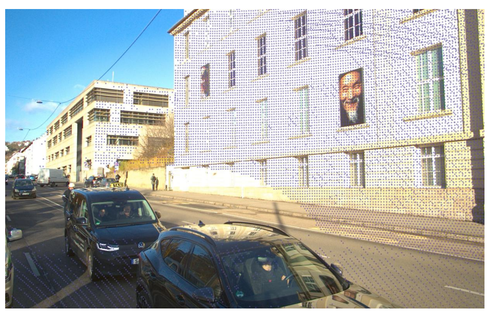
IEEE IV 2025
CLUJ-NAPOCA, ROMÂNIA
36th IEEE Intelligent Vehicles Symposium

HOCHSCHULE ESSLINGEN

CoopScenes: Multi-Scene Infrastructure and Vehicle Data for Advancing Collective Perception in Autonomous Driving

Marcel Vosshans, Alexander Baumann, Matthias Drueppel, Omar Ait-Aider, Youcef Mezouar, Thao Dang and Markus Enzweiler





Specification

The Ego-Vehicle

The ego-vehicle is a small public transport bus equipped with 10 sensors arranged symmetrically around a stereo camera and top-mounted Ouster LiDAR. Two angled front LiDARs reduce blind spots caused by the vehicle's height and overall dimensions. Wide-angle Basler cameras to the front and back enhance lateral coverage, while an INS with RTK-corrected GNSS ensures precise localization.

	Camera		LiDAR		INS
	SL, SR, BL, BR	FL, FR	TOP	L, R	GNSS/IMU
Model	a2A1920-51gcPRO	a2A1920-51gcPRO	OS1	OS0	3DM-GQ7
Resolution	1920 × 1200	1920 × 1200	1024 × 128	1024x128	-
Frequency	10 Hz	10 Hz	10 Hz	10 Hz	1 kHz
HFoV/ VFoV	57.6°/ 37.7°	79.1°/ 54.3°	360°/ 45°	360°/ 90°	-
Details	Exposure Time: 800μs, Focal Length: 6 mm, Aperture: f/4.0	Exposure Time: 800μs, Focal Length: 4 mm, Aperture: f/4.0	Range (10%): 90 m	Range (10%): 35 m	Accuracy: 1.0 cm RTK, 0.2°heading, 0.05°pitch/roll

CoopScenes

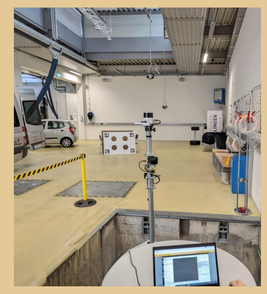
A novel dataset that captures diverse driving scenarios from two complementary viewpoints: an intelligent vehicle and an infrastructure observer.

Specification

The Infrastructure Tower

The sensor tower serves as a static reference for the ego-vehicle, emphasizing modularity and autonomy. It features a 3 m pole on a weighted base with battery and solar power. It supports LTE/5G and V2X communication. The tower includes a central 360° LiDAR with GNSS, and two adjustable camera-LiDAR units for scene-specific alignment. Precise positioning is achieved via GNSS-based point positioning.

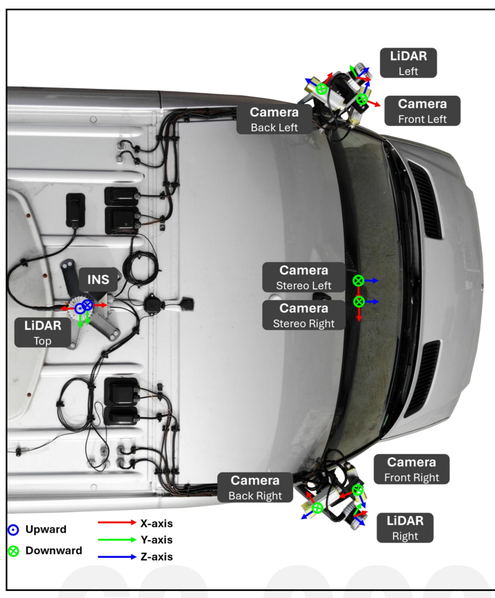
	Camera	LiDAR		GNSS
	V1, V2	V1, V2	UP	RTK Base
Model	a2A1920-51gcPRO	Cube 1 Outdoor	OS2	C099-F9P
Resolution	1920 × 1200	400 × 51	1024 × 128	-
Frequency	10 Hz	10 Hz	10 Hz	-
HFoV/ VFoV	57.6°/ 37.7°	70°/ 30°	360°/ 22,5°	-
Details	Exposure Time 800μs, Focal Length 6 mm, Aperture f/4.0	Range (10%): 30 m	Range (10%): 200 m	PPP with > 1 h data, accuracy variable



Calibration

Extrinsic

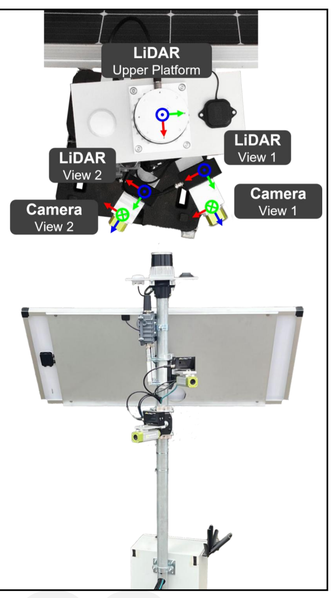
We used velo2cam[1] calibration to estimate the extrinsic transformation on a calibration target.

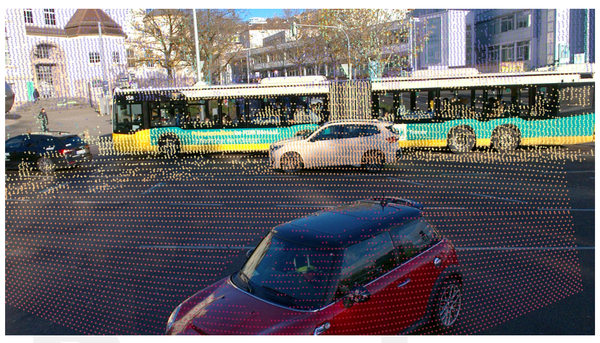
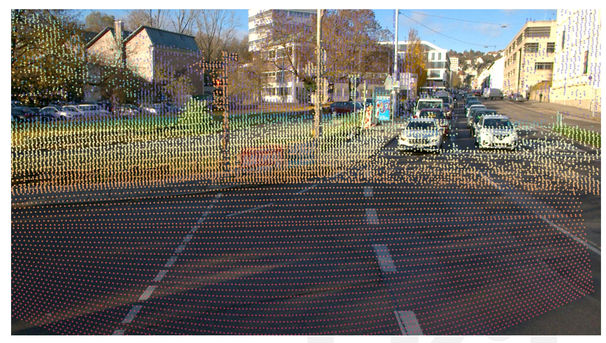


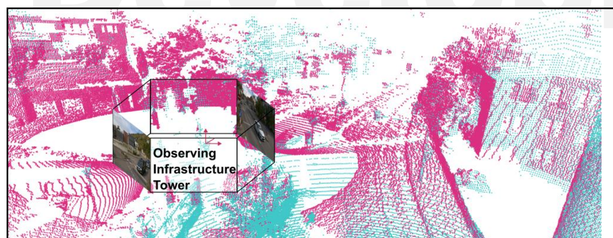
Specification

The Setup

Focusing on public transport, the dataset features a perception-equipped bus and a mobile sensor tower, deployed in real-world scenarios. It provides spatially and temporally synchronized data at 10 Hz, using PTP and GNSS-based agent time synchronization with a mean deviation of 2.3 ms, and features precise point cloud registration and automated annotations.





Purpose & Tasks

Collective Perception

The dataset supports common tasks like detection, segmentation, and SLAM. Its scale and diversity make it particularly suited for *deep learning* applications. CoopScenes enables research in collective perception—enhancing vehicle awareness via infrastructure support—sensor fusion and transformation confidence, sensor system design optimization, and behavior prediction through repeated observation of real-world situations.

Recording

Scene Book

The dataset includes 104 minutes of data from 9 locations in the Stuttgart region, covering 100 maneuvers across diverse scenarios. It comprises 62,465 frames (527 GB) in a compressed *.4mse* format.

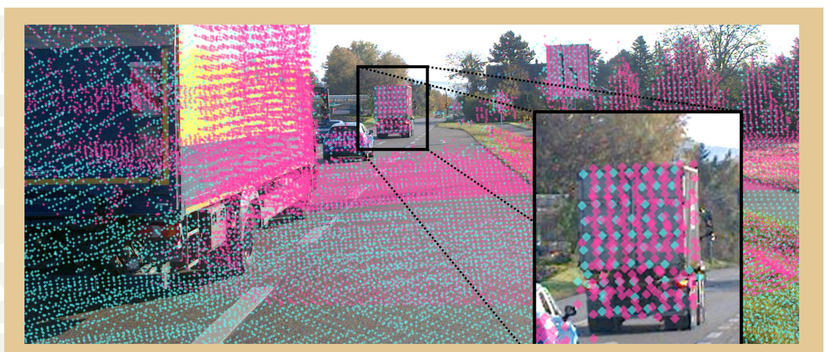
Ground Truth

Reference Labels

To ensure scalability, CoopScenes integrates automated pipelines for annotation and anonymization. The dataset includes 5.54 million detected instances, averaging 12.3 objects per frame. Initial annotations include 2D bounding boxes generated using a Co-DETR[2] model, achieving 90.6% precision and 89.2% recall for vehicles (based on a hand labeled test dataset).

Anonymization

For anonymization, we publish *BlurScene*, a Faster R-CNN-based model that outperforms EgoBlur[3], achieving 96% recall for faces and 98% for license plates. By introducing an additional label distinction between identifiable and non-identifiable instances, we optimized the detection threshold to minimize unnecessary blurring. The model, including code and pretrained weights, is freely available for all purposes.

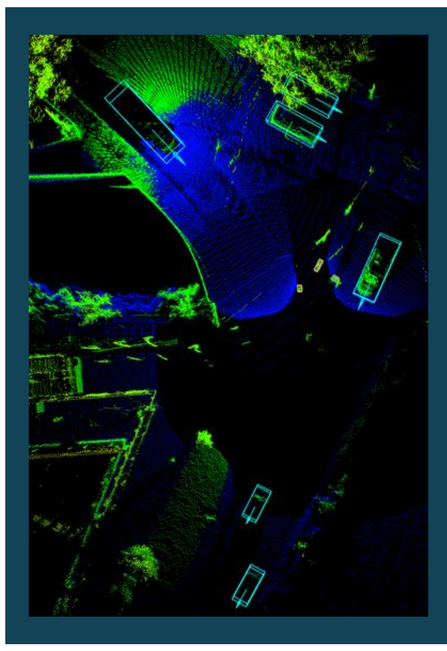


Calibration

LiDAR-Camera Projection

We estimate the inter-agent transformation via LiDAR-based KISS-ICP[4] , supported by Fast Point Feature Histograms (FPFH)[5].

Assuming the tower as global origin, the initial alignment is selected from GNSS-proximal frames using RANSAC and refined with ICP, then clustered with DBSCAN. Transformations are propagated, enabling estimates even without direct sensor overlap.



CONCLUSIONS

Conclusion & Future Works

We present CoopScenes, the first large-scale European dataset for cooperative perception, featuring synchronized infrastructure-vehicle data across diverse public transport scenarios. The dataset includes automatic annotation, anonymization pipelines, and spatial registration via agent transformations, enabling neural network training and benchmarking. Future extensions will include 3D labels to support broader tasks such as real-time registration and intention prediction.

Department of Computer Science and Engineering, Institute for Intelligent Systems, University of applied Science Esslingen, Germany

Institut Pascal ISPR (Image, Systems of Perception, Robotics), Universite Clermont Auvergne INP / CNRS, France

[1] J. Beltrán, C. Guindel, A. de la Escalera, and F. García, "Automatic ExtrinsicCalibration Method for LiDAR and Camera Sensor Setups," IEEE Transactions onIntelligent Transportation Systems, vol. 23, no. 10, pp. 17 677–17 689, Oct. 2022
[2] Z. Zong, G. Song, and Y. Liu, "DETRs with Collaborative Hybrid AssignmentsTraining," in 2023 IEEE/CVF International Conference on Computer Vision(ICCV). Paris, France: IEEE, Oct. 2023
[3] N. Raina, G. Somasundaram, K. Zheng, S. Miglani, S. Saarinen, J. Meissner,M. Schwesinger, L. Pesqueira, I. Prasad, E. Miller, P. Gupta, M. Yan, R. Newcombe,C. Ren, and O. M. Parkhi, "EgoBlur: Responsible Innovation in Aria," 2023
[4] I. Vizzo, T. Guadagnino, B. Mersch, L. Wiesmann, J. Behley, and C. Stachniss, "KISS-ICP: In Defense of Point-to-Point ICP – Simple, Accurate, and Robust Registration IfDone the Right Way," IEEE Robotics and Automation Letters, 2023
[5] R. B. Rusu, N. Blodow, and M. Beetz, "Fast Point Feature Histograms (FPFH)for 3D registration," in 2009 IEEE International Conference on Robotics andAutomation. Kobe: IEEE, May 2009, pp. 3212–3217

Baden-Württemberg Ministerium für Verkehr
VOLKMANN
DHBW Duale Hochschule Baden-Württemberg
kaitos.
ITSS IEEE Intelligent Transportation Systems Society