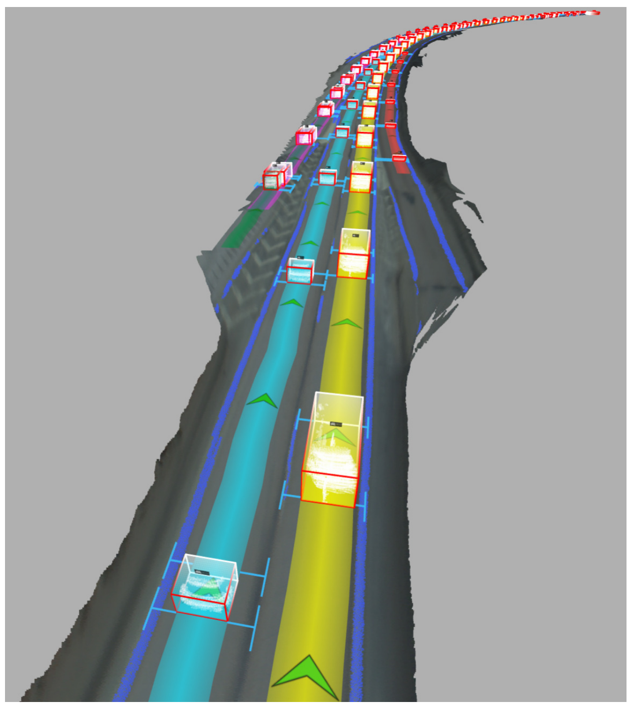
Safety-driven data labelling platform to enable safe and responsible AI

Public report



Author: Yinan Yu, Samuel Scheidegger, Jörg Bakker
Date: 2021-10-31
Project within: FFI, Traffic safety and automated vehicles

FFI Fordonsstrategisk Forskning och Innovation

Contents

1 Sammanfattning på Svenska **3**

2 Executive summary **3**

3 Background **4**

4 Purpose, research questions and method **6**

5 Objective **8**

6 Results and deliverables **8**

7 Dissemination and publication **19**

8 Conclusion and further research **21**

9 Participants and contact persons **23**

1 Sammanfattning på Svenska

Djup maskininlärning och artificiell intelligens används flitigt i AD/ADAS-system. Dess oförutsägbara beteende är dock ett stort problem i säkerhetskritiska system. Ett problem är att AI-systemet vanligtvis utvecklas isolerat, vilket innebär att algoritmerna lär sig av ganska standardiserade annoteringsriktlinjer, såsom rektanglar och semantisk segmentering, som inte är specifika för trafiksäkerhetskraven. På grund av denna utvecklingsprocess är det resulterande AI-systemet inte tränat för att vara medvetet om säkerhetsspecifikationerna.

Givet denna begränsning är målet med detta projekt tvåfaldigt. Först strävar vi efter att ta fram en uppsättning säkerhetskodade riktlinjer för annotering för djup maskininlärning. För det andra, baserat på dessa riktlinjer, skapas en proof-of-concept säkerhetsdrivet benchmark-dataset för träning och validering av djup maskininlärning. Detta dataset innehåller tidssynkroniserad sensordata som samlats in runt Göteborgsområdet i Sverige. Det är sammansatt av kamerabilder, lidarpunktmoln, GPS-koordinater, hastighetsmätning och IMU-information med säkerhetskodade annoteringar på all sensordata. För att kunna validera det tränade AI-systemets prestanda och effektivitet krävs vanligtvis en stor mängd annoterad data, där en manuell annoteringsprocess kanske inte är genomförbar. Därför är datasetet i detta projekt annoterat av en skalbar automatisk annoteringsprocess med vår AI-dataplattform SnapXS. Dessutom är detta dataset helt sökbar givet säkerhetsrelaterade attribut för datafiltrering, urval och forskningsändamål.

Det finns två avsedda användningsfall: 1) det annoterade datasetet är redo att användas för AI-utveckling och validering ur ett trafiksäkerhetsperspektiv; 2) de säkerhetskodade annoteringsriktlinjerna kan användas som referens för säkerhetsdriven AI-utveckling. Detta initiativ syftar till att öka Sveriges satsning på att bygga tillförlitliga AI-system och bidra till arbetet mot nollvisionen.

2 Executive summary

Deep learning and artificial intelligence (AI) is widely used in AD/ADAS systems. However, their unpredictable behavior is a major concern in safety critical systems. One problem is that the AI system is typically developed in isolation, meaning that the algorithms are learning from rather standardized

annotation guidelines, such as bounding boxes and semantic segmentations, which are not specific to the road safety requirements. Due to this development process, the resulting AI system is not trained to be aware of the safety specifications.

Given this limitation, the goal of this project is twofold. First, we aim to produce a set of safety-encoded deep learning annotation guidelines. Secondly, based on these guidelines, a proof-of-concept safety-driven benchmark dataset is created for deep learning training and validation. This dataset contains time synchronized sensor data recorded around the Gothenburg area in Sweden. It is composed of camera images, LiDAR point clouds, GPS coordinates, speedometer recordings and IMU information with safety-encoded ground truth labels on all sensor data. Typically, to be able to validate the performance and efficiency of the trained AI system, a large amount of labelled data is needed, where a manual annotation process may not be feasible. Therefore, the dataset in this project is labelled by a scalable automated labelling process using our AI data platform SnapXS. Moreover, this dataset is made fully searchable given safety related attributes for data filtering, curation and research purposes.

There are two intended use cases: 1) the annotated dataset is ready-to-use for AI development and validation from a traffic safety's perspective; 2) the safety-encoded annotation guidelines can be used as a reference for safety-driven AI development. This initiative aims to boost the effort of building reliable AI systems and contribute to the road map towards Sweden's Vision Zero.

3 Background

Deep learning algorithms are being widely used in ADAS and AD systems due to their impressive capabilities. However, the reliability of these algorithms is known to be challenging due to their data-driven nature. In this project, we aim to bridges this gap by incorporating safety requirements into the AI algorithm development in order to improve their interpretability and verifiability.

To gain a better insight into the problem, we need to take a closer look at the development process of these AI algorithms. Typically, there are multiple steps involved in this process. First, a large amount of data is collected by development vehicles with multiple sensors, such as video cameras, LiDARs, radars, positioning devices. The amount of data depends on the sensor setup, where a typical number is 10TB per day per vehicle. To be able to develop AI algorithms using these data, a core dataset needs to be curated from this large, increasing pool of incoming data.

This step is typically called *data curation* and the outcome is a subset of interests. The level of magnitude for this core dataset depends on multiple factors, but a typical size is estimated to be around 20TB. At the next step, these terabytes of data need to be labelled as the ground truth given some guidelines. This ground truth is then used as the teachers in the data-driven AI development. That is, a deep neural network architecture is selected and trained using this core dataset. These trained networks are then evaluated on a validation dataset that contains terabytes to petabytes of data, which need to be labelled as well. Once the network is trained and validated, it is passed to the subsequent functionalities, such as object tracking and decision making modules.

These steps are repeated throughout the research and development cycle until the requirements of the product are satisfied. Each of these steps has their own requirements, and these requirements should compose to the final requirements of the end product.

However, due to the distinct natures of these steps, they are typically carried out by multiple teams with different specifications, which in turn makes the composition of these requirements challenging. As a result, it is difficult to interpret the behavior of neural networks from the perspective of traffic safety. In particular, we identify three main challenges in the common approach: 1) safety requirements are not systematically encoded into the data annotation guideline; 2) training data curation is not taking safety requirements into its consideration; 3) manual data labelling is highly time consuming and costly, which makes it challenging to iterate and to follow detailed safety guidelines. For instance, one of the state-of-the-art approaches for data curation, active learning, is to select a subset of data that the network is highly uncertain about. If we look at the problem in isolation, this approach makes sense, since the more data varieties we expose the network to, the better data diversity the network may be able to handle. In theory, this heuristic and greedy approach expands the capacity of the network. However, it does not necessarily guarantee safety of the system. Therefore, additional measures need to be taken for this purpose.

The primary goal of this project is to create a set of safety-driven deep learning annotation guidelines and produce a proof-of-concept benchmark dataset for AI algorithm development and validation. Benchmark datasets have played significant roles in the progress of AI development. It is one of the most important ways to contribute to the domain and stay relevant to the state-of-the-art. There are existing datasets such as Kitti, nuScenes, Waymo Open Dataset, etc., created by various research and development organizations. These datasets mainly focus on standard deep learning annotations instead of traffic safety specifications, which is the aspect we focus on in this project. The resulting benchmark dataset from this project consists of annotated and time synchronized sensor measurements (camera images, LiDAR point clouds, GPS coordinates, IMU information, etc) recorded in Sweden. The data is collected using the research vehicle from the laboratory of Resource for vehicle research at Chalmers (REVERE). Data processing and automated annotation is implemented using our AI data platform SnapXS.

4 Purpose, research questions and method

4.1 Research questions

This project is aimed at bridging the gap between deep learning and traffic safety research. The research questions being studies are stated as follows:

1. What are the road safety specifications that are relevant to a typical automotive perception system (**relevant traffic safety requirements**)?
2. What are the important parameters in the data curation process given these specifications (**data curation criteria**)?
3. How to encode these safety specifications into the data labelling process to make the deep learning training driven by the safety requirements (**annotation guideline encoding**)?

4.2 Scope

Safety research and artificial intelligence are both very broad research areas. In order to produce a set of meaningful annotations that reflect traffic safety requirements, research question 1 needs to be addressed as the foundation of the annotation guideline. In this feasibility study, we limit our focus to the training and validation of perception AI systems with the purpose of crash avoidance. Given this scope, we further divide the research question 1 into the following subquestions:

- How many different crash types are there?
- What type of initiation and dynamics causes the crash within each type?
- What are the contributing parameters for each crash type?

Moreover, in order to translate traffic safety specifications into deep learning annotation guidelines, we limit ourselves to answers with the following criteria:

- **Relevance**: the specifications shall follow established standards that can be composed with standardized downstream analysis;
- **Tractability**: the specifications shall not cause combinatorial explosion by adding a few new parameters;
- **Measurability**: clear key performance indicators (KPIs) can be derived from these specifications;
- **Feasibility**: the specifications shall result in annotation guidelines that can be easily implemented or translated into standard deep learning annotations.

4.3 Other technical considerations

There are other technical questions that need to be answered as prerequisites before the annotations can be produced, for instance:

- How to automate and scale up this annotation process?
- How to enable the data filtering machemism on terabytes or even petabytes of data using a reasonable amount of computional resources?

In this project, we utilize our AI data platform SnapXS to achieve these prerequisites. The automated annotation functionality of SnapXS is improved thanks to this project.

5 Objective

AI algorithms are data-driven techniques and therefore it is crucial that the safety specifications are encoded into the data set that the AI algorithms are based upon. In this project, we aim to create such a safety-driven benchmark data set that is ready-to-use for AI algorithm training and validation. The resulting dataset contains sensor data logged in Sweden, together with ground truth labels encoded with safety specifications. There are two intended use cases:

1) deep learning algorithm development using the annotation guidelines as meta data for filtering and data curation purposes; and

2) building end-to-end AI systems to capture the safety-driven annotation guidelines.

In both cases, the annotation is aimed at being used for evaluations of the algorithm given traffic safety KPIs, which in turns incorporates safety specifications in the AI development and validation process.

6 Results and deliverables

The project is divided into three work packages (WP):

WP 1: Data collection activities

WP 2: Safety-driven annotation guideline

WP 3: Automated annotation

6.1 WP1: Data collection activities

The data collection activity is divided into two parts: 1) controlled experiments at the AstaZero test track, and 2) data collected on public roads in urban and rural areas. Both collections are using the research vehicle from the laboratory of Resource for vehicle research at Chalmers (REVERE).

6.1.1 Controlled experiments at the AstaZero test track

The first part of the data collection is conducted at the AstaZero test track. The test track and driving days were funded by the Open Research Programme by Chalmers, RISE, and SAFER in collaboration with AstaZero. The processing and analysis of the data is part of this project.

This data collection strives to provide edge cases for testing AI based perception systems. Currently available public data is limited to normal driving situations (e.g. KITTI or Waymo Open Dataset). Data that covers edge cases like safety critical scenarios to test these systems is still lacking. More specifically, these edge cases are derived from real world scenarios and crash test configurations that lead to accidents caused by human drivers. The created dataset provides ground truth and is aware of safety requirements by also delivering KPIs to evaluate the perception system of AD/ADAS functions.

In total, 164 tests that simulate accident pre-crash scenarios along with non-crash reference tests are recorded at the AstaZero track. The selected scenarios are based on accident data analysis, EuroNCAP and USNCAP test configurations as well as real world accident pre-crash data from the US ensuring relevance for real life safety. The specification of each scenario describes the pre-crash phase which then is simulated in the tests. The dynamic objects in the tests are real objects (e.g. no dummy cars) to mimic realistic sensor data for visual perception systems in real world usage of production cars.

The tests are comprised of:

- 71 city scenarios at a crossing, including different variations of turning (42 scenarios) and crossing accidents (29 scenarios), with and without occlusion by other cars or buildings;
- 93 rural scenarios of head-on accidents with different degrees of overlap on roads with curvature (63 scenarios) and without curvature (30 scenarios).

As a first step, virtual simulations of all pre-crash phases of the test scenarios are created. Examples can be found in Fig. 1. Professional drivers from AstaZero drove these scenarios according to the specification of the virtual simulations in slow motion that is at half speed in the city and a quarter of the speed on the rural road of the test track. Later, the data is time warped

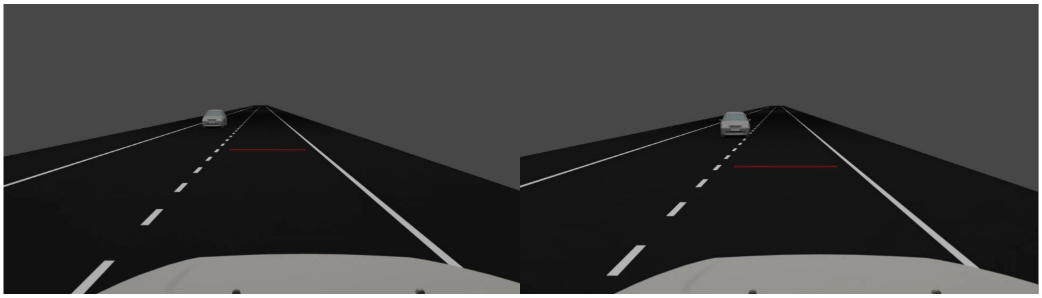


Figure 1: Virtual scenarios as a basis for the real tests: left reference non-crash, right pre-crash.

to run at the original speed of the crash scenario. The cars stops very close to the supposed *collision* of the ego vehicle. Each scenario varies with different parameters such as speeds, angles, etc. One example setup can be found in Fig. 2.

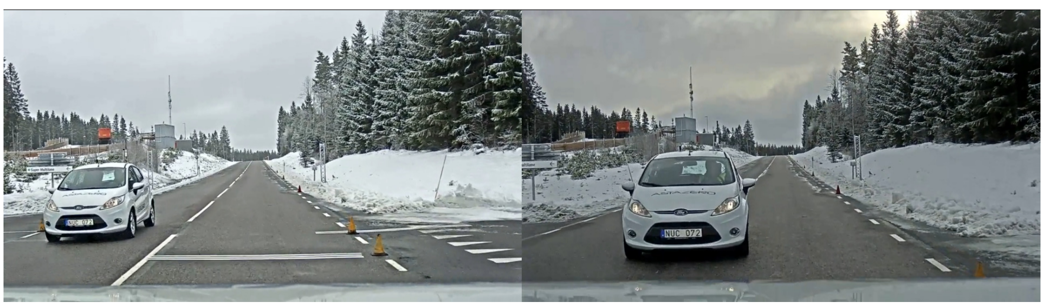


Figure 2: Examples of rural straight road scenarios: reference case (left) versus collision case (right).

6.1.2 Data collected on public roads in urban and rural areas

The second part of the data collection is to collect data on public roads around the Gothenburg area in Sweden. With this funding, we have collected about 38 hours of data that amount to approximately 23 TB of raw recordings.

The sensor data include 32-layer LiDAR point clouds, 5 image cameras that covers around 200 degrees field-of-view mainly on the front, right and rear side of the ego vehicle. One alternative forward looking camera mounted inside the windshield. A GNSS receiver, an IMU platform and a speedometer for positioning and ego motion measurements.

These recorded trips are designed to be diverse in terms of the environmental conditions, ego behaviors and target object properties. These variables are derived from the safety specifications developed in WP2.

To protect personal information, the collected data is anonymzied using our automated anonymization tools.

6.2 WP2: Safety-driven annotation guideline

To be able to define a set of safety-driven annotation guidelines that satisfy the criteria in Sec. 4.2, the second work package is to research and experiment with different sources of safety specifications. These specifications are also discussed with various safety experts and stakeholders. As a result, the main sources we have adopted are 1) the EuroNCAP/USNCAP test configurations, 2) the crash analysis from the International Road Assessment Programme (iRAP) and 3) the accident data from the iGLAD database, where Asymptotic AI is a member and an active contributor. These specifications are then translated into deep learning annotation guidelines so that KPIs can be evaluated from the algorithm output. The annotations produced in this project are based on the guidelines developed in this work package. The annotation guidelines can also be used as a stand alone reference in safety-driven AI development. More detailed descriptions and analysis of the specifications and guidelines will be presented in our publication.

6.3 WP3: Automated annotation

Given the established annotation guideline from WP2, the labels are automatically generated using our AI data platform SnapXS. In this process, we avoid human labors to increase efficiency and we also aim to minimize computational resources for the automation.

6.3.1 Data preparation

To be able to select and search for any data point from any sensor for analysis and annotation from terabytes of raw recordings, first, the collected data needs to be preprocessed and structured. This is enabled by SnapXS.

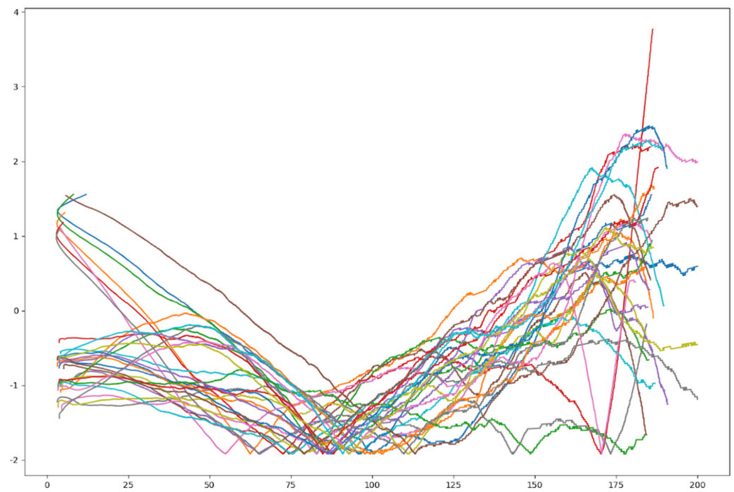


Figure 3: KPI lateral offset in the rural straight road scenarios.

6.3.2 AstaZero test track (cf. Sec. 6.1.1)

The test data collected at the AstaZero test track is annotated with the precise ground truth that is required to derive the safety KPIs. The purpose of the KPIs is to detect a collision threat as early and reliably as possible.

As an example, Fig. 3 shows the KPI "lateral offset" in the rural straight road scenarios. The x-axis denotes the distance to the meeting point of the cars and the y-axis represents the lateral offset, where zero distinguishes between critical scenarios (below zero) and reference scenarios (greater than zero).

Early warnings based on this KPI can already take place at 100 meters distance to the meeting point. A late and urgent warning at up to approximately 40 meters indicates that the collision is not preventable on the current course and immediate actions need to be taken which include emergency braking, steering maneuver, signals (horn, light) or a pre-crash system activation.

Example annotations can be found in Fig. 4 - 6. Fig. 4 shows the front camera view of the ego vehicle in a rural straight road reference scenario, where the target car is passing the ego vehicle in a straight line without any interference. The KPI "collwarn" is based on the lateral offset of the ego and the target car and shows constantly no warning throughout the sequence.

Fig. 5 shows a corresponding pre-crash scenario where the target car is

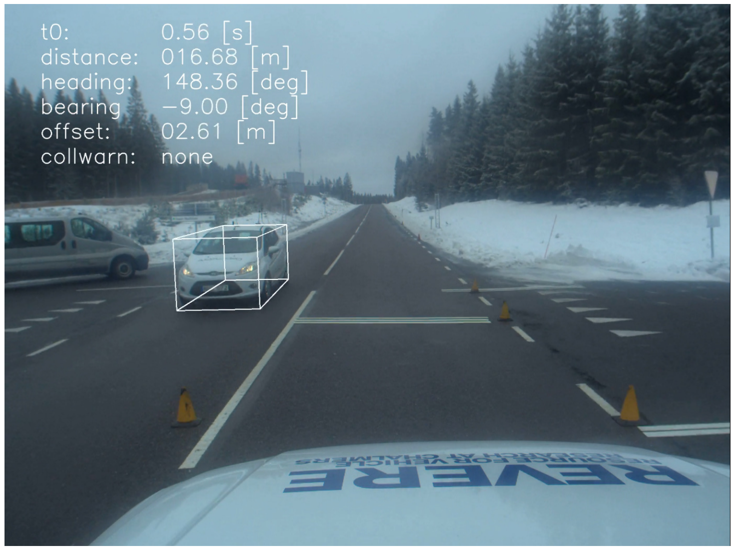


Figure 4: Annotated reference scenario with KPI "collwarn" for collision warning.

slowly starting to drift into the lane of the ego vehicle. The KPI "collwarn" is already issued on a medium level at a distance of around 100 meters and is then raised to a critical level at approximately 45 meters with a remaining reaction time of around one second, which is shown in Fig. 6.

When benchmarking an AI visual perception system for the KPIs, the given KPIs derived from the ground truth in the dataset need to be replicated as closely as possible based on the corresponding video annotations (3D bounding boxes) generated by the AI system. The accuracy of the AI generated KPIs is then benchmarked against the ones calculated from the ground truth.

6.3.3 Public road (cf. Sec. 6.1.2)

As opposed to the controlled experiments at the AstaZero test track, when collecting data on public roads, we have no control over the precise location and movement of dynamic objects in the surroundings. In this set up, the labels on the data collected are automatically generated by aggregating information from all available sensors. The labels are generated based on the following analysis:

Figure 5: Annotated collision scenario with early and medium collision warning.

Figure 6: Annotated collision scenario with urgent collision warning, actions need to be taken.

- Road estimation: a 3D model of road surfaces and drivable areas are estimated and extracted; one example is shown in Fig. 7[1], where the reachable road of the ego vehicle is extracted from the collected data, where static and dynamic objects are identified and removed.

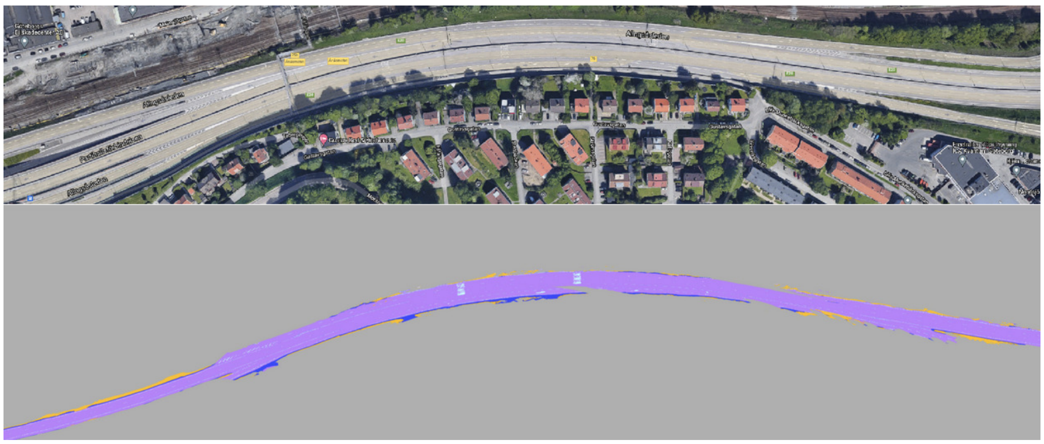


Figure 7: The image on the top is a satellite imaging of a road segment from Google maps (at approximately 57.7239327, 12.0173449). The image on the bottom is the extracted ego road in 3D (with static and dynamic objects removed) of that segment.

- Lane detection and modelling: to understand the traffic rules and safety constrains, lane markings and road edges are detected and modelled;
- Dynamic object localization and classification: each dynamic object within the range of the ego perception system is localized, classified and modelled. The travelling *heading* and *velocity* of the target object are annotated for evaluating safety related properties;
- Static 3D global map: the static surroundings are modelled as a 3D map with semantic annotations; this will be used in future projects for road infrastructure analysis;
- Multi-object tracking: the trajectory of each dynamic object is tracked individually for KPI evaluations.

[1] Data providers: Google, Aerodata Internatinal Sruveys, CNES/Airbus, Lantmäteriet/Metria, Maxar Technologies

To create these labels, we use a combination of *model-based techniques* (e.g. signal processing, geometric modelling) and *data-driven algorithms* (e.g. deep learning).

To demonstrate the annotations and how they are related to the safety specifications, some examples are illustrated below.

According to the definition of a *traffic crash*, *at least one vehicle* is involved and *at least one person is injured or killed*. Therefore, our primary goal is to analyze the level of threats a vehicle poses to other road users and to itself. To this end, we identify two main categories in terms of safety-driven annotations: 1) *occupancy* and 2) *trend*. Occupancy indicates the space that a target vehicle occupies and trend refers to the information that can be used for predicting future occupancies. In this section, we briefly describe some examples of occupancy and trend annotations.

Occupancy In this dataset, the occupancies of target vehicles are annotated. In particular, the occupancy in relation to the lane marking is strongly related to the intention and state of the road user.

One example can be found in Fig.8, where we demonstrate the following annotations:

1) A 3D bounding box that encloses the vehicle is annotated, e.g. the bounding boxes in Fig. 8. The white surface of the box indicates the travelling direction. A close in example of the bounding box is shown in Fig. 9.

2) A trajectory is annotated to indicate the location of each vehicle over time in the global coordinate system, e.g. the band under each bounding box in Fig. 8.

3) The drivable lanes are identified and modelled in 3D for further road safety analysis, e.g. the blue markings in Fig. 8.

4) At each point in time on its trajectory, the occupancy of a vehicle in relation to the nearest lanes is annotated, e.g. the cyan line next to each bounding box in Fig. 8 and Fig. 10:

   - If a vehicle is driving within a single lane, the occupancy is annotated as the distance from the vehicle to the nearest lane markings both to its left and right within the current lane, e.g. Fig. 10a.

– If a vehicle is crossing a lane marking, such as a single broken white line, the occupancy is annotated as the distance from the vehicle to the lane markings that include both lanes, e.g. Fig. 10b.

The occupancy is annotated for various environmental factors, target ve-

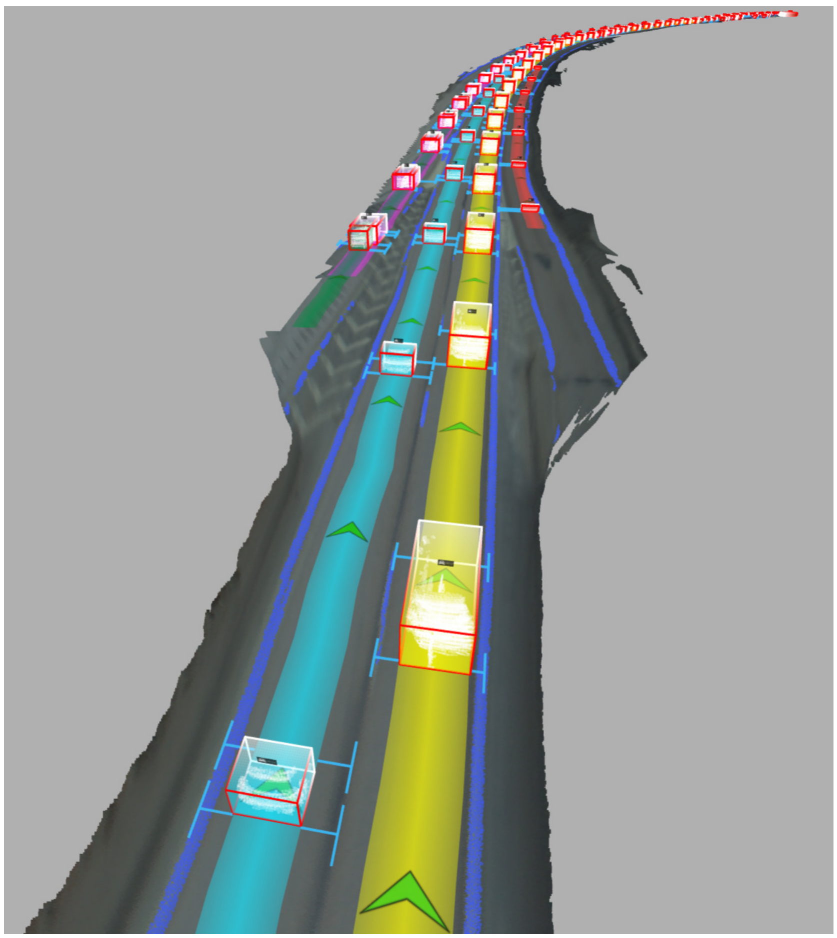


Figure 8: Annotations of all recorded tracks on one road segment. Each colored band represents the annotated trajectory of one vehicle. The blue lanes are lane markings and the perpendicular lines indicate the occupancy of the vehicle. In this example, the box is designed to enclose visible LiDAR points to the sensors mounted on the ego vehicle. Annotations with estimated box size are also available. The white surface indicates the travelling direction of the car.

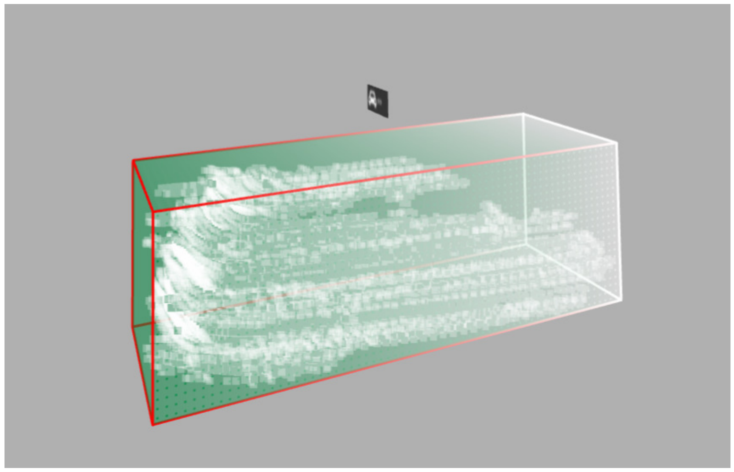
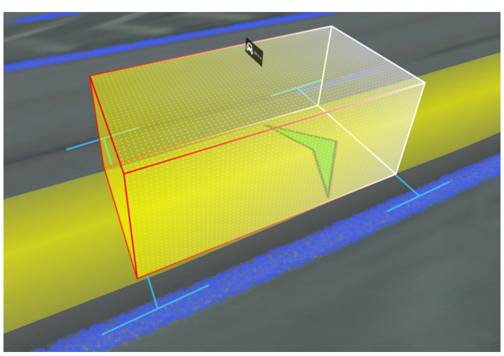
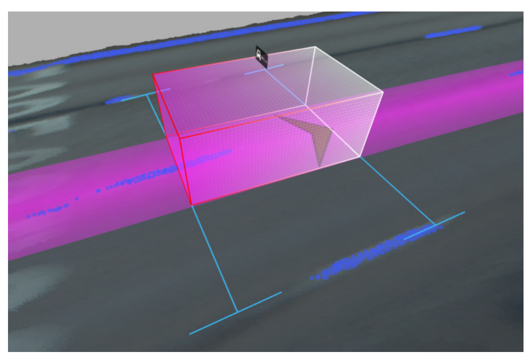


Figure 9: Example of the bounding around visible LiDAR points from a moving target vehicle. Note that here the points are from all LiDAR scans registered into one object.

hicle conditions and ego behaviors to create different scenarios. These scenarios represent multiple threat levels from a traffic safety's perspective. In Fig. 11, one example of annotated lane change is presented from the birdeye view. Examples shown in this section are extracted from the collected sensor data described in Sec. 6.1.2.



(a) Vehicle travelling within a single lane: the occupancy is denoted as the distance to the nearest markings within the current lane.



(b) Vehicle performing a lane change: the occupancy is annotated as the distance from the vehicle to the lane markings that include both lanes

Figure 10: Occupancy of a vehicle in relation to lanes.

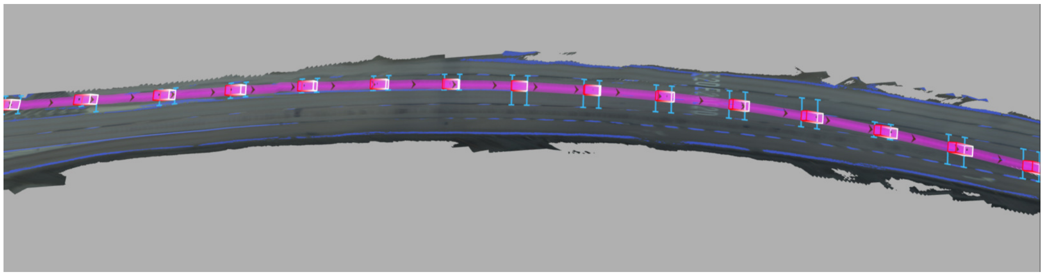


Figure 11: Birdeye view of a target vehicle performing a lane change.

Trend This refers to a set of annotations that are aimed at aiding prediction tasks and deriving the braking time and braking distance. Trend annotations include information such as vehicle heading and velocity. This information is provided in various forms including: 1) its absolute unit (e.g. heading in the world coordinate system, velocity in meters per hour, etc), 2) relative to the lanes, and 3) relative to the ego vehicle. Furthermore, environmental factors, such as weather condition and road slipperiness, are also part of the annotation. Users can then utilize these annotations to model the trajectories of dynamic objects and predict their future occupancies.

Some examples of this type of annotations are demonstrated in Fig. 12, 13 and 14, where the example safety KPI "collwarn" is indicated [2]. The description and interpretation of this KPI can be found in Sec. 6.3.2.

More details on the modelling and specifications including the resources used for the annotation process will be presented in our publication.

7 Dissemination and publication

7.1 Dissemination

The dissemination is shown in Tab. 1.

[2]The trend annotations are available in the same format as occupancy, but they are best demonstrated in the image domain.

Figure 12: Example of medium and urgent warning are demonstrated in this image. The example traffic safety KPI "collwarn" depends on the relative speed and reachability of the target vehicle in relation to the ego vehicle.

Figure 13: Example of reference case (no warning) is demonstrated in the this image.

Figure 14: A sequence of a target vehicle performing a lane change in front of the ego vehicle. The sequence of images are annotated with the heading and relative velocity to the ego vehicle. The “collwarn” signal (described in Sec. 6.3.2) is indicated in the annotations.

7.2 Publication

A safety-driven benchmark dataset for deep learning training and validation for crash avoidance, Yinan Yu, Samuel Scheidegger, Jörg Bakker, in preparation.

8 Conclusion and further research

In this project, we have created a set of safety-driven annotation guidelines and a benchmark dataset for deep learning training and validation. This dataset is automatically annotated using our AI data platform SnapXS. The annotation guidelines are derived from traffic safety specifications. We are currently investigating the licensing and resources available for hosting the release of the dataset. We are also looking into various possibilities of data sharing mechanisms between different actors in a more generic setting. This dataset will be used for our future collaboration with the REVERE lab in automotive research. Moreover, this dataset is served as a basis in an ongoing

Table 1: Dissemination

How has / is the project result to be used and disseminated?	Mark with x	Comment
Increase knowledge in the field	x	Safety-driven annotation guidelines; automated annotation tool chain improvement for automotive use cases.
Be passed on to other advanced technological development projects	x	The result is used as a basis for multiple purposes. For instance, it is part of the ongoing research applications on the subject of roadside infrastructure modelling for improving road safety. In addition, the result will be used in Vinnova project DecarbonAIte (2021-02759) for energy efficient urban planning and future collaborations between REVERE and Asymptotic AI on traffic safety research.
Be passed on to product development projects	x	The outcome of this project will be used as a basis for the upcoming Vinnova project: a quality validation toolbox for automotive perception data towards trustworthy AI (2021-02577). Furthermore, as a result of this project, we have improved the automated annotation module of our AI data platform SnapXS.
Introduced to the market		The automated annotation functionality of SnapXS is introduced to the market, but the result of this project has mainly been used for research and development so far.
Used in investigations / regulatory / licensing / political decisions	x	We are investigating the licensing for the dataset release.

research application on the subject of roadside infrastructure detection and maintenance. Furthermore, the dataset and the annotation technology will be applied in projects such as DecarbonAIte (Vinnova 2021-02759) for energy efficient urban planning and data quality control for reliable AI (Vinnova 2021-02577).

9 Participants and contact persons

Yinan Yu - Project leader
yinan.yu@asymptotic.ai

Samuel Scheidegger
samuel.scheidegger@asymptotic.ai

Jörg Bakker
jorg.bakker@asymptotic.ai