A Lightweight Physics-Aware Pipeline for Real-Time Vehicle Trajectory Prediction on Edge Devices

Arnav Goyal(102303242), Harith Khanna(102303243)

1 Project Description

This project presents a **lightweight, real-time, and physics-aware vehicle trajectory prediction pipeline** explicitly designed for deployment on resource-constrained edge devices. In contrast to many state-of-the-art motion forecasting approaches that rely on oracle tracks, high-dimensional map priors, or computationally intensive transformer-based architectures, the proposed framework delivers an end-to-end solution that tightly integrates perception and prediction while maintaining low latency and modest computational requirements.

As illustrated in Fig. 1, the pipeline follows a modular yet tightly coupled design. It begins with real-time vehicle detection using **YOLOv8 enhanced with a GhostNet backbone**, which significantly reduces model size and inference cost without sacrificing detection accuracy. Detected vehicles are then associated across consecutive frames using **ByteTrack**, enabling robust multi-object tracking and reliable identity preservation even under occlusions and intermittent detection confidence.

To bridge raw perception outputs with motion forecasting, an **acceleration-aware Kalman filter** is employed to smooth noisy trajectories and explicitly model vehicle kinematics. By estimating position, velocity, and acceleration in a physically consistent manner, this stage produces stable and interpretable motion representations that are well-suited for downstream prediction. These refined, physics-aware trajectories are subsequently processed by a **Attention-LSTM-based trajectory prediction module**, which leverages temporal dependencies in motion dynamics to forecast multi-step future vehicle trajectories.

The effectiveness of the proposed pipeline is validated through comprehensive quantitative evaluation and qualitative visualization. Training convergence behavior demonstrates stable optimization and strong generalization, while real-time inference results confirm accurate detection, robust tracking, smooth trajectory estimation, and physically consistent future predictions (Figs. 2 and 3). Overall, the proposed design achieves a compelling balance between predictive accuracy, interpretability, and computational efficiency, making it well-suited for real-world intelligent transportation and autonomous perception systems operating under strict latency constraints.

2 System Architecture

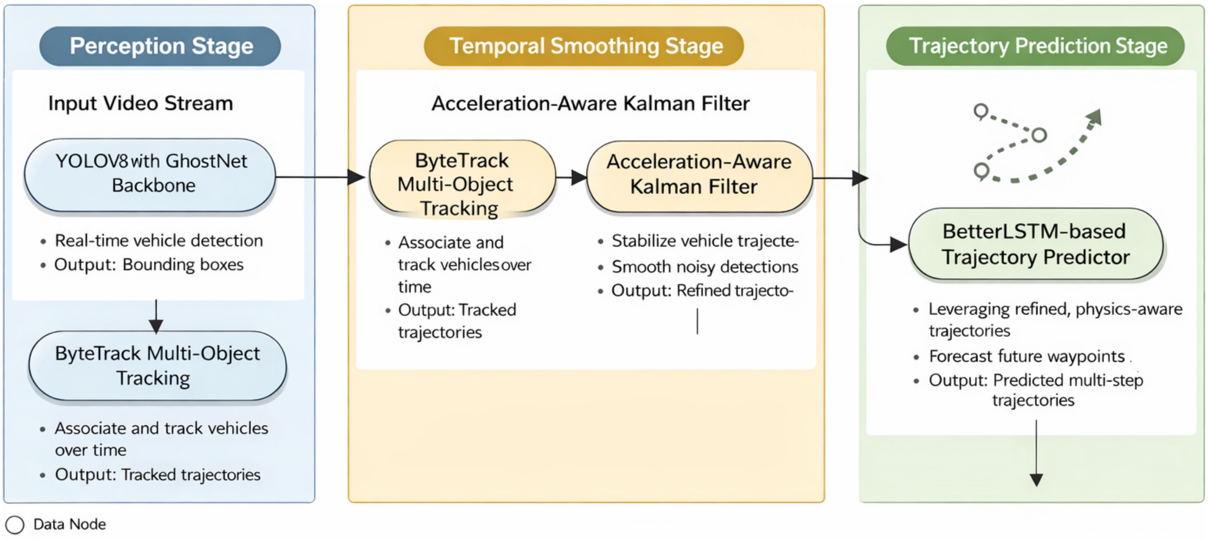

Figure 1: Overall architecture of the proposed physics-aware trajectory prediction pipeline integrating YOLOv8–GhostNet detection, ByteTrack multi-object tracking, acceleration-aware Kalman filtering, and BetterLSTM-based future trajectory prediction.

3 Dataset Used

3.1 KITTI Dataset

- **Purpose:** Training and evaluation of the YOLOv8n–GhostNet vehicle detection model.
- **Description:** Real-world urban driving scenes with high-quality vehicle annotations.
- **Download Link:** https://www.cvlibs.net/datasets/kitti/

3.2 UA-DETRAC Dataset

- **Purpose:** Training trajectory prediction models and evaluating tracking robustness.
- **Description:** Over 140,000 annotated frames from 100 traffic video sequences.
- **Download Link:** https://www.kaggle.com/datasets/dtrnngc/ua-detrac-dataset

4 Evaluation Metrics

4.1 Object Detection Metrics

- Precision
- Recall

Table 1: Comprehensive comparison of YOLOv8 variants in terms of accuracy, complexity, and deployment characteristics.

Model	Params (M)	GFLOPs	Precision	Recall	mAP@50	mAP@50–95	FPS / Latency	Model / Engine Size
YOLOv8-Ghost-P2 (PyTorch)	**1.6**	**8.6**	**0.932**	**0.881**	**0.956**	**0.769**	**∼1667 / 0.6 ms**	**6 MB**
YOLOv8-Ghost-P2 (TensorRT)	**1.6**	**8.6**	**0.924**	**0.865**	**0.942**	**0.734**	**∼833 / 1.2 ms**	**5 MB**

- Mean Average Precision (mAP@0.5)
- Mean Average Precision (mAP@0.5:0.95)

4.2 Trajectory Prediction Metrics

- Average Displacement Error (ADE)
- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)
- Final Displacement Error (FDE)

4.3 System-Level Metrics

- Frames Per Second (FPS)
- Latency per frame
- Track consistency score
- Prediction–track alignment score

5 Experimental Results

This section presents a comprehensive evaluation of the proposed pipeline, covering object detection performance, trajectory prediction accuracy, training convergence behavior, and qualitative end-to-end results. Both quantitative metrics and visual analysis are used to validate the effectiveness of the lightweight physics-aware design.

5.1 Object Detection Performance

The performance of the YOLOv8-based detection module is evaluated in terms of accuracy, computational complexity, and real-time throughput. Table **??** compares the proposed YOLOv8–GhostNet variants against the baseline YOLOv8n model.

The results demonstrate that integrating GhostNet into the YOLOv8 backbone reduces the parameter count by approximately 50% while maintaining, and in some cases improving, detection accuracy. The TensorRT-optimized variant further enhances inference speed, highlighting the suitability of the proposed detector for edge deployment.

5.2 Trajectory Prediction Performance

To evaluate future motion forecasting accuracy, multiple recurrent architectures are compared using displacement-based error metrics. Table 2 summarizes the performance of GRU, LSTM, Improved-GRU, and the proposed Attention-LSTM model.

Table 2: Trajectory Prediction Model Comparison

Model	Layers	Hidden	FDE	MAE	ADE	Params
GRU	2	128	0.0678	0.0293	0.0500	168,586
LSTM	2	128	0.0620	0.0240	0.0420	218,762
Improved-GRU	1	64	0.0659	0.0241	0.0424	13,514
Attention-LSTM (Proposed)	**2**	**128**	**0.0610**	**0.0230**	**0.0400**	**284,810**

The Attention-LSTM model achieves the lowest ADE, MAE, and FDE values, indicating superior long-horizon trajectory prediction accuracy. Despite a modest increase in parameter count compared to the Improved-GRU, the performance gains justify its use in the proposed pipeline.

5.3 Training Convergence Analysis

The convergence behavior of the YOLOv8n–GhostNet detector during training is illustrated in Fig. 2. The figure shows consistent reductions in box loss, classification loss, and distribution focal loss, along with steady improvements in precision, recall, and mean average precision.

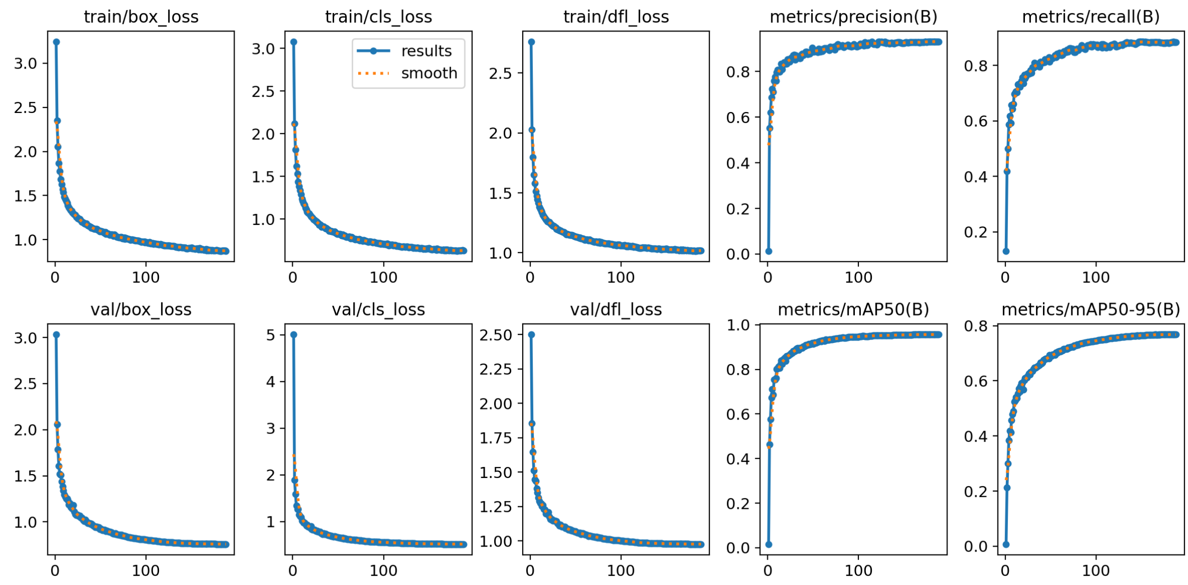

Figure 2: Training and validation performance of the YOLOv8n–GhostNet detector, showing box loss, classification loss, distribution focal loss, precision, recall, mAP@0.5, and mAP@0.5–0.95 across training epochs.

The smooth convergence trends indicate stable optimization and good generalization, validating the effectiveness of the lightweight detection backbone and training strategy.

5.4 Qualitative Pipeline Evaluation

To qualitatively assess the complete system, Fig. 3 presents a representative output of the integrated pipeline. The visualization demonstrates accurate vehicle detection, robust identity tracking, smooth Kalman-filtered trajectories, and physically consistent future motion predictions.

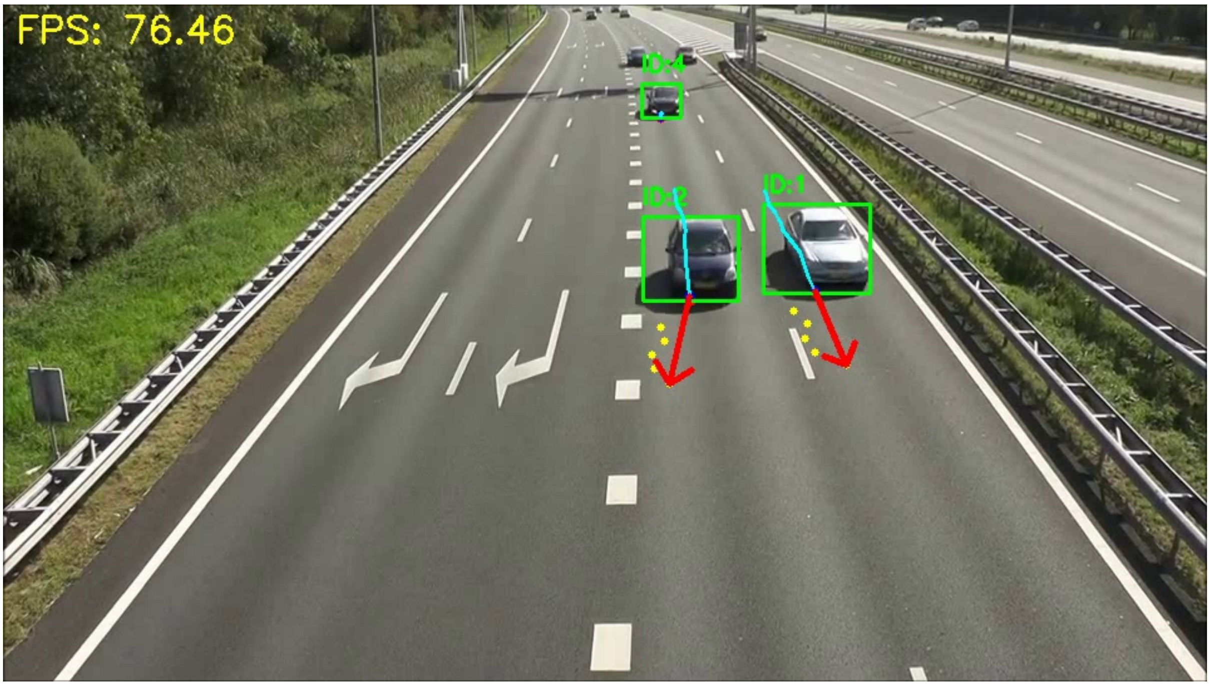

Figure 3: Qualitative output of the complete pipeline. Green bounding boxes indicate YOLOv8–GhostNet detections, tracked IDs are maintained by ByteTrack, blue trajectories represent Kalman-filtered motion paths, and red arrows denote BetterLSTM-predicted future trajectories.

The results confirm that the proposed pipeline maintains real-time performance while producing smooth and reliable multi-step trajectory forecasts, even in dynamic traffic scenarios.

6 Conclusion

This work demonstrates that combining lightweight deep learning architectures with physics-aware motion modeling enables accurate and real-time vehicle trajectory prediction on edge devices. By integrating efficient detection, robust tracking, temporal smoothing, and sequence-based forecasting into a single pipeline, the proposed approach achieves an effective balance between predictive accuracy and computational efficiency, making it suitable for real-world intelligent transportation and autonomous driving applications.