NEURAL INFORMATION PROCESSING SYSTEMS

MAN TruckScenes: A Multimodal Dataset for Autonomous Trucking in Diverse Conditions

Felix Fent, Fabian Kuttenreich, Florian Ruch, Farija Rizwin, Stefan Juergens, Lorenz Lechermann, Christian Nissler, Andrea Perl, Ulrich Voll, Min Yan, Markus Lienkamp

The Thirty-Eighth Annual Conference on Neural Information Processing Systems

December 13th, 2024 | 01:30 – 04:30PM

Vancouver Convention Center

Introduction

Why do we need another dataset?

- Autonomous trucking can change the future of logistics
- Perception development needs large-scale datasets
- No large-scale dataset exists for autonomous trucks

  … different operational design domain

  … different sensor mounting positions

  … movable truck-trailer combination

- The world's first dataset for autonomous trucking

	KITTI	nuScenes	Waymo	Argo2	Ours
Scenes	22	1000	1150	1000	747
Sample	1,5 k	40 k	230 k	150 k	30 k
Duration	1,5 h	5,5 h	6,4 h	4,2 h	4,2 h
Coverage	-	4 km²	76 km²	17 km²	100 km²
Camera	4	6	5	9	4
Lidar	1	1	5	2	6
Radar	0	5	0	0	6
GNSS	1	1	0	0	1
IMU	1	1	0	0	2
Map	no	yes	no	yes	no
Range	91 m	141 m	80 m	214 m	226 m
Classes	3	23	4	30	27
Vehicle	car	car	car	car	truck

A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets robotics: The KITTI dataset. The InternationalJournal of Robotics Research, 32(11):1231–1237, Aug. 2013.

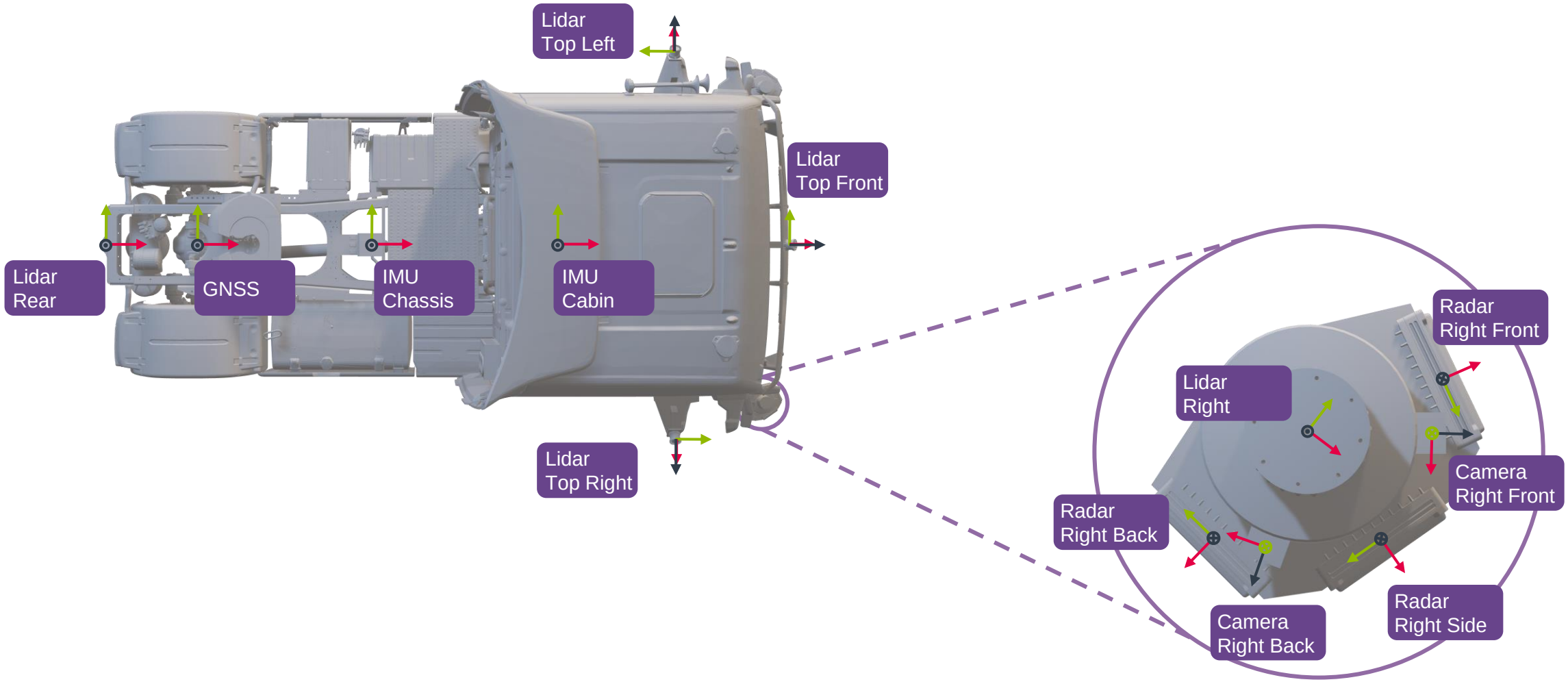
H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom. nuScenes: A Multimodal Dataset for Autonomous Driving. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, June 2020.

P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, Y. Zhang, J. Shlens, Z. Chen, and D. Anguelov. Scalability in Perception for Autonomous Driving: Waymo Open Dataset. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, June 2020.

B. Wilson, W. Qi, T. Agarwal, J. Lambert, J. Singh, S. Khandelwal, B. Pan, R. Kumar, A. Hartnett, J. K. Pontes, D. Ramanan, P. Carr, and J. Hays. Argoverse 2: Next Generation Datasets for Self-Driving Perception and Forecasting. In Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks (NeurIPS Datasets and Benchmarks 2021), 2021.

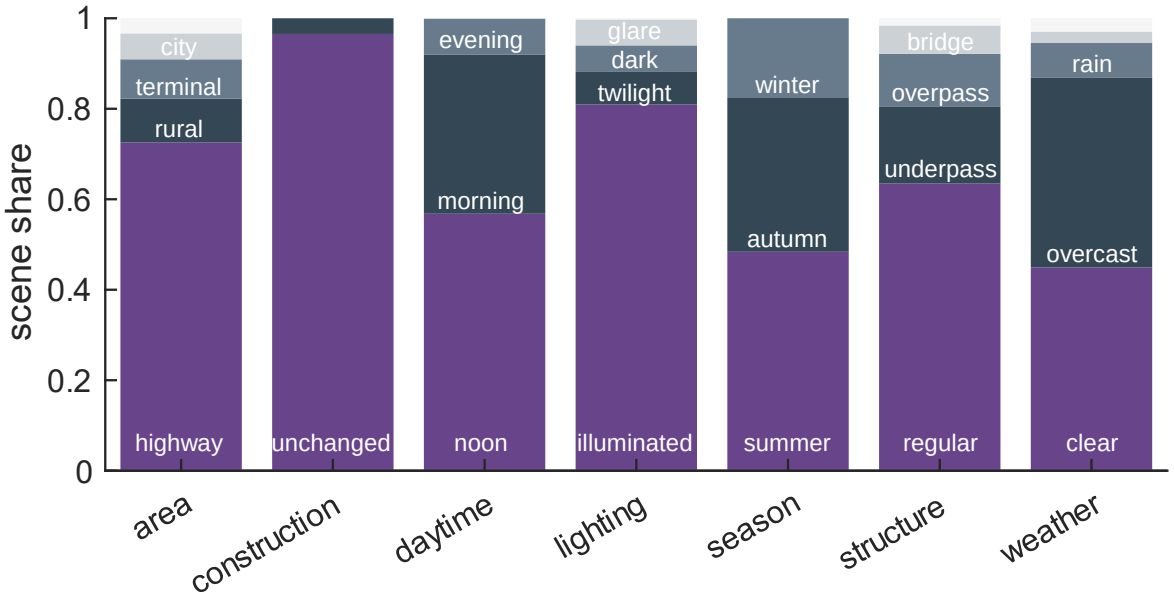
The Autonomous Truck

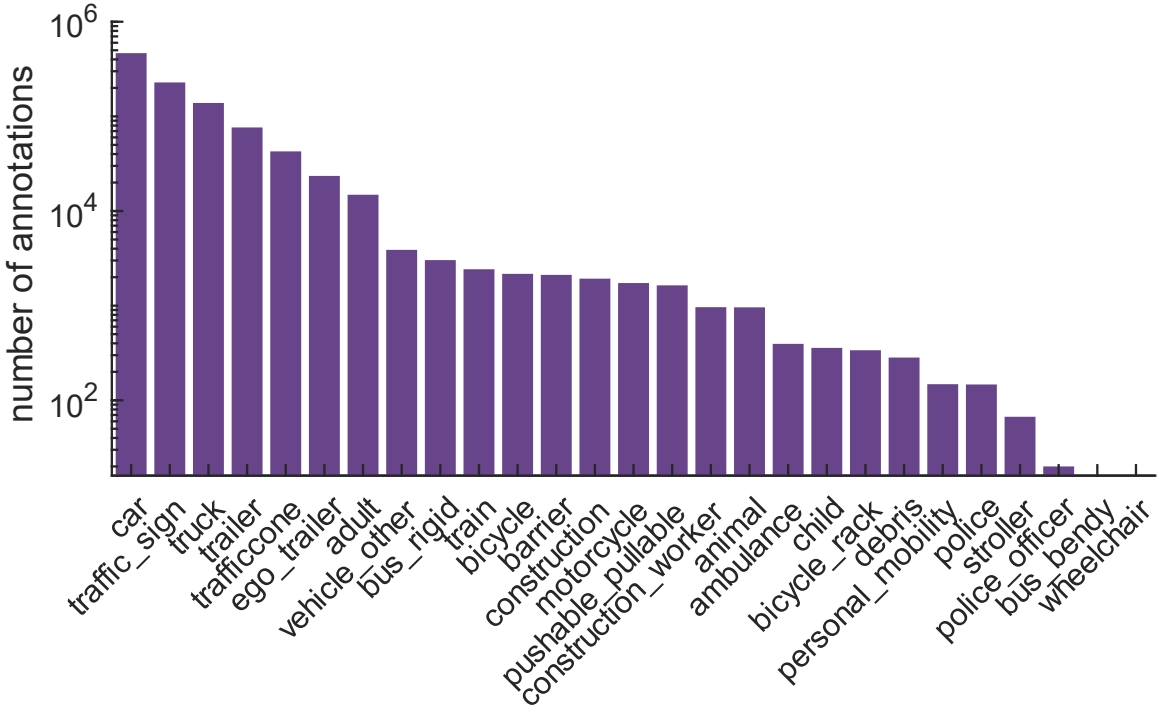
Multimodal data collection



MAN TruckScenes

The world's first dataset for autonomous trucking





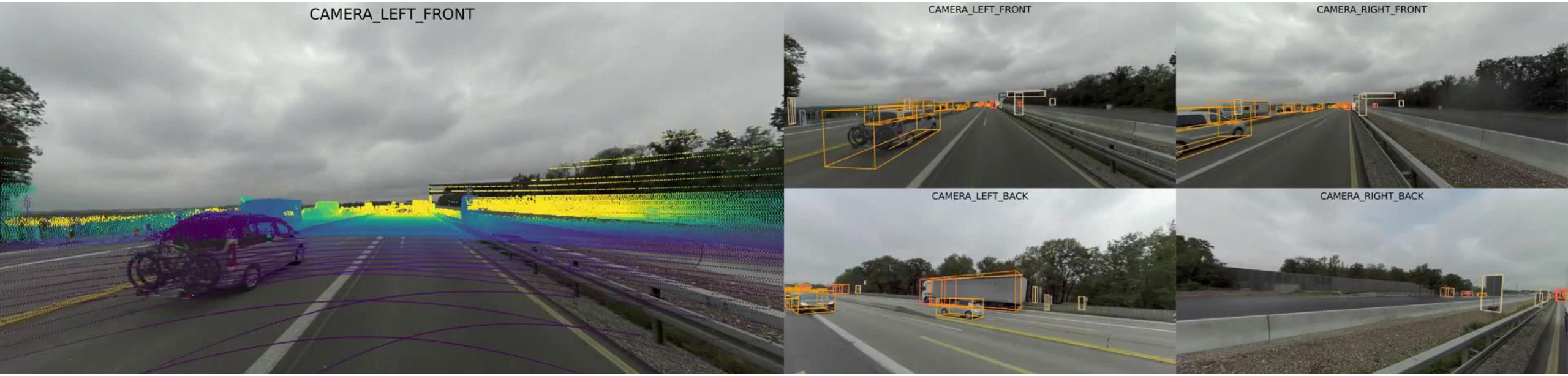
› Diverse driving scenes

› A multitude of object classes

MAN TruckScenes

The world’s first dataset for autonomous trucking

NEURAL INFORMATION PROCESSING SYSTEMS



Results

A baseline to beat

Method	Modality	mAP	NDS	ATE	ASE	AOE	AVE	AAE
PETR	Camera	0.02	0.12	1.13	0.69	0.65	1.50	0.56
RadarGNN	Radar	0.07	0.11	0.89	0.81	1.13	8.00	0.57
CenterPoint	Lidar	0.27	0.41	0.41	0.35	0.28	2.73	0.20

- Long-range detection remains challenging
- The detection quality of diverse object classes is insufficient
- No single sensor modality is sufficient for the diverse set of environmental conditions

Y. Liu, T. Wang, X. Zhang, and J. Sun. PETR: Position Embedding Transformation for Multi-view 3D Object Detection, page 531–548. Springer Nature Switzerland, 2022.
F. Fent, P. Bauerschmidt, and M. Lienkamp. Radargnn: Transformation invariant graph neural network for radar-based perception. In 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, June 2023.
T. Yin, X. Zhou, and P. Krahenbuhl. Center-based 3D Object Detection and Tracking. In 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, June 2021.