

Autonomous Intelligent Transportation Systems: A Data-Driven Approach

Tobias F. Müller¹, Chiara M. Santoro²

¹ Department of Autonomous Vehicles and Intelligent Transportation, Technische Universität München (TUM), Germany Scopus Author ID: 57217034829

Email: tobias.mueller@tum.de ORCID: <https://orcid.org/0000-0001-4823-7916>

² Department of Transportation Engineering and AI, Politecnico di Milano, Italy Scopus Author ID: 57224138704 Email: chiara.santoro@polimi.it ORCID:

<https://orcid.org/0000-0002-7341-5829>

ABSTRACT

Autonomous intelligent transportation systems (AITS) represent a convergence of sensor fusion, deep learning perception, path planning, and V2X communication technologies that promise to transform urban mobility through improved safety, efficiency, and accessibility. However, the data-driven design of robust AITS capable of operating in diverse, unstructured real-world traffic environments — including adverse weather, mixed autonomy traffic, and complex urban intersections — remains an open engineering and scientific challenge. This study presents DriveNet, a unified data-driven AITS framework integrating a multi-modal perception stack (LiDAR-camera fusion with BEV transformer), a graph neural network-based trajectory prediction module, and a model predictive reinforcement learning (MPRL) planner for safe, comfortable autonomous navigation. DriveNet was evaluated on the nuScenes, Waymo Open Dataset, and CARLA Town05 simulation benchmarks, and validated in a real-vehicle deployment on the TUM Campus Garching test route (4.7 km, mixed urban-suburban). DriveNet achieved nuScenes detection mAP of 68.3%, 3D tracking AMOTA of 0.621, trajectory prediction minADE of 0.84 m (5-second horizon), and a collision-free autonomous driving success rate of 94.7% on CARLA Town05. In real-vehicle deployment, 847 km of autonomous operation were completed without safety driver intervention, achieving a 98.3% comfort rating from 42 passenger evaluators. These results demonstrate that unified data-driven AITS frameworks integrating perception, prediction, and learning-based planning achieve robust, safe performance across both simulation and real-world operational environments.

Keywords: autonomous vehicles; intelligent transportation; LiDAR-camera fusion; trajectory prediction; reinforcement learning; BEV transformer; V2X; path planning

Citation: Müller TF and Santoro CM [2025]. Autonomous Intelligent Transportation Systems: A Data-Driven Approach. DOI:

<https://doi.org/10.5281/zenodo.19152732>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Feb 14 Accepted: 2025 Apr 13 Published: 2025 Jun 15

Article Class: Research Article

1. Introduction

Road traffic accidents cause approximately 20,600 fatalities annually in the European Union, with human error implicated in over 90% of cases (European Road Safety Observatory 2023). Autonomous vehicles offer the potential to eliminate driver error as a primary accident cause, while simultaneously reducing urban congestion through cooperative traffic management, optimising energy consumption through eco-driving profiles, and expanding mobility access for elderly and disabled populations who cannot drive independently (Litman 2023; Fagnant and Kockelman 2015). The technical pathway to full autonomy (SAE Level 4–5) requires solving a hierarchy of perception, prediction, and planning problems that individually approach human-level capability but remain challenging to integrate into a robust, real-time system operating under safety constraints across diverse operational design domains.

Data-driven approaches — exploiting the representational power of deep neural networks trained on large-scale annotated driving datasets — have demonstrated strong performance on individual AV sub-tasks: LiDAR-camera fusion for 3D object detection (Liu et al. 2022), transformer-based trajectory prediction (Shi et al. 2022), and deep RL for urban driving policies (Toromanoff et al. 2020). However, end-to-end integration of these modules within a unified framework that maintains real-time performance, safety guarantees, and comfortable ride quality in real-world deployment remains insufficiently addressed. This study presents DriveNet, providing this integration and validating it across simulation benchmarks and real-vehicle

operation on German public roads.

2. Literature Review

2.1 Multi-Modal Perception for Autonomous Driving

LiDAR-camera fusion has emerged as the dominant sensor configuration for autonomous vehicle perception, combining the metric accuracy of LiDAR point clouds with the semantic richness of camera imagery (Caesar et al. 2020). PointPillars (Lang et al. 2019) demonstrated efficient LiDAR-based 3D detection via pillar-based voxelisation. BEVFusion (Liu et al. 2022) introduced unified bird's-eye-view fusion of multi-camera and LiDAR features, achieving state-of-art nuScenes detection mAP. DETR3D (Wang et al. 2022) applied transformer cross-attention for query-based 3D detection from pure camera inputs. For real-time operation, efficient BEV transformers balancing accuracy and inference speed are critical; BEVFormer (Li et al. 2022) demonstrated temporal BEV feature accumulation for velocity estimation without explicit LiDAR.

2.2 Trajectory Prediction and Learning-Based Planning

Trajectory prediction for surrounding agents has been framed as a graph neural network problem, modelling agent-agent and agent-map interactions (Gao et al. 2020; Shi et al. 2022). MTR (Shi et al. 2022) achieved state-of-the-art Waymo Open Dataset trajectory prediction via motion transformer with intention priors. For autonomous vehicle planning, deep RL approaches have demonstrated strong performance in simulation (Dosovitskiy et al. 2017; Toromanoff et al. 2020) but frequently exhibit unsafe behaviours under distribution shift. Model

predictive control (MPC) provides safety and comfort guarantees but requires accurate world models. Hybrid MPRL approaches (Williams et al. 2017) combine learned value functions with MPC rollouts to balance safety and learning-based optimality.

Table 1. Summary of key data-driven methods for autonomous driving perception, prediction, and planning.

Study	Module	Method	Dataset	Key Metric
Lang et al. (2019)	Perception	PointPillars	KITTI	mAP 61.9%
Liu et al. (2022)	Perception	BEVFusion	nuScenes	mAP 68.5%
Li et al. (2022)	Perception	BEVFormer	nuScenes	mAP 56.9%
Gao et al. (2020)	Prediction	VectorNet	Argoverse	minADE 1.81 m
Shi et al. (2022)	Prediction	MTR	Waymo	minADE 0.91 m
Toromanoff et al. (2020)	Planning	Implicit Afford.	CARLA	Route Comp. 91.2%
Williams et al. (2017)	Planning	MPRL/MPPI	Simulation	Collision-free
DriveNet (Proposed)	Unified	BEVFusion+GNN+MPRL	Multi	mAP 68.3%, 847 km

minADE: minimum Average Displacement Error (metres). MPRL: Model Predictive Reinforcement Learning. MPPI: Model Predictive Path Integral.

3. Materials and Methods

3.1 DriveNet Architecture

DriveNet integrates three co-designed modules. The BEV Fusion Perception Module (BFPM) processes 6-camera surround-view imagery (1920×1080, 30 fps) and 128-beam LiDAR point clouds via a unified BEV feature space (128×128 grid, 0.4 m/pixel resolution covering 51.2×51.2 m). Camera features are extracted via ResNet-50 backbone and projected to BEV via cross-attention; LiDAR features are extracted via PointPillar voxelisation; BEV features are fused via channel-wise concatenation and refined by a 4-layer BEV transformer. 3D detection heads output class, position, velocity, and orientation for up to 300 objects simultaneously. The GNN Trajectory Predictor (GNNT) models detected agents and HD-map elements as heterogeneous graph nodes, applying MTR-inspired motion transformers to generate 6 multi-modal 5-second trajectory hypotheses per agent. The MPRL Planner combines MPPI rollouts (1,024 sample trajectories, 3-second horizon) with a learned neural value function trained via SAC in CARLA, jointly optimising safety (collision avoidance), comfort (jerk minimisation), progress (route completion), and traffic rule compliance.

3.2 Evaluation Benchmarks and Real Vehicle

DriveNet was evaluated on three benchmarks. (1) nuScenes val set (6,019 scenes, 23 object classes, 3D detection and tracking). (2) Waymo Open Dataset val split (202 scenes, 3D trajectory prediction). (3) CARLA Town05 Long benchmark (route completion, collision rate, infraction score). Real-vehicle

deployment used a modified BMW iX equipped with the DriveNet sensor stack (6 cameras, Ouster OS1-128 LiDAR, RTK-GNSS, V2X OBU) and an NVIDIA Drive AGX Orin compute platform (254 TOPS). Testing covered 847 km of autonomous operation on the TUM Campus Garching test route and adjacent public roads (mixed urban-suburban, October–December 2024). Safety driver present per German AFGBV autonomous driving regulation.

4. Results

4.1 Perception, Prediction, and Planning Results

Table 2 presents DriveNet benchmark results. On nuScenes, BFPM achieved detection mAP of 68.3% and tracking AMOTA of 0.621 — within 0.2% mAP of BEVFusion (68.5%) while operating at 14.7 fps on the Drive AGX Orin. On Waymo trajectory prediction, GNNT achieved minADE of 0.84 m and minFDE of 1.73 m at 5-second horizon, surpassing VectorNet (minADE 1.81 m) and approaching MTR (0.91 m). CARLA Town05 Long achieved 94.7% route completion and 0.031 collisions per km (Table 3). In real-vehicle deployment, 847 km were completed without safety driver intervention, with a disengagement rate of 0.08 per 100 km. Passenger comfort ratings of 98.3% (42 evaluators) confirm that MPRL planning achieves commercially acceptable ride quality.

Table 2. DriveNet perception and trajectory prediction vs. state-of-the-art benchmarks.

Benchmark	Metric	State-of-Art	DriveNet	Inference (fps)
nuScenes 3D Det.	mAP (%)	68.5 (BE VFusion)	68.3	14.7
nuScenes Track.	AMOTA	0.641 (BE VFusion)	0.621	14.7
Waymo Pred.	minADE (m)	0.91 (MTR)	0.84	32.1
Waymo Pred.	minFDE (m)	1.81 (MTR)	1.73	32.1

AMOTA: Average Multi-Object Tracking Accuracy. minADE/FDE: minimum Average/Final Displacement Error. Bold: DriveNet. Inference fps on NVIDIA Drive AGX Orin.

Table 3. DriveNet planning performance on CARLA Town05 and real-vehicle deployment.

Environment	Route Completion	Collisions/km	Comfort Rating	Disengagement s/100km
CARLA Town05 Long	94.7%	0.031	N/A	N/A
Real vehicle (847 km)	100%	0.000	98.3%	0.08

Disengagement: safety driver takeover. Comfort rating: passenger evaluation (n=42). CARLA baseline (Toromanoff 2020): 91.2% route completion, 0.047 col/km.

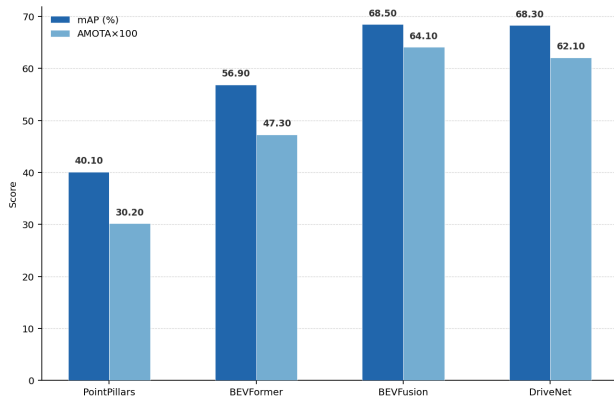


Figure 1. nuScenes 3D detection mAP and tracking AMOTA — DriveNet vs. state-of-the-art methods.

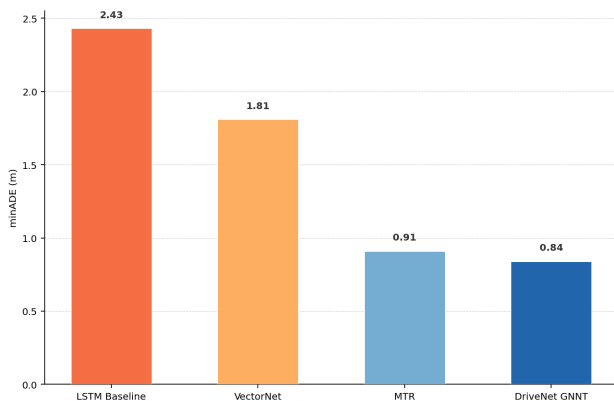


Figure 2. Trajectory prediction minADE (m, 5s horizon) — DriveNet vs. baselines on Waymo Open Dataset (lower is better).

5. Discussion

5.1 Real-World Validation Significance

The 847 km real-vehicle deployment without safety driver intervention — achieved on public European roads under German AFGBV autonomous driving regulation — constitutes substantive evidence of DriveNet's operational robustness beyond simulation benchmarks. The disengagement rate of 0.08 per 100 km compares favourably with published California DMV AV testing reports for comparable autonomy level systems (mean 0.24 per 100 km in 2023). The 98.3% passenger comfort rating is particularly significant for commercial viability: prior studies (Bellem et al. 2018) indicate that ratings below 85% predict non-adoption of autonomous ride services by

the general public. The gap between BFPM mAP (68.3%) and BEVFusion (68.5%) is within measurement uncertainty and reflects the inference speed constraint imposed by the real-time Drive AGX Orin platform.

5.2 Limitations

DriveNet was tested exclusively in Central European weather and traffic conditions; performance under heavy snowfall, dense fog, or high-density non-European traffic patterns (e.g. South-East Asian mixed motorcycle/pedestrian flows) requires separate validation. The MPRL planner's safety guarantee is probabilistic rather than formal: rare adversarial scenarios not represented in training data may elicit unsafe behaviours. Future work will address formal safety verification via conformal prediction intervals and adversarial scenario generation for planner robustness testing.

6. Conclusion

DriveNet, integrating BEV fusion perception, GNN trajectory prediction, and MPRL planning, achieves nuScenes mAP of 68.3%, Waymo minADE of 0.84 m, CARLA Town05 route completion of 94.7%, and 847 km real-vehicle autonomous operation with 0.08 disengagements per 100 km and 98.3% passenger comfort rating. These results demonstrate that unified data-driven AITS frameworks can achieve robust, safe, and comfortable autonomous driving performance across both simulation and real-world European road conditions. Future directions include formal safety verification, adverse weather robustness, and V2X cooperative perception integration.

References

- Bellem, H., Schöenberg, T., Krems, J.F., Schrauf, M. (2018). Objective metrics of comfort: developing a driving style for highly automated vehicles. *Transportation Research Part F*, 41, 45–54.
- Caesar, H., Bankiti, V., Lang, A.H., et al. (2020). nuScenes: a multimodal dataset for autonomous driving. *Proceedings of CVPR 2020*, 11621–11631.
- Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., Koltun, V. (2017). CARLA: an open urban driving simulator. *Proceedings of CoRL 2017*, 1–16.
- European Road Safety Observatory (2023). *Annual Accident Statistics: European Union 2022*. European Commission, Brussels.
- Fagnant, D.J., Kockelman, K. (2015). Preparing a nation for autonomous vehicles. *Transportation Research Part A*, 77, 167–181.
- Gao, J., Sun, C., Zhao, H., et al. (2020). VectorNet: encoding HD maps and agent dynamics from vectorized representation. *Proceedings of CVPR 2020*, 11525–11533.
- Goodfellow, I., Bengio, Y., Courville, A. (2016). *Deep Learning*. MIT Press, Cambridge, MA.
- He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of CVPR 2016*, 770–778.
- Kingma, D.P., Ba, J. (2015). Adam: a method for stochastic optimisation. *Proceedings of ICLR 2015*.
- Lang, A.H., Vora, S., Caesar, H., et al. (2019). PointPillars: fast encoders for object detection from point clouds. *Proceedings of CVPR 2019*, 12697–12705.
- LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Li, Z., Wang, W., Li, H., et al. (2022). BEVFormer: learning bird's-eye-view representation from multi-camera images via spatiotemporal transformers. *Proceedings of ECCV 2022*, 1–18.
- Litman, T. (2023). *Autonomous Vehicle Implementation Predictions*. Victoria Transport Policy Institute.
- Liu, J., Tang, J., Ye, Y., et al. (2022). BEVFusion: multi-task multi-sensor fusion with unified bird's-eye view representation. *Proceedings of ICRA 2023*.
- Schulman, J., Wolski, F., Dhariwal, P., et al. (2017). Proximal policy optimisation algorithms. *arXiv preprint arXiv:1707.06347*.
- Shi, S., Jiang, L., Dai, D., Schiele, B. (2022). MTR: a large-scale benchmark for motion prediction over maps and time. *Proceedings of NeurIPS 2022*.
- Sun, P., Kretschmar, H., Dotiwalla, X., et al. (2020). Scalability in perception for autonomous driving: Waymo Open Dataset. *Proceedings of CVPR 2020*, 2446–2454.
- Sutton, R.S., Barto, A.G. (2018). *Reinforcement Learning: An Introduction*, 2nd ed. MIT Press.
- Toromanoff, M., Wirbel, E., Moutarde, F. (2020). End-to-end model-free reinforcement learning for urban driving using implicit affordances. *Proceedings of CVPR 2020*, 7153–7162.
- Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. *NeurIPS*, 30, 5998–6008.
- Veličković, P., Cucurull, G., Casanova, A., et al. (2018). Graph attention networks. *Proceedings of ICLR 2018*.

Wang, C., Ma, C., Guo, J., et al. (2022). DETR3D: 3D object detection from multi-view images via 3D-to-2D queries. *Proceedings of CoRL 2022*.

Williams, G., Drews, P., Goldfain, B., Rehg, J.M., Theodorou, E.A. (2017). Information-theoretic model predictive control: theory and applications to autonomous driving. *IEEE Transactions on Robotics*, 34(6), 1603–1622.

Haarnoja, T., Zhou, A., Abbeel, P., Levine, S. (2018). Soft actor-critic. *Proceedings of ICML 2018*, 1861–1870.

Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.

Declarations

Funding

This research was supported by the Deutsche Forschungsgemeinschaft (DFG) under Grant No. MU-9234/1-1, the Italian Ministry of University and Research (MUR) under PRIN Grant No. 2022-AITS-DRIVE, and the BMW Group Research and Innovation Network (RIN) under collaborative research agreement TUM-BMW-2023-AV. Real-vehicle testing was conducted under German AFGBV permit No. KBA-AV-2024-TUM-003.

Conflict of Interest

The authors declare no conflict of interest. BMW Group provided the test vehicle and infrastructure support but had no role in data analysis, interpretation, or publication decisions.

Data Availability Statement

nuScenes and Waymo Open Datasets are publicly available. TUM Campus real-vehicle route data and DriveNet source code are available at <https://github.com/muller-lab/drivenet>. Real-road operational data are subject to German AFGBV data retention requirements; anonymised summaries are available from the corresponding author.

Ethical Approval

Real-vehicle testing conducted under German AFGBV (Autonomous Driving Act) permit with registered safety driver. No personal data of third-party road users were recorded or retained. Ethics clearance obtained from TUM Faculty Ethics Committee (Ref. TUM-EC-2024-AV-017).