KI ABSICHERUNG

Safe AI for Automated Driving

October 11 2021, ITS World Congress

KI Absicherung

Lydia Gauerhof, Robert Bosch GmbH

1

Project vision and goals

Making the safety of AI-based function modules for highly automated driving verifiable
KI ABSICHERUNG
Safe AI for Automated Driving
Pedestrian detection

Challenge

AI Land

Promising new technology with unimagined possibilities

Established safety processes cannot be applied

KI ABSICHERUNG
Safe AI for Automated Driving

Safety Land

Safe, trustworthy driving function

Consensus in industry and academia on the Safety of AI

2

Methodological and conceptual approach

KI ABSICHERUNG

Safe AI for Automated Driving

Provide the AI functions for pedestrian detection

AI Function-Pedestrian detection

Function

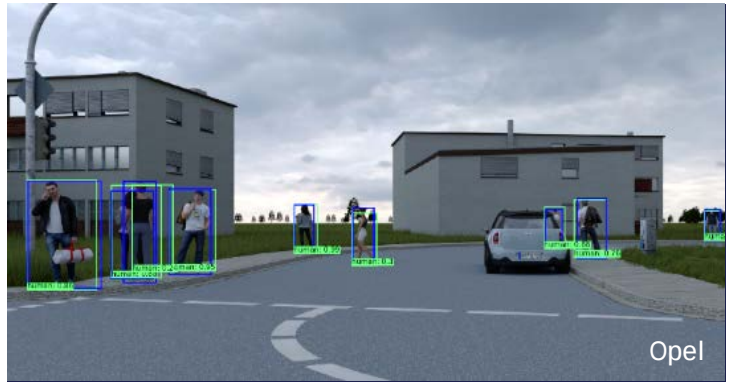
Environment

System architecture

AI Function

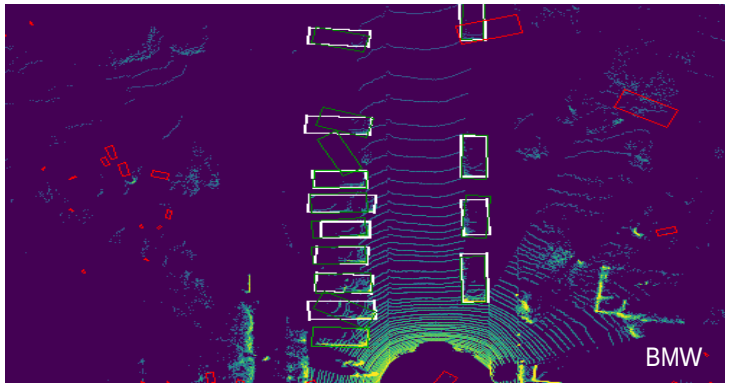
Semantic Segmentation

2D Bounding Box Detection

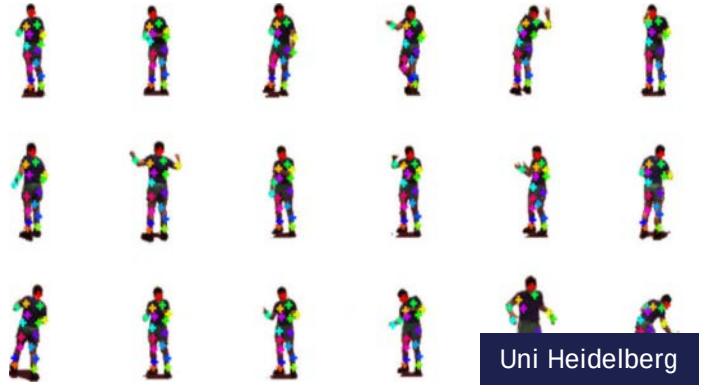


Instance Segmentation

3D Bounding Box Detection



3D Pose estimation



KI ABSICHERUNG

Safe AI for Automated Driving

- **Provide the AI functions for pedestrian detection**
- **Generate synthetic data**

ML-Lifecycle-Validation data

ML LIFECYCLE

- ML-Data
- Model design
- Pre-Processing
- Training
- Post-Processing
- **Test/ V&V incl. Daten**
- Monitoring
- Maintenance

KI ABSICHERUNG

Safe AI for Automated Driving

- Provide the AI functions for pedestrian detection
- Generate synthetic data
- Develop methods and measures

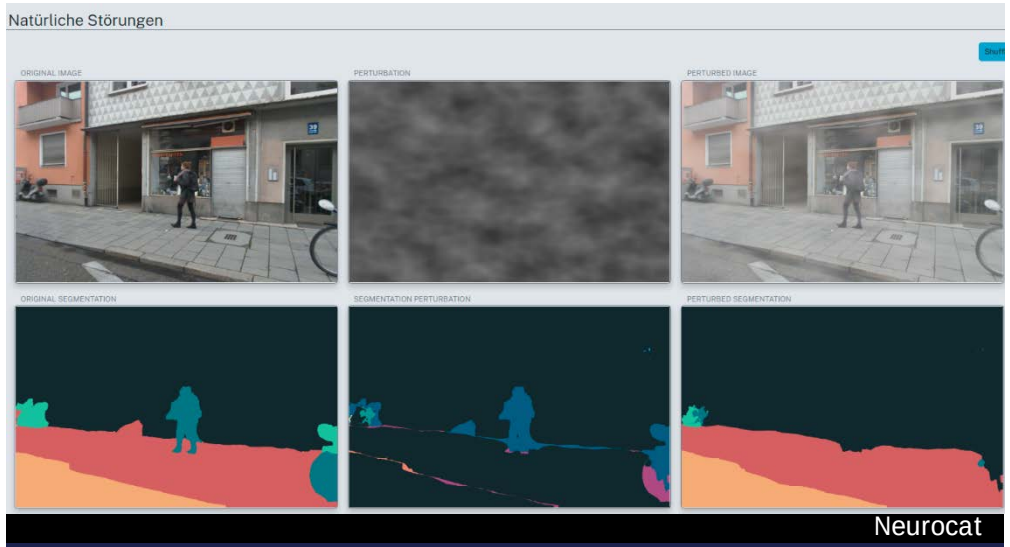
Identify, Measure and & Counteract „DNN-specific Safety Concerns”

DNN-specific safety concern:

- False positive / negative: Pedestrian detection is incorrect resp. not robust enough

Method:

- Assessment of uncertainty: Stochastic evaluation of a multitude of model variations (Monte Carlo Dropout)



DNN-specific safety concern:

- Adversarial error: leads to overlooking (false negative)

Method:

- Systematic analysis of adversarial errors
- Methods to evaluate adversarial resilience
- Counter mechanisms during training or operation

KI ABSICHERUNG

Safe AI for Automated Driving

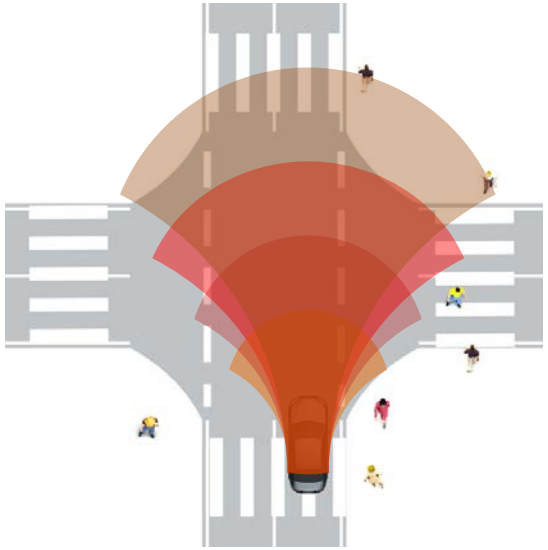
- Provide the AI functions for pedestrian detection
- Generate synthetic data
- Develop methods and measures
- Establish an overall safety argument for the AI function

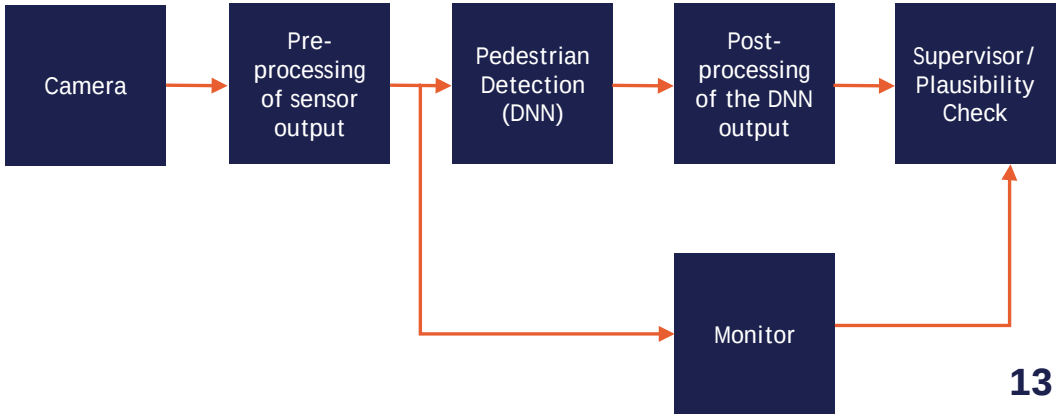
Specification

Function

Environment

System Architecture





Our Approach: Establishment of a Holistic Safety strategy

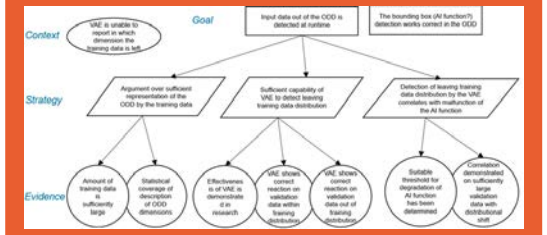
Identification: DNN specific insufficiencies & "safety concerns"

Quality metric and data: Methods for the evaluation

Countermeasure: to reduce the remaining risk

Effectiveness: Contribution to safety? Yes / no

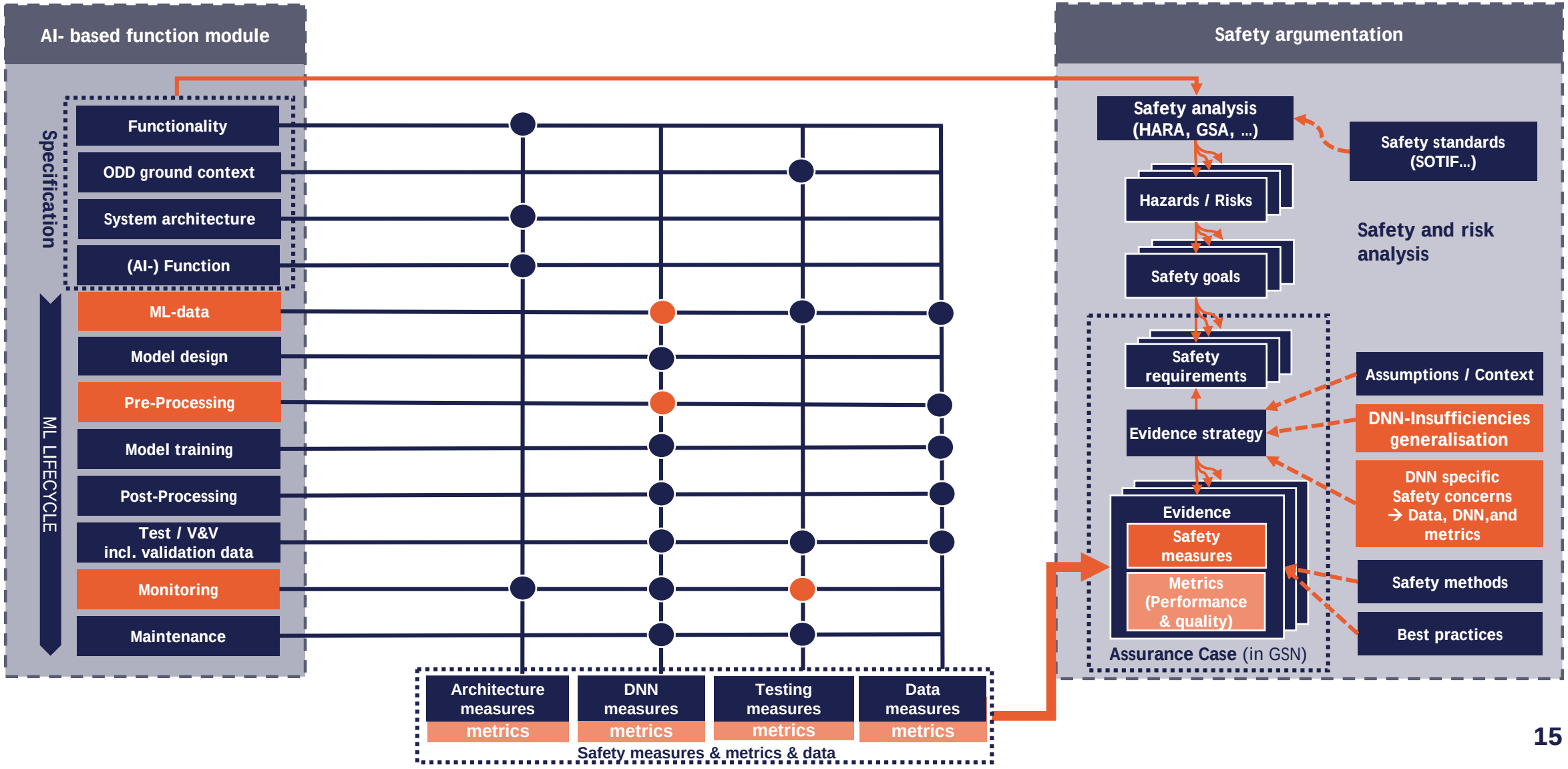
Safety Argument: GSN-based Notation



Structured safety argument

Assurance Case

AI Specific Evidence-Based Safety Argumentation



KI ABSICHERUNG

Safe AI for Automated Driving

- Provide the AI functions for pedestrian detection
- Generate synthetic data
- Develop methods and measures
- Establish an overall safety argument for the AI function
- Communicate the results

KI ABSICHERUNG

Safe AI for Automated Driving

- Provide the AI functions for pedestrian detection
- Generate synthetic data
- Develop methods and measures
- Establish an overall safety argument for the AI function
- Communicate the results

KI ABSICHERUNG

Safe AI for Automated Driving

Project coordination: Dr. Stephan Scholz | Volkswagen AG

Co-lead and scientific coordination: PD. Dr. Michael Mock | Fraunhofer IAIS

ki-absicherung-konsortialfuehrung@eict.de

KI Absicherung is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Energy.

KI FAMILIE

Supported by:

Federal Ministry for Economic Affairs and Energy

on the basis of a decision by the German Bundestag

www.ki-absicherung-projekt.de @KI_Familie KI Familie