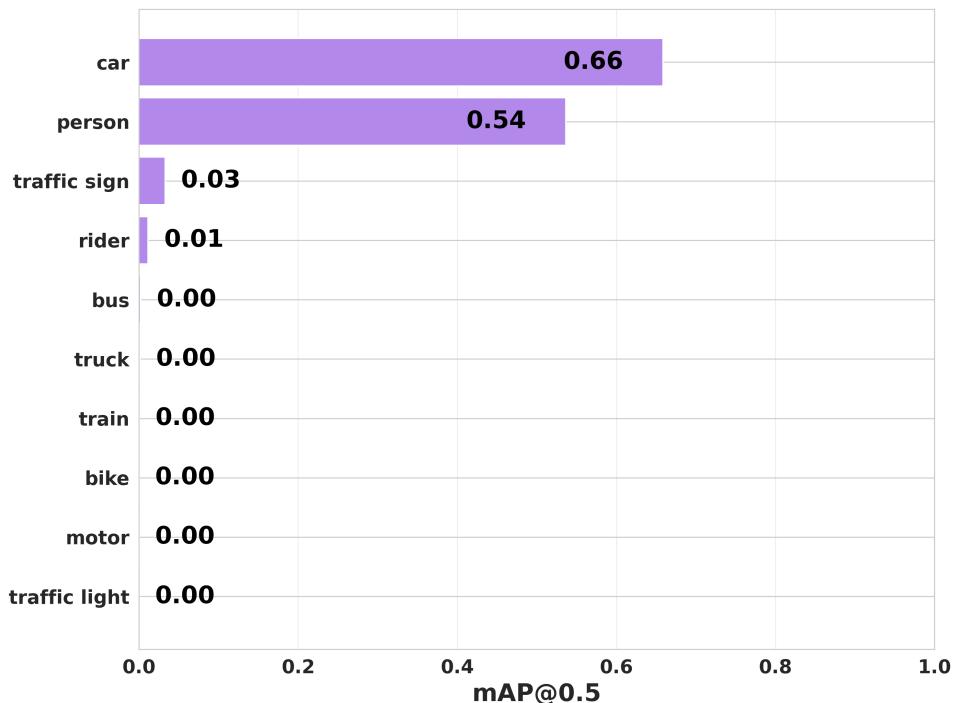
### F1-Score by Class



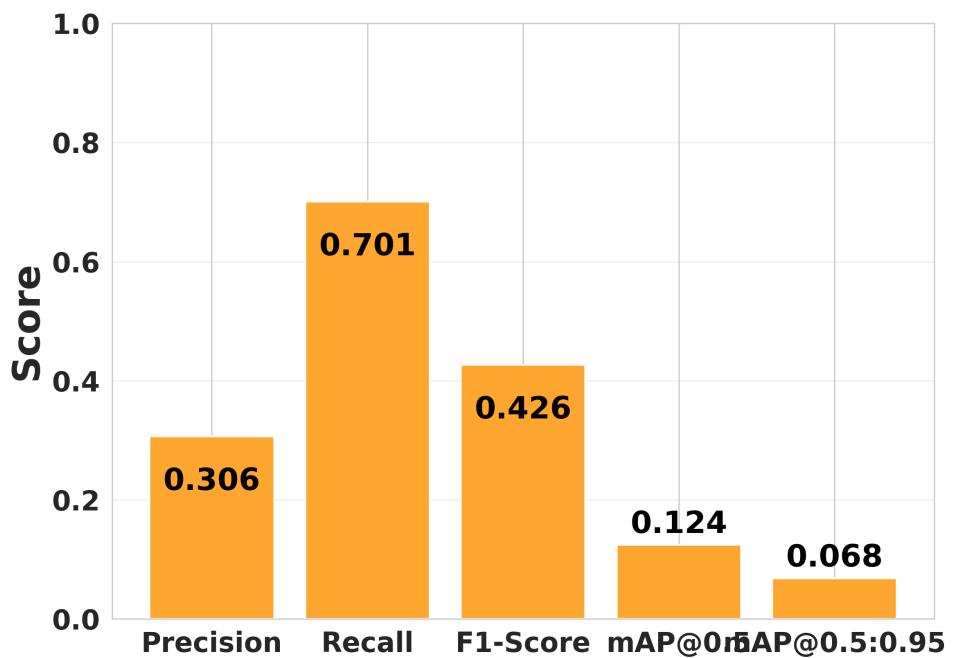
## Overall Detection Outcomes



## mAP@0.5 by Class



## Overall Metrics



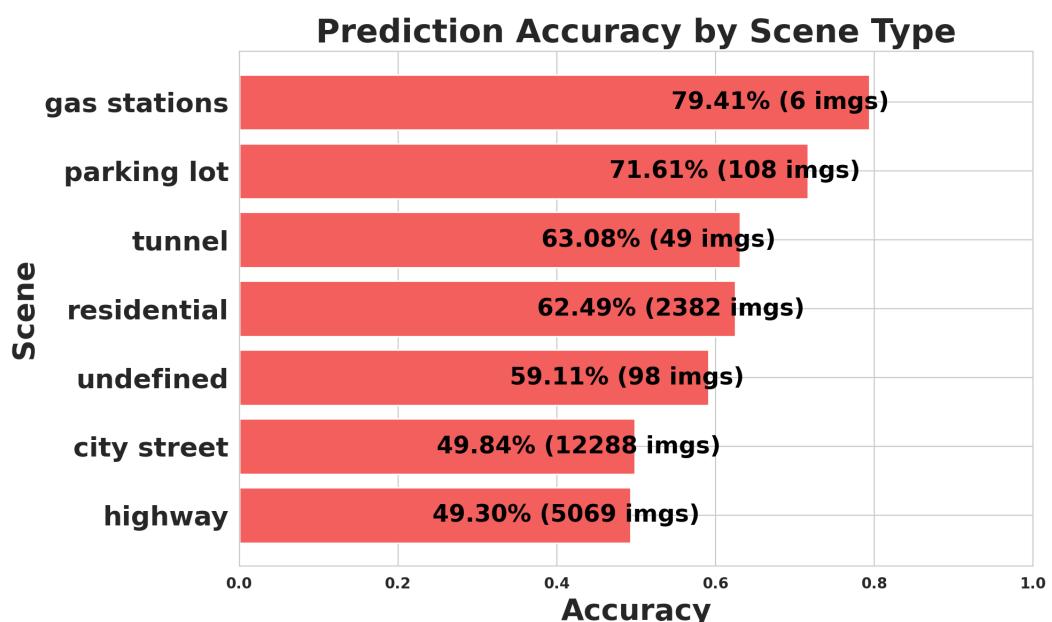
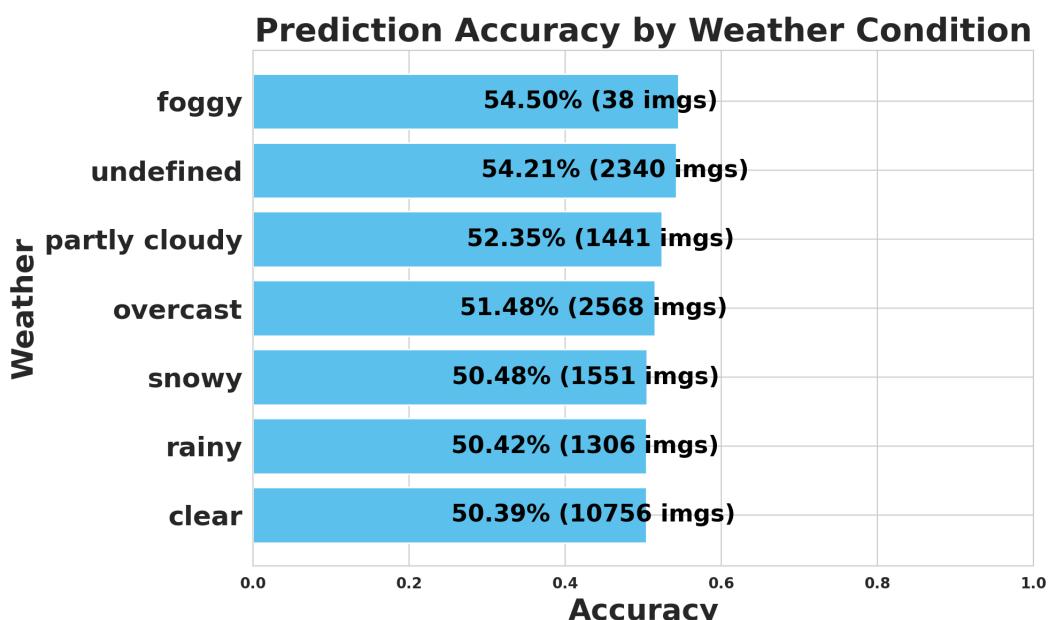
## Per-Class Performance

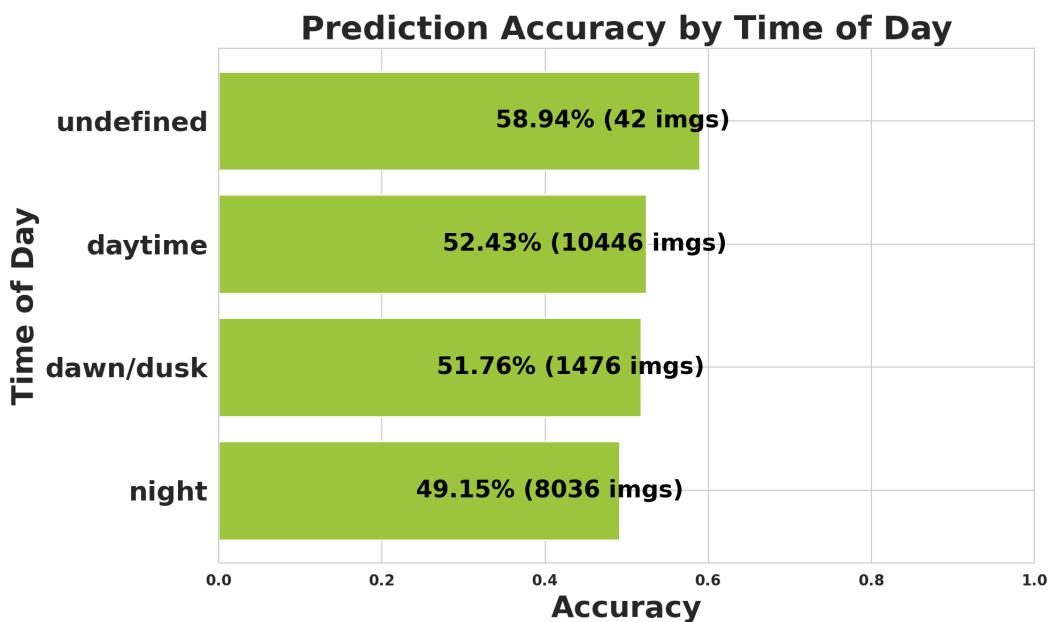
Class	TP	FP	FN	Precision	Recall	F1-Score	mAP@0.5
person	10748	13902	2398	0.4360	0.8176	0.5687	0.5364
rider	1	1293	754	0.0008	0.0013	0.0010	0.0114
car	101441	103708	15803	0.4945	0.8652	0.6293	0.6588
truck	0	8704	431	0.0000	0.0000	0.0000	0.0013
bus	0	3217	17	0.0000	0.0000	0.0000	0.0020
train	1	27	2790	0.0357	0.0004	0.0007	0.0004
motor	0	841	228	0.0000	0.0000	0.0000	0.0001
bike	2	1996	11390	0.0010	0.0002	0.0003	0.0001
traffic light	0	52840	17	0.0000	0.0000	0.0000	0.0000
traffic sign	415	68591	14250	0.0060	0.0283	0.0099	0.0327

# Comprehensive Failure Analysis

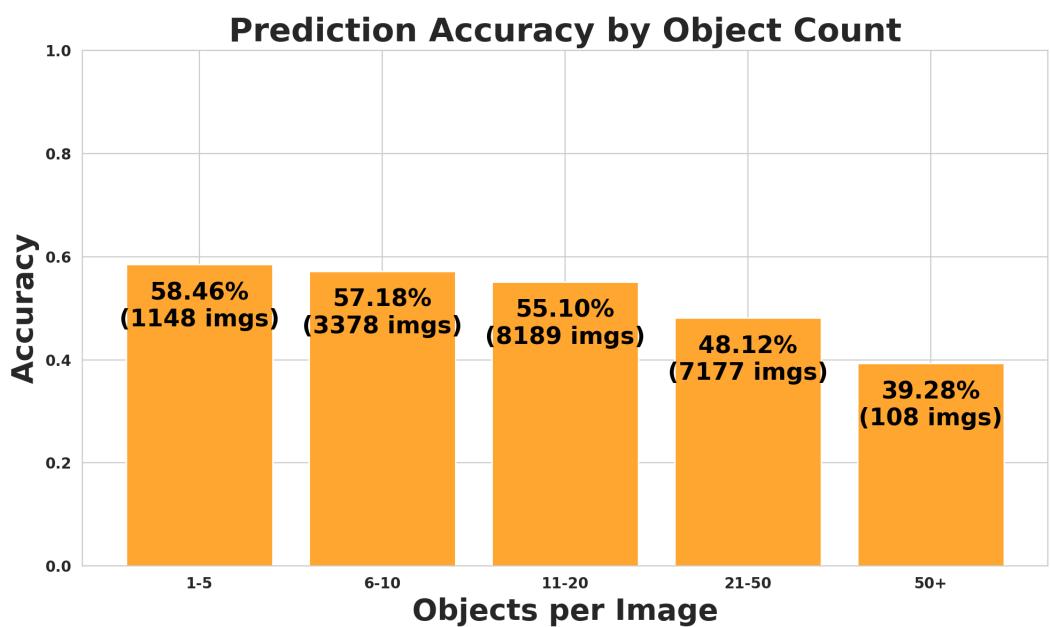
Metric	Value
Total Images Analyzed	20000
Total Expected Objects	367728
Total Matched Objects	188301
Overall Accuracy	51.21%

## Accuracy by Environmental Attributes

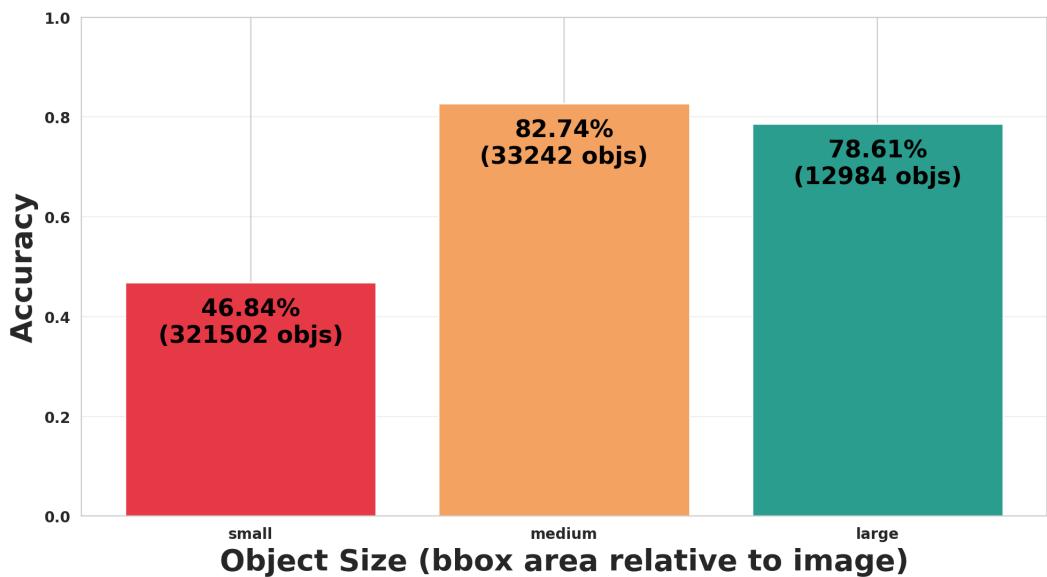




## Accuracy by Object Characteristics

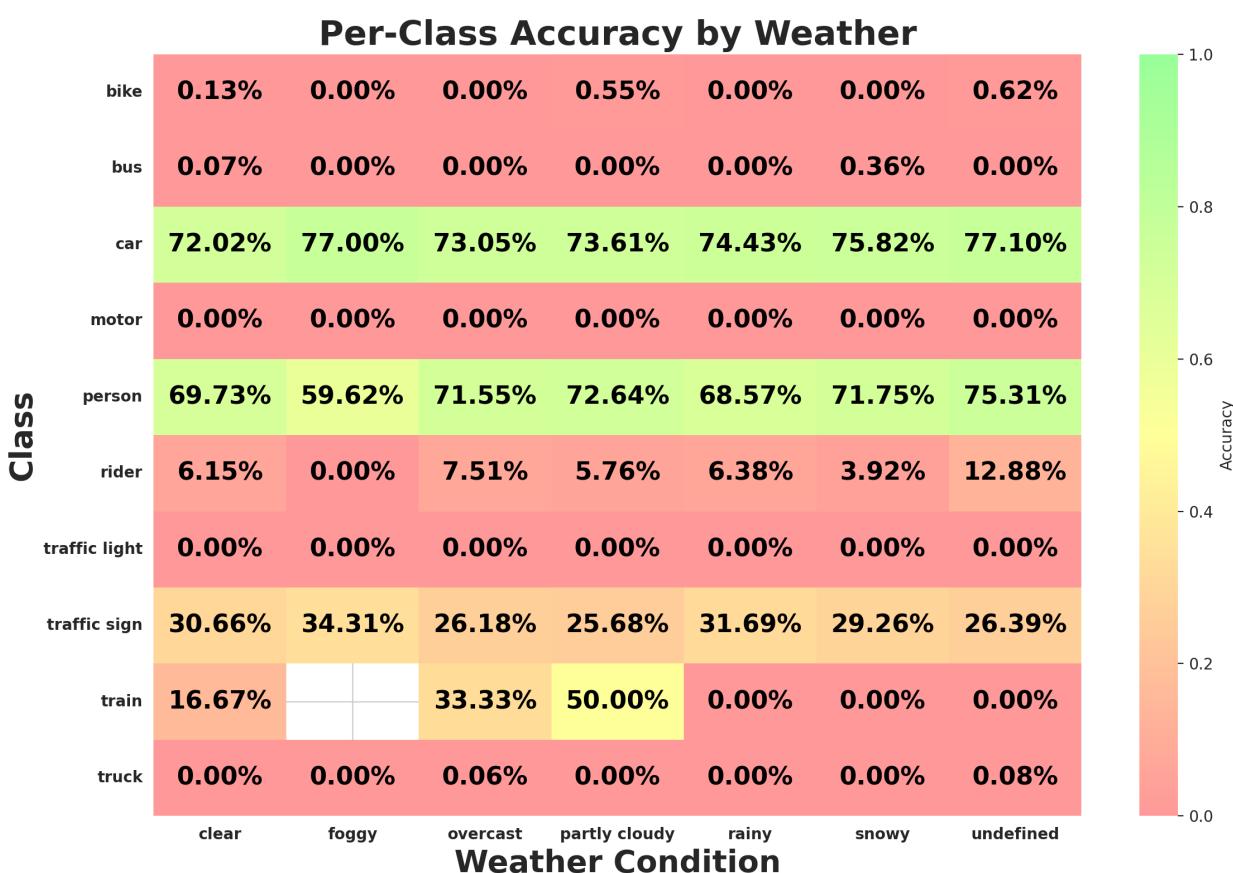
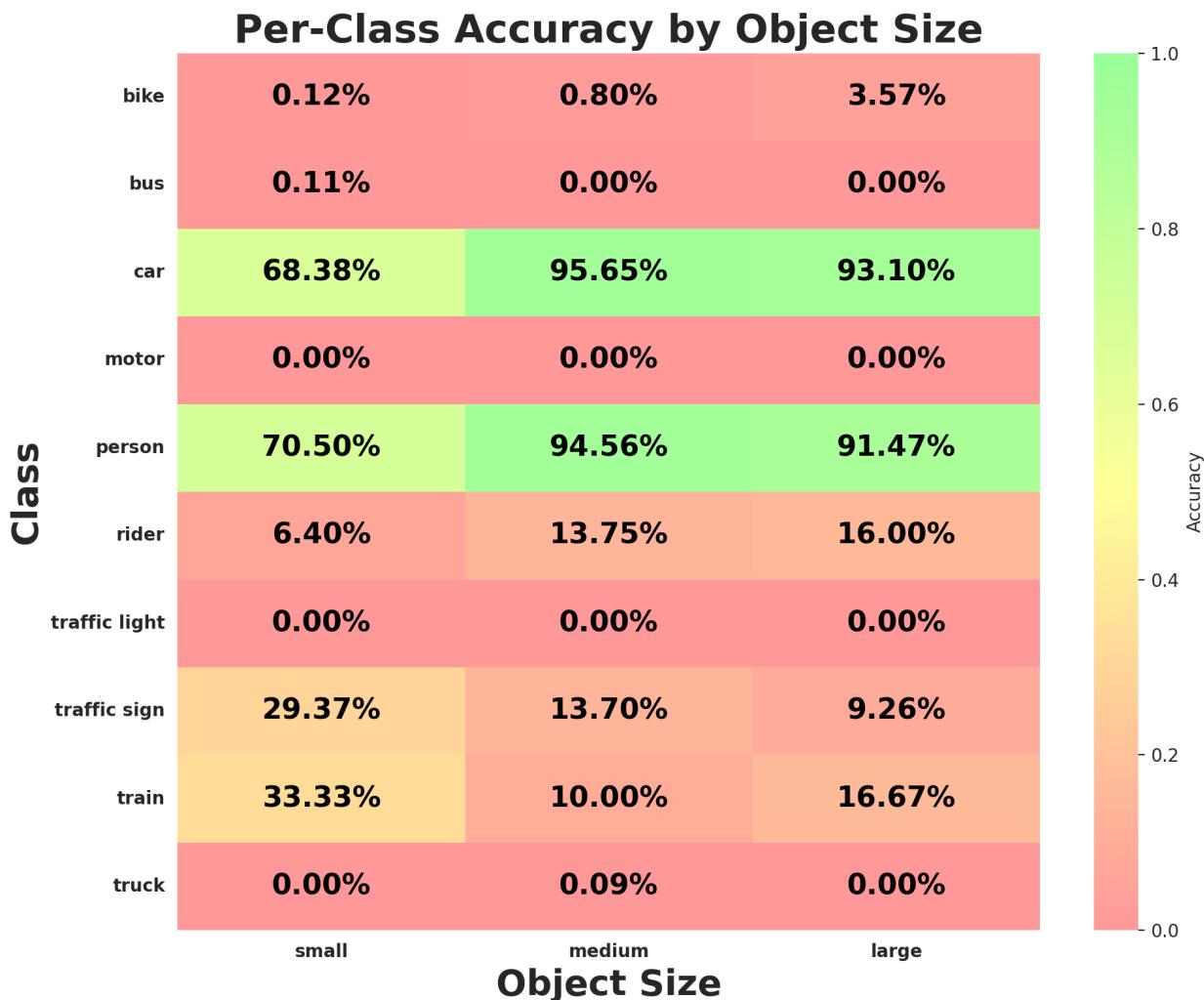


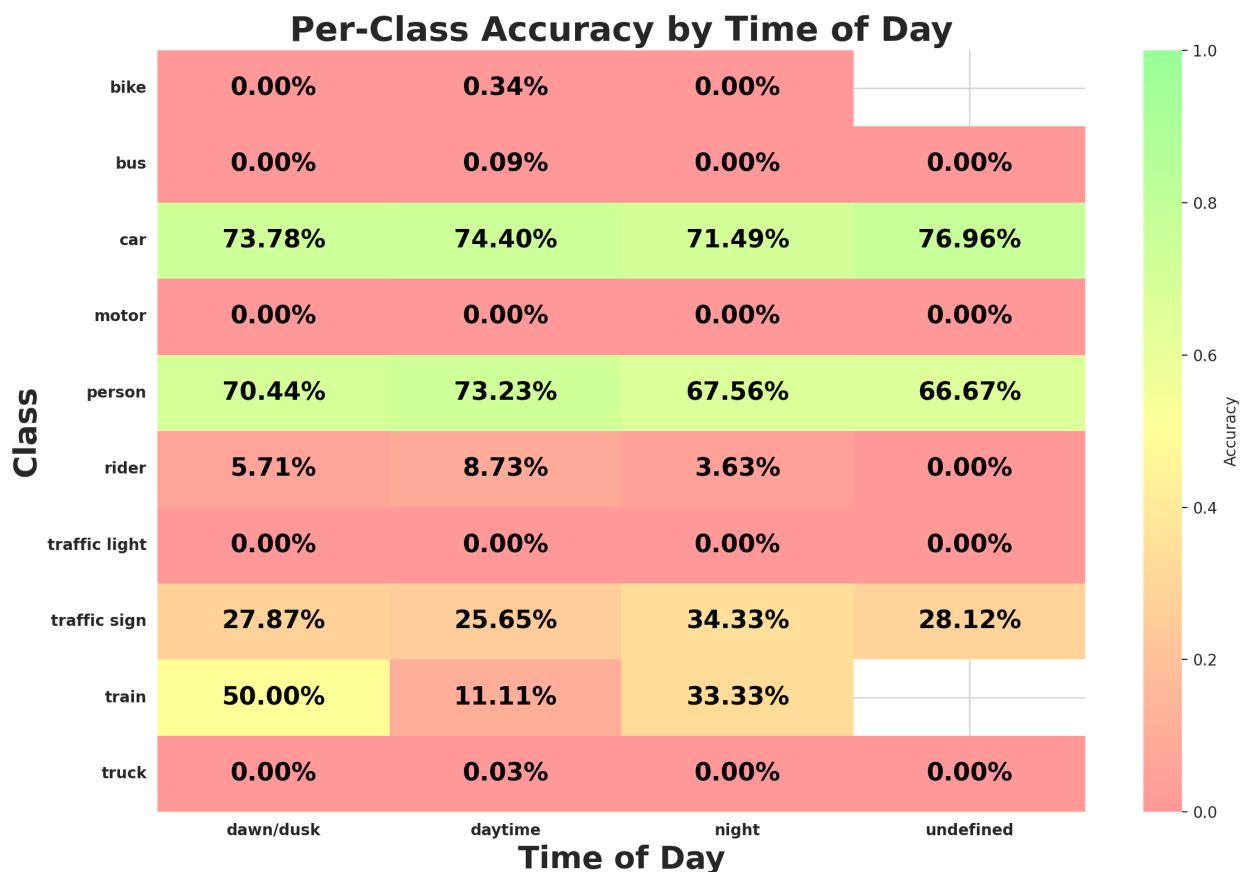
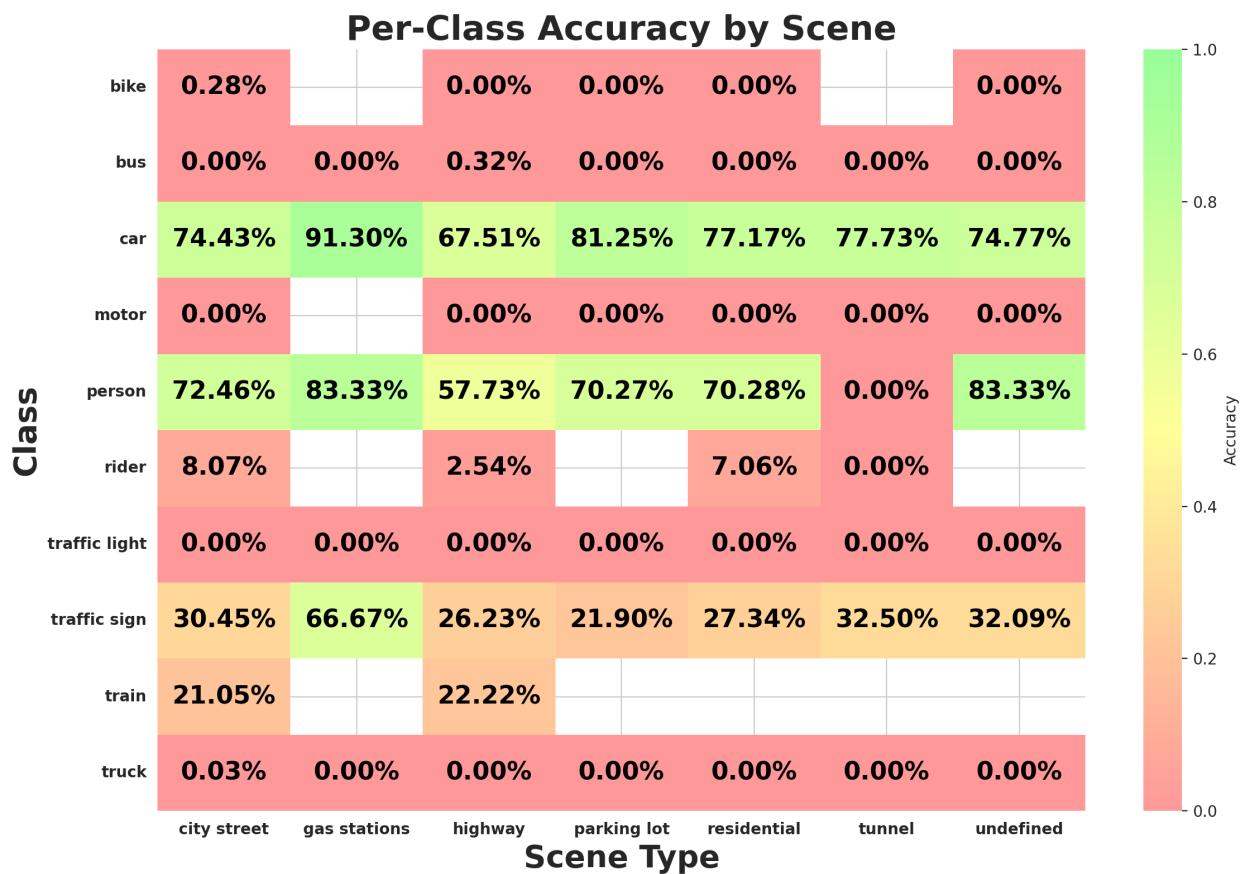
### Prediction Accuracy by Object Size



**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: f0654fea-c1e46ef7.jpg

Ground Truth Objects: 13 | Predicted Objects: 4  
Weather: clear | Scene: city street | Time of Day: night



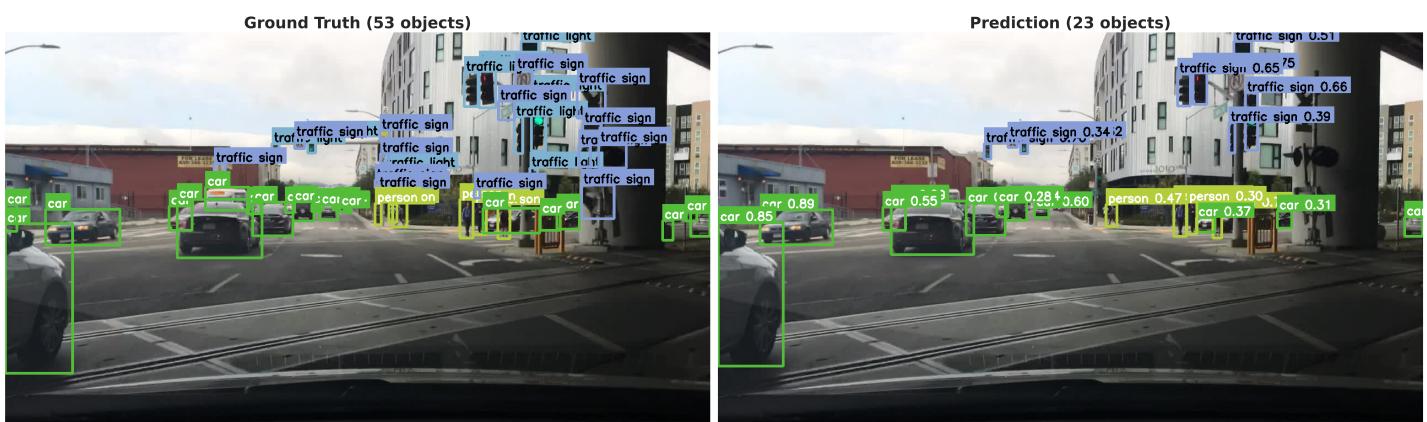
## Sample #2: ef515db3-4b227202.jpg

Ground Truth Objects: 11 | Predicted Objects: 8  
Weather: undefined | Scene: city street | Time of Day: daytime



## Sample #3: fa5aeb9f-40cfcd6e5.jpg

Ground Truth Objects: 53 | Predicted Objects: 23  
Weather: overcast | Scene: city street | Time of Day: daytime



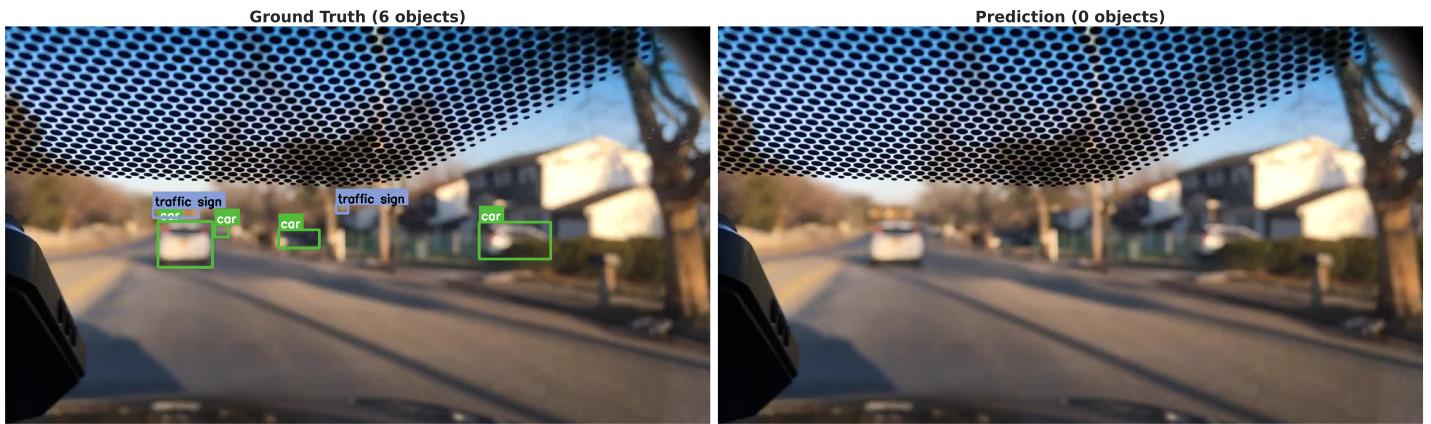
## Sample #4: dc3bd27d-11febcd2c.jpg

Ground Truth Objects: 11 | Predicted Objects: 4  
Weather: clear | Scene: highway | Time of Day: daytime



**Sample #5: def5426b-fe19fd0c.jpg**

Ground Truth Objects: 6 | Predicted Objects: 0  
Weather: clear | Scene: residential | Time of Day: daytime



**Sample #6: d2d0f26d-94a4dc07.jpg**

Ground Truth Objects: 16 | Predicted Objects: 11  
Weather: rainy | Scene: city street | Time of Day: dawn/dusk



Generated by YOLO Testing Pipeline