openDD: A Large-Scale Roundabout Drone Dataset

Antonia Breuer¹, Jan-Aike Termöhlen², Silviu Homoceanu¹, Tim Fingscheidt²

Abstract—Analyzing and predicting the traffic scene around the ego vehicle has been one of the key challenges in autonomous driving. Datasets including the trajectories of all road users present in a scene, as well as the underlying road topology are invaluable to analyze the behavior of the various traffic participants. The interaction between the traffic participants is especially high in intersection types that are not regulated by traffic lights, the most common one being the roundabout. We introduce the openDD dataset, including 84,774 accurately-tracked trajectories and HD map data of seven different roundabouts. The openDD dataset is annotated using images taken by a drone in 501 separate flights, totalling in over 62 hours of trajectory data. As of today the openDD is by far the largest publicly available trajectory dataset recorded from a drone perspective, while comparable datasets span 17 hours at most. The data is available, for both commercial and non-commercial use, at: http://www.l3pilot.eu/openDD.

I. INTRODUCTION

In recent years autonomous driving has become one of the major applications for numerous fields of research. A main challenge faced in autonomous driving is the prediction of the traffic scene surrounding the ego vehicle. Predicting the surrounding traffic scene is particularly difficult in urban and rural scenarios due to the high inter-dependency between the involved road users. One way to face this challenge is to use increasing amounts of data, causing a surge of popularity of data-driven approaches that rely on large-scale trajectory datasets in recent years [1], [2], [3], [4]. Other applications of trajectory datasets include the modeling and analysis of driving behavior [5] and the analysis of safety of the autonomous driving function [6].

Datasets recorded from a ground view instead of a bird’s eye perspective are limited by occlusions and the restricted field of view of the recording device at the ground. Labeling the trajectories with the help of image data captured by an aerial drone ensures a complete overview of the traffic situation and enables algorithms that use the dataset to take all present road users into account. The interaction between different road users is particularly high in intersections that are not regulated by traffic lights, the most common being the roundabout. In the openDD dataset presented in this work, seven roundabouts with different topologies are covered, one of them shown in Fig. 1. The dataset includes trajectories of all recorded road users, shapefiles and an extensible markup language (XML) file, describing the road topology of the underlying intersections. One reference image taken by the drone is provided per intersection. An exemplary visualization of the data included in the dataset can be seen in Fig. 1. As shown there, all dynamic traffic participants are accurately tracked in the relevant area surrounding the roundabout, and also vehicles hard to see for the human eye, such as the grey car on top of the picture are accurately detected. The introduced openDD dataset spans more than 62 hours in total, covers 84,774 trajectories, and can be accessed on the following website: http://www.l3pilot.eu/openDD.

II. RELATED WORK

Table 1 gives an overview of trajectory datasets recorded from a drone perspective and their characteristics.

The Stanford drone dataset [7] was the first publicly available trajectory dataset recorded from a drone’s perspective and is tailored to the analysis of pedestrian trajectories. It consists of 9 hours of data over eight unique locations on the campus and has a high percentage rate of labeled pedestrians and cyclists, while only about 7% of the labeled targets are cars. The DUT and CITR datasets [9] are especially designed for the analysis of the behavior of pedestrians when interacting with vehicles and span less than half an hour in total. One of the first large-scale trajectory datasets based on the footage of an aerial drone is the highD dataset [8]. It includes trajectory data from 110,000 cars on German highways and spans 5,600 lane changes over 16.5 hours of data. During the creation of the presented openDD dataset, descriptions...
of the \textit{INTERACTION} [11] and the \textit{inD} [10] datasets have been published. The \textit{INTERACTION} dataset spans about 16.5 hours and covers data from 11 intersections, including 5 roundabouts, 3 unsignalized intersections, 2 merging and lane change situations, and 1 signalized intersection. The \textit{InD} dataset [10] distinguishes pedestrians, bicycles, cars, trucks, and buses and includes 10 hours of data recorded by a drone. At the time of writing of this publication, the \textit{InD} dataset has not been released yet, thus no further description than the one stated in the publication can be given.

III. DATASET

This work introduces the openDD dataset, a trajectory dataset recorded from a drone perspective. The dataset includes $R = 501$ recordings, each representing one coherent drone flight, capturing one of the $I = 7$ roundabouts covered in the dataset, depicted in Fig. 2. Each recording indexed by $r \in R = \{1, \ldots, R\}$ spans 5 to 15 minutes in total and was taken from the drone perspective with a camera capturing 30 fps. The used drone is a DJI Phantom 4, a high-end consumer drone, recording at a resolution of $3840 \times 2160$ pixels, being slightly below 4K. The video footage taken by the drone is stabilized and rectified before it is used to detect and track all traffic participants in the given scene.

For each recording $r$ we define $N_r$ to be the number of time instants included in the recording, equal to the number of frames captured in the recorded video.

The openDD dataset defines which objects, each with unique object index $j$ are present at time instant $n \in \{1, \ldots, N_r\}$. The state vector $s_n^{(j)}$ of an object $j$ at a time instant $n$ is defined by (1) $s_n^{(j)} = [x_n^{(j)}, y_n^{(j)}, v_n^{(j)}, a_n^{(j)}, w_n^{(j)}, l_n^{(j)}, \alpha_n^{(j)}, a_{\alpha,n}^{(j)}, a_{y,n}^{(j)}, c_n^{(j)}, \ldots, C, V, T, B, P, R, M, Y]$. The vector $[x_n^{(j)}, y_n^{(j)}]^T$ describes the Universal Transverse Mercator (UTM) coordinates of the object’s bounding box center. The orientation of the bounding box is given by its yaw $\alpha_n$ in radians relative to the x-axis of the UTM reference coordinate system, whereas the dimension of the bounding box is given by the width $w_n^{(j)}$ and length $l_n^{(j)}$. The dynamic state of the object is described by the velocity $v_n^{(j)}$ defined in m/s, whereas $a_{\alpha,n}^{(j)}, a_{y,n}^{(j)}$, and $c_n^{(j)}$ describe the lateral, tangential, and total acceleration of the object in m/s$^2$. Additionally, the class $c_n \in C = \{C, V, T, B, P, R, M, Y\}$ of the object is defined, with each object either being a passenger car, van, truck, bus, pedestrian, trailer, motorcycle, or bicyclist. A visualization of the included bounding box information and the color-encoded class labels for a given scene can be seen in Fig. 1.

We provide the underlying HD map for each roundabout rd, with $i \in \{1, \ldots, I\}$ included in the dataset. The map data is provided as shapefiles and an XML file and distinguishes three logical elements of the road topology:
lane centerlines, lane boundaries, and drivable areas. For each such logical element, three shapefiles are provided, a .shp, a .shx, and a .dbf, resulting in a total of nine shapefiles included in the dataset per roundabout. The .shp file defines the underlying geometry of the logical element, such as the points of the lanes. The .dbf attribute files are in dBase format and define element-specific attributes. For the lane centerlines, the .dbf file contains, among other attributes, a unique identifier for each lane centerline and a list of succeeding lane centerlines, preceding lane centerlines, and parallel lane centerlines. The .dbf file for the lane boundaries defines the corresponding lane identifier, as well as the material of the lane boundary, such as CONCRETE for a curbside and NONE for an implicit lane boundary. The .shx file is an index file of the shape geometry .shp file, providing a way to quickly iterate over the defined geometry. An exemplary visualisation of the provided shapefiles is given in Fig. 3. In addition to the shapefiles we also provide an .xml file for each intersection i, representing the information contained in the shapefiles in a non-binary format that can be easily parsed by most programming languages.

Beyond the HD map and object states, the dataset includes a geo-referenced and anonymized example picture taken from the drone perspective of each intersection i. This picture both provides a way of visualizing the trajectory data, as well as a supplementary input to the provided HD map.

A. Dataset Statistics

The dataset was recorded at different times of day, for each roundabout including at least 2 h at the rush hour times in the morning and afternoon, as well as regular intervals in between rush hours.

The dataset spans 62.7 h, of which 18.8 h cover the first roundabout rdb1 and the remaining 43.9 h are distributed among the remaining six roundabouts rdb2 to rdb7, with varying lengths around 7 h for each. An example picture of each roundabout is depicted in Fig. 3. In total 84,774 trajectories are included in the dataset, covering 8,501.14 km.

A detailed overview of the openDD dataset, distinguishing between the seven different roundabouts included in the dataset, is provided in Table II. Here, the number of trajectories, the average trajectory duration, length, velocity, and total acceleration is stated for each object class, as well as for each roundabout. The average trajectory duration over all classes and all data subsets is 17.64 s, with an average trajectory length of 100.28 m.

The average velocity is 6.63 m/s and the average total acceleration is 1.42 m/s².

The relatively high amount of vehicles, 81,372 across the whole dataset, compared to the 3,402 pedestrians and bicyclists, is caused by the high percentage of covered rush hour times, as well as the remote locations of some of the covered roundabouts. The roundabout rdb1 has an especially high traffic load with 13,644 unique vehicles passing the roundabout in 6.9 h.

Roundabout rdb2 has a very high average trajectory duration of pedestrians, with 65.42 s compared to the average pedestrian trajectory duration of 87.64 s. The high pedestrian trajectory duration of rdb2 is caused by several pedestrians idling in the recordings of rdb2.

IV. USING THE DATASET

Publications of trajectory prediction models that use this dataset should be evaluated in a uniform fashion. To this end we define metrics to evaluate predicted trajectories, different splits of the openDD dataset, and propose several challenges using this dataset in the following.

A. Distance Metrics

Similar to our previous work [3], we define several distance metrics $D(T^{(j)}, \hat{T}^{(j)})$ that can be used to evaluate the accuracy of a trajectory prediction algorithm. For a given object with index j, this distance metric D compares the predicted trajectory $T^{(j)}$ with the actual ground truth trajectory $\hat{T}^{(j)}$ of the object, as given in the dataset. In the example scripts that are made available with the dataset, implementations of the used metrics are provided.

Euclidean displacement at time $t_n$: The Euclidean point-to-point distance between the n-th trajectory point of $T^{(j)}$ and the m-th trajectory point of $\hat{T}^{(j)}$ is defined as

$$D_{E} (T^{(j)}(n), \hat{T}^{(j)}(m)) = \sqrt{(x_n - x_m)^2 + (y_n - y_m)^2}.$$  

Mean-squared Euclidean distance: Given two trajectories $T^{(j)}, \hat{T}^{(j)}$, spanning the same sequence of time instants $n \in N = \{0, 1, ..., N-1\}$, the mean squared Euclidean distances between the two entire trajectories is defined as the normalized sum of the squared Euclidean distances between the points corresponding to the same time instant n:

$$D_{MSE} (T^{(j)}, \hat{T}^{(j)}) = \frac{1}{N} \sum_{n \in N} D_{E}^2 (T^{(j)}(n), \hat{T}^{(j)}(n)).$$
Modified Hausdorff (MH) distance: The following definition of the modified Hausdorff (MH) distance is adopted from a work on object matching by Dubuisson et al. [13].

The definition of the point-to-set distance between the n-th point of a trajectory \( T^{(j)}(n) \), and another entire trajectory \( \overline{T}^{(j)} \), is:

\[
D_{PS}(T^{(j)}(n), \overline{T}^{(j)}) = \min_{m \in N} \left( D_{EL}(T^{(j)}(n), T^{(j)}(m)) \right).
\]

The directed modified Hausdorff (DMH) distance is defined by Dubuisson et al. [13] as

\[
D_{DMH}(T^{(j)}, \overline{T}^{(j)}) = \frac{1}{N} \sum_{n \in N} D_{PS}(T^{(j)}(n), \overline{T}^{(j)}).
\]
TABLE III: Description of the three training data splits $R_1, R_2, R_3$, as well as of the test data splits $R_A, R_B, R_C$, as defined for the challenges described in Section IV.

| Subset | $R_{123}$ | $R_1$ | $R_2$ | $R_3$ | $R_{ABC}$ | $R_A$ | $R_B$ | $R_C$ |
|--------|-----------|-------|-------|-------|-----------|-------|-------|-------|
| Recorded time | 47.8 h | 38.5 h | 8.5 h | 42.1 h | 34.0 h | 15.4 h | 8.5 h | 19.0 h | 9.4 h |
| # recordings from rdb | 130 | 107 | 23 | 0 | 0 | 107 | 23 | 23 | 0 | 23 |
| rdb | 48 | 40 | 8 | 40 | 8 | 40 | 8 | 8 | 8 | 8 |
| rdb3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 54 | 0 | 54 |
| rdb4 | 59 | 49 | 10 | 49 | 10 | 49 | 10 | 10 | 0 | 10 |
| rdb5 | 51 | 42 | 9 | 42 | 9 | 42 | 9 | 9 | 0 | 9 |
| rdb6 | 44 | 36 | 8 | 36 | 8 | 0 | 0 | 8 | 8 | 0 |
| rdb7 | 48 | 39 | 9 | 39 | 9 | 39 | 9 | 9 | 0 | 9 |

distances:

$$D_{MH}(\tau^{(i)}, \hat{\tau}^{(i)}) = D_{MH}(\hat{\tau}^{(i)}, \tau^{(i)})$$

$$= \max \left( D_{DMH}(\tau^{(i)}, \hat{\tau}^{(i)}), D_{DMH}(\hat{\tau}^{(i)}, \tau^{(i)}) \right).$$

(6)

The $MH$ distance captures the spatial similarity of the trajectories without considering temporal misalignments. For example, two trajectories encoding the same path traveled at different velocities would have a low $MH$ distance, while their Euclidean displacement $DMH$ and mean squared Euclidean distance $D_{MSE}$ would be high.

B. Dataset Splits

We divide the drone recordings of the seven roundabouts rdb, into subsets for training, validation, and testing, as specified in Table III. The exact assignment of recordings to the different subsets introduced in the following, defining which recording belongs to which subset, is provided in the dataset.

For testing purposes, we use all recordings $r$ from rdb1, as well as 15% of the recordings from the other roundabouts making up the total test set $R_{ABC} \subset R$. We define three different subsets of the total test set $R_{ABC}$ to evaluate algorithms on: $R_A$, $R_B$, $R_C \subset R_{ABC}$.

The subset $R_A$ includes all recordings in $R_{ABC}$, but the ones of rdb3, thus it includes 15% from all roundabouts rdb, with $i \in \{1, 2, 4, 5, 6, 7\}$. The second subset $R_B$ covers all recordings of rdb3. Lastly, $R_C$ includes 15% from the recordings of rdb3, as well as all recordings from the other roundabouts included in $R_{ABC}$ such that in total 15% of each roundabout are covered by this test set.

For training purposes, from the entire data $R = R_{ABC} \cup R_{123}$ we split the recordings $R_{123}$ not included in $R_{ABC}$ into three different splits $R_k = R_{k\text{train}} \cup R_{k\text{val}}, k \in \{1, 2, 3\}$. To define the splits $R_k$, we split the data included in $R_{123}$ for the roundabouts rdb$i, i \in \{1, 2, 4, 5, 6, 7\}$, such that around 18% of the recordings of each roundabout, except rdb1, form the validation set and the remaining 82% the training set. For the first split $R_1$, we include the training and validation subsets formed in this way for all the roundabouts included in $R_{123}$. The second split $R_2$, and third split $R_3$ are equal to $R_1$, but leave out all recordings from rdb1 and rdb6, respectively.

An important feature of this proposed division of the dataset is that the splits $R_k$ can be combined with any of the test sets $R_A, R_B, R_C$ to analyze different aspects of the learning process. The test set $R_A$ only includes recordings for topologies that are also included in $R_{123}$, allowing for an evaluation of the model’s ability to predict trajectories for previously seen roundabout topologies. The test set $R_B$ includes only recordings from an unseen roundabout, thus it is suitable to evaluate the generalization capability of the learned model. Especially a combination with the split $R_3$ is of interest, since $R_3$ does not include the data of rdb6, which is the only other roundabout with a similar topology including lanes to skip the center roundabout lane. Lastly, the test set $R_C$ is a mixture between the other two test sets, enabling an evaluation of both the learning capacity of the model, as well as its generalization ability.

C. Challenges

We encourage all publications using the openDD dataset to report their results using the Euclidean displacement as defined in [2] and the $MH$ as defined in (6) at the maximal prediction horizon they are reporting, and after 1 s, 3 s, and 6 s, if applicable.

Along the lines of our previous work [3], we propose to investigate the utility of the various information of the environment provided in the dataset. Thus, we invite parties interested in using the dataset to adopt their algorithms for trajectory prediction in such a way that they can work with variable input data.

Out of the possible nine combinations between the three test sets $R_A, R_B$, and $R_C$ and the training and validation splits $R_K$, we encourage researchers using the dataset to report the results of the following four combinations, with the first two being the most relevant.

**Train on $R_1$, test on $R_A$:** All roundabout topologies tested upon have been seen during the training. Thus the results on this combination will measure the general capability of the model to solve the task at hand.

**Train on $R_1$, test on $R_B$:** The test set only includes data covering a previously unseen roundabout, this combination can be used to assess the generalization capability of the evaluated model.
Train on $\mathcal{R}_3$, test on $\mathcal{R}_B$: Only recordings from a previously unseen roundabout, rdb$_3$ are included in the test set. Additionally, no recordings from the roundabout with the road topology most similar to the one of the test roundabout, rdb$_B$ is included in $\mathcal{R}_3$. This combination represents an even more difficult generalization evaluation.

Train on $\mathcal{R}_2$, test on $\mathcal{R}_C$: No recordings from rdb$_1$, the roundabout which covers almost 20h, is included in $\mathcal{R}_2$. Recordings from both seen and unseen roundabouts is included in the test set. This allows for an evaluation of both the capability of the model as well as the generalization included in the test set. This makes it appealing to both research institutions, as well as to companies. In a future publication we plan to provide a baseline on the given dataset, addressing the challenges introduced in Section IV.

Beyond the reporting of the aforementioned measures and training/testing splits, we propose three further research challenges for trajectory prediction using the openDD dataset:

1) Evaluation of the benefit of knowledge of the movement of other objects up to time instant $t$.
2) Evaluation of the benefit of the provided map data, divided by centerlines, lane boundaries and drivable areas. Here, a separate evaluation with the same algorithm using all of the provided information, only centerlines, and only drivable areas is desirable.
3) Given the image of each intersection $i \in \{1, \ldots, I\}$ provided in the dataset is used by the trajectory prediction algorithm, we would like the authors to evaluate the benefit of the image. The image can be for example used as an input for a convolutional neural network (CNN) and be combined with a birds-eye view of the current traffic scene, implicitly encoding the structure of the surroundings [3].

If challenge 3) shows that the information provided by the image gives additional benefits for the trajectory prediction task, interesting follow-up research would deal with how to integrate this information into the HD map.

V. FUTURE WORK

The openDD dataset introduced in this work is the biggest published trajectory dataset recorded from a drone perspective as of today. In addition to its length of over 62 h, it covers varying roundabout topologies, which makes it also valuable to study the generalization of trajectory prediction algorithms trained on it. The license of the provided openDD dataset, covering commercial and non-commercial usage, makes it appealing to both research institutions, as well as to companies. In a future publication we plan to provide a baseline on the given dataset, addressing the challenges introduced in Section IV.

VI. ACKNOWLEDGMENT

The authors would like to thank all partners within the H2020 co-funded project L3Pilot — grant agreement number 723051 — for their cooperation in making this dataset available online. Responsibility for the information and views set out in this publication lies entirely with the authors.

REFERENCES

[1] D. Ridel, E. Rehder, M. Lauer, C. Stiller, and D. Wolf, “A Literature Review on the Prediction of Pedestrian Behavior in Urban Scenarios,” in Proc. of ITSC, Maui, HI, USA, Nov. 2018, pp. 3105–3112.
[2] H. Cui, V. Radosavljevic, F.-C. Chou, T.-H. Lin, T. Nguyen, T.-K. Huang, J. Schneider, and N. Djuric, “Multimodal Trajectory Predictions for Autonomous Driving Using Deep Convolutional Networks,” arXiv:1906.03732, Sep. 2018. [Online]. Available: http://arxiv.org/abs/1906.03732.
[3] A. Breuer, S. Elfelein, T. Joseph, J. Bolte, S. Homoeanu, and T. Fingscheidt, “Analysis of the Effect of Various Input Representations for LSTM-Based Trajectory Prediction,” in Proc. of ITSC, Auckland, New Zealand, Oct. 2019, pp. 2728–2735.
[4] H. Cui, T. Nguyen, F.-C. Chou, T.-H. Lin, J. Schneider, D. Bradley, and N. Djuric, “Deep Kinematic Models for Physically Realistic Prediction of Vehicle Trajectories,” arXiv:1908.00219, Aug. 2019. [Online]. Available: http://arxiv.org/abs/1908.00219.
[5] K. Driggs-Campbell, V. Govindarajan, and R. Bajcsy, “Integrating Intuitive Driver Models in Autonomous Planning for Interactive Maneuvers,” in Proc. of ITSC, vol. 18, Yokohama, Japan, Dec. 2017, pp. 3461–3472.
[6] Enable-S3 Consortium. [Online]. Available: https://www.enable-s3.eu/.
[7] A. Robicquet, A. Sadeghian, A. Alahi, and S. Savarese, “Learning Social Etiquette: Human Trajectory Understanding in Crowded Scenes,” in Proc. of ECCV, Amsterdam, The Netherlands, Oct. 2016, pp. 549–565.
[8] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, “The highDD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems,” arXiv:1810.05642v1, Oct. 2018. [Online]. Available: https://arxiv.org/abs/1810.05642v1.
[9] D. Yang, L. Li, K. Redmill, and U. Özgüner, “Top-View Trajectories: A Pedestrian Dataset of Vehicle-Crowd Interaction from Controlled Experiments and Crowded Campus,” in Proc. of IV, Paris, France, Jun. 2019, pp. 899–904.
[10] 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,” 2019. [Online]. Available: https://arxiv.org/abs/1911.07602.
[11] L. Wei, L. Sun, D. Wang, H. Shi, A. Clasusse, M. Naumann, J. Kummerle, H. Konigshof, C. Stiller, A. d. L. Fortelle, and M. Tomizuka, “INTERACTION Dataset: An INTERNational, Adversarial and Cooperative moTION Dataset in Interactive Driving Scenarios With Semantic Maps,” arXiv:1910.03088, Sep. 2019. [Online]. Available: https://arxiv.org/abs/1910.03088.
[12] F. Poggenhans, J. Pauls, J. Janosovits, S. Orf, M. Naumann, F. Kuhnt, and M. Mayr, “Lanelet2: A High-Definition Map Framework for the Future of Automated Driving,” in Proc. of ITSC, Maui, HI, USA, Nov. 2018, pp. 1672–1679.
[13] M. Dubuisson and A. K. Jain, “A Modified Hausdorff Distance for Object Matching,” in Proc. of ICPR, vol. 1, Jerusalem, Israel, Oct. 1994, pp. 566–568.