The DLR Urban Traffic Dataset (DLR-UT): A Comprehensive Traffic Dataset from the AIM Research Intersection

Clemens Schicktanz*, Lars Klitzke*, Kay Gimm*, Giancarlo Rizzo*, Karsten Liesner*, Henning Hajo Mosebach*, Sascha Knake-Langhorst*

***Abstract*— Current trajectory datasets of traffic participants often lack detailed environmental information, which is crucial for developing effective data-driven methods for future mobility solutions. To address this gap, we introduce the comprehensive DLR Urban Traffic dataset version 1.2.0. The dataset includes 32,296 trajectories of traffic participants, along with traffic light data, local weather data, air quality data, and road condition data collected at the Application Platform for Intelligent Mobility Research Intersection during a single day of recording. A comparison with other publicly available datasets reveals that our dataset offers more comprehensive information about the traffic environment than existing alternatives. An analysis of our dataset shows that trajectories are available for all possible 16 routes at the intersection, with the number of trajectories per route varying significantly, ranging from 11 to 3,344. Most interactions between motorized road users occur during unprotected left turns with oncoming traffic. However, there are also interactions between motorized and vulnerable road users, particularly during right turns. All in all, the dataset provides researchers with the resources needed to improve urban mobility solutions. Available for non-commercial use, the dataset can be directly downloaded from `https://doi.org/10.5281/zenodo.14773161`.**

***Index Terms*— Trajectory Dataset, Urban Traffic Data, Automated Vehicles, Traffic Research**

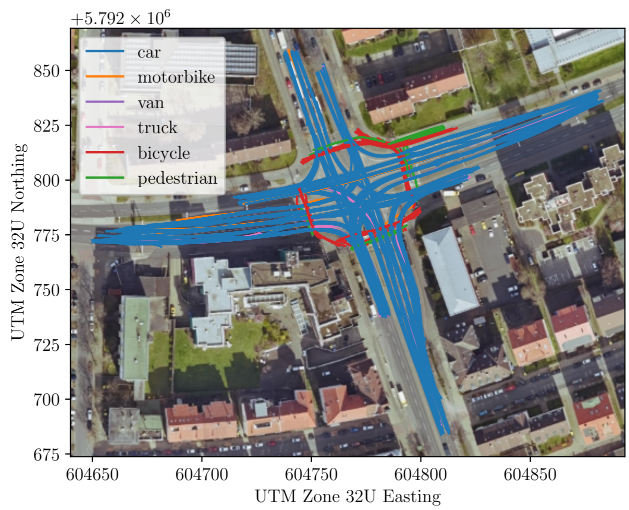

Fig. 1: Visualization of trajectories from the DLR-UT dataset over a 15-minute period. Satellite image in the background: © The City of Braunschweig, Department Geographic Information.

I. Introduction

Road traffic poses a significant problem in modern society, as it contributes not only to fatalities and injuries but also to the rising levels of carbon dioxide emissions that drive climate change [1], [2]. To address these issues, transportation research aims to enhance both the safety and efficiency of traffic systems. Central to this research is the collection and analysis of traffic data, which are crucial for quantifying the behavior of traffic participants and developing effective solutions [3].

In recent years, there has been an increasing emphasis on gathering and sharing traffic data, particularly trajectory datasets, which have proven valuable for a wide range of analyses [4]. The use of stationary cameras for recording videos has proven to be highly effective for collecting traffic information [5]. However, many of the collected stationary-mounted cameras are demounted after the end of the related research project. That limits a comprehensive measurement which is necessary for a holistic description of traffic situations according to the *6 Layer Model* [6]. With recent technological advancements, drones have emerged in the field of traffic data collection as an alternative [5]. However, drone datasets are often recorded during good weather or consisting of trajectory data only, lacking information about the environmental context. That can hinder the development of models capable of operating in challenging environmental conditions [7].

To overcome these limitations, future datasets for automated driving should encompass extensive descriptions of real-world traffic situations and environmental conditions, enabling a variety of research questions [7]. Testbeds designed and financed for long-term operation are suitable for generating such datasets. The *German Aerospace Center (DLR)* initiated its first activities in testbeds for automated driving in 2009 [8]. That includes the establishment of the *Application Platform for Intelligent Mobility (AIM)* [9] and the *Testbed Lower Saxony* [10]. A key component of the AIM is the *AIM Research Intersection* [11], which has been

* Clemens Schicktanz, Lars Klitzke, Kay Gimm, Giancarlo Rizzo, Karsten Liesner, Henning Hajo Mosebach, and Sascha Knake-Langhorst are with the German Aerospace Center (DLR), Institute of Transportation Systems, Braunschweig, Germany, {clemens.schicktanz, lars.klitzke, kay.gimm, giancarlo.rizzo, karsten.liesner, henning.mosebach, sascha.knake-langhorst}@dlr.de.

TABLE I: Comparison of our dataset with relevant publicly available datasets (”x” indicates data available, ”-” indicates data not available, ”*” indicates approximate value).

Dataset Name	Duration in Hours	Number of Trajectories	Digital Map	Traffic Lights	Weather	Air Quality	Road Condition
inD	10	13,599	x	x	-	-	-
Waterloo	1	x	x	x	-	-	-
CitySim dataset	19	x	x	x	-	-	-
City-scale Data	x	5,000,000*	x	-	-	-	-
Interaction	1	3,775	x	-	-	-	-
TUMDOC-MUC	42	x	-	x	-	-	-
pNEUMA	x	500,000*	-	-	-	-	-
TUMTraf Intersection	x	506	-	-	-	-	-
VRU Trajectory	x	1,532	-	-	-	-	-
Ko-PER	x	350	-	-	-	-	-
DLR-UT (ours)	24	35,090	x	x	x	x	x

continuously used and developed across various projects [9], [11]. At the intersection, video recordings, trajectory data, traffic light data, weather data, road condition data, and air quality data are collected.

In this publication, we introduce the *DLR Urban Traffic dataset v1.2.0* [12], a publicly available dataset recorded at the *AIM Research Intersection* in Braunschweig, Germany. This dataset provides researchers with a valuable resource for exploring real-world road traffic, positioning itself as a stationary recorded dataset for comprehensive mobility research. Position values of the trajectories are depicted in Fig. 1.

The structure of this publication is as follows. Section II will provide an overview of other relevant publicly available datasets and highlight our contribution. Subsequently, Section III will delve into the data collection and processing methodologies. Section IV will present the dataset, including the number of trajectories and interactions captured. Finally, the discussion in Section V will reflect on the dataset's contributions, critically assess its limitations, and explore potential avenues for future enhancements. The publication will conclude with a summary of our findings in Section VI.

II. Related Work

A. Dataset Comparison

In recent years, various datasets have been released in the field of traffic research. This section highlights how the dataset presented in this publication differs from others and outlines its contribution.

Trajectory datasets can be categorized into three main types based on their data collection method. One approach involves collecting trajectory data directly from vehicles, as demonstrated by the *Waymo Open Dataset* [13]. For this literature review, however, we focus on datasets collected via drones or stationary setups mounted on roadside infrastructure. These types of datasets cover specific areas, making them particularly suitable for direct comparison with our dataset. Among the first widely used datasets was the *Next Generation Simulation (NGSIM)* [14] dataset released in 2006. The *highD* [15] dataset, introduced in 2018, was the first large-scale dataset collected from public roads using drones and has since been extensively utilized . An increasing number of datasets have been developed for urban environments to address more complex traffic scenarios than those typically found on highways. The *inD* [16] dataset is one such example. Similar to the *Waterloo* [17] and *CitySim* [18] datasets, it provides digital maps along with traffic light data. Other datasets, like the *City-scale vehicle trajectory* [19] dataset and the *Interaction* [20] dataset, only offer digital maps without additional traffic data. The *TUMDOT-MUC* [21] dataset, collected via drones, goes a step further by providing traffic light data from traffic light systems and traffic volume information from inductive loops. In contrast, the larger *pNEUMA* [22] dataset includes trajectory data from over 100 intersections but lacks additional contextual information. Smaller datasets, such as *TUMTraf Intersection* [23], *VRU Trajectory* [24], and *Ko-PER* [25], also focus solely on trajectory data. In Ingolstadt, Germany, live traffic data are available, including information on traffic volume and local weather conditions [26].

Table I compares our dataset with the most similar datasets, highlighting the information included in each of them. Table I emphasizes that our dataset stands out by providing additional information for a more detailed description of the traffic situation. It is important to note that the table does not represent all existing datasets. The table is based on the contents of surveys [7], [27], [28] and the referenced dataset publications.

Starting from dataset version 1.2.0, traffic volume data for the lanes on the road (excluding sidewalks) and files containing all data in OpenSCENARIO format are also included in our dataset. However, since these are not raw data but results extracted from the raw data, they are not included in Table I. The presence of this metadata highlights the intention behind our dataset release. In the future, additional annotations and metadata will be added to the dataset, including critical braking maneuvers, lane changes, and red-light violations, to provide users with a more comprehensive traffic representation and enable higher-level analyses.

B. Data Usage

To illustrate potential applications of this dataset, several examples from related works are presented below.

In [29], the potential of collecting a large dataset from the

AIM Research Intersection for scenario-based testing was highlighted. The study demonstrated that a rare illegal U-turn maneuver at the intersection could be extensively described due to continuous data recording. In [30], the digital map, trajectory data, traffic light data, and weather data were used to assess the impact of traffic congestion on safety and efficiency at a signalized intersection. Safety was evaluated by measuring time gaps between intersecting vehicles using the Post-Encroachment Time. Additionally, the data were combined to calculate metrics such as traffic volume and delay to assess traffic efficiency. In [31], the data were used to present a real-world example of a corner case trajectory and categorize it within the established taxonomy of corner case trajectories. In [32], the correlation between traffic density and air quality data was analyzed and compared with recordings from a mobile measurement vehicle. In [33], it was shown that even at a signalized intersection with many protected maneuvers, rare critical situations can occur. The study illustrated how such situations can be captured in a dataset covering almost half a year, and what knowledge can be derived from this data. It was also shown that combining weather data with trajectory data provides a more realistic representation of traffic conditions, making the interpretation of the traffic situation more accurate.

In summary, the presented examples demonstrate how the dataset can be utilized to explore various aspects of traffic safety, efficiency, and the analysis of traffic in relation to environmental factors, highlighting the dataset's broad applicability in transportation research.

III. Method

The dataset was recorded using a data processing toolchain, which is continuously developed on a project basis over several years. The following section describes the system architecture and the data processing procedures.

A. Technical Setup

The distributed object tracking system is composed of 14 multi-sensors mounted on 10 streetlight poles. 8 of 14 installations are strategically placed on the four central traffic islands of the intersection, with each sensor aimed at the opposite side, as illustrated in the bird's eye view in Fig. 2 (blue aims inside the intersection, green aims outside the intersection). This setup creates a redundant system that effectively captures all relevant objects within the central area of the intersection, significantly reducing the risk of occlusion. 4 other systems (red) aim at the crossings of vulnerable road users (VRU) to increase their quality of detection. Another 2 systems (yellow) are placed at the northeastern area of the intersection to enable analysis of VRU behavior on the sidewalks.

As depicted in Fig. 3, each multi-sensor system includes a pair of stereo sensors (orange) and active infrared lighting (red) to enhance scene visibility. Each stereo sensor (*GiGEVision Prosillica GT2750*) is connected to a signal processing server which requests the primary camera with 10 Hz frequency. The primary camera triggers the secondary camera and infrared lighting to provide synchronized images from both cameras.

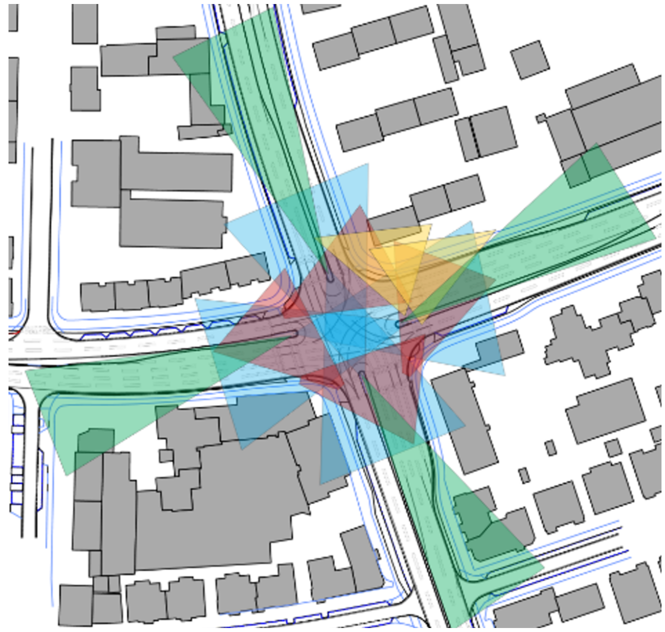

Fig. 2: Camera Field of Views (FOVs) at AIM Research Intersection.

The traffic light data is recorded from the traffic light system, which consists of 30 individual signals at the intersection. The state of each signal is recorded every second. As other data of the dataset, these data are then stored in a database, where each entry is indexed by a timestamp and unique identification number (ID) of the signal.

The weather, air quality, and road condition data are collected from sensors mounted on a traffic light pole as shown in Fig. 4. The sensor used for air quality measurements is the *Air Quality Transmitter AQT530*. General weather information is collected using the *Weather Transmitter WXT536* sensor. The *Silicon Pyranometer SP Lite2* sensor is used to measure sunshine intensity, while the *Present Weather and Visibility Sensor PWD22* is responsible for determining visibility. The *Remote Road State Sensor DSC211* and *Remote Road Surface Temperature Sensor DST111* sensors are employed to assess the condition of the road surface.

B. Data Processing Pipeline

The cameras are calibrated using a high-accuracy digital map from a bird's-eye perspective, where image points are manually marked and linked to corresponding points in the digital map. In addition, an optical flow field is derived to incorporate visual movement information into the process of object detection. Optical measurements are converted into the Universal Transverse Mercator (UTM) coordinate system to enable the implementation of an object detection algorithm based on visual real-time data from all installations. That includes, that objects are identified by their movement patterns at the pixel level in consecutive images. From the resulting flow field, 3D voxels are generated, which are then

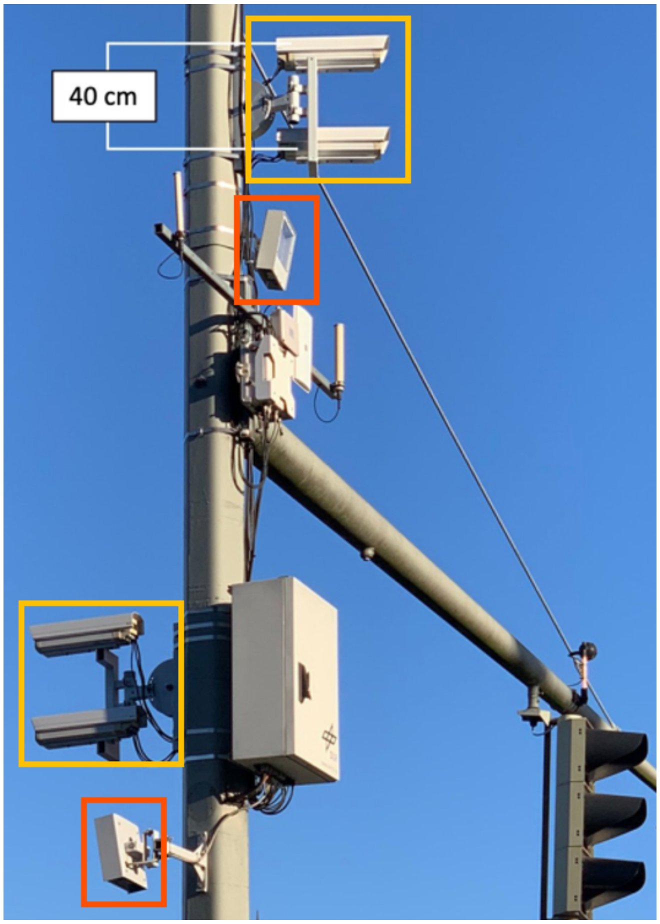

Fig. 3: Multi-sensor system on the pole: The upper system is directed outward, while the lower system faces the intersection.

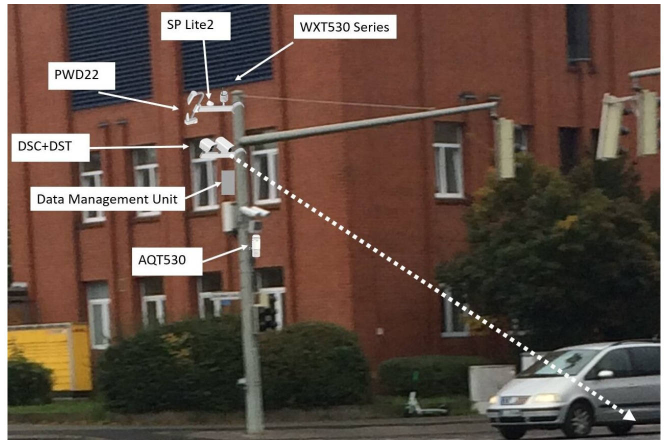

Fig. 4: Weather station mounted on the pole, including sensor labels and the measurement range of the DSC and DST sensors (dashed lines).

Fig. 5: Images of cameras facing the intersection. Top left: facing south. Top right: facing west. Bottom left: facing north. Bottom right: facing east.

aggregated into objects in the form of bounding boxes. That results in multi-object detection on the whole intersection.

Low-resolution images from the view of the central area of the intersection (blue FOVs of Fig. 2) are shown in Fig. 5. The raw high-resolution video data is processed directly at the AIM Research Intersection in a server house to receive anonymized trajectory data. High-resolution images are converted to anonymized low-resolution images for later analysis.

For each detected object in every image, a new entry, also referred to as *object pose*, is created in the form of a new row in the trajectory dataset. Since an ID is generated for each detected object, all data associated with a specific object ID represents the trajectory of that object. In a subsequent step, the trajectory data is processed offline to improve data quality. Trajectories with a duration of less than 2 s, a length of less than 4 m, or a distance between the start and end points of less than 2 m are excluded from the dataset. Additionally, the medians of the object classification probabilities and the size measurements are computed for each trajectory and written into all object poses of the respective trajectory. Furthermore, a Kalman Filter is applied for smoothing and interpolating the position to 20 Hz. Finally, the velocity, acceleration, and heading values are derived from the position data. The heading is kept constant when the velocity drops below 1.5 m/s to ensure that the heading remains stable for objects with minimal movement.

Weather, road condition, air quality, and traffic light data are exported directly from a database to the dataset without further processing. However, the weather, road condition, and air quality data represent measurements from a specific area as the sensors are mounted on a traffic light pole in the southeastern part of the intersection. It is important to note this limitation when analyzing the data, as measurements like water layer thickness cannot be generalized to the entire intersection but for that specific location as shown in Fig. 4.

C. Tooling

The dataset is designed to ensure that its usage is as user-friendly as possible. Therefore, we have provided a Python library, *Traffic Analysis and Situation Interpretation (TASI)* [34], which simplifies both the download and analysis of the data. The library is actively maintained and continuously updated with new functions for analyzing traffic data. It can also be used to load data from the *DLR Highway Traffic* dataset [35], as it is provided in the same format.

IV. Dataset Description

The DLR UT v1.2.0 [12] dataset was recorded using the setup presented in the previous section. This chapter provides statistics about the dataset to give potential users an overview of whether the dataset is suitable for their specific use case. A comprehensive description of the dataset, including its format, can be found in the accompanying documentation, which is available on Zenodo together with the dataset.

A. Key Facts

The dataset comprises trajectory data of traffic participants, along with traffic light data, local weather data, air quality data, and road condition data from the AIM Research Intersection, and was recorded on Sunday, September 24, 2023.

The trajectory data is indexed by object ID and timestamps, including detailed information about the position, velocity, acceleration, dimensions, and classification of each object. The dataset contains 32,296 trajectories including 15,328,297 object poses, covering both motorized road user (MRU) and VRU of object classes pedestrian, bicyclist, motorbike, car, van, and truck.

The traffic light data captures the current state of all 30 traffic lights at the intersection. The weather data provide information on wind, sunlight, precipitation, visibility, and more. The air quality data represent concentrations of five different gases and fine, and coarse particle concentrations in the atmosphere. The road condition data provide information on surface temperature, water layer thickness, and more.

The data were collected at the AIM Research Intersection, an urban intersection located at northeastern corner of the inner city ring road near the Technical University of Braunschweig. Buses from the local public transport system regularly pass through the intersection, as well as large industrial vehicles. Therefore, interactions between heavy vehicles and VRUs are included in the dataset.

B. Trajectory Data

1) Object Classes: The dataset contains a total of 31,592 trajectories, with the majority being trajectories of cars (25,894). The next most common categories are bicycles (3,153), motorbikes (1,105), trucks (838), pedestrians (765), and vans (541).

2) Length of Trajectories: The total length of the trajectories is approximately 3,401 km, with an average trajectory length of 105 meters. Among motorized vehicles, cars (118 m), vans (117 m), trucks (99 m), and motorbikes (77 m) have longer trajectories compared to VRU trajectories, with bicycles (31 m) and pedestrians (15 m) having shorter average lengths. The detailed distribution of trajectory length per object class is shown in Fig. 6.

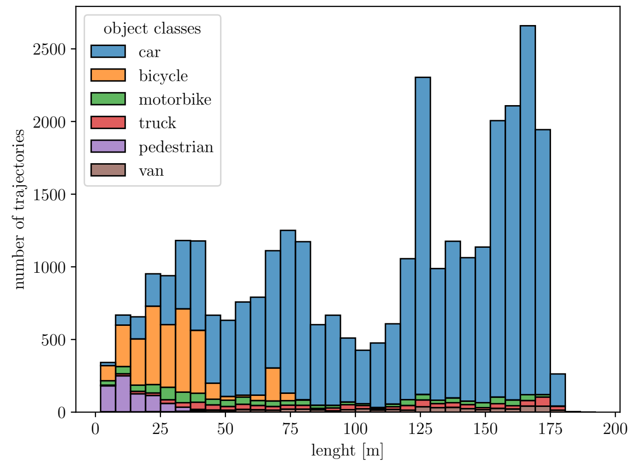

Fig. 6: Distribution of trajectory length per object class.

3) Interactions: As is typical for signalized intersections, vehicles going straight do not interact with other road users. However, right-turning vehicles frequently come into contact with VRUs. Therefore, additional cameras were installed specifically for the right turn from east to north. The fields of view (FOV) of these cameras are highlighted in yellow in Fig. 2. Detailed information about this interaction is provided in the studies [36] and [37].

The left-turning vehicles from the north, east, and south do a protected left turn as they have dedicated left turn arrows, which generally prevent interactions with other road users. However, the left-turn from west to north is an unprotected left turn, and thus, vehicles must yield to oncoming traffic before proceeding. This scenario results in the most interactions in motorized traffic, which is why it has been studied in [33].

U-turns are allowed for all left-turning vehicles except those from the west. Despite this restriction, U-turns are frequently performed, and as such, they were investigated in [29]. VRUs interact throughout the periphery of the entire intersection. During special events, such as traffic jams, demonstrations, or police blockades, unusual interactions take place. Their effects on traffic safety and efficiency were explored in [30].

4) Traffic Volume: This section presents the traffic volume on the road, while traffic volume on the sidewalks and bike paths is not detailed. The data was extracted from the digital map of the intersection [38]. In the dataset, traffic volumes are provided for individual lanes, but here they are summarized by route to give a general overview. A route is defined by an entry from one of the four directions and an exit in one of the four directions at the intersection, resulting in 16 possible routes. The routes of the trajectories are estimated using the method presented in [39]. In the dataset, trajectories for all possible routes at the intersection are available. The most common trajectories are found on the routes from east to west and west to east. U-turns, with 106 trajectories, are the least frequent. The exact number

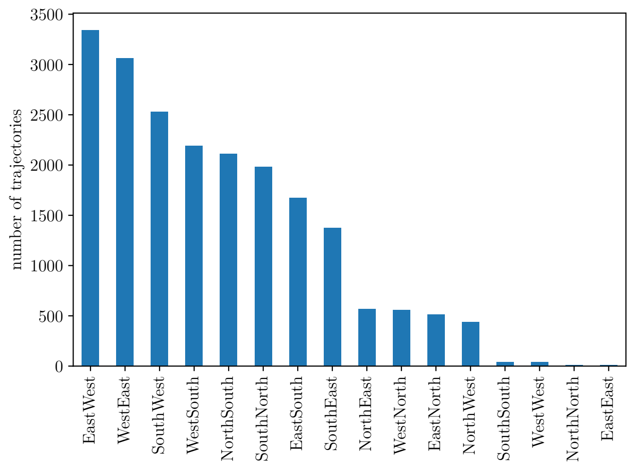

Fig. 7: Total number of objects per route.

of trajectories per route can be found in Fig. 7, while Fig. 8 shows the traffic volume per lane and hour. The highest traffic volume per hour is recorded on the route from east to west at 12:00 UTC+0, which corresponds to 14:00 local time.

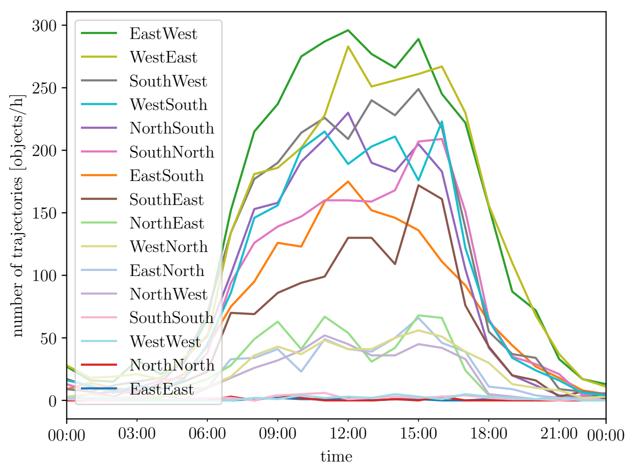

Fig. 8: Number of objects per hour and route.

C. Traffic Light Data

The status of the 30 traffic signals is recorded every second. Therefore, the dataset contains 2,585,610 signal states. General analyses of the traffic signal states at the time when vehicles enter the intersection are presented in [33].

D. Weather Data

Weather data is recorded every 10 s, resulting in 8,640 records in the dataset. The dataset was recorded under very favorable weather conditions in late summer. There was no precipitation, snow, or hail, and visibility remained predominantly at the maximum sensor value of 20 km, only decreasing to aproximately 13 km for a duration of 10 min. The wind blew from various directions at speeds ranging from 0 to 4 m/s. The air temperature fluctuated between 12.9 and 20.2 °C, and the sun shone almost throughout the entire day as shown in Fig. 9. The figure also clearly shows that the air temperature remained high from 15:00 UTC to 22:00 UTC, even after solar radiation had significantly decreased.

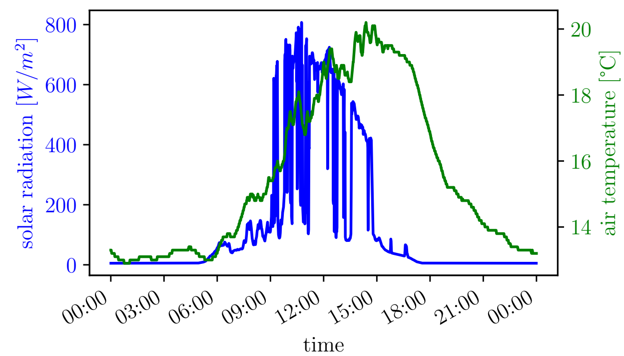

Fig. 9: Solar radiation and air temperature data from the dataset.

E. Road Condition Data

The data is recorded every 30 s, resulting in 2,880 records in the dataset. Due to the favorable weather conditions, the road condition remained consistently good. Since there was no precipitation and the road surface temperature did not reach critical levels of negative temperature as depicted in Fig. 10, no influence on driving behavior is assumed.

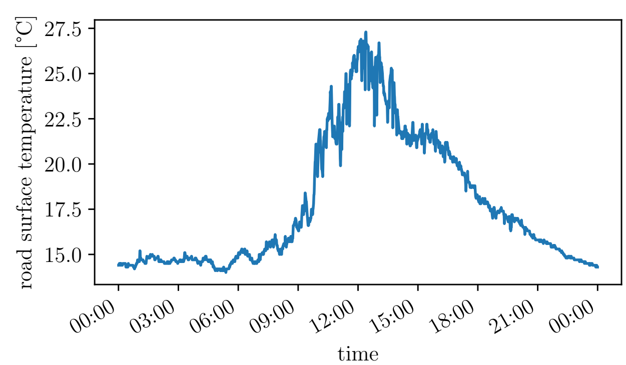

Fig. 10: Road surface temperature from the dataset.

F. Air Quality Data

Weather data is recorded every 60 s, resulting in 1,440 records in the dataset. The data trend throughout the day is shown in Fig. 11. A more detailed analysis of the weather station data in relation to traffic volume can be found in [32].

V. Discussion

In this discussion, we will highlight the key contributions of our dataset, address its limitations, and outline potential directions for future work.

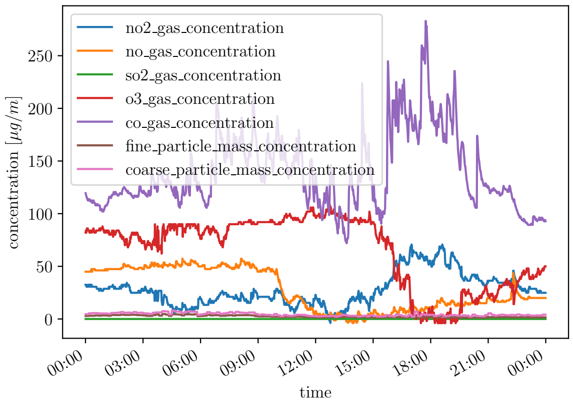

Fig. 11: Air quality data from the dataset.

A. Dataset Contributions

The analysis in Table I shows that our dataset stands out from existing publicly available datasets by providing richer contextual information essential for accurately modeling traffic situations. Many datasets focus solely on trajectory data, often recorded under optimal weather conditions, lacking the comprehensive environmental context necessary for developing robust models that can operate in challenging traffic scenarios [7].

The presented DLR-UT dataset addresses these limitations by integrating trajectory data from various traffic participants with critical environmental factors. This multifaceted approach enhances understanding of real-world traffic dynamics and supports a wide range of research questions related to automated driving and urban mobility.

For example, the local weather data in the DLR-UT dataset offer advantages over publicly available sources like the *German Weather Service*. Parameters such as ice layer thickness and road grip are vital for traffic research because they significantly influence the safety of traffic situations. Moreover, air quality data is essential for studying traffic emissions in urban areas as demonstrated in [32].

B. Dataset Limitations

Our dataset has some limitations that impact the analysis of traffic behavior. Although pose estimation data for VRUs is feasible and already available in existing datasets, our dataset does not provide pose estimation data. This absence constrains the understanding of interactions, such as the hand gestures of a turning bicyclist. Additionally, objects are categorized into just six classes, whereas a more nuanced classification could be achieved using neural networks, as supported by many other datasets [7]. Image data are also absent from our dataset. While other datasets include traffic images [7], we lack them due to limited upload capacity on the Zenodo platform. However, these images can be provided upon request.

C. Future Work

To address these limitations, we recommend expanding the AIM Research Intersection and releasing additional datasets to provide a broader data foundation for researchers, facilitating studies on traffic safety and efficiency. We encourage users to share analyzed metadata to create a more comprehensive picture of traffic dynamics. Our future plans include publishing datasets from different time periods, including winter or autumn, to capture varied environmental conditions.

The dataset's versioning allows for continuous expansion, including potential raw data like audio signals. In addition, higher-level metadata, for instance, red light violations can be identified from raw data and added to the dataset. This facilitates efficient analyses of red-light violations, without users needing to detect these events on their own.

Despite including more environmental data than other datasets, as shown in Table I, the data foundation in all publicly available datasets is still insufficient for a complete understanding of traffic conditions. Additional data types are required, which could potentially enhance the 6 Layer Model. Future work could focus on collecting data from individuals inside vehicles, such as age, or emotional state facilitated by the growing use of wearables. This could improve analyses of human experiences and predictions of driver behavior.

VI. Conclusion

This study presents the DLR Urban Traffic dataset, which contains traffic data collected at the AIM Research Intersection in Braunschweig, Germany, from Sunday, September 24, 2023. To provide a comprehensive representation of traffic conditions, the dataset includes not only trajectory data, but also traffic light, weather, road condition, and air quality data. Additionally, the dataset features processed information such as traffic volume and data in OpenSCENARIO format, enabling direct simulation of the data. The provided metadata simplifies the process of working with the raw data and offers a detailed description of traffic conditions. Furthermore, the study outlines potential extensions to the dataset and the 6 Layer Model, which can be used to describe traffic in greater detail. Overall, it can be concluded that the dataset distinguishes itself from already published datasets by including a wealth of environmental data. This provides a broad data foundation for researchers to study traffic safety and efficiency. Currently, the dataset contains data from a single day, but plans are already in place for future releases of one week of data from a winter period.

References

[1] H. Naci, D. Chisholm, and T. D. Baker, "Distribution of road traffic deaths by road user group: a global comparison," *Injury Prevention*, vol. 15, no. 1, pp. 55–59, 2009. [Online]. Available: https://injuryprevention.bmj.com/content/15/1/55.short

[2] J. Lu, B. Li, H. Li, and A. Al-Barakani, "Expansion of city scale, traffic modes, traffic congestion, and air pollution," *Cities*, vol. 108, p. 102974, 2021.

[3] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 383–398, 2018.

[4] Y. Zheng, "Trajectory data mining," *ACM Transactions on Intelligent Systems and Technology*, vol. 6, no. 3, pp. 1–41, 2015.

[5] E. N. Barmpounakis, E. I. Vlahogianni, and J. C. Golias, "Unmanned aerial aircraft systems for transportation engineering: Current practice and future challenges," *International Journal of Transportation Science and Technology*, vol. 5, no. 3, pp. 111–122, 2016.

[6] M. Scholtes, L. Westhofen, L. R. Turner, K. Lotto, M. Schuldes, H. Weber, N. Wagener, C. Neurohr, M. H. Bollmann, F. Kortke, J. Hiller, M. Hoss, J. Bock, and L. Eckstein, "6-layer model for a structured description and categorization of urban traffic and environment," *IEEE Access*, vol. 9, pp. 59 131–59 147, 2021.

[7] M. Liu, E. Yurtsever, J. Fossaert, X. Zhou, W. Zimmer, Y. Cui, B. L. Zagar, and A. C. Knoll, "A survey on autonomous driving datasets: Statistics, annotation quality, and a future outlook," *IEEE Transactions on Intelligent Vehicles*, pp. 1–29, 2024.

[8] S. Bauer, E. Bretschneider, J. Rataj, and T. Schumann, "Aim – application platform intelligent mobility," 2009. [Online]. Available: https://elib.dlr.de/60757/

[9] S. Knake-Langhorst, K. Gimm, T. Frankiewicz, and F. Köster, "Test site aim – toolbox and enabler for applied research and development in traffic and mobility," *Transportation Research Procedia*, vol. 14, pp. 2197–2206, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352146516302381

[10] German Aerospace Center, "Test bed lower saxony: Innovation laboratory for the mobility of the future," 2024. [Online]. Available: https://www.dlr.de/en/ts/research-transfer/research-services/test-bed-lower-saxony

[11] S. Knake-Langhorst and K. Gimm, "Aim research intersection: Instrument for traffic detection and behavior assessment for a complex urban intersection," *Journal of large-scale research facilities JLSRF*, vol. 2, p. A65, 2016. [Online]. Available: https://jlsrf.org/index.php/lsf/article/view/122

[12] C. Schicktanz, L. Klitzke, K. Gimm, S. Knake-Langhorst, G. Rizzo, H. H. Mosebach, and K. Liesner, "Dlr urban traffic dataset (dlr ut)," Jan. 2025. [Online]. Available: https://doi.org/10.5281/zenodo.14773161

[13] 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, 2020.

[14] U.S. Department of Transportation, Federal Highway Administration, "Next generation simulation (ngsim)," 2006. [Online]. Available: http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm

[15] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, "The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018, pp. 2118–2125.

[16] J. Bock, R. Krajewski, T. Moers, S. Runde, L. Vater, and L. Eckstein, "The ind dataset: A drone dataset of naturalistic road user trajectories at german intersections," in *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 1929–1934.

[17] University of Waterloo, "Waterloo multi-agent traffic dataset: intersection," 2024. [Online]. Available: https://uwaterloo.ca/waterloo-intelligent-systems-engineering-lab/datasets/waterloo-multi-agent-traffic-dataset-intersection

[18] O. Zheng, M. Abdel-Aty, L. Yue, A. Abdelraouf, Z. Wang, and N. Mahmoud, "Citysim: A drone-based vehicle trajectory dataset for safety-oriented research and digital twins." [Online]. Available: http://arxiv.org/pdf/2208.11036

[19] F. Yu, H. Yan, R. Chen, G. Zhang, Y. Liu, M. Chen, and Y. Li, "City-scale vehicle trajectory data from traffic camera videos," *Scientific Data*, vol. 10, no. 1, p. 711, 2023. [Online]. Available: https://www.nature.com/articles/s41597-023-02589-y

[20] W. Zhan, L. Sun, Di Wang, H. Shi, A. Clausse, M. Naumann, J. Kummerle, H. Konigshof, C. Stiller, A. de La Fortelle, and M. Tomizuka, "Interaction dataset: An international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps." [Online]. Available: http://arxiv.org/pdf/1910.03088

[21] A. Kutsch, M. Margreiter, and K. Bogenberger, *TUMDOT–MUC: Data Collection and Processing of Multimodal Trajectories Collected by Aerial Drones*, 2024.

[22] E. Barmpounakis and N. Geroliminis, "On the new era of urban traffic monitoring with massive drone data: The pneuma large-scale field experiment," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 50–71, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X19310320

[23] W. Zimmer, C. Creß, H. T. Nguyen, and A. C. Knoll, "Tumtraf intersection dataset: All you need for urban 3d camera-lidar roadside perception," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2023, pp. 1030–1037.

[24] M. Goldhammer, S. Kohler, S. Zernetsch, K. Doll, B. Sick, and K. Dietmayer, "Intentions of vulnerable road users—detection and forecasting by means of machine learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 3035–3045, 2020.

[25] E. Strigel, D. Meissner, F. Seeliger, B. Wilking, and K. Dietmayer, "The ko-per intersection laserscanner and video dataset," in *ITSC : 17th International IEEE Conference on Intelligent Transportation Systems : 8-11 October 2014*, . 17th IEEE Conference on Intelligent Transportation Systems Qingdao, China, Ed. IEEE, 2014, pp. 1900–1901.

[26] G. Stanzl and J. Sichermann, "Savenow - verkehrsdaten von ingolstadt," 2024. [Online]. Available: https://savenow.github.io/sta-docs/index.html

[27] J. Guo, U. Kurup, and M. Shah, "Is it safe to drive? an overview of factors, challenges, and datasets for driveability assessment in autonomous driving." [Online]. Available: http://arxiv.org/pdf/1811.11277

[28] Y. Kang, H. Yin, and C. Berger, "Test your self-driving algorithm: An overview of publicly available driving datasets and virtual testing environments," *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 2, pp. 171–185, 2019.

[29] C. Schicktanz and K. Gimm, "Detection and analysis of critical interactions in illegal u-turns at an urban signalized intersection," *Data Science for Transportation*, vol. 7, no. 1, p. 3, 2025.

[30] C. Schicktanz, L. Klitzke, and K. Gimm, "Microscopic analysis of the impact of congestion on traffic safety and efficiency at a signalized intersection: A case study," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2023, pp. 2827–2834.

[31] K. Rosch, F. Heidecker, J. Truetsch, K. Kowol, C. Schicktanz, M. Bieshaare, B. Sick, and C. Stiller, "Space, time, and interaction: A taxonomy of corner cases in trajectory datasets for automated driving," in *2022 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2022, pp. 86–93.

[32] T. Schripp, T. Grein, C. Schicktanz, K. Gimm, and S. Weber, "Monitoring of ultrafine particles and road traffic at aim research intersection and its vicinity," 2021. [Online]. Available: https://elib.dlr.de/148097/

[33] C. Schicktanz and K. Gimm, "Detection and analysis of corner case scenarios at a signalized urban intersection," *Accident Analysis & Prevention*, vol. 210, p. 107838, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S000145752400383X

[34] L. Klitzke and C. Schicktanz, "Tasi: A python library for traffic data analysis and situation interpretation," 2024.

[35] C. Schicktanz, L. Klitzke, K. Gimm, R. Lüdtke, H. H. Mosebach, K. Liesner, F. Heuer, A. Wodtke, and L. Asbach, "Dlr highway traffic dataset (dlr ht)," 2025. [Online]. Available: https://doi.org/10.5281/zenodo.14811064

[36] M. Zhang, M. Dotzauer, and C. Schießl, "Analysis of implicit communication of motorists and cyclists in intersection using video and trajectory data," *Frontiers in psychology*, vol. 13, p. 864488, 2022.

[37] H. Saul, M. Junghans, M. Dotzauer, and K. Gimm, "Online risk estimation of critical and non-critical interactions between right-turning motorists and crossing cyclists by a decision tree," *Accident Analysis & Prevention*, vol. 163, p. 106449, 2021.

[38] M. Scholz, "Opendrive dataset of the inner ring road in brunswick."

[39] L. Klitzke, C. Schicktanz, K. Gimm, C. Koch, and F. Köster, "Scenario mining in the urban domain: Exploiting the topology of a road network for maneuver annotation and scenario extraction," *Authorea Preprints*, 2023.