BDD100K Dataset Analysis, Model, and Evaluation

Author: Abhinav Gadge

Abstract

This report presents the analysis, modeling, and evaluation conducted on the BDD100K dataset for object detection. Due to model compatibility constraints, the focus is on 9 categories overlapping with YOLOv8’s pretrained classes. The project includes dataset analysis, end-to-end inference pipeline development using OpenVINO-quantized YOLOv8, and detailed evaluation across quantitative and qualitative metrics. An end-to-end containerized solution is provided for reproducibility and deployment.

Table of Contents

1. Introduction

The BDD100K dataset is one of the largest and most diverse autonomous driving datasets. This assignment focuses on the object detection task, restricted to bounding-box annotations across 10 categories. Due to model compatibility with YOLOv8 pretrained on COCO, the 'rider' class was excluded, leaving 9 categories. The goal is to analyze the dataset, build an inference pipeline, and evaluate model performance on a validation subset.

2. Data Analysis

The dataset exhibits strong imbalance across categories. Cars dominate the dataset, while classes like motor, bus, and train are very rare. Small objects (traffic lights and signs) are frequent but challenging to detect. The validation subset used here consists of 300 images.

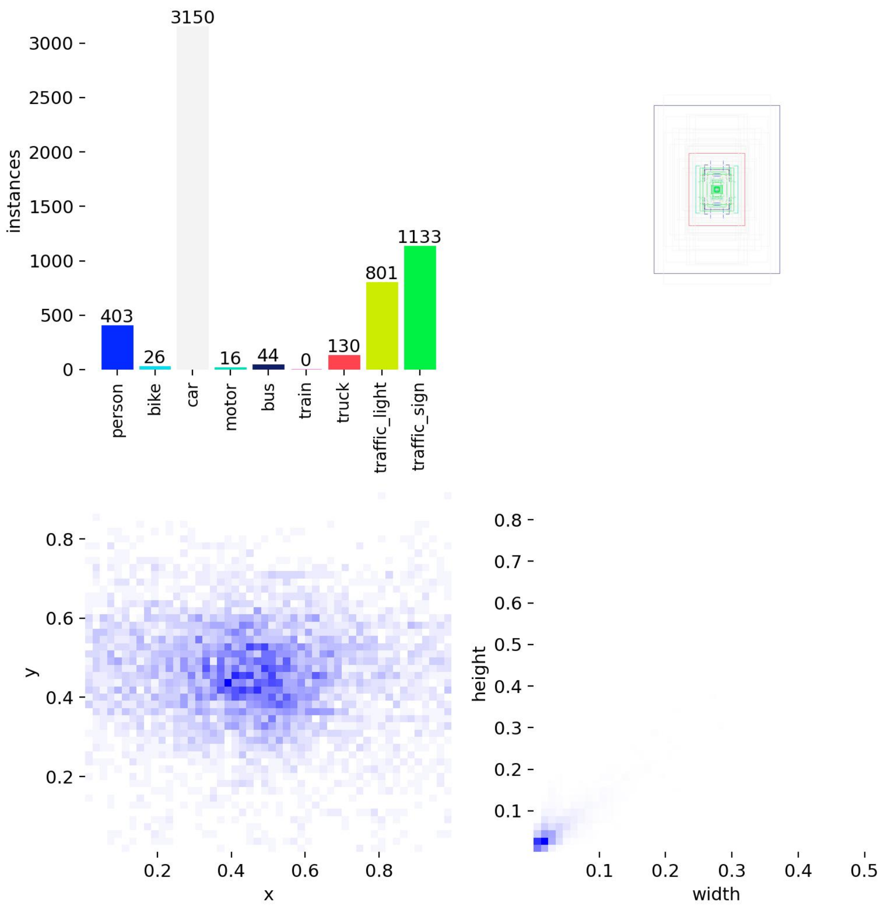


Figure 1: Distribution of object instances across 9 classes in the 300-image validation subset.

3. Model

We used YOLOv8n as the backbone, exported to OpenVINO for efficient inference on CPU and Intel iGPU. Quantization to INT8 enabled real-time performance on limited resources. The model was chosen due to overlap with BDD100K classes, lightweight architecture, and ease of integration into OpenVINO pipelines.
Pipeline Enhancements: - Preprocessing: aspect ratio preserving resize. - Multi-view inference: augment input with horizontal flips. - Multiprocessing: distribute tasks across CPU and iGPU in parallel. - Postprocessing: NMS pre- and post-aggregation.

4. Evaluation and Results

Validation subset size: 300 images
Explanation of Metrics:
- Detection Accuracy: Fraction of ground-truth objects detected as any class. - Classification Accuracy: Fraction of ground-truth objects detected with correct class. - False Positive Instances: Fraction of detections that did not match any ground truth. - Missed Instances: Fraction of ground-truth objects not detected at all. - False Positive Images: Percentage of images containing at least one false positive detection.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	2	8	0	0	0	0	0	4

	2	0	0	0	0	0	0	0
	0	1094	0	4	0	13	5	8
	1	0	0	0	0	0	0	0
	0	9	0	10	0	2	2	1
	0	0	0	1	0	1	0	2
	1	41	0	5	0	25	1	7
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	31	1874	11	29	1	66	865	1189

Metric	Value
Detection Accuracy	24.30%
Classification Accuracy	21.87%
False Positive Instances	6.12%
Missed/FN Instances	75.70%
False Positives Images	55.67%
mAP@0.5	9.11%

Class	Det Acc	Class Acc	FP	Missed	Precision	Recall	AP
person	35.77%	30.35%	13.82%	69.65%	68.71%	30.35%	20.86%
bike	16.22%	5.41%	2.70%	94.59%	66.67%	5.41%	3.60%
car	38.07%	36.15%	10.61%	63.85%	77.31%	36.15%	27.95%
motor	0.00%	0.00%	27.27%	100.00%	0.00%	0.00%	0.00%
bus	40.82%	20.41%	51.02%	79.59%	28.57%	20.41%	5.83%
train	0.00%	0.00%	500.00%	100.00%	0.00%	0.00%	0.00%
truck	38.32%	23.36%	74.77%	76.64%	23.81%	23.36%	5.56%
traffic_light	0.92%	0.00%	0.00%	100.00%	NaN	0.00%	NaN
traffic_sign	1.82%	0.00%	0.00%	100.00%	NaN	0.00%	NaN

5. Error Analysis

The evaluation shows frequent misclassification between visually similar categories such as car and truck. Rare classes such as train and motor were often completely missed. Small objects like traffic lights and traffic signs were particularly challenging, leading to very low recall. False positives occurred mainly in crowded nighttime scenes.

6. End-to-End Pipeline

The complete inference pipeline integrates preprocessing, inference, and evaluation steps to enable reproducible, real-time object detection. Steps include: 1. Preprocessing: aspect ratio preserving resize. 2. Inference: YOLOv8n with OpenVINO, multi-view (original + flipped). 3. Parallel Execution: CPU + iGPU multiprocessing. 4. Postprocessing: NMS for box refinement. 5. Evaluation: per-class AP, FP/FN logging. 6. Outputs: bounding box overlays and YOLO-format .txt files.

7. Suggestions for Improvement

To further improve model performance, we recommend: - Oversample or augment rare classes to reduce imbalance. - Use higher resolution inputs or tiling for small object detection. - Apply modern augmentations (Mosaic, Copy-Paste) during fine-tuning. - Semi-supervised training on unlabeled validation images. - Employ class-balanced loss functions. - Explore advanced architectures (YOLOv8x, DETR) for improved small object detection.

8. Deliverables

The repository contains: - Source code under python_inference_client/ (analysis, inference, utils). - Quantized YOLOv8 model in yolov8n_openvino_model/. - requirements.txt and Dockerfile for containerized execution. - Analysis plots and evaluation results in outputs/. - This detailed report (BDD100K_Assignment_Report_Complete.pdf). - Original dashboard PDF (data_analysis.pdf) appended as Appendix.

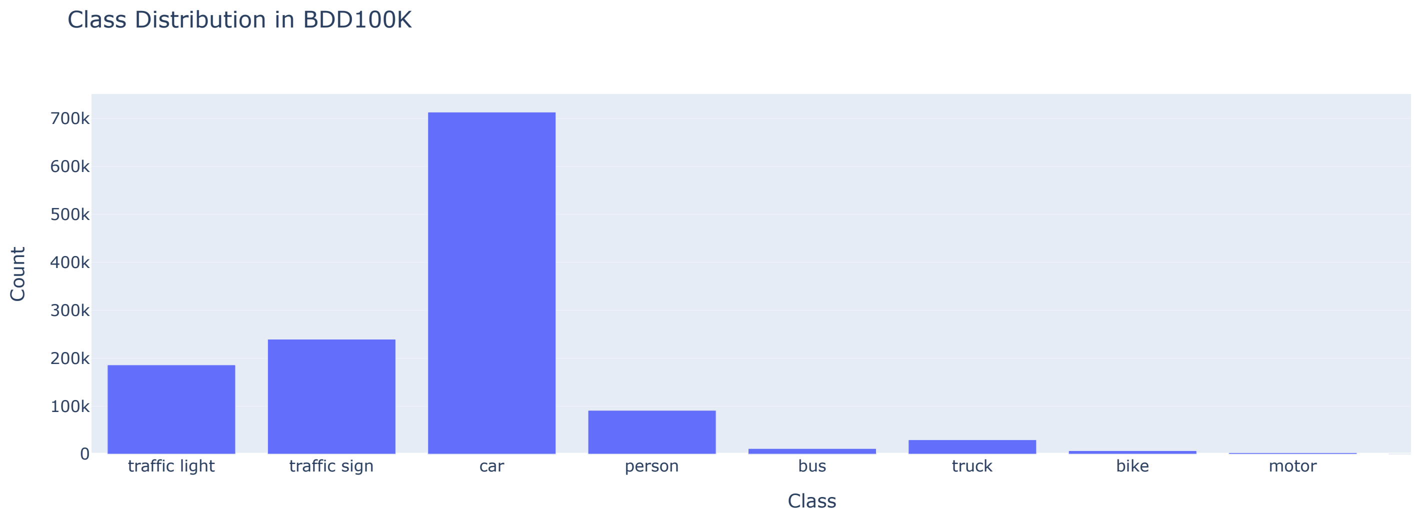
9. Conclusion

This project demonstrates end-to-end handling of the BDD100K dataset for object detection, from analysis to inference and evaluation. Despite constraints such as limited hardware and class imbalance, a lightweight, real-time pipeline was developed using OpenVINO-quantized YOLOv8. Evaluation revealed strong performance on common categories (car, person) and highlighted challenges in rare and small object classes. Future improvements should focus on rebalancing and advanced augmentation techniques to better handle edge cases.

10. Appendix

The original analysis dashboard is appended in the following pages.

BDD100K Object Detection Dashboard



Bounding Box Anomalies

Class	Avg W	Avg H	Min Area	Max Area	#Boxes
traffic light	15.87	25.29	0.94	302654.31	186117
traffic sign	32.28	25.20	3.57	917709.77	239686
car	74.62	58.02	0.87	612645.18	713211
person	27.67	66.71	2.40	344937.76	91349
bus	145.66	127.38	8.97	775744.23	11672
truck	127.58	115.12	3.00	828772.18	29971
bike	60.63	67.37	3.75	273027.06	7210
motor	67.12	68.28	30.98	316440.64	3002
train	267.08	83.87	51.55	500339.24	136

Interesting / Unique Samples

	Category	Image	Class	Area	BBox
○	rare_classes	002d290d-89f4e5c0.jpg	train	9184.2	(0, 443.579455, 363.915519, 25.237117000000012)
○	rare_classes	00225f53-67614580.jpg	motor	4939.7	(1091.715047, 419.552463, 69.19320700000003, 71.389817)
○	rare_classes	000f8d37-d4c09a0f.jpg	bike	814.4	(739.356518, 351.337253, 21.937484999999924, 37.124975000000006)
○	smallest	20a773e1-a9268e6b.jpg	car	0.9	(416.055098, 554.300932, 3.49281400000001, 0.24948700000004465)
○	smallest	7bc6abdd-6a5c7b75.jpg	traffic light	0.9	(434.033484, 295.578561, 0.30604600000003757, 3.0604529999999954)
○	smallest	3bb2f5c2-1463d3a1.jpg	traffic light	1.4	(453.770513, 117.66316, 0.10662500000000819, 12.78947199999999)
○	smallest	a34c36f3-22c7e2c3.jpg	traffic light	1.5	(382.828683, 288.816451, 2.448360999999977, 0.6120890000000259)
○	smallest	84643961-e2e4a747.jpg	car	2.0	(565.291458, 412.249803, 3.9917859999999337, 0.49897300000003497)
○	largest	414fe350-4f0f6e64.jpg	bus	775744.2	(198.709109, 0, 1079.8343570000002, 718.391881)
○	largest	70cc96e1-aa1781b5.jpg	truck	778199.5	(0, 0, 1138.777474, 683.364003)
○	largest	4efe84c6-93313761.jpg	truck	828772.2	(86.39997, 22.949993, 1192.049582, 695.249757)
○	largest	2e5a3ced-c7603d0d.jpg	traffic sign	906063.4	(0, 2.495509, 1273.956878, 711.2198070000001)
○	largest	1d33c83b-71e1ea1c.jpg	traffic sign	917709.8	(0, 1.836272, 1279.269409, 717.370215)

Select a row above to preview image