

8. Supplementary material

8.1. Drives

The 29 drives are collected in two cities, namely Gothenburg, Sweden and Paris, France. The average duration is 3 minutes, with the shortest being around 1 minute and the longest about 5 minutes. The combined duration is longer than 1 hour and 30 minutes. Further details can be seen in Figure 8 and Figure 9.

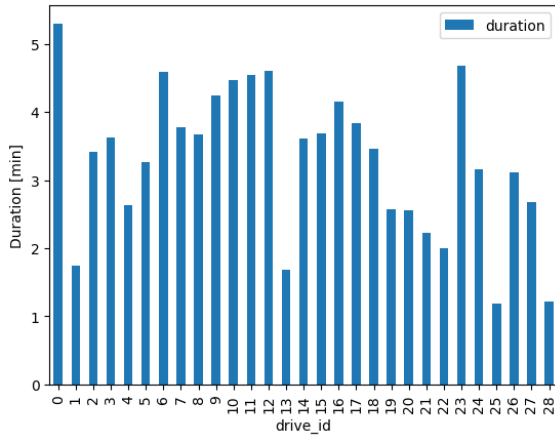


Figure 8: Duration of ZOD Drives.

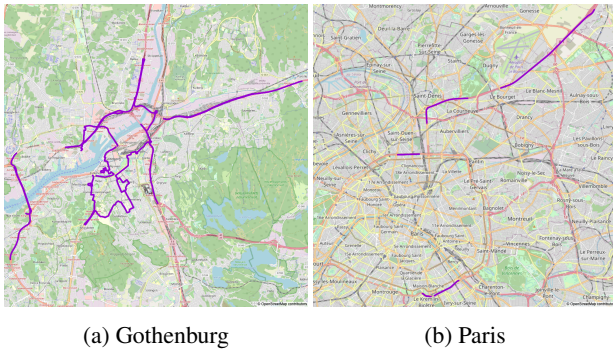


Figure 9: GNSS of ZOD Drives projected onto the map.

8.2. Annotation details

We provide complete multi-task annotations for *Sequences*, but *Frames* are only partially annotated for traffic sign and ego road annotations, as shown in Figure 10. All annotations are provided in the *GeoJSON* format [2], which encodes a variety of geometries in a *JSON* format with additional properties. Annotation files can be easily read and visualized by the provided development kit. Table 9 details high-level classes, sub-classes, and additional attributes assigned to each annotated object or polygon for different annotation tasks in ZOD.

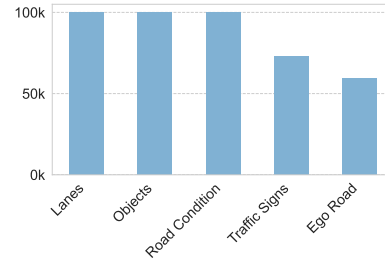


Figure 10: Number of annotated frames per project.

8.3. Additional dataset statistics

We illustrate the number of annotated dynamic and static objects per class in ZOD *Frames* in Figure 11. Note that some classes are much more represented than others. For example, vehicles are naturally orders of magnitude more prevalent than animals in traffic scenarios. The different colors represent the levels of detail that an object annotation can have, from paired 2D and 3D bounding boxes, to simply a 2D box, to a rough “unclear” region that indicates the possibility of one or more objects in the given image region. We also show the number of annotated 3D cuboids per ZOD *Frame* in Figure 12, for three top-level classes – namely *Vehicle*, *Vulnerable Vehicle*, and *Pedestrian*.

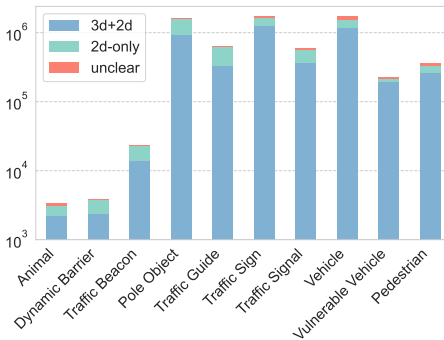


Figure 11: Number of annotated objects, broken down by class and annotation detail.

Similarly, we present the number of lane annotations per class in ZOD *Frames* in Figure 13. These counts correspond to single polygons, not associated instances across many polygons. We note that ZOD contains many dashed and solid lines, as expected, but also a significant number of road paintings and shaded areas. The latter may be particularly of interest, since it corresponds to areas on the road where shadows are cast in a way that very strongly resembles lane markings. These are one of the primary causes for false positives, and having explicit annotations makes it possible to both train the network to be resistant to these distractors and also to perform target evaluations.

Train	Eval	AP	AP_{50}	AP_{75}	AP_s	AP_m	AP_l	AP_{veh}	AP_{VV}	AP_{ped}
blur	blur	30.116	54.562	28.693	7.273	30.479	50.981	42.393	25.753	22.203
	DNAT	30.131	54.554	28.733	7.264	30.476	50.998	42.389	25.748	22.257
	original	30.117	54.558	28.714	7.266	30.476	50.975	42.398	25.738	22.217
DNAT	blur	30.277	54.988	28.870	7.210	30.645	51.126	42.452	26.028	22.350
	DNAT	30.305	54.927	28.912	7.200	30.662	51.255	42.458	26.050	22.406
	original	30.308	54.931	28.918	7.208	30.657	51.281	42.463	26.046	22.416
original	blur	30.317	54.746	28.897	7.288	30.649	51.142	42.479	26.186	22.285
	DNAT	30.352	54.860	28.933	7.295	30.676	51.323	42.499	26.186	22.372
	original	30.352	54.863	28.921	7.289	30.661	51.329	42.497	26.182	22.378

Table 6: Experimental results when training and evaluating on all possible combinations of anonymization methods.

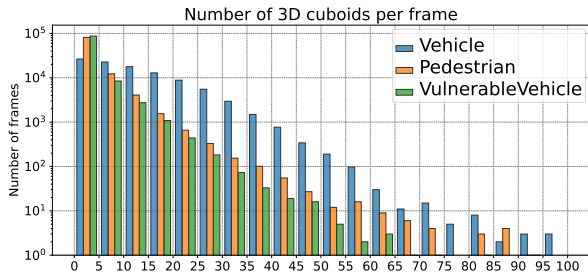


Figure 12: Annotated 3D cuboids per frame and class.

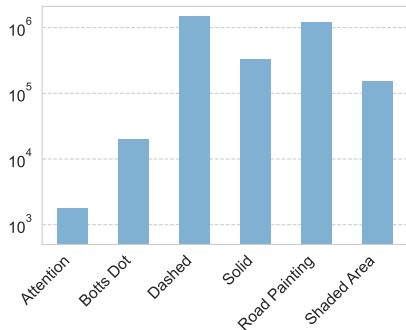


Figure 13: Number of annotated polygons for lanes, broken down by class, including road paintings and shaded areas.

8.4. Additional anonymization results

To provide a more comprehensive view of the anonymization experiments, we provide the complete experimental results in Table 6. Here, we show the performance of the Faster-RCNN [24] pipeline when trained and evaluated across all possible combinations of anonymization methods.

8.5. Compression results

ZOD contains high-resolution images of every scene. For each image, we do JPG compression of the original PNG image to reduce the memory footprint of ZOD, increasing the usability of the dataset. As the JPG compression is lossy, we analyze the effect this has on downstream

tasks. To evaluate this effect, we train the Faster-RCNN object detection pipeline on both JPG and PNG images, and see no significant performance degradation due to compression when evaluating both trained networks on the PNG images, see Table 7.

Metric	PNG	JPG
AP	30.23 ± 0.09	30.02 ± 0.03
AP_{50}	54.79 ± 0.06	54.34 ± 0.10
AP_{75}	28.72 ± 0.15	28.51 ± 0.02
AP_s	7.23 ± 0.04	7.14 ± 0.06
AP_m	30.49 ± 0.14	30.30 ± 0.04
AP_l	51.23 ± 0.07	50.95 ± 0.02
AP_{veh}	42.41 ± 0.07	42.26 ± 0.04
AP_{VV}	25.96 ± 0.15	25.71 ± 0.06
AP_{ped}	22.32 ± 0.04	22.08 ± 0.07

Table 7: Impact of image compression. We report AP (computed according to COCO evaluation protocol [18]) when training Faster-RCNN on compressed vs. uncompressed images. The metrics are computed using the original (uncompressed) images, and are presented as the mean and standard deviation across three separate runs.

8.6. Panoptic segmentation benchmark

We use Panoptic-DeepLab [8] as the panoptic segmentation baseline. Here, we set lanes as *things* while ego-road and cross-walks are set as *stuff*. Table 8 shows Panoptic-Quality (PQ), segmentation quality component (SQ), and recognition quality component (RQ). For context, SOTA PQ results for the BDD100K [36] and CityScapes [10] datasets are 23.90 and 70.1, respectively.

Table 8: Panoptic segmentation

Class	PQ	SQ	RQ
<i>Things</i>	40.0	77.6	51.5
<i>Stuff</i>	67.8	88.3	72.7
All	60.9	85.6	67.6

Annotation tasks	High-level classes	Sub-classes	Additional attributes
Ego road	Road	N/A	N/A
	Debris	N/A	N/A
Lane markings and road paintings	Solid	N/A	InstanceID, Coloured, MultipleLaneMarkings
	Dashed	N/A	
	Botts dot	N/A	
	Shaded area	Split, Merge, Undefined	N/A
	Road paintings	ContainsArrow, ContainsPictogram, ContainsMarker, ContainsTrafficSigns, ContainsCrossWalks, ContainsText, ContainsOther, Odd, Unclear	N/A
Dynamic and static objects	Vehicle	Car, Van, Truck, Trailer, Bus, HeavyEquip, TramTrain, Other, Inconclusive	Unclear, OcclusionRatio, Emergency, RelativePosition, IsPullingOrPushing
	Vulnerable vehicle	Bicycle, Motorcycle, Wheelchair, Stroller, PersonalTransporter, Other, Inconclusive	Unclear, OcclusionRatio, Emergency, IsPullingOrPushing, WithRider
	Pedestrian	N/A	Unclear, OcclusionRatio, Emergency, RelativePosition, IsPullingOrPushing
	Animal	N/A	N/A
	TrafficSign	See below for more granular categorization of Traffic signs	Unclear, OcclusionRatio, InOnDynamicObject, IsForConstruction, TrafficContentVisible, IsForEgoRoad, IsForEgoLane, IsForOtherTrafficParticipants, ComplementaryToLandmark, IsSticker
	TrafficSignal	N/A	Unclear, OcclusionRatio, InOnDynamicObject, IsForConstruction, TrafficContentVisible, IsForEgoRoad, IsForEgoLane, IsForOtherTrafficParticipants
	TrafficGuide	Reflector, Attention, SnowMarker, Bollard, Other, Inconclusive	Unclear, OcclusionRatio, InOnDynamicObject, IsForConstruction, TrafficContentVisible, IsSticker
	PoleObject	LampPole, LandmarkPole, LargeLandmarkPole, Other, Inconclusive	Unclear, OcclusionRatio, InOnDynamicObject, IsForConstruction, TrafficContentVisible
	TrafficBeacon	N/A	
DynamicBarrier	N/A		
Traffic signs	MandatorySigns	PassOnThisSideLeft, PassOnThisSideRight, PassOnEitherSide, ProceedStraightOrTurnLeft, ProceedStraight, ProceedStraightOrTurnRight, TurnLeftAhead, TurnRightAhead, TurnAhead, TurnRight, TurnLeft, Roundabout	Unclear, OcclusionRatio, ComplementaryToLandmark, ContainsInnerSigns, ContentContainsText, IsForConstruction, IsForEgoRoad, IsForEgoLane, IsElectronic, IsForOtherTrafficParticipants, IsOnDynamicObjects
	PrioritySigns	GiveWay, GiveWayOncoming, PriorityStop, PrioOverOncoming, PriorityRoadBegin, PriorityRoadEnd	
	ProhibitorySigns	NoEntry, NoParking, NoStopping, NoUTurn, NoTurn, RoadClosed, NoOvertakingBegin, NoOvertakingEnd, MaximumSpeedLimitXBegin, MaximumSpeedLimitXEnd, SpeedLimitZoneXBegin, SpeedLimitZoneXEnd	
	RoadTypeSigns	MotorwayBegin, MotorwayEnd	
	SpecialSigns	VulnerableRoadUserCrossing, VulnerableRoadUserPathWay, IndicationCameraSurveillance	
	WarningSigns	Children, Crossing, Cyclists, Animal, Curve, RoadWorkBegin, RoadWorkEnd, Roundabout, TrafficSignalAhead, RoadNarrows, RoadBump, RoughRoad, Slippery, GenericWarning, CongestionAhead, TwoWayTraffic, MergingTraffic, Crossroads, DoubleCurve, TunnelAhead	
	NotListed	N/A	
Unclear	N/A	N/A	
Road condition	N/A	N/A	Wetness, SnowCoverage

Table 9: Annotation details for different tasks in ZOD.