BOSCH

AI数据驱动工程 - 如何在大模型时代保障高阶辅助驾驶感知系统的安全性

季康, 博世车辆运动智控事业部

2024 MathWorks
中国汽车年会

Agenda


Motivation & Introduction


Data-Driven Testing


Data in Machine Learning


Data-Driven Engineering (DDE) as Solution


Automation of DDE Process Workflows


Demo on DDE Process Workflows & Outlook

Credits to: Andreas Albrecht (XC-AS/EDL2) & Zhuang Lin (XC-AS/EDL1)

BOSCH

Classic SW Development versus Machine Learning

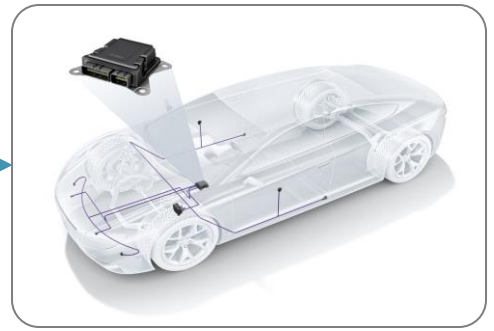
Requirements given explicitely in textual form, e.g.:

„If the steering angle does not match the driving direction, then start active compensation …

Requirements Specification

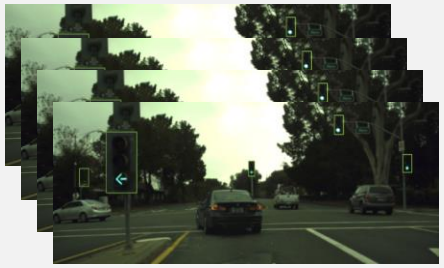
SW Coding

SW Units

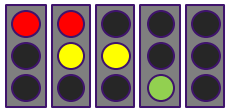


e.g. ESP

Deep Neural Networks:
First time to implement implicit requirements!
=> New fault models needed!



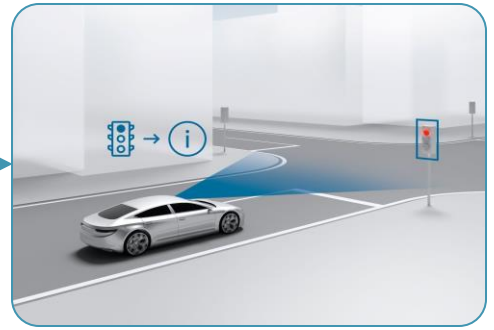
Requirements given implicitely by the data contents …



Data Sets

ML Model Training

SW Units



e.g. video-based traffic light detection

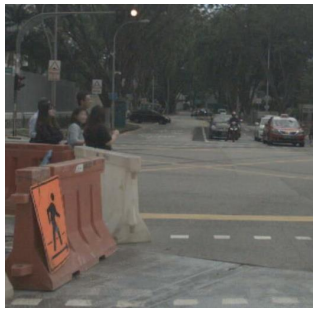
ML models learn their I/O-behavior from the training data. We need to design our data sets to implicitely define the ML function!

AD: Automated Driving DDE: Data-Driven Engineering DNN: Deep Neural Network FoV: Field of View GT: Ground Truth ML: Machine Learning (connectionistic ML – esp. DNNs) ODD: Operational Design Domain TL: Traffic Light

Challenging Scenarios for Perception Systems

Construction sites with workers

Occluded pedestrians



Pedestrians with low contrast (at night)

Metal structures + rainy weather + bad lighting conditions

Complex urban scenarios

Rare unusual human poses

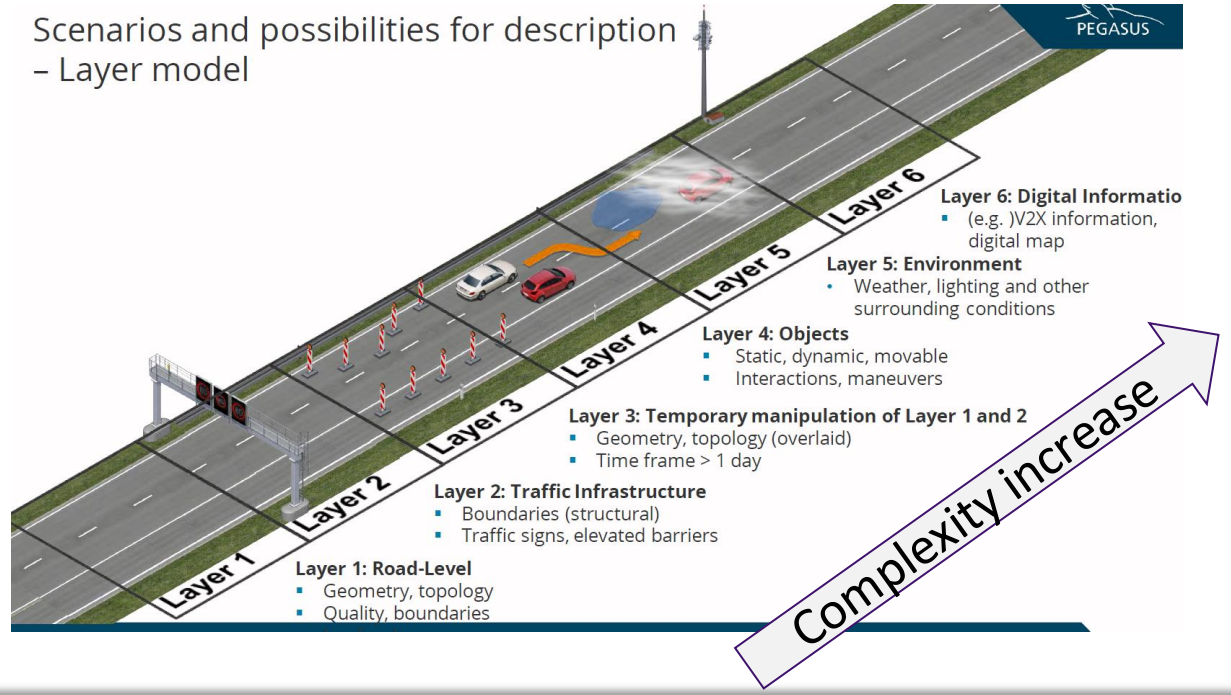
Dark tunnels + cyclists or pedestrians

Steel bridges, guard rails

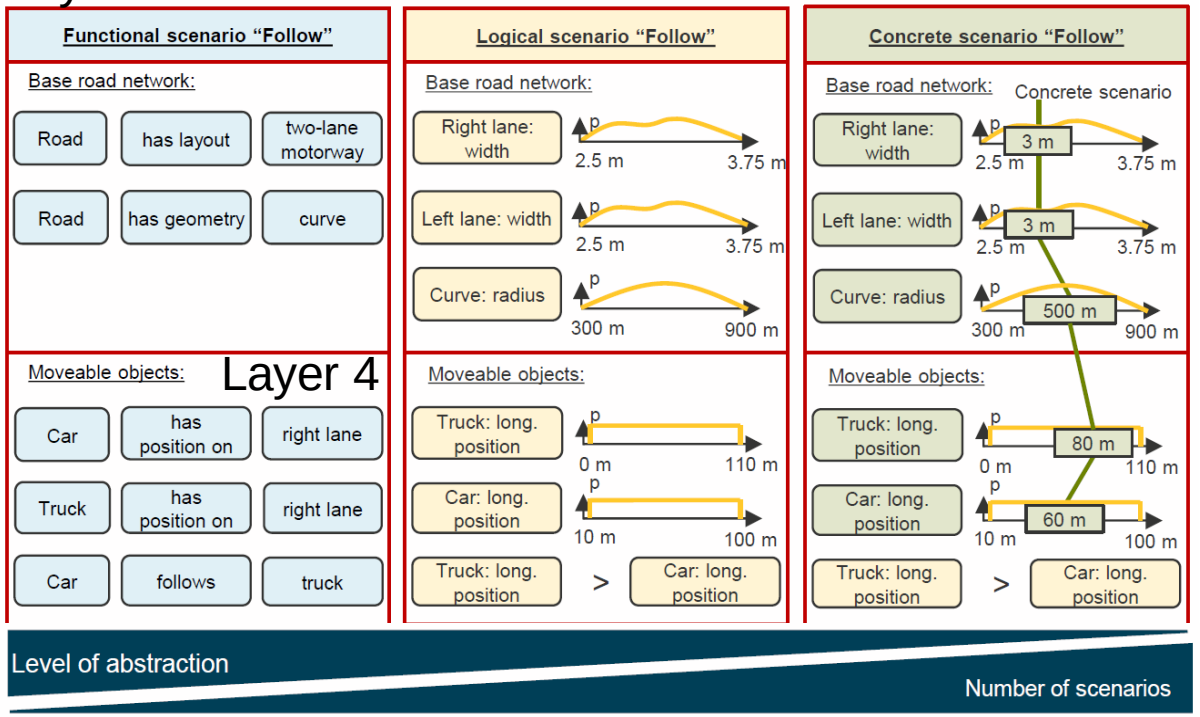
Specific scenario constellations and environment conditions may trigger functional insufficiencies of our system under test.

ODD Model & Ontology are the Basis for Data Space Description

Defined Layer of influence parameter (Environment, Infrastructure, other road users)



Concept phase	System development phase	Test phase
Human experts shall be able to formulate scenarios in the fields terminology in **natural language.**	Scenarios shall include the **parameter ranges** of the **state values** used for scenario representation.	Scenarios shall be modeled via a **single representative** for each state value to ensure reproducibility.



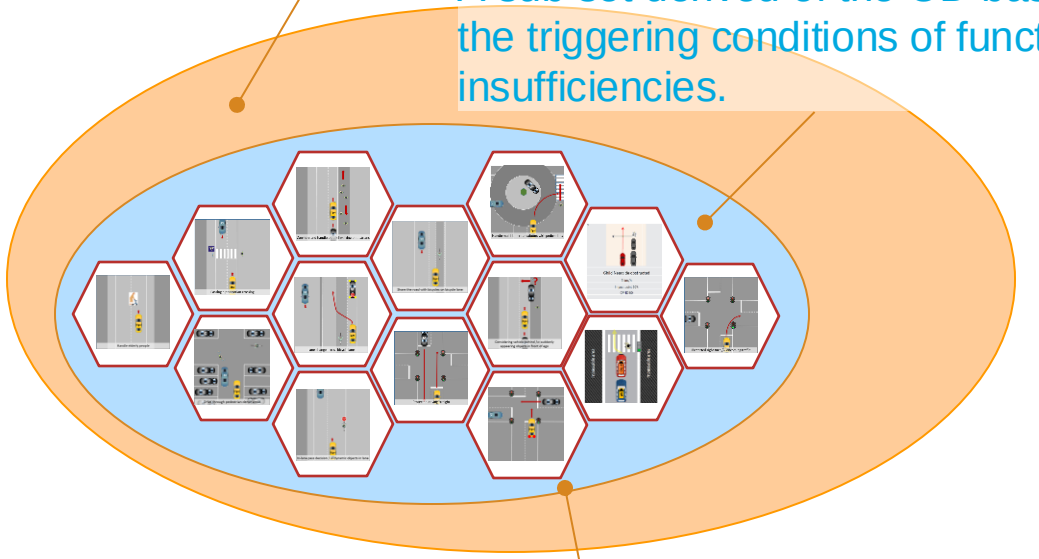
Ideally we will derive our domain-specific languages to describe our data space from a common ODD model & ontology

BOSCH

Gain Deep Understanding of the Operational (Design) Domain

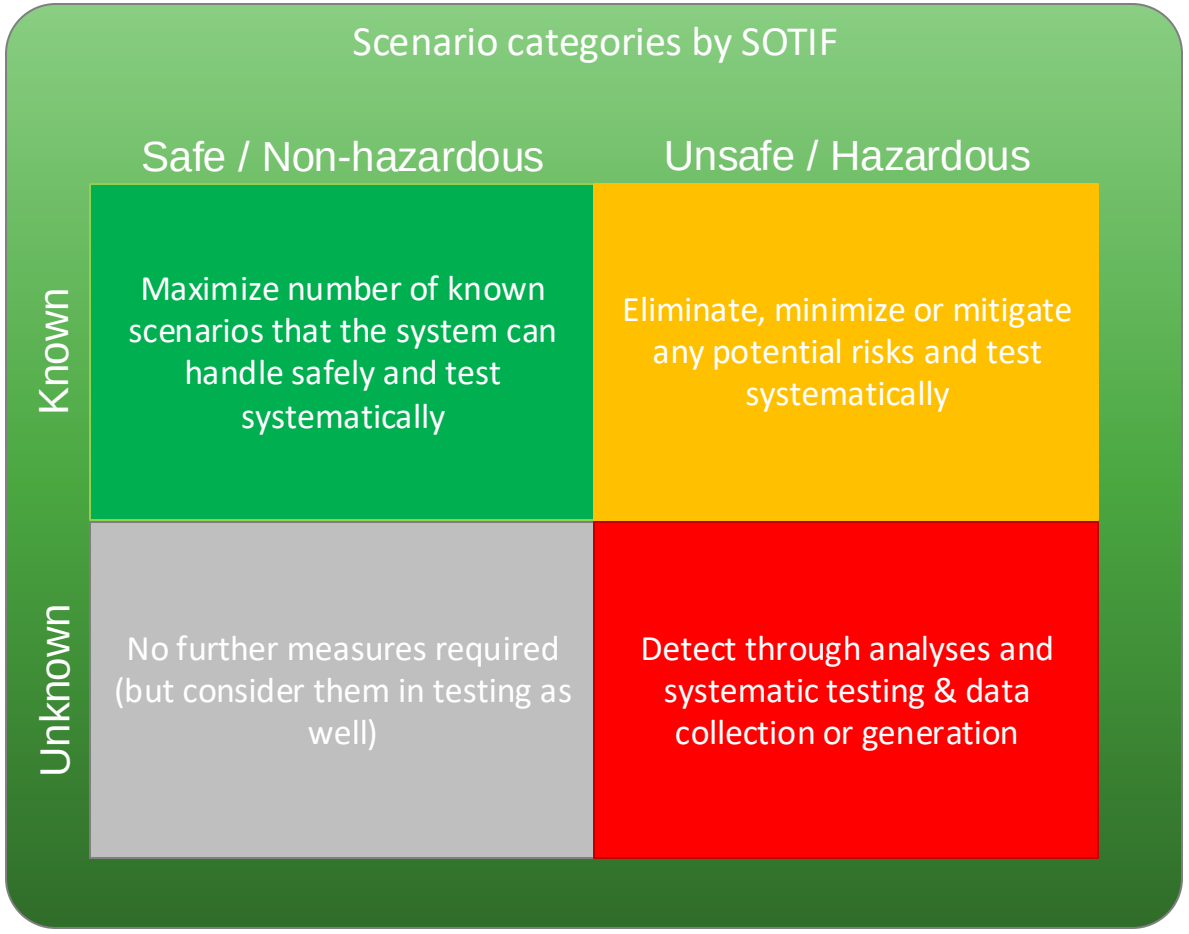
Operational Domain (OD) for which an automated driving function is designed.

Operational Design Domain (ODD): A sub-set derived of the OD based on the triggering conditions of functional insufficiencies.



The border between OD and ODD has to be defined by SOTIF: What is in & what is out of ODD bounds.

ODD description by decomposition into ODD categories & items.

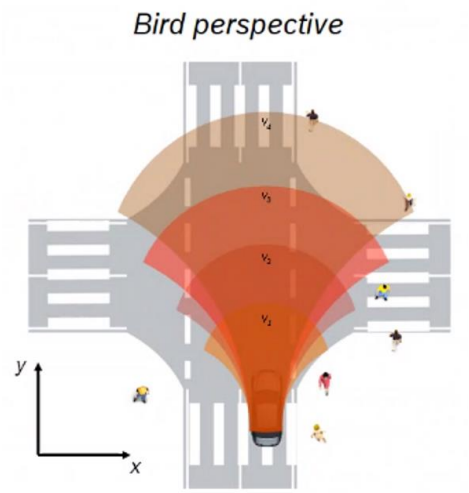


The ODD is decomposed into scenario classes & categories, which are analysed if they can trigger functional insufficiencies.

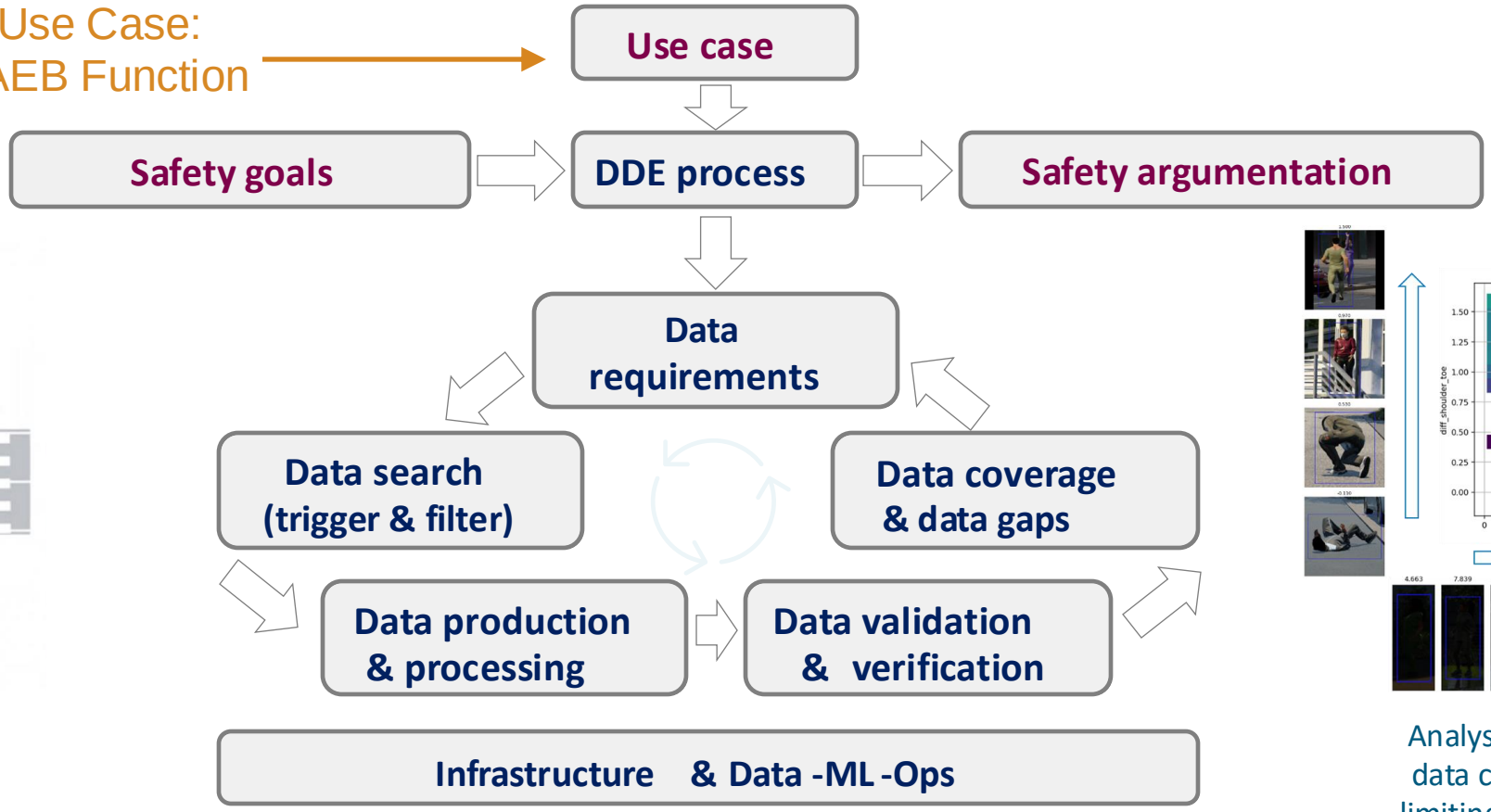
BOSCH

Systematically Derive Safety Arguments based on Data Coverage

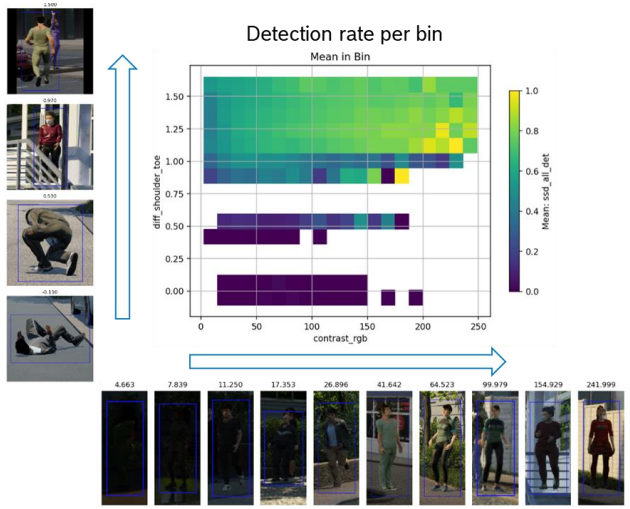
First DDE PoC Use Case:
Vision-baed L2+ AEB Function



Derive safety-related data requirements from system use case and Operational Design Domain (ODD)



Filter data pools for suitable samples in order to systematically increase test data coverage w.r.t. performance limiting factors to support a safety argumentation.



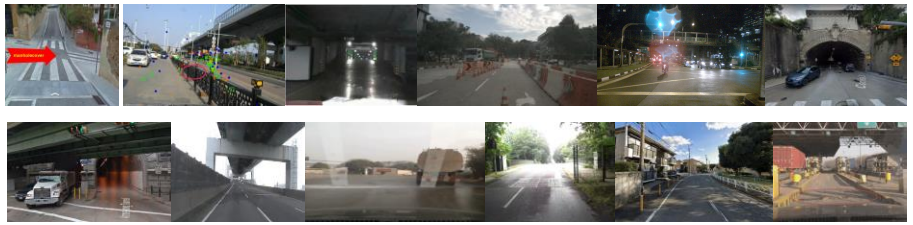
Analyse DNN performance versus data coverage w.r.t. performance limiting factors to find weaks spots

The DDE process supports systematic derivation & iterative refinement of data requirements to support a safety argumentation.

Verified Datasets with defined Coverage for Data-driven Testing

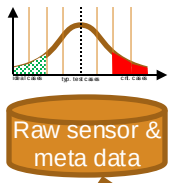
Selective search for desired scenarios or data features to compile representative test datasets with defined data coverage

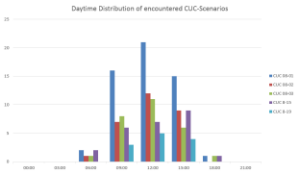
Design datasets with defined coverage

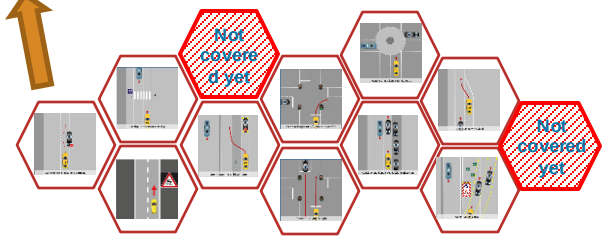


Filter for data items with desired features

Data pool with raw or partially labeled data

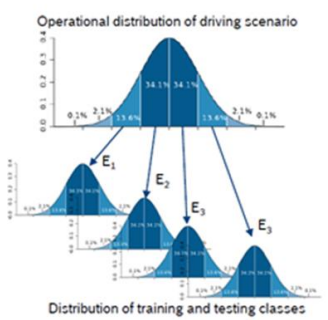






Verified test dataset for a specific test goal

Data coverage statistics



Test Plan

- Test Data
  - **Test Data Link:** … &
    Contents: …
- Test Suite 1
- Test Suite 2

Execute test suites on the given test dataset

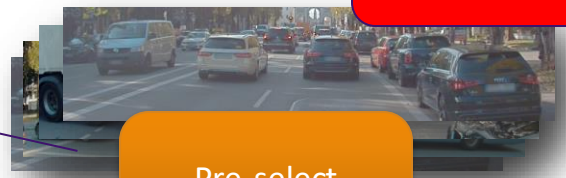
We want to have verified test datasets with defined coverage that are representative for the Operational Design Domain (ODD).

BOSCH

Verified Datasets with defined Coverage

To provide sufficient ODD coverage we must consider a top-down system & safety engineering perspective.

Interact with data management system



Pre-select images

reduce labelling effort, cost and time

Scan & score existing & incoming data

DNN-/DuT-based scoring

Scoring is biased focusing on the current decision boundaries of the DuT.

If an AI model learns its behavior from data why not let the model decide itself what training data it needs?

Labelling & Label Quality Control

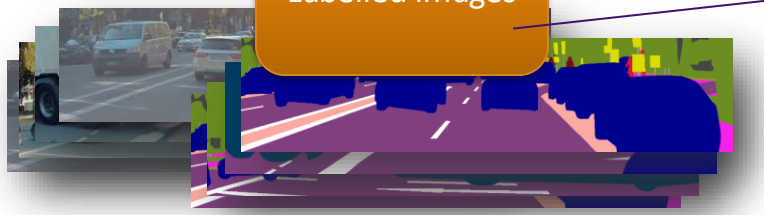
Continuous Training

AL does not apply for testing.

AL only considers the perspective of the AI model, which knows nothing about system architecture, ODD, or safety.

Integrates closely with Continuous Training

Labelled images



Iteratively improves models / datasets while limiting required labelling effort.

AL focuses only on the AI model's perspective, but does not consider the ODD or any system & safety engineering aspects.

AL: Active Learning AI: Artificial Intelligence DNN: Deep Neural Network DuT: Detector under Test

Iterative, Systematic Data-Driven Engineering Process Needed

AI Life Cycle Model with Data Life Cycle Model required

ISO PAS 8800 2022-09-20

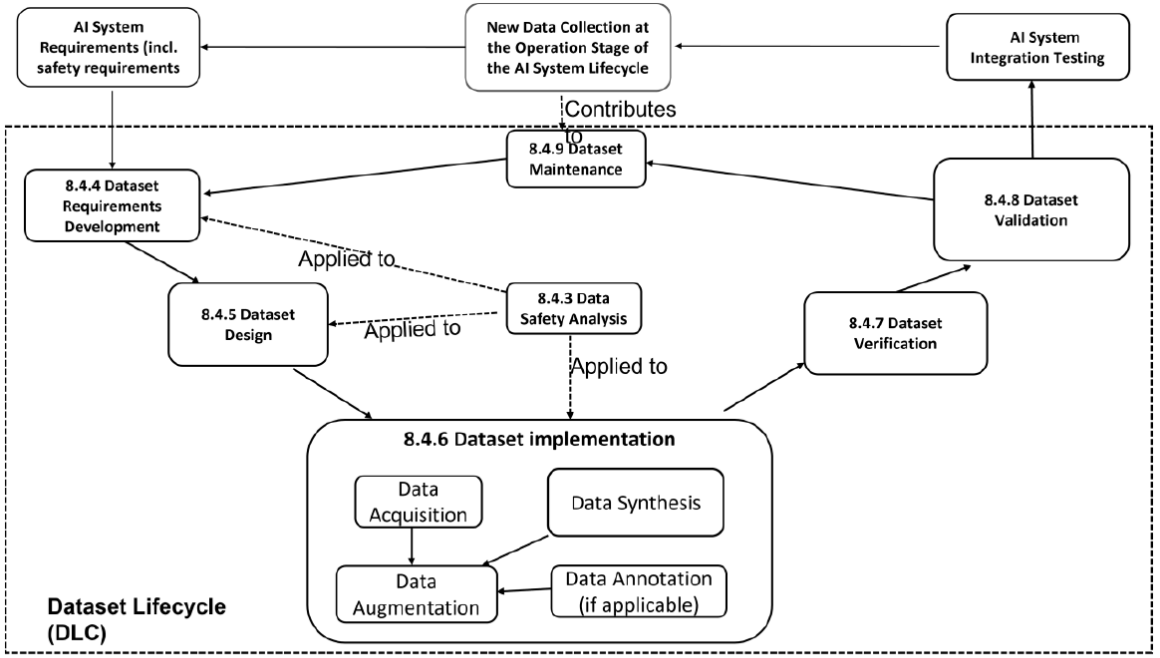
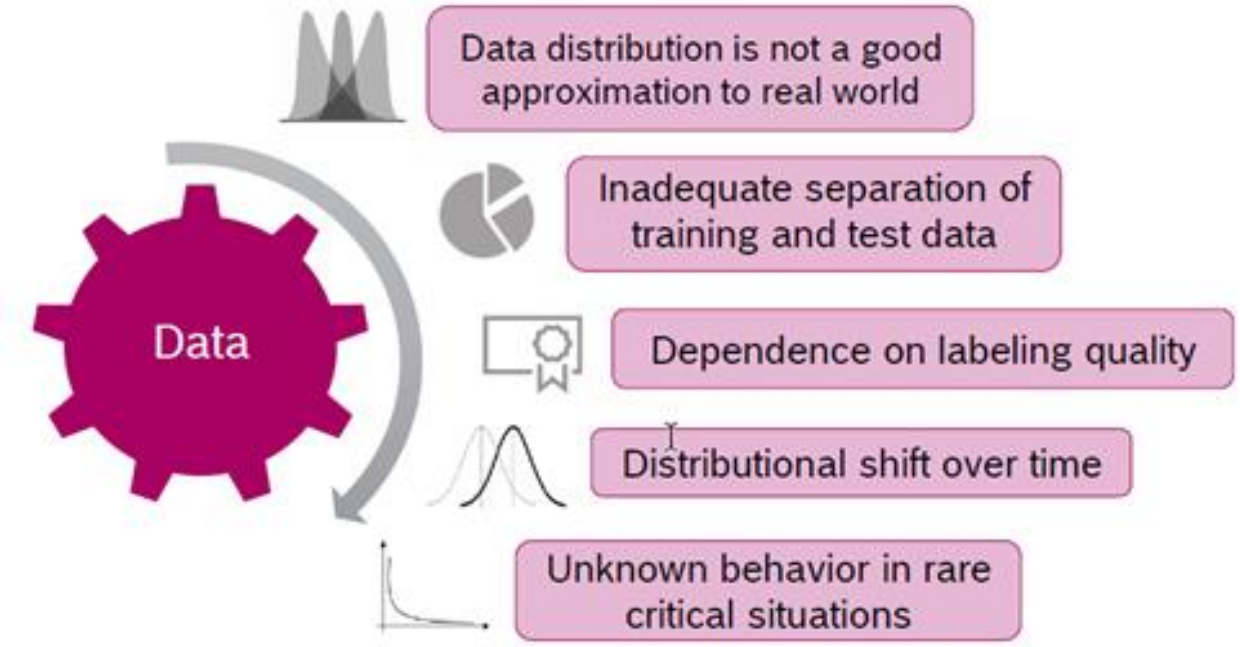


Figure 10–2 — Dataset Lifecycle Model

Source: ISO/AWI PAS 8800:2023(E), Road Vehicles — Safety and artificial intelligence

AI safety conerns raised on data must be addressed



Source: KI Absicherung

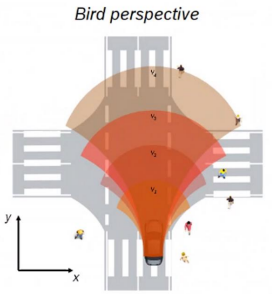
Future safety argumentation must address safety concerns amongst others on data aspects raised for the respective AI system.

AI: Artificial Intelligence

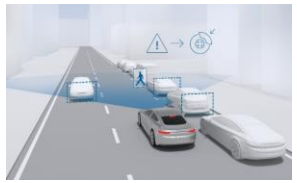
Data-Driven Engineering brings in Top-Down SYS, ODD & Safety

Feedback from validation & verification and monitoring w.r.t. actual data distributions, corner cases, …

Drive from A to B



Detect collision-relevant objects

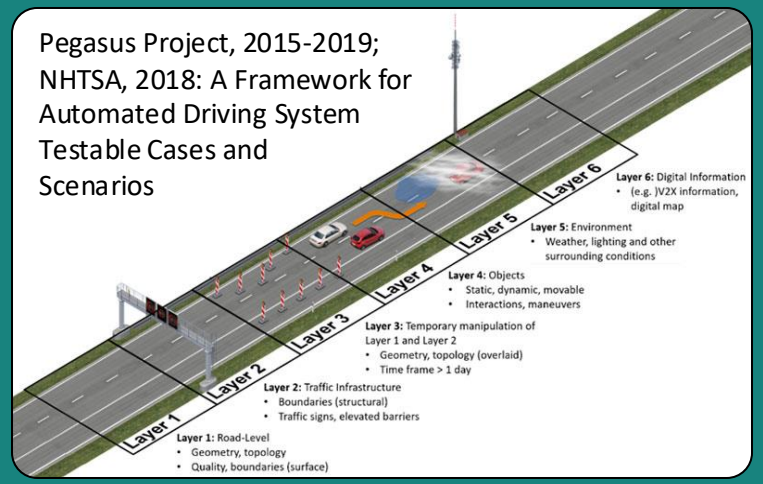


Emergency braking scenario: Occluded pedestrian steps Into driving corridor

- Overall mission goals
- Use cases per mission
- Base scenario classes
- Scenario collections / variants to be covered

Operational Design Domain (ODD) Model

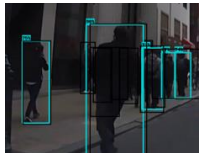
Pegasus Project, 2015-2019; NHTSA, 2018: A Framework for Automated Driving System Testable Cases and Scenarios



- SYS/SW/ML development, verification & result analysis
- Data set verification, data set content & gap analysis
- Requirement-based data set implementation
- Data requirement analysis & data set design

Unusual poses

Occlusion



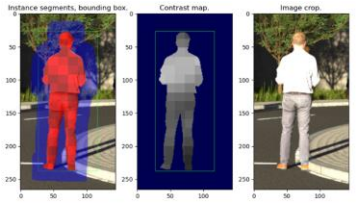
Low contrast or brightness

Input: Desired scenarios, ODD elements with attributes incl. parameter ranges as search criteria

An aligned ODD model & ontology is the basis for deriving data requirements & criteria for data coverage on all design levels.

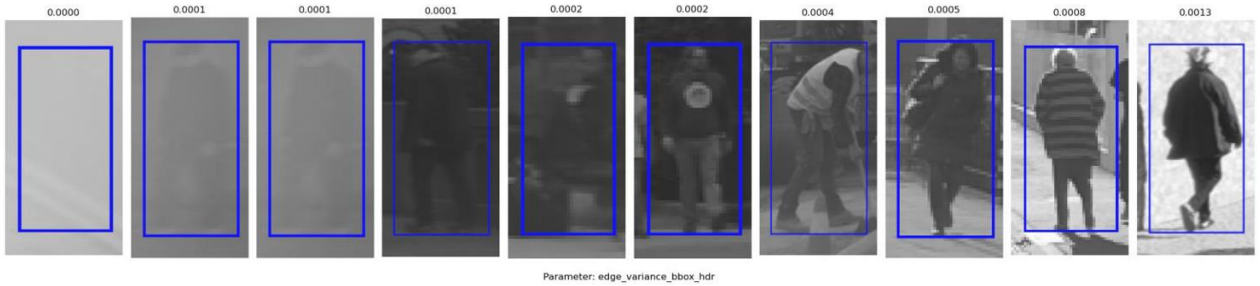
AD: Automated Driving DDE: Data-Driven Engineering DNN: Deep Neural Network FoV: Field of View GT: Ground Truth ML: Machine Learning (connectionistic ML – esp. DNNs) ODD: Operational Design Domain TL: Traffic Light

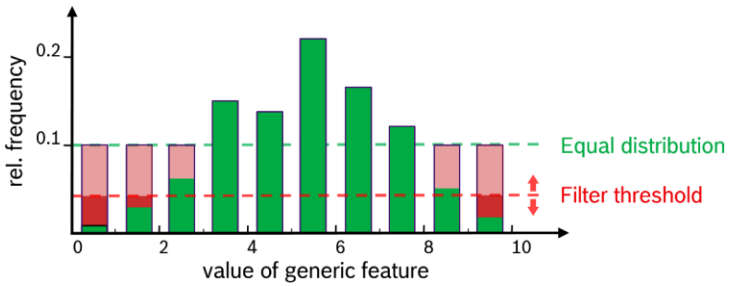
How to Specify Data Requirements in DDE



Example: RMS contrast along object contour as potential data feature candidate to be specified by a data requirement (focus on data contents)

Specify the relevant data feature space and cut it into discrete bins





Define a desired statistical distribution per data feature class or combinatorial class (e.g. min / max) to make it testable & find existing gaps

DDE Data Requirement

- Scenario class ID / data feature class ID
- **Qualitative** (data content-based) specification of ODD coverage => which scenario / data feature classes => derive filter criteria
- **What kind of scene/scenario** or data sample are we looking for?
- **Quantitative** specification of ODD coverage
- **How many frames or samples** of each scenario class resp. data feature class do we need?

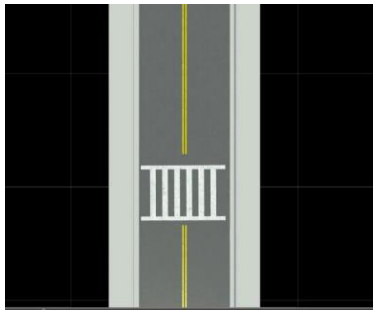
Data Set Design = set of data filter configurations (or configs for data augmentation / synthetization & simulation)

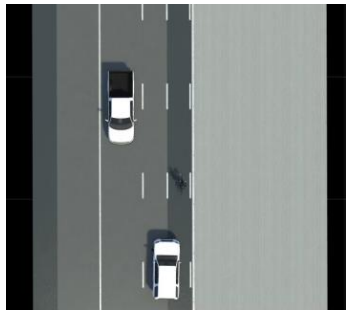
Verify data set / coverage with DDE Data Test Cases

Requirement-based data set implementation & verification is a key step that can be automated using configurable data filters.

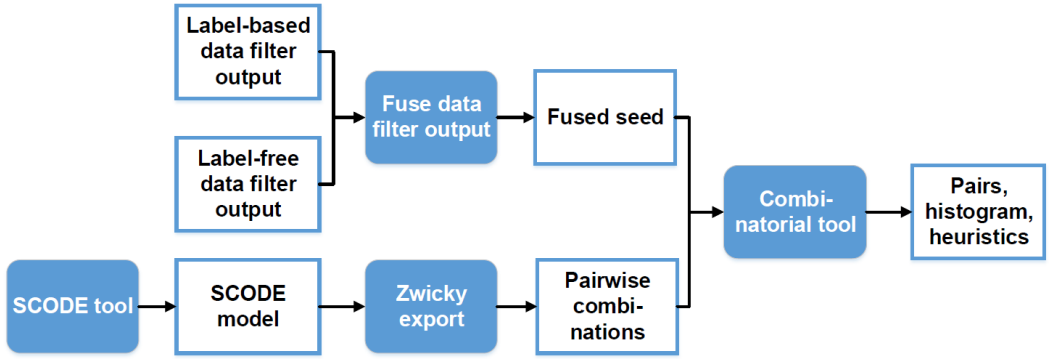
DDE: Data-Driven Engineering ID: identifier ODD: Operational Design Domain RMS: Root mean square
Software Val: (Cross-)Validation

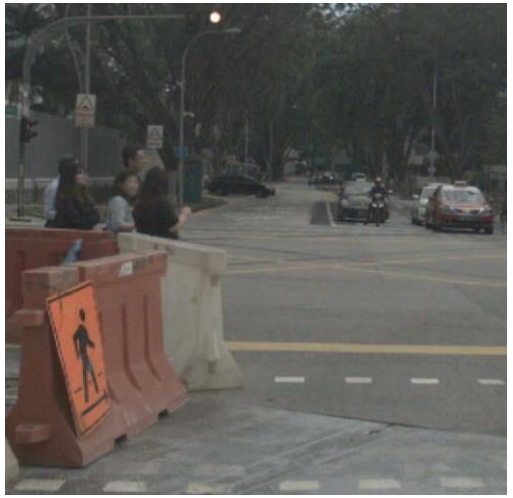
Define Scenario Variations using Combinatorial Methods



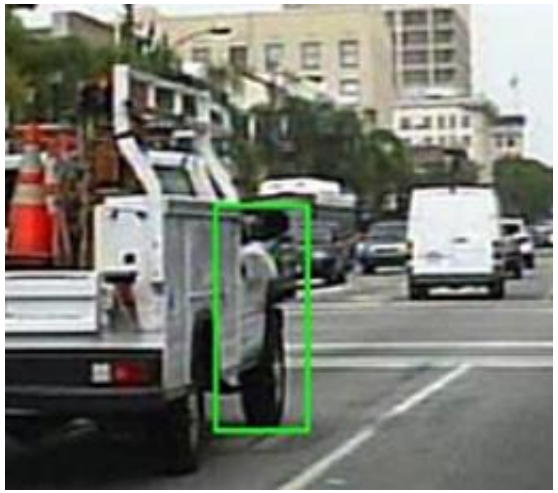


Each scenario in the catalogue may occur in different variants – safety-relevant aspects can be derived analytically by combinatorial methods.





FN VRU candidate, source: https://www.nuscenes.org/



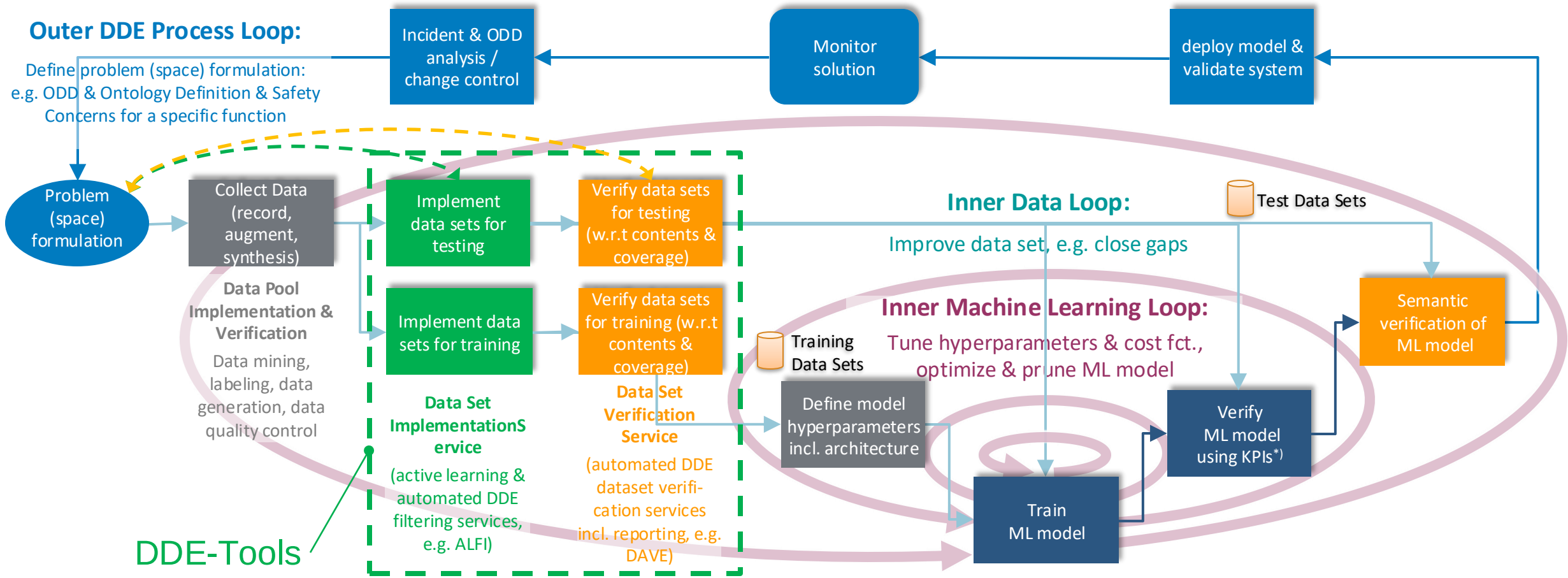
FP VRU candidate, source: https://www.frontiersin.org/articles/10.3389/fnbot.2018.00064/full

Data feature	Feature class combinations		
Longitudinal distance	Near	Short	Far
Lateral distance	Left lane	Ego lane	Right lane
Occlusion ratio	0…33%	33…66%	66…100%
Contrast	Low	Medium	High
Brightness	Dark	Medium	Bright

Systematically apply combinatorial methods to improve ODD coverage in important / safety-relevant data space regions.

BOSCH

What‘s in Machine Learning & Data-Driven Engineering Workflows

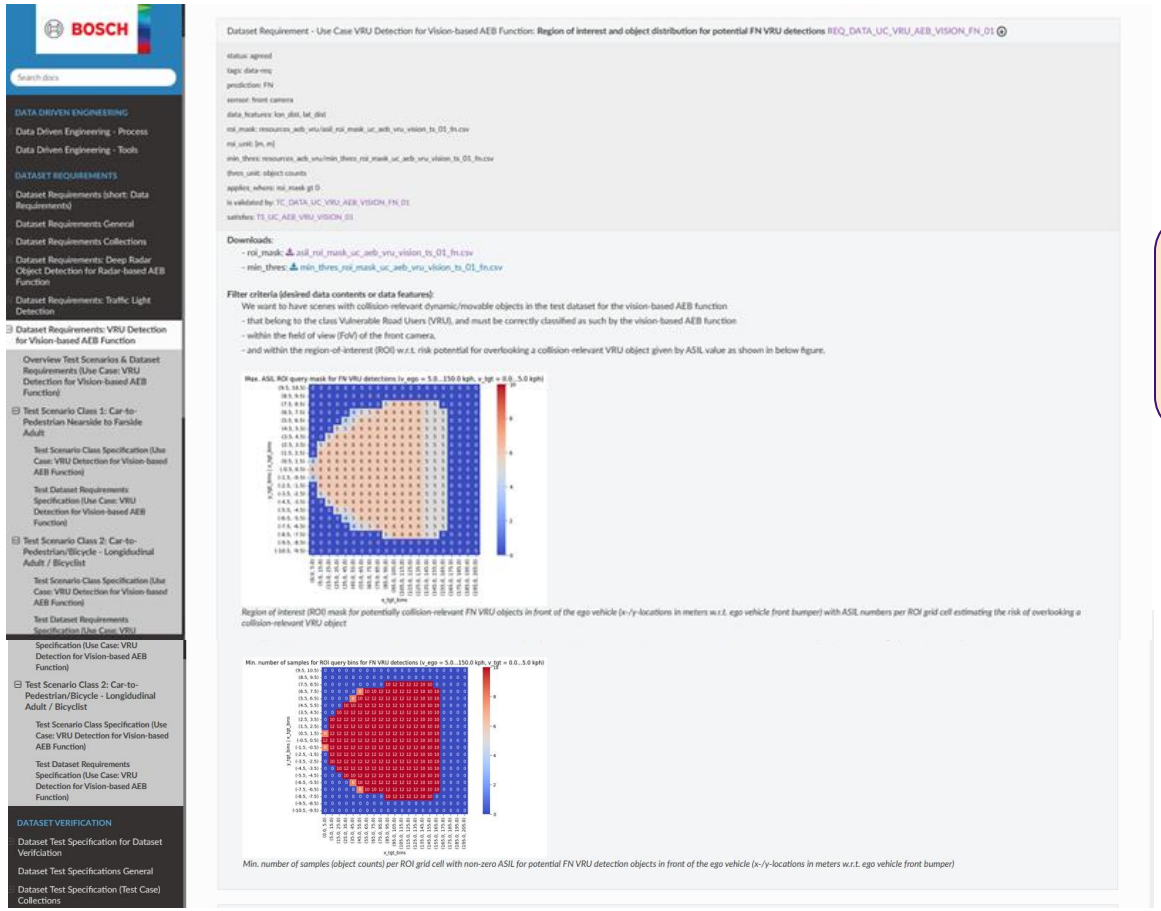


*) KPIs: metrics for certain characteristics applied to specifically compiled data sets

The DDE process closes an outer loop for systematic refinement of data requirements & development of verified ODD coverage.

AD: Automated Driving DDE: Data-Driven Engineering DNN: Deep Neural Network FoV: Field of View GT: Ground Truth ML: Machine Learning (connectionistic ML – esp. DNNs) ODD: Operational Design Domain TL: Traffic Light

Data Requirements on Safety-related Data Features or Properties



Reading data requirements from req. eng. tool

Define a set of scenarios & data requirements as refrence for a desired data set design

Requirements engineering tool (e.g. sphinx needs)

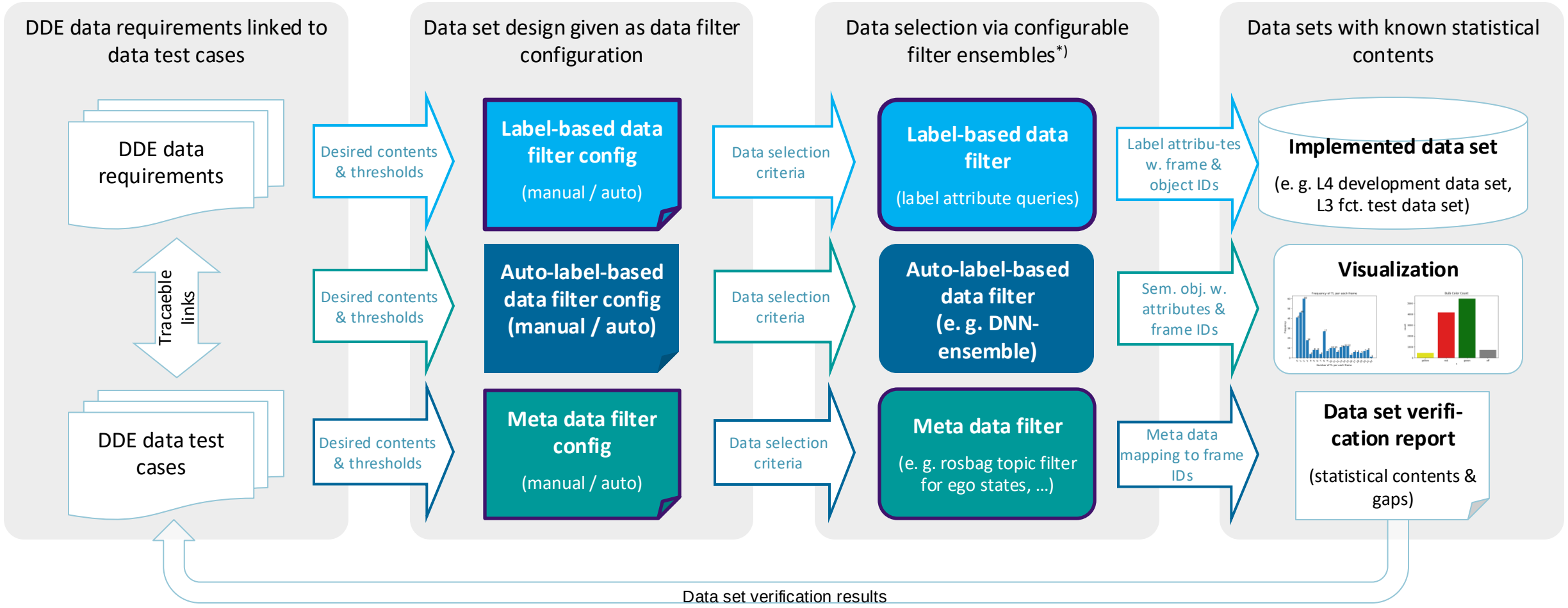
Read data requirements from req. eng. tool (e.g. via REST API or from needs.json url)

Filter configuration: For real data => Which data contents are required?

Test cases: For real data => How many samples do we need per feature class?

If we define data requirements in a machine-readable way we can automate the data set implementation & verification workflow.

Automated Data Set Implementation & Verification using Filter Ensembles

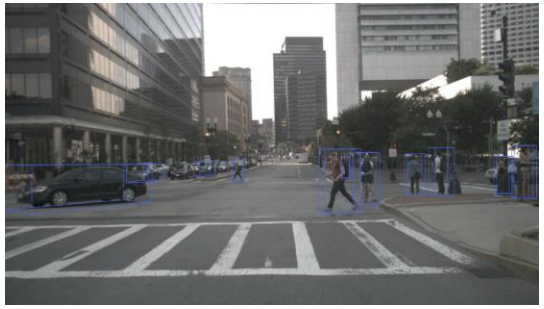


Configurable auto-/label-based filters are key enablers to automatically compile data sets tailored to data requirements

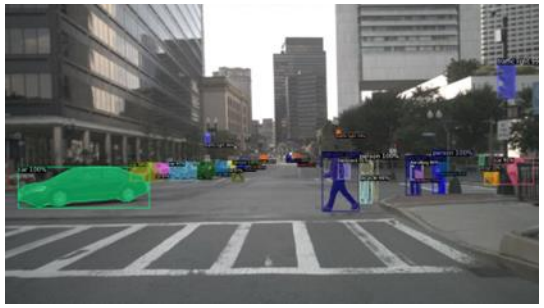
*) DDE data content filters are inspired from Business Values Assessment (BVA) ingest filter concept in Data Analytics Framework (DANF), but carry the idea further towards an autom. req.-based data set implementation & verification.

Available ALFI Task Modules (DNN Ensembles or Single Models)

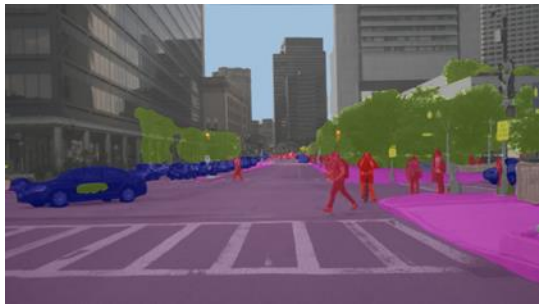
1. 3D object bounding boxes + distance & orientation



2./3. Instance segmentation mask + tight-fit 2D bboxes

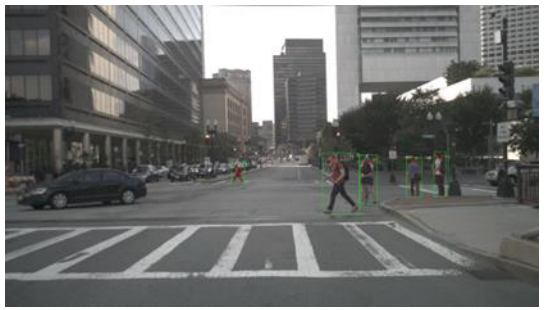


4. Semantic segmentation + free drivable space detection



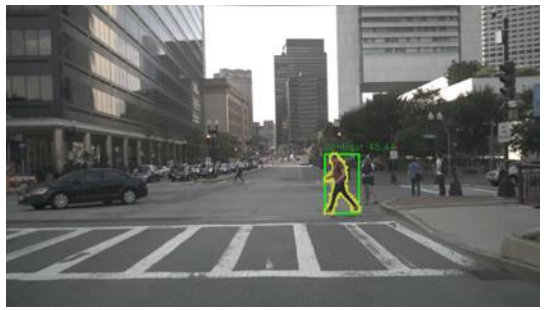
5. Depth map (relative depth)

6. Skeleton keypoint estimation for VRU



7. Occclusion detection (occluder & occludee + occlusion ratio)

8. Visual appearance meta data

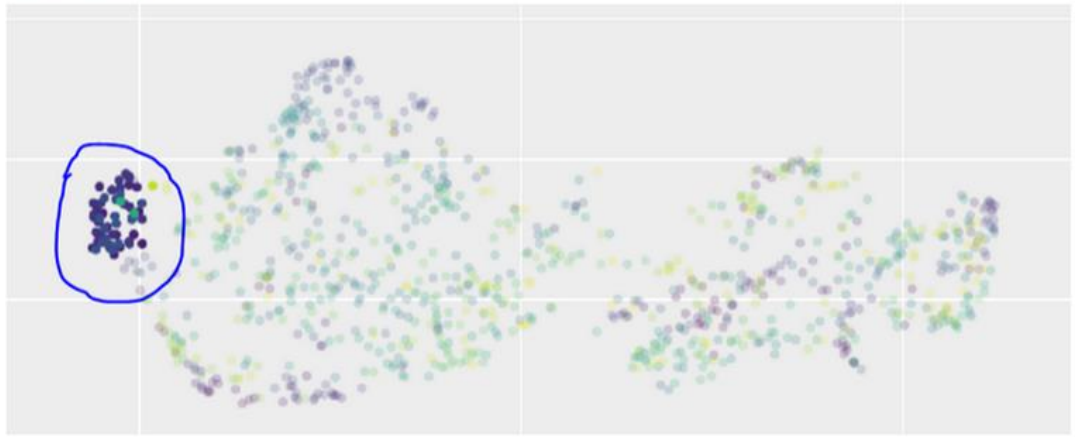


9. Mono camera distance estimation using calibration data (long./lat. dist.)

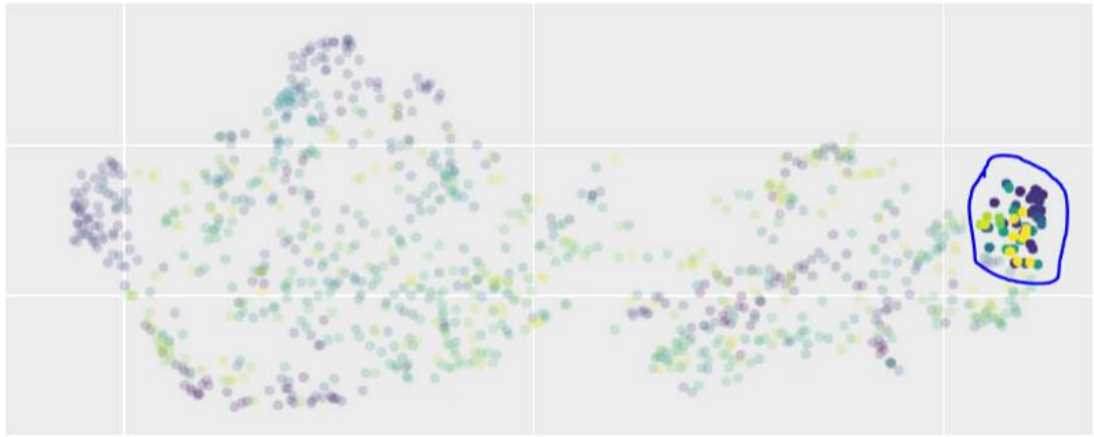
9 / 10 ALFI modules are available in dockerized form and partially as AML ensemble (truncation detection is not available yet)

Embedded Feature Space Dataset Analysis using t-sne

next selected region:



and one more region:



images in selected region:

images inside selected region:

Using a combined embedded feature & semantic feature space analysis allows us to merge both human & model perspective.

Labeler Apps

- Label ground truth for image, video, lidar and medical data

- Important for training networks for:
  - Classifiers
  - Object Detectors
  - Segmentation

- Features:
  - Create label definitions and attributes.
  - Semi automated or automated labeling with built-in or custom algorithms
  - Blocked processing support (image)
  - Superpixel automation (Image, Video)

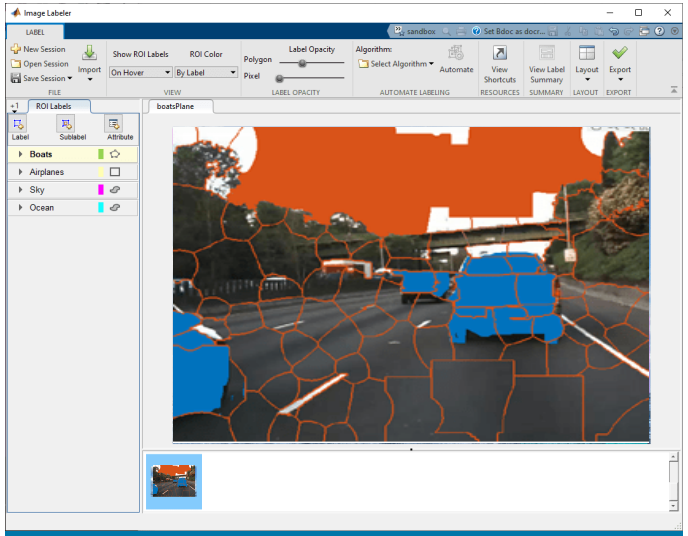
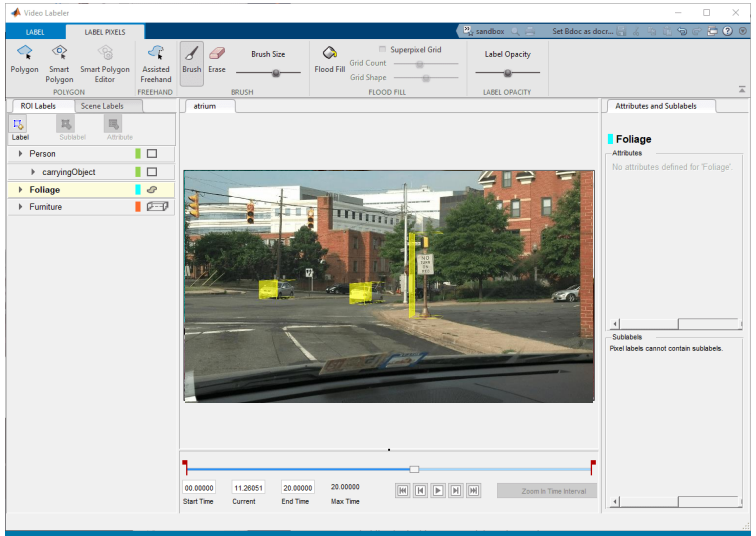
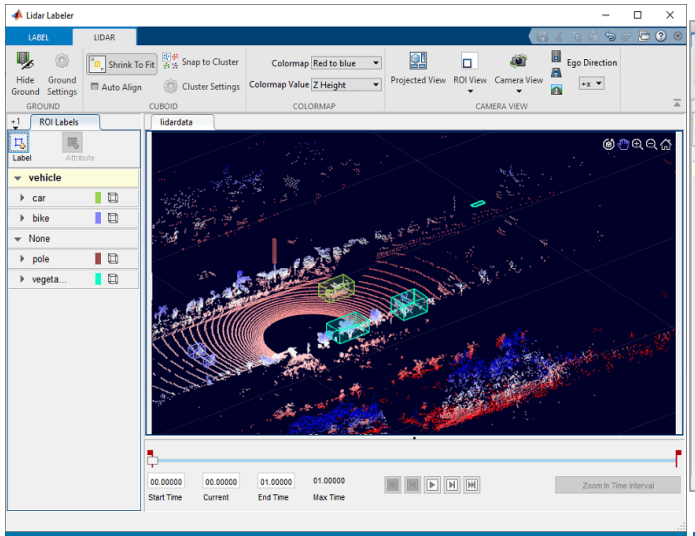


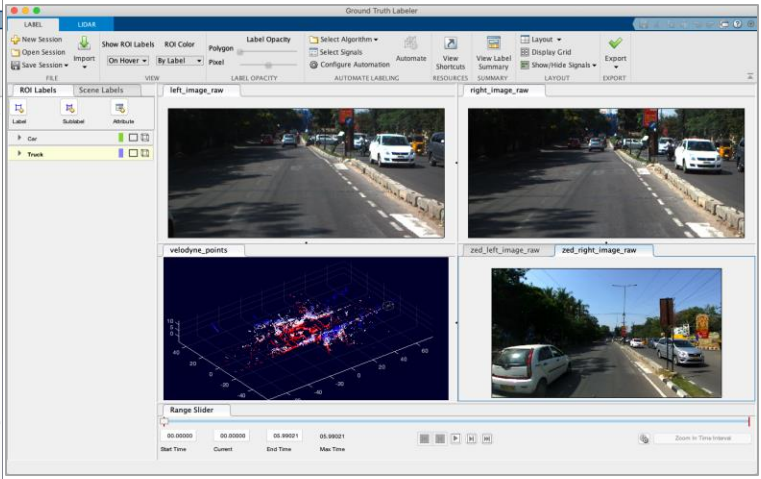
Image Labeler (Computer Vision Toolbox)



Video Labeler (Computer Vision Toolbox)



Lidar Labeler (Lidar Toolbox)



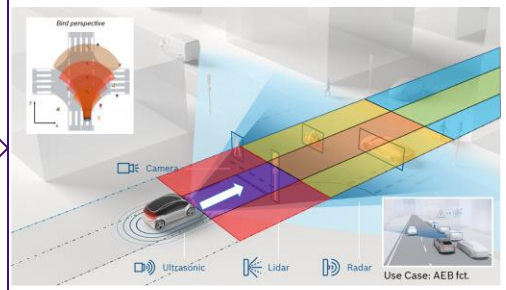
Ground Truth Labeler (Automated Driving Toolbox)

- **Convert Image Labeler ground truth data to OpenLABEL Format**

Example how to Design & Verify AI Test Datasets using DDE

System & ODD & UC analysis

Test scenario catalogue

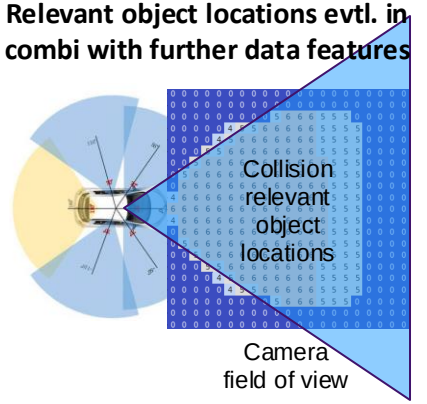


Hazard & risk analysis

Collision simulation

Requirements elicitation

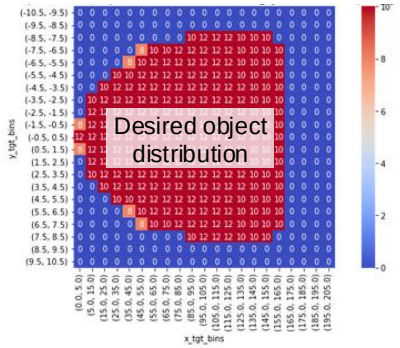
Relevant object locations evtl. in combi with further data features



AI requirements elicitation

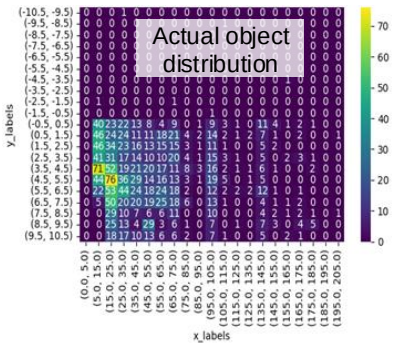
AI dataset requirements

Region of interest & desired object distribution (BEV)



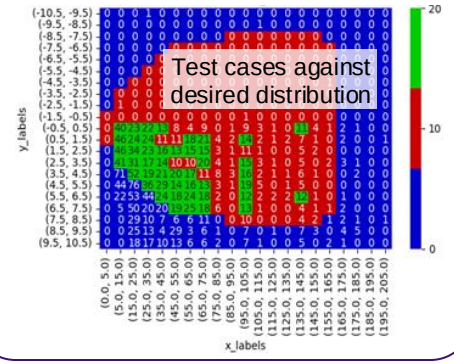
Query against auto-/label database

Dataset implementation by requirement-based filtering



Desired vs. actual distribut.

Datase verification incl. coverage & gap analysis



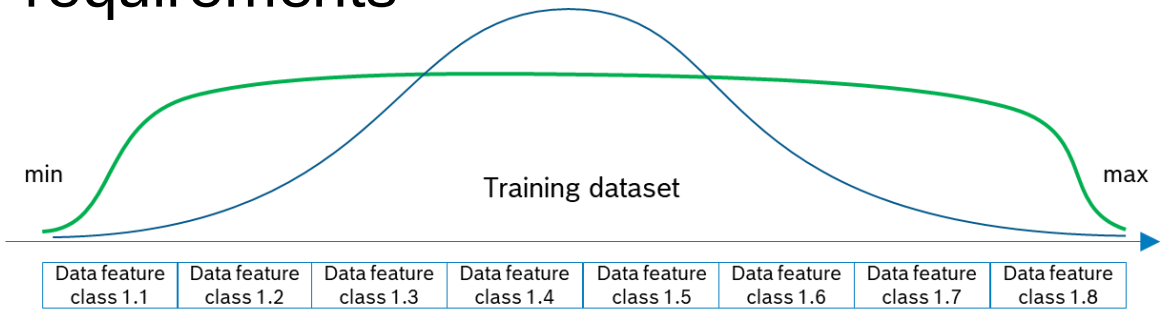
AI perform. eval. vs. data features

We need quality control for our AI datasets w.r.t. data coverage, contents, diversity, balance & integrity – not only label quality!

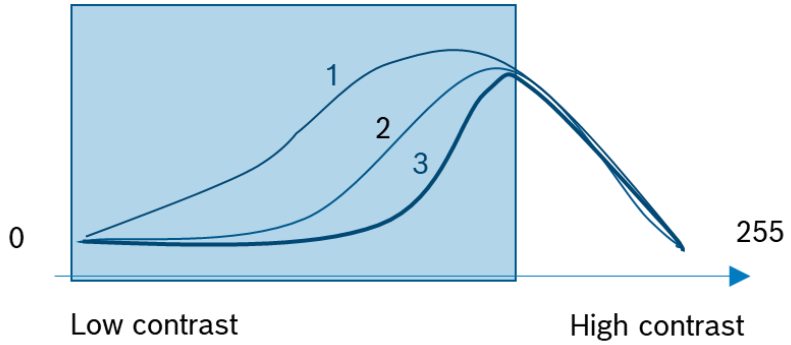
AI: Artificial Intelligence BEV: Bird‘s eye view DDE: Data Driven Engineering FoV: Field of View GT: Ground Truth ML: Machine Learning ODD: Operational Design Domain ROI: Region-of-interest

Goal: Tailor Datasets Content-wise against Dataset Requirements

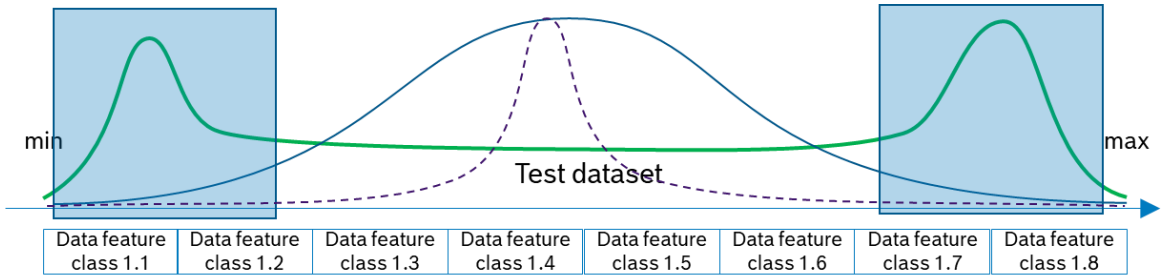
Tailor training datasets against data requirements



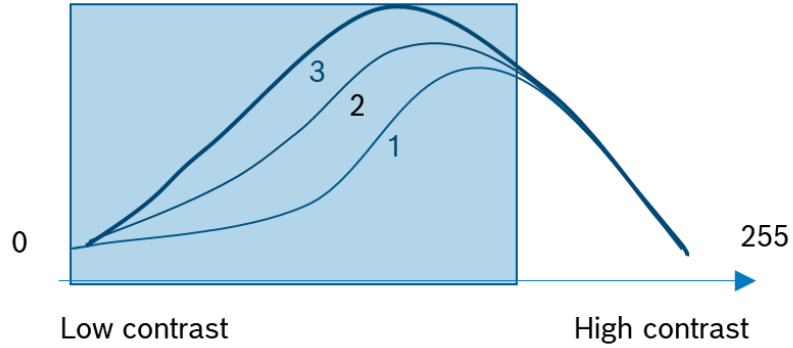
Reduce training / test data coverage in certain regions



Tailor test datasets against data requirements

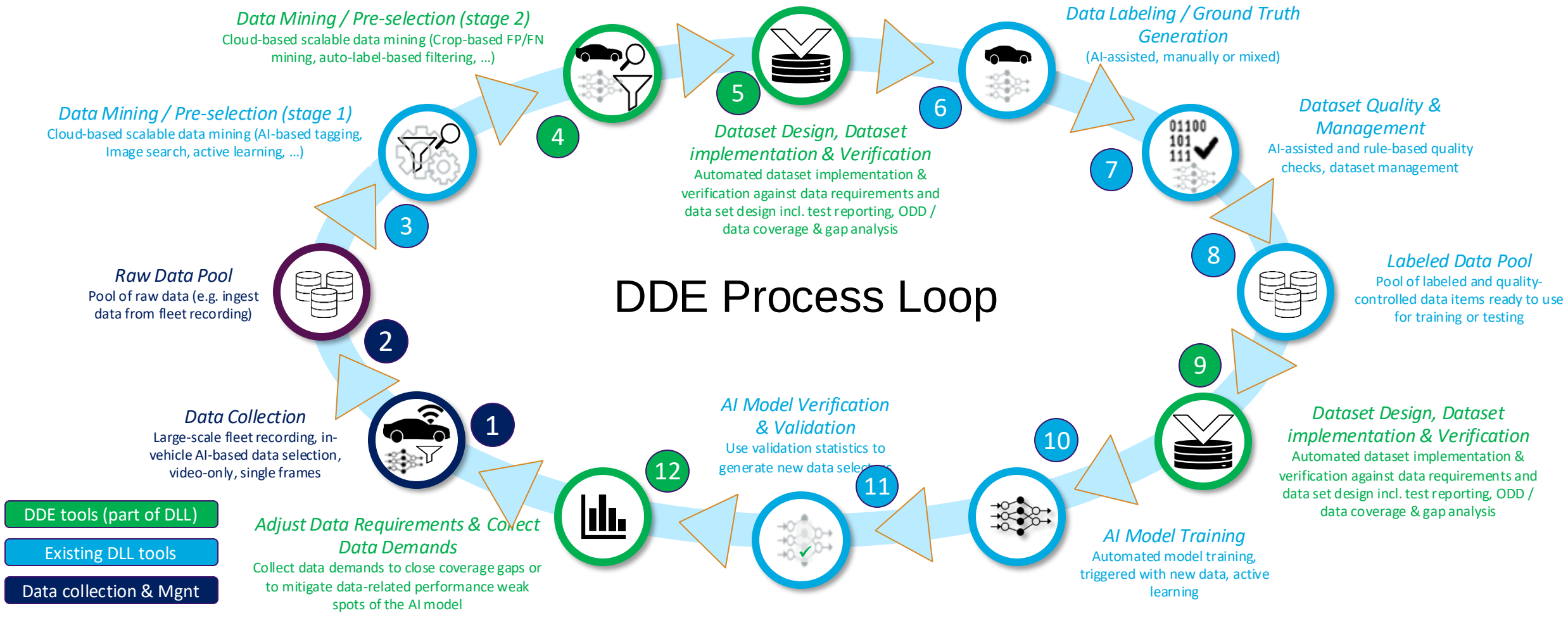


Increase training / test data coverage in certain regions



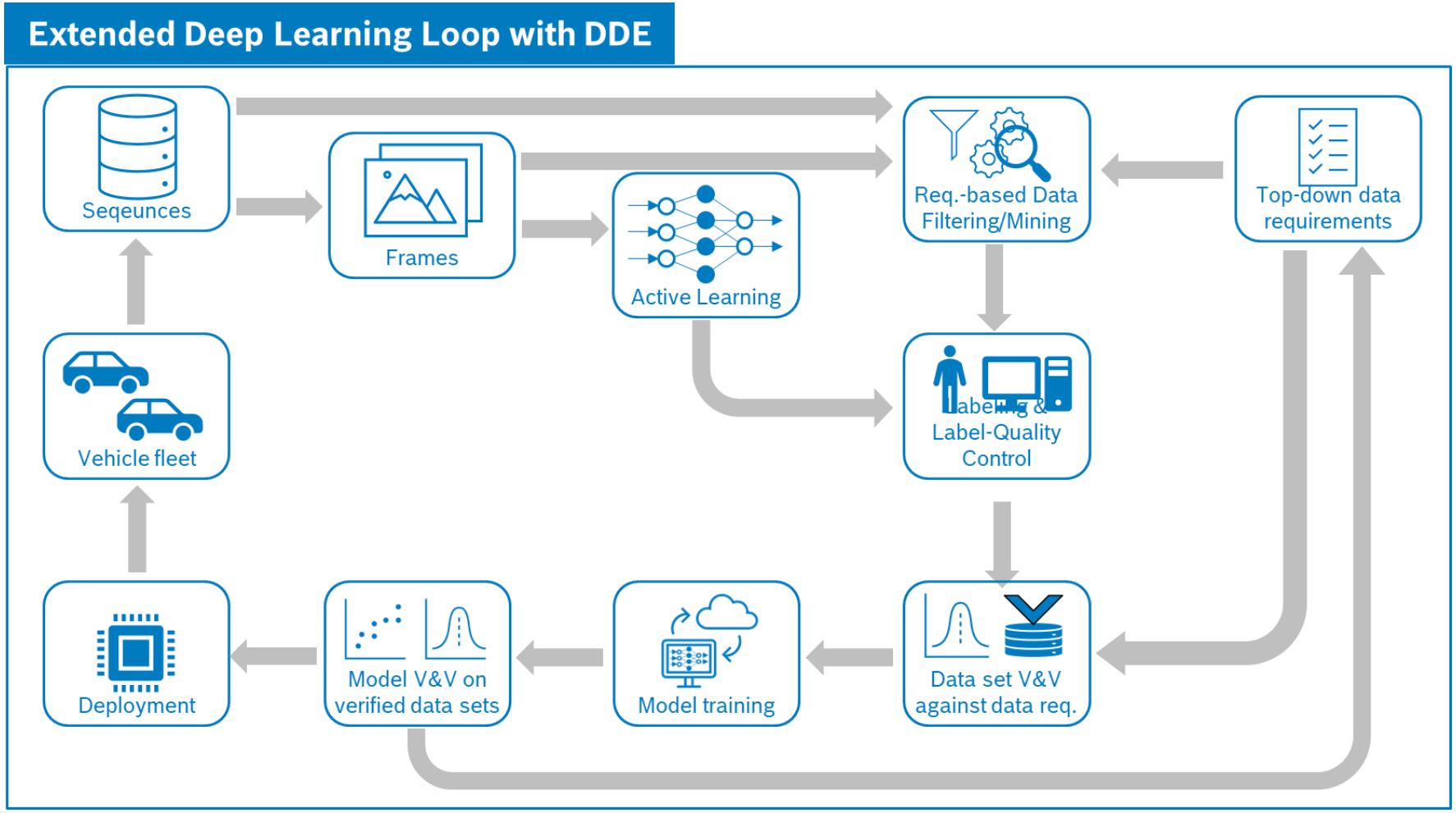
DDE-tools like ALFI & DAVE allow tailoring & verification of real datasets w.r.t. contents/features against dataset requirements.

Interactions of DDE Process & DLL Tools along the Process Flow



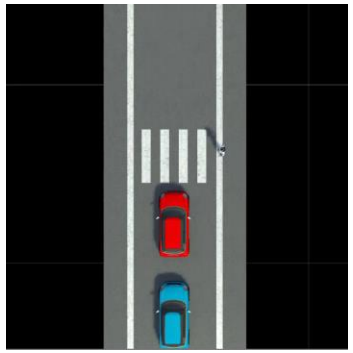
A mini DDE loop has been closed including the steps 3, 4, 5, 9, 10, 11, 12, which we can operate to execute training experiments.

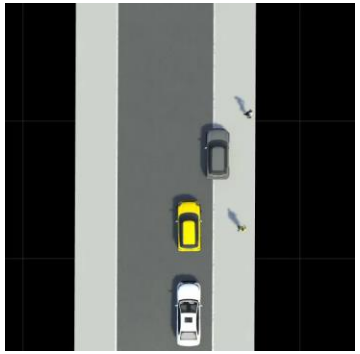
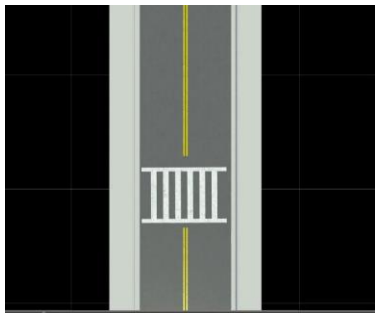
Combine Data-Driven Engineering & Active Learning synergetically

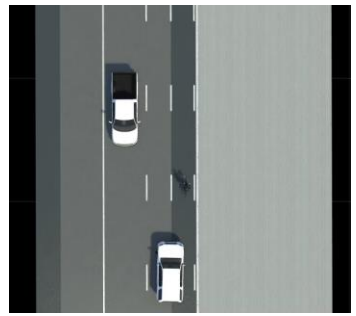


Data-Driven Engineering (top-down view) & Active Learning (bottom-up view) complement one another in a deep learning loop.

Outlook: Scenario Class Coverage given a Scenario Catalogue





...

Scenario ID 00X-000
Class name: …
Config file: …
Description: …
Parameters: …
Data Coverage: …

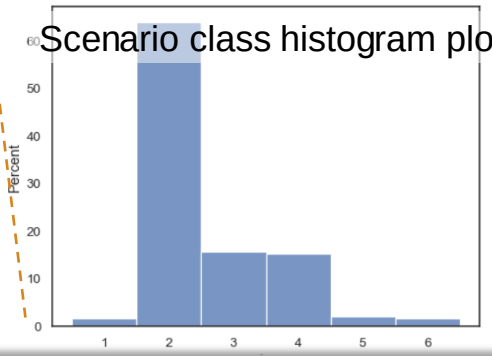
Scenario ID 00X-001
Class name: …
Config file: …
Description: …
Parameters: …

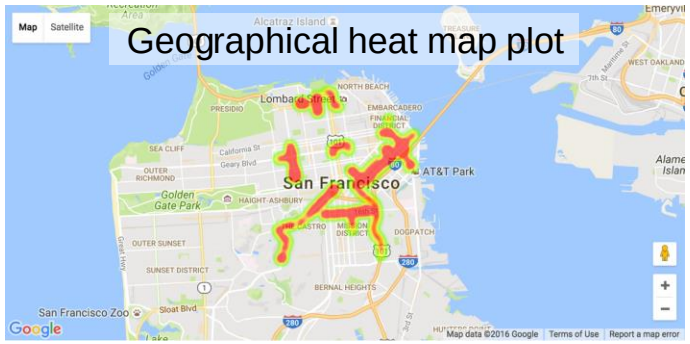
Scenario ID 00X-002
Class name: …
Config file: …
Description: …
Parameters: …

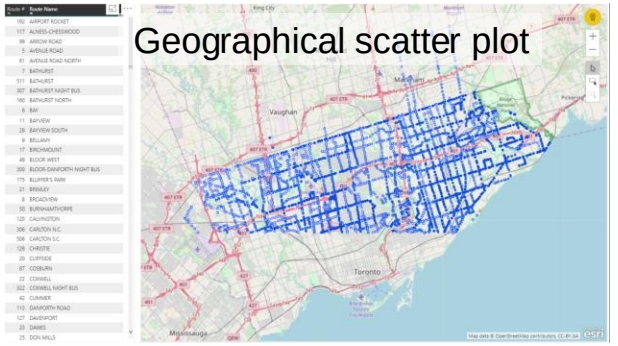
Scenario ID 00X-003
Class name: …
Config file: …
Description: …
Parameters: …

Scenario ID 00X-004
Class name: …
Config file: …
Description: …
Parameters: …

Scenario ID 00X-005
Class name: …
Config file: …
Description: …
Parameters: …







...

We cut our Operational Design Domain (ODD) into a collection of abstract scenario classes for which we doe a safety analysis.

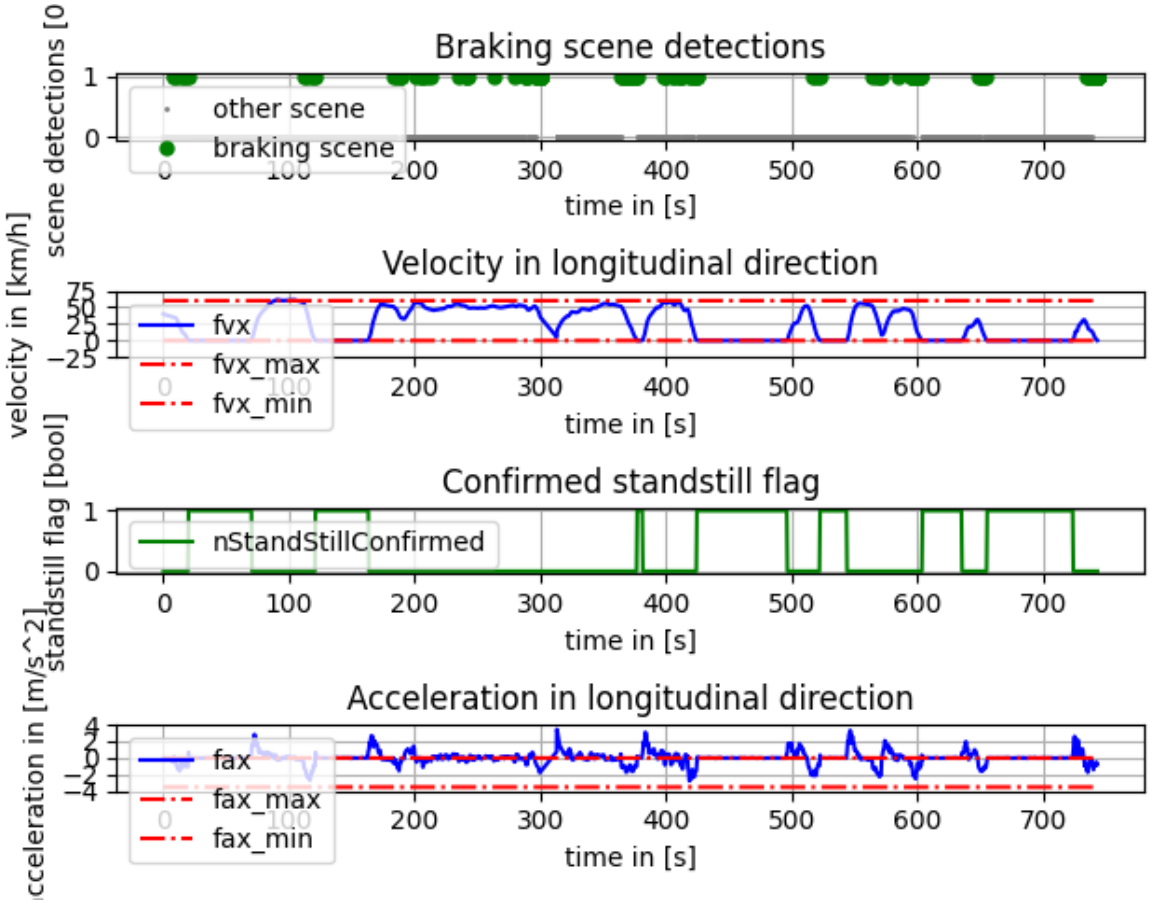
Outlook: Classify Scenarios using Sequential Multi-Sensor Data

Framewise detect braking scenes with lead vehicle

Configurable filter parameters:

- Object classes, lane bins, distance bins w. tolerance, driving state min/max thresholds with tolerance

Detect braking scenes with lead vehicle over time



For scenario identification we need to look into (multi-sensor) sequential data considering spatio-temporal effects.

2024 MathWorks
中国汽车年会

Thank you

MathWorks®