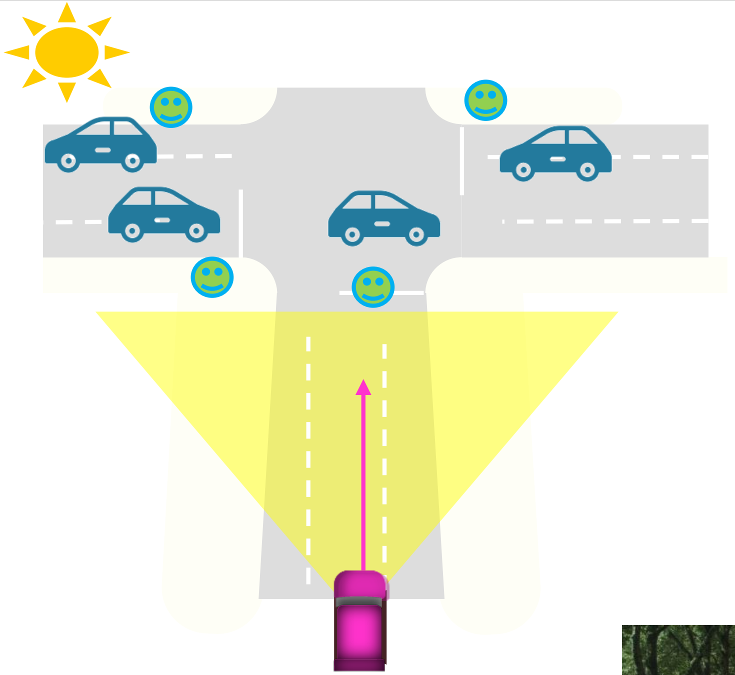
Technische Universität Braunschweig

Institut für Nachrichtentechnik

Corner Cases and Occlusions in Environment Perception for Automated Driving

Jasmin Breitenstein
Institute for Communications Technology
Technische Universität Braunschweig

1. Motivation



Aim: Reliable perception for highly automated driving

- The ego-vehicle drives without a driver being able to intervene
- Easy example: no other traffic participants
  …this seems doable, right?

BUT we know: the tasks becomes harder as other traffic participants crowd the scenario

- Many datasets exist that allow to train and evaluate perception methods for this case:

Cityscapes

BDD100k

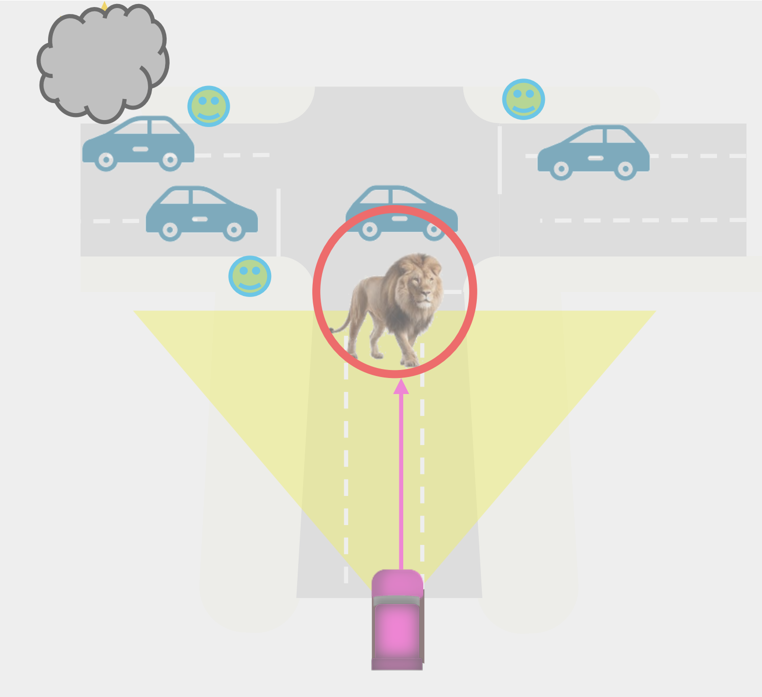
KITTI

[M. Cordts et al., “The Cityscapes Dataset for Semantic Urban Scene Understanding”, in Proc. of CVPR, 2016]

[F. Yu et al., “BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning”, in Proc. of CVPR, 2020]

[A. Geiger et al., “Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite”, in Proc. of CVPR, 2012]

1. Motivation



..but how about the lion?
- The lion will be recognized as a known class

…what happens when it is foggy?
- Perception performance will likely degrade considerably

…what happens when there is an accident?
- The recognition of the turned over car will degrade

Problem:
- Perception methods only know what was contained in the training data
- Which corner cases are there?
- How to treat so-called corner cases?

What do we need corner case detection for:
- Select suitable data for training and evaluation during development
- Recognize corner cases online while they happen

Outline

Problem:

There exists an infinite amount of situations that can possibly occur while driving!

Training data cannot cover all critical situations.

⇒ We need to treat such unexpected and possibly dangerous situations safely.

⇒ We need reliable corner case detectors.

First definition [Bolte et al.,2019]:

A corner case is given, if there is a non-predictable relevant object/class in a relevant location.

Extension of this definition:

A more detailed definition is necessary to facilitate specific and targeted development of corner case detectors.

⇒ We need to introduce structure into this infinite possible amount of corner cases.

[J.-A. Bolte, A. Bär, D. Lipinski, T. Fingscheidt, “Towards Corner Case Detection for Autonomous Driving”, in Proc. of IV, 2019]

2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence. Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects*, but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object*, but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance*, but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, "Systematization of Corner Cases for Visual Perception in Automated Driving", in Proc. of IV, 2020]

2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects,* but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object,* but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance,* but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

overexposure in the scene makes perception harder

Technische Universität Braunschweig

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, "Systematization of Corner Cases for Visual Perception in Automated Driving", in Proc. of IV, 2020]

[J. Geyer et al., "A2D2: Audi Autonomous Driving Dataset", arXiv, 2020]

Institut für Nachrichtentechnik

2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects,* but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object,* but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance,* but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

sticker in window hinders perception on the left

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, “Systematization of Corner Cases for Visual Perception in Automated Driving”, in Proc. of IV, 2020]

[F. Yu et al., “BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning”, in Proc. of CVPR, 2020]

2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects*, but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object*, but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance*, but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

snowy streets can be an appearance shift

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, “Systematization of Corner Cases for Visual Perception in Automated Driving”, in Proc. of IV, 2020]

[F. Yu et al., “BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning”, in Proc. of CVPR, 2020]

2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence. Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects*, but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object*, but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance*, but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

dog on the street is unknown

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, "Systematization of Corner Cases for Visual Perception in Automated Driving", in Proc. of IV, 2020]

[K. Li et al., "CODA: A Real-World Road Corner Case Dataset for Object Detection in Autonomous Driving", in Proc. of ECCV, 2022]

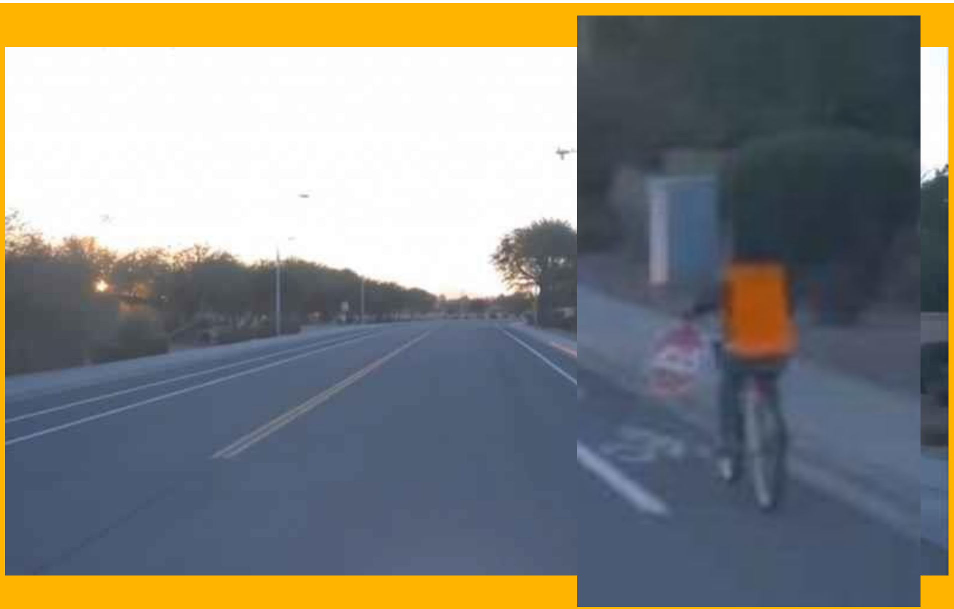
2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects,* but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object,* but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance,* but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases



stop sign carried by rider

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, “Systematization of Corner Cases for Visual Perception in Automated Driving”, in Proc. of IV, 2020]

[D. Anguelov, “Taming the Long-tail of Autonomous Driving Challenges”, 2019]

2. Corner Case Definitions
Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence. Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects,* but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object,* but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance,* but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

unseen crowd of people

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, “Systematization of Corner Cases for Visual Perception in Automated Driving”, in Proc. of IV, 2020]
[K. Li et al., “CODA: A Real-World Road Corner Case Dataset for Object Detection in Autonomous Driving”, in Proc. of ECCV, 2022]

2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects*, but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object*, but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance*, but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

Overtaking a cyclist

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, "Systematization of Corner Cases for Visual Perception in Automated Driving", in Proc. of IV, 2020]

2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level	Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence Recognition requires scene understanding	**Anomalous Scenario** Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly** *Multiple known objects*, but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly A *known object*, but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)** An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift** A large, constant *shift in appearance*, but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier** *One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier *All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

cars making way for an emergency vehicle

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, "Systematization of Corner Cases for Visual Perception in Automated Driving", in Proc. of IV, 2020]

[Video source: EnjoyFirefighting - International Emergency Response Videos]

2. Corner Case Definitions

Systematization of Corner Cases: Camera Sensor

Corner Cases

Level		Description:	Examples:
Scenario Level Patterns are observed during the course of an image sequence. Recognition requires scene understanding	**Anomalous Scenario**	Pattern that *was not observed* during the training process and has *high potential for collision*	• *Person suddenly walking onto the street* • *Car accident*
	Novel Scenario	Pattern that *was not observed* during the training process, but *does not increase the potential for collision*	• *Truck appears from a side road (but is going to stop)* • *Accessing the freeway*
	Risky Scenario	Pattern that *was observed* during the training process, but *still contains potential for collision*	• *A car is coming towards me. (Potentially short time to collision)* • *Overtaking a cyclist*
Scene Level Non-conformity with expected patterns in a single image	**Collective Anomaly**	*Multiple known objects,* but in an *unseen quantity*	• *Demonstration, e.g., critical mass ride* • *Traffic jam*
	Contextual Anomaly	A *known object,* but in an *unusual location*	• *Tree on the street* • *Barrier, e.g., a fence on the street*
Object Level Instances that have not been seen before	**Single-Point Anomaly (Novelty)**	An *unknown object*	• *Bear, tiger, etc.* • *Lost objects* • *Rollator*
Domain Level World model fails to explain observations	**Domain Shift**	A large, constant *shift in appearance,* but *not in semantics*	• *Weather conditions, rain, fog, snow* • *Traffic sign appearance* • *Location (Europe – U.S.A.)*
Pixel Level (Perceived) errors in data	**Local Outlier**	*One or few pixels* fall outside of the expected range of measurement	• *Pixel errors (dead pixels)* • *Dirt on the windshield*
	Global Outlier	*All or many pixels* fall outside of the expected range of measurement	• *Lighting conditions* • *Overexposure*

Corner Case detection complexity increases

car accident seen from the ego-vehicle

[J. Breitenstein, J.-A. Termöhlen, D. Lipinski, T. Fingscheidt, “Systematization of Corner Cases for Visual Perception in Automated Driving”, in Proc. of IV, 2020]

[Video source: Youtube, Wham Baam Dashcam]

2. Corner Case Definitions
Systematization of Corner Cases: Extension to Other Sensors

Systematization of corner cases extended to LiDAR and RADAR:

	Sensor Layer		Content Layer			Temporal Layer
	Hardware Level	**Physical Level**	**Domain Level**	**Object Level**	**Scene Level**	**Scenario Level**
LiDAR-based corner cases	Laser Error • *Broken mirror* • *Misaligned actuator*	Beam-Based Corner Case • *Black cars disappear* • ...	Domain Shift on Single Point Cloud • *Shape of Road markings*	Single-Point Anomaly on Single Point Cloud • *Dust cloud* • ...	Contextual/Collective Anomaly on Single Point Cloud • *Sweeper cleaning the sidewalk*	Corner Cases on Multiple Point Clouds and Frames • *Person breaks traffic rule* • *Overtaking a* • ...
Camera-based corner cases	Pixel Error	Pixel-Based Corner Case	Domain Shift on Single Frame	Single-Point Anomaly on Single Frame	Contextual/Collective Anomaly on Single Frame	
RADAR-based corner cases	Impulse Error • *Low voltage* • *Low temperature*	Impulse-Based Corner Case • *Interference* • ...	Domain Shift on Single Point Cloud • *Weather, e.g., snow, rain, etc.*	Single-Point Anomaly on Single Point Cloud • *Lost objects* • ...	Contextual/Collective Anomaly on Single Point Cloud • *Demonstration* • *Tree on street*	

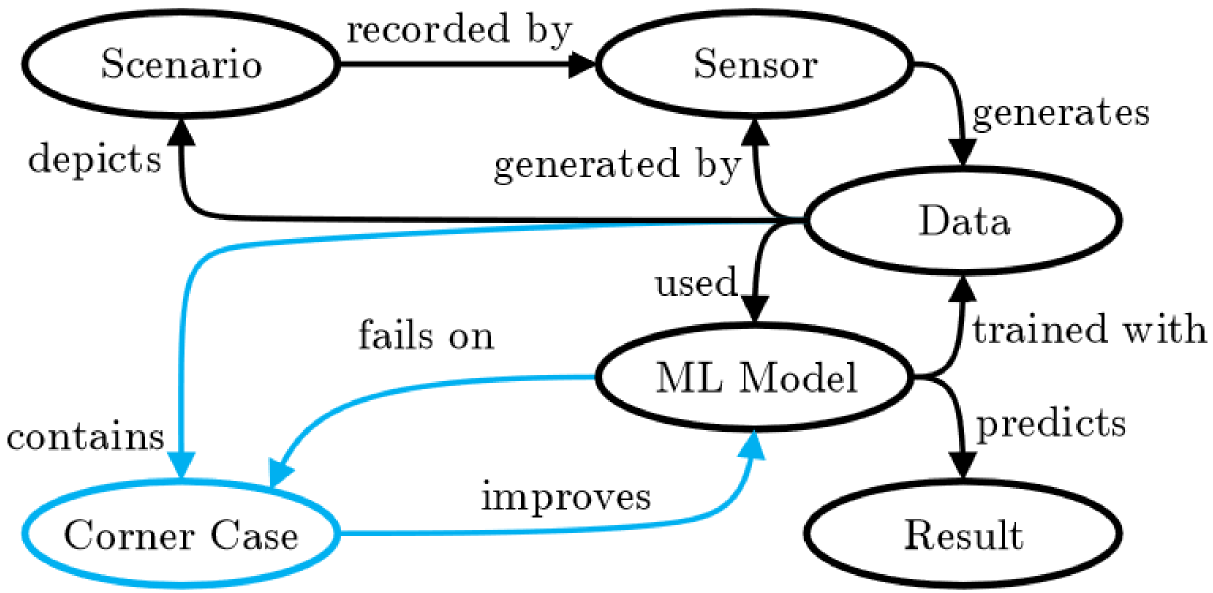
Camera sensor corner cases according to previous systematization

- Corner case levels can be extended to LiDAR and RADAR
- Additional method layer: Corner cases that arise through method intrinsics

[F. Heidecker, J. Breitenstein, K. Rösch, J. Löhdefink, M. Bieshaar, C. Stiller, T. Fingscheidt, B. Sick, “An Application-Driven Conceptualization of Corner Cases for Perception in Highly Automated Driving”, in Proc. of IV, 2021]

2. Corner Case Definitions
From Systematization to Descriptions

Machine Learning (ML) pipeline in automated driving:



- Emphasis on corner cases and relation to data and ML models
- Need for corner case descriptions
- Systematization provides description on different levels of abstraction

Research Directions for corner case description:
- Automatic corner case description generation
- Corner case coverage in data
- Validation and verification of ML models using corner case descriptions
- Targeted corner case data generation, see for example: [Bogdoll et al., 2022]

[D. Bogdoll, J. Breitenstein, F. Heidecker, M. Bieshaar, B. Sick, T. Fingscheidt, J. M. Zöllner, “Description of Corner Cases in Automated Driving: Goals and Challenges”, in Proc. of ICCV-Workshops, 2021]
[D. Bogdoll, S. Guneshka, J. M. Zöllner, “One Ontology to Rule Them All: Corner Case Scenarios for Autonomous Driving”, in Proc. of ECCV-Workshops, 2022]

Relevance Estimation of Corner Cases for Semantic Segmentation

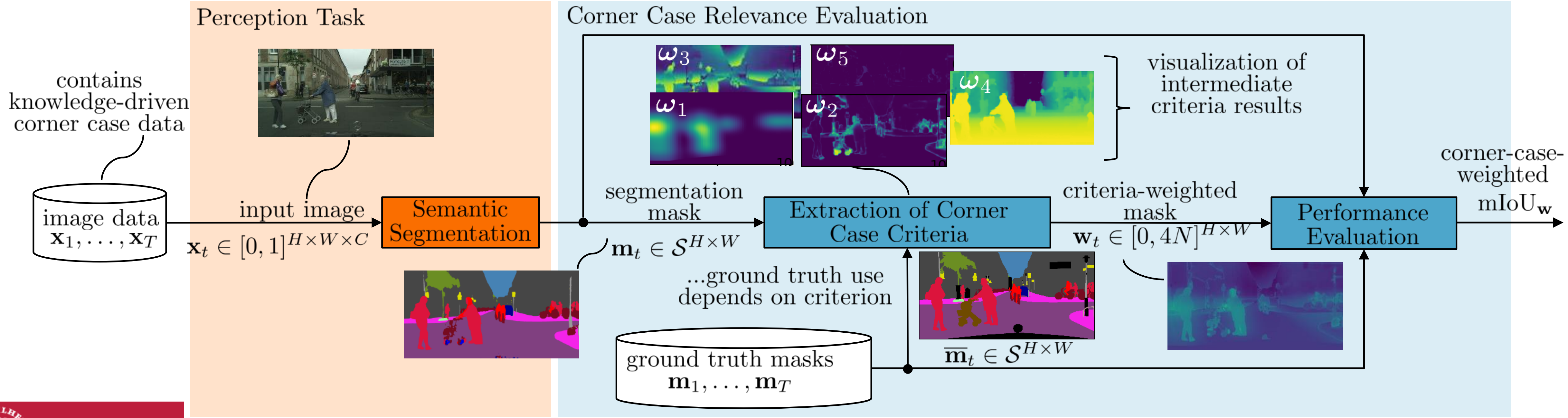
Validation and verification of ML models on corner cases is an important aspect

- How relevant is a certain corner case type for semantic segmentation?

Performance measure in semantic segmentation: mIoU
…but mIoU is insensitive if small parts of the image are affected by, e.g., a corner case

→ corner-case-weighted mIoU uses pixel-wise weightings to measure effect of corner cases on performance

Pipeline for corner case relevance estimation:

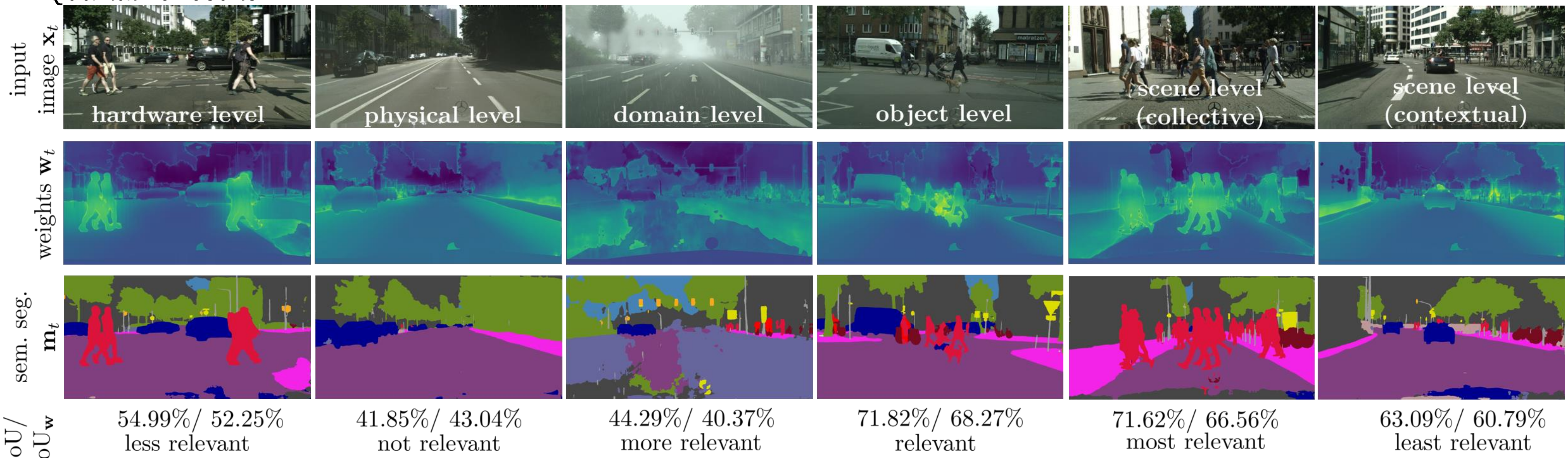


[J. Breitenstein, F. Heidecker, M. Lyssenko, D. Bogdoll, M. Bieshaar, J. M. Zöllner, B. Sick, T. Fingscheidt, “What Does Really Count? Estimating Relevance of Corner Cases for Semantic Segmentation in Automated Driving”, in Proc. of ICCV-Workshops, 2023]

2. Corner Case Definitions
Relevance Estimation of Corner Cases for Semantic Segmentation

- Evaluation on specific corner case data according to systematization
- Semantic segmentation: OCRNet

Qualitative results:



Collective anomalies (scene level corner cases) have the highest relevance (=highest impact on weighted mIoU) for OCRNet
Physical level corner cases have no relevance for OCRNet

Technische Universität Braunschweig

[J. Breitenstein, F. Heidecker, M. Lyssenko, D. Bogdoll, M. Bieshaar, J. M. Zöllner, B. Sick, T. Fingscheidt, “What Does Really Count? Estimating Relevance of Corner Cases for Semantic Segmentation in Automated Driving”, in Proc. of ICCV-Workshops, 2023]

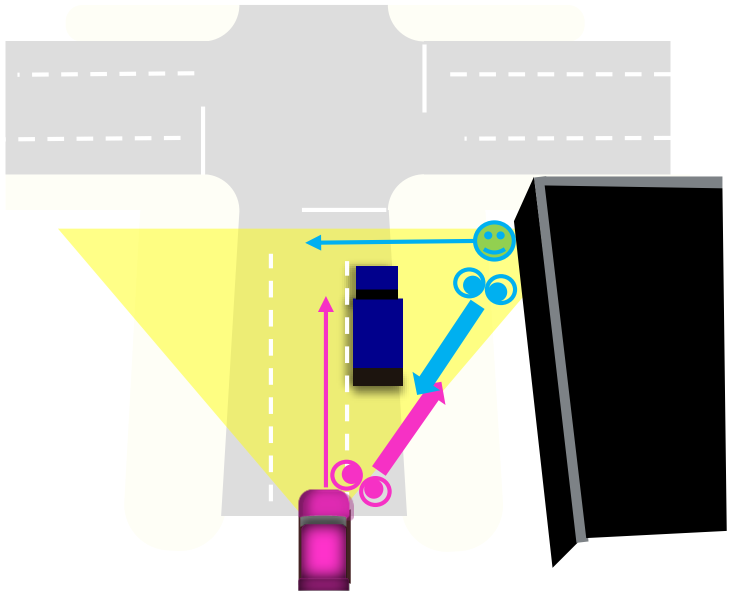
Institut für Nachrichtentechnik

Outline

3. Corner Cases Caused By Occlusions
Motivation to Investigate Amodal Segmentation Methods

- Amodal perception is the ability to hallucinate the full shape of (partially) occluded objects
- Automated driving: perceiving occludees behind occlusions is crucial for safe environment perception
- Humans are good at this, perception methods in general not

Example scenario for occlusion in automated driving:

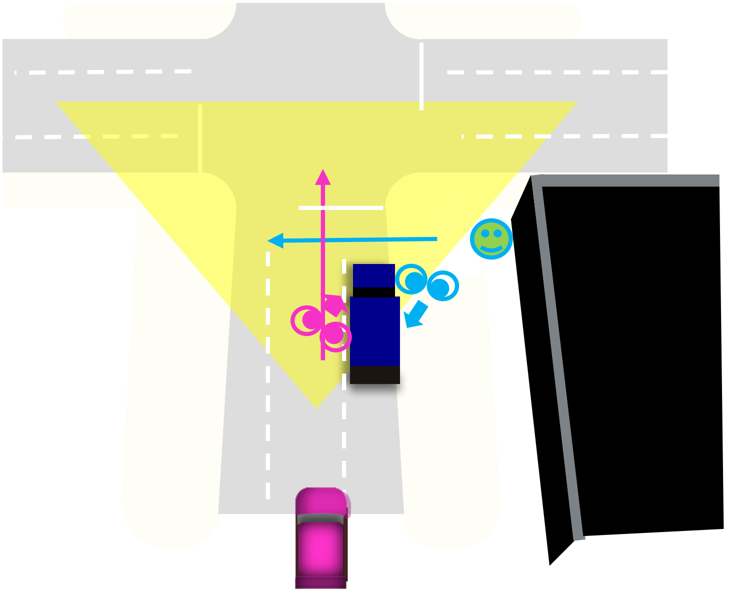


A truck is parked on the side of the street.
Pedestrian is walking towards the street.
Ego-vehicle is driving towards the intersection.
Pedestrian and ego-vehicle see each other.

Motivation to Investigate Amodal Segmentation Methods

- Amodal perception is the ability to hallucinate the full shape of (partially) occluded objects
- Automated driving: perceiving occludees behind occlusions is crucial for safe environment perception
- Humans are good at this, perception methods in general not

Example scenario for occlusion in automated driving:

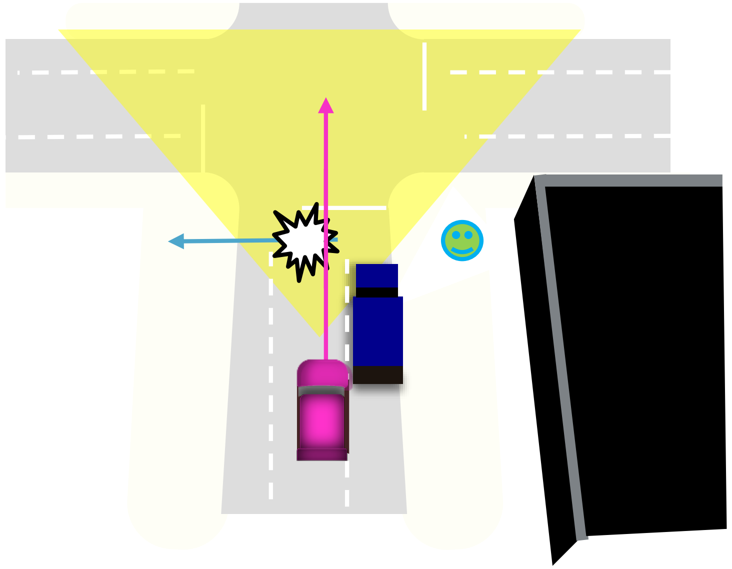


A truck is parked on the side of the street.
Pedestrian is walking towards the street.
Ego-vehicle is driving towards the intersection.
Ego-vehicle is now next to the truck.
Pedestrian and ego-vehicle do not see each other.

Motivation to Investigate Amodal Segmentation Methods

- Amodal perception is the ability to hallucinate the full shape of (partially) occluded objects
- Automated driving: perceiving occludees behind occlusions is crucial for safe environment perception
- Humans are good at this, perception methods in general not

Example scenario for occlusion in automated driving:



A truck is parked on the side of the street.
Pedestrian is walking onto the street.
Ego-vehicle is driving towards the intersection.
Ego-vehicle and pedestrian reach the end of the truck.
Pedestrian and ego-vehicle see each other again but it is too late.

3. Corner Cases Caused By Occlusions

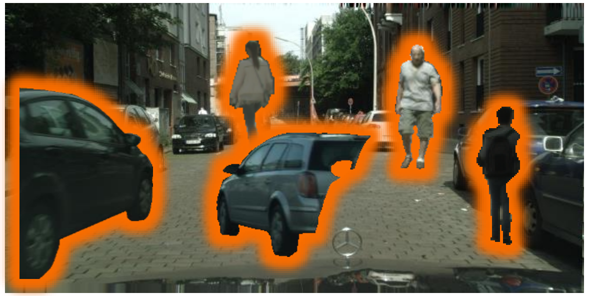
Amodal Segmentation on Images: Amodal Cityscapes

- Training requires an amodal ground truth
- Copy-paste of instances to synthesize Amodal Cityscapes based on Cityscapes

Creation process for one image of the dataset:

❶ Select a target image

❷ Select instances for pasting



❸ Paste instances in target image

- Each image from the standard Cityscapes dataset is used once as target image for the amodal dataset
- The target image is selected from the specified dataset split

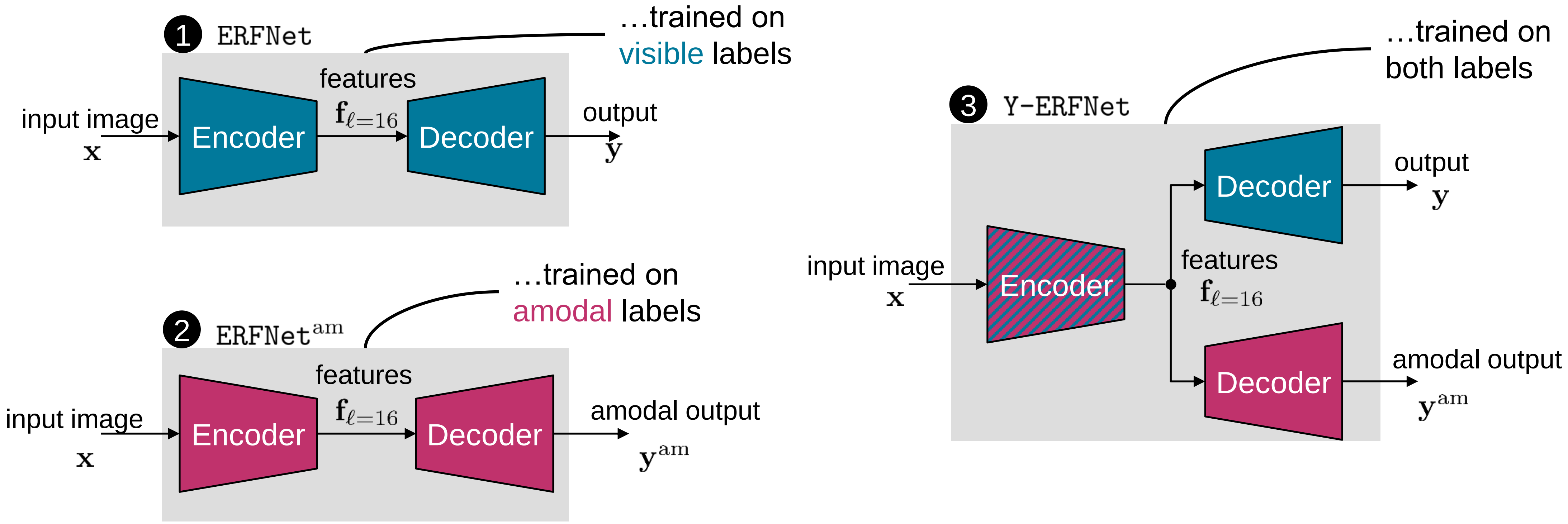
- Instances are selected from source images in the same dataset split minus the current target image
- Randomly choose occlusion level 0%-10%
- Number of instances to paste is selected based on the occlusion level

- Pasting location: paste instances onto the same horizontal line (y-coordinate) that they originate from, x-coordinate is chosen randomly from the range of the image width
- Gaussian blurring of edges
- No overlap with other pasted instances

3. Corner Cases Caused By Occlusions
Joint Training Methods for Visible and Amodal Segmentation: Method Overview

Drawback of previous amodal semantic segmentation: classes from the same group cannot occlude each other!
Idea: Train an entirely amodal semantic segmentation predicting all 19 classes separately

Overview of the three investigated methods:



[J. Breitenstein, J. Löhdefink, T. Fingscheidt, “Joint Prediction of Amodal and Visible Semantic Segmentation for Automated Driving”, In Proc. of ECCV-Workshops 2022]

3. Corner Cases Caused By Occlusions
Joint Training Methods for Visible and Amodal Segmentation: Results

Results on $\mathcal{D}^{\mathrm{amCS}}_{\mathrm{test}}$ (training on $\mathcal{D}^{\mathrm{amCS}}_{\mathrm{train}}$ for 120 epochs):

Method	Dataset	mIoU	mIoU$^{\mathrm{inv}}$
ERFNet	$\mathcal{D}^{\mathrm{amCS}}_{\mathrm{test}}$	62.99%	5.00%
ERFNet$^{\mathrm{am}}(K=4)$	$\mathcal{D}^{\mathrm{amCS}}_{\mathrm{test}}$	62.76%	23.60%
ERFNet$^{\mathrm{am}}$	$\mathcal{D}^{\mathrm{amCS}}_{\mathrm{test}}$	20.16%	36.48%
Y-ERFNet	$\mathcal{D}^{\mathrm{amCS}}_{\mathrm{test}}$	**63.32%**	**43.32%**

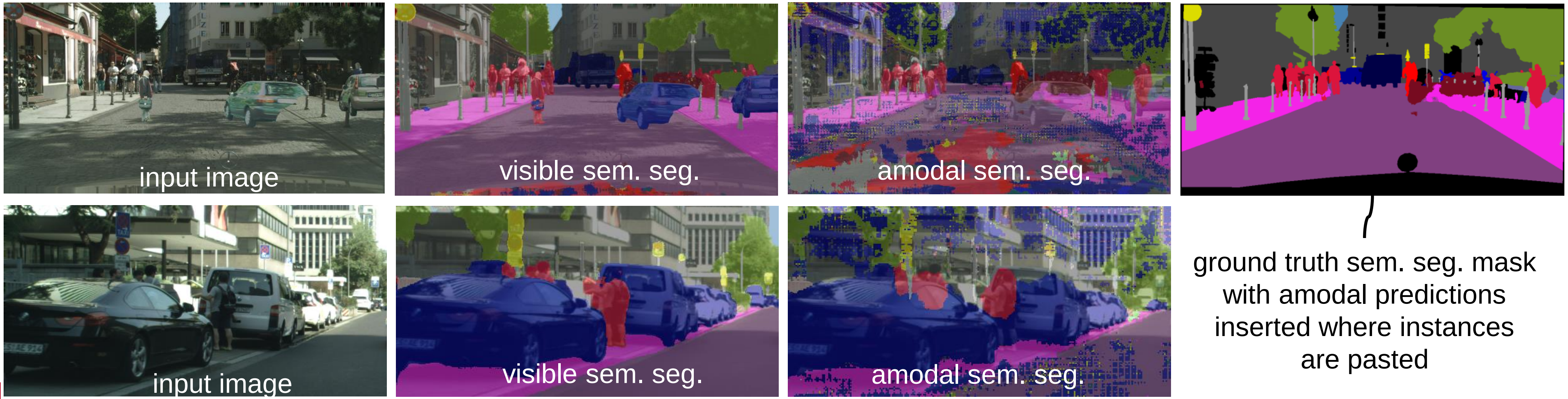
groupwise training technique by [Purkait et al.,2019]

our training technique

Qualitative results of Y-ERFNet

$\mathcal{D}^{\mathrm{amCS}}_{\mathrm{test}}$

$\mathcal{D}^{\mathrm{CS}}_{\mathrm{val}}$



ground truth sem. seg. mask with amodal predictions inserted where instances are pasted

[P. Purkait, C. Zach, I. Reid, “Seeing Behind Things: Extending Semantic Segmentation to Occluded Regions”, In Proc. of IROS 2019]

[J. Breitenstein, J. Löhdefink, T. Fingscheidt, “Joint Prediction of Amodal and Visible Semantic Segmentation for Automated Driving”, In Proc. of ECCV-Workshops 2022]

3. Corner Cases Caused By Occlusions
From Images to Videos

Amodal semantic segmentation on images:
- Amodal Cityscapes: synthetic copy-paste dataset based on Cityscapes with amodal masks in plausible locations
- Joint multi-task training of visible and amodal semantic segmentation improves performance in both tasks

scene from the SAIL-VOS dataset

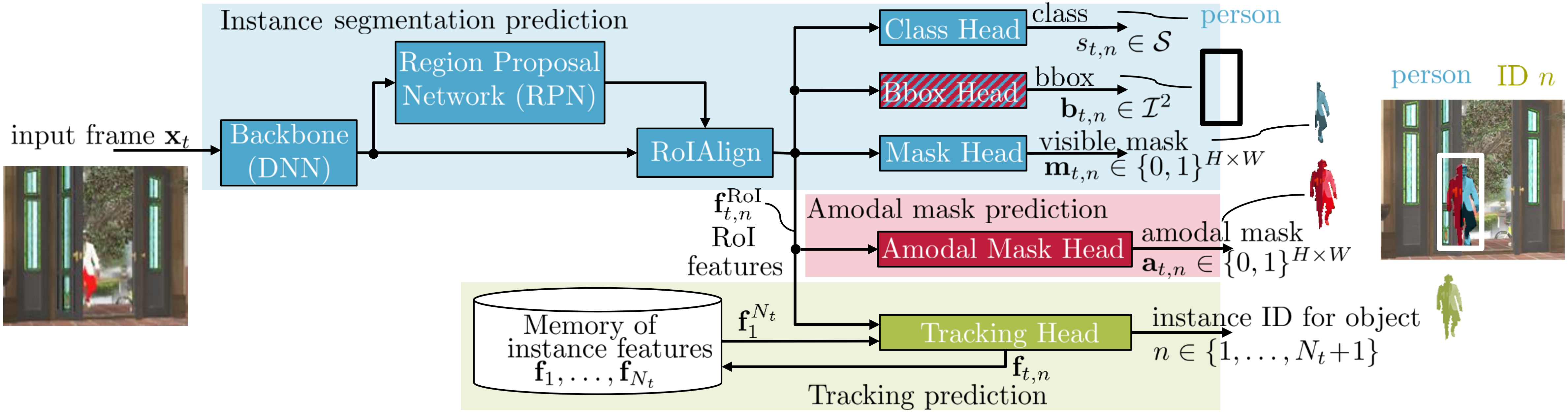
Towards video-based methods:
- Overall aim: detect pedestrians behind occlusions → we need video-based methods
- Amodal segmentation methods only exist for single images
- There is only one dataset for amodal video instance segmentation: SAIL-VOS
- Methods for video instance segmentation exist

→ We want to provide baseline methods for amodal video instance segmentation

[Y.-T. Hu, H.-S. Chen, K. Hui, J.-B. Huang, A. G. Schwing, “SAIL-VOS: Semantic Amodal Instance Level Video Object Segmentation - A Synthetic Dataset and Baselines,” In Proc. CVPR, 2019.]

3. Corner Cases Caused By Occlusions
End2end Amodal Video Instance Segmentation

Functioning of our proposed VATrack during inference:



- Simultaneous prediction of amodal and visible instance masks
- Additional tracking of the (amodal) instance masks throughout the sequence
- Instance segmentation prediction is based on Mask R-CNN [He et al., 2017]
- Investigation of two tracking methods: QDTrack (QD) [Pang et al., 2021] and MaskTrack R-CNN (MT) [Yang et al., 2019]

[K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” In Proc. ICCV, 2017.]
[J. Breitenstein, K. Jin, A. Hakiri, M. Klingner, T. Fingscheidt, “End-to-End Amodal Video Instance Segmentation”, in Proc. of BMVC-Workshops, 2023]
[J. Pang, L. Qiu, X. Li, H. Chen, Q. Li, T. Darrell, F. Yu, “Quasi-dense Similarity Learning for Multiple Object Tracking“, In Proc. of CVPR, 2021.]

End2end Amodal Video Instance Segmentation

Amodal Results on Image- and Video-Level on the SAIL-VOS-cut dataset

image-based methods

video-based methods

- SAIL-VOS is the only dataset with annotated amodal video instance labels → But it contains jump cuts
- SAIL-VOS-cut splits a video at a jump cut into 2 videos (still same content!)

amodal results on image-level (in AP and derivatives)

Method	DNN	Visible	Amodal	AP	AP_{50}	AP^{P}_{50}	AP^{H}_{50}	AP^{L}_{50}	AP^{M}_{50}	AP^{S}_{50}
MaskAmodal [Hu et al., 2019]	?		✓	13.0	23.0	24.3	16.7	36.6	21.5	6.1
MaskJoint [Hu et al., 2019]	?	✓	✓	14.1	24.8	24.3	18.9	37.8	21.5	5.7
MaskAmodal*	RX101		✓	16.3	25.6	27.4	17.1	35.2	24.2	10.1
MaskJoint*	RX101	✓	✓	16.7	25.6	26.9	17.3	33.0	22.3	9.0
AmodalTrack (MT)	RX101		✓	15.9	25.7	24.9	17.8	36.8	22.8	11.2
VATrack (MT)	RX101	✓	✓	16.4	26.0	24.9	18.0	**38.6**	22.5	10.6
AmodalTrack (QD)	RX101		✓	17.8	27.4	29.2	18.6	34.7	26.8	11.4
VATrack (QD)	RX101	✓	✓	**18.3**	**28.6**	**29.7**	**20.1**	38.1	**26.9**	**15.7**

temporal context (TC) can improve results

joint prediction mostly improves image-level results

amodal results on video-level (in vAP and derivatives)

Method	DNN	Visible	Amodal	vAP	vAP_{50}	vAP^{P}_{50}	vAP^{H}_{50}	vAP^{L}_{50}	vAP^{M}_{50}	vAP^{S}_{50}
AmodalTrack (MT)	ResNext101		✓	2.4	3.1	3.8	1.7	3.8	1.4	0.4
VATrack (MT)	ResNext101	✓	✓	2.3	3.1	3.8	1.7	3.7	1.5	0.3
AmodalTrack (QD)	ResNext101		✓	13.1	20.5	21.0	10.7	29.4	14.7	**8.9**
VATrack (QD)	ResNext101	✓	✓	**14.1**	**22.3**	**22.0**	**12.8**	**32.8**	**15.6**	8.8

joint (QD) prediction mostly improves video-level results

MT-based cannot provide meaningful results

[Y.-T. Hu, H.-S. Chen, K. Hui, J.-B. Huang, A. G. Schwing, “SAIL-VOS: Semantic Amodal Instance Level Video Object Segmentation - A Synthetic Dataset and Baselines,” In Proc. CVPR, 2019.]

[J. Breitenstein, K. Jin, A. Hakiri, M. Klingner, T. Fingscheidt, “End-to-End Amodal Video Instance Segmentation”, in Proc. of BMVC-Workshops, 2023]

End2end Amodal Video Instance Segmentation

Amodal Results on Image- and Video-Level on the SAIL-VOS-cut dataset

image-based methods

video-based methods

- SAIL-VOS is the only dataset with annotated amodal video instance labels → But it contains jump cuts
- SAIL-VOS-cut splits a video at a jump cut into 2 videos (still same content!)

Great, and now what?

amodal results on image-level (in AP and derivatives)

Method	DNN	Visible	Amodal	AP	AP_{50}	[illegible]	[illegible]	AP^{L}_{50}	AP^{M}_{50}	AP^{S}_{50}
MaskAmodal [Hu et al., 2019]	?		✓	13.0	[illegible]	[illegible]	[illegible]	[illegible]	21.5	6.1
MaskJoint [Hu et al., 2019]	?	✓	✓	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	21.5	5.7
MaskAmodal*	RX101		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	24.2	10.1
MaskJoint*	RX101	✓	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	22.3	9.0
AmodalTrack (MT)	RX101	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	22.8	11.2
VATrack (MT)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**38.6**	22.5	10.6
AmodalTrack (QD)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	18.6	34.7	<u>26.8</u>	<u>11.4</u>
VATrack (QD)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**29.7**	**20.1**	<u>38.1</u>	**26.9**	**15.7**

temporal context (TC) can improve results

joint prediction mostly improves image-level results

amodal results on video-level (in vAP and derivatives)

Method	[illegible]	[illegible]	[illegible]	[illegible]	vAP_{50}	vAP^{P}_{50}	vAP^{H}_{50}	vAP^{L}_{50}	vAP^{M}_{50}	vAP^{S}_{50}
AmodalTrack (MT)	[illegible]	[illegible]	[illegible]	2.4	3.1	3.8	1.7	3.8	1.4	0.4
VATrack (MT)	[illegible]	[illegible]	✓	2.3	3.1	3.8	1.7	3.7	1.5	0.3
AmodalTrack (QD)	[illegible]	[illegible]	✓	<u>13.1</u>	<u>20.5</u>	<u>21.0</u>	<u>10.7</u>	<u>29.4</u>	<u>14.7</u>	**8.9**
VATrack (QD)	ResNext101	✓	✓	**14.1**	**22.3**	**22.0**	**12.8**	**32.8**	**15.6**	8.8

joint (QD) prediction mostly improves video-level results

MT-based cannot provide meaningful results

[Y.-T. Hu, H.-S. Chen, K. Hui, J.-B. Huang, A. G. Schwing, “SAIL-VOS: Semantic Amodal Instance Level Video Object Segmentation - A Synthetic Dataset and Baselines,” In Proc. CVPR, 2019.]

[J. Breitenstein, K. Jin, A. Hakiri, M. Klingner, T. Fingscheidt, “End-to-End Amodal Video Instance Segmentation”, in Proc. of BMVC-Workshops, 2023]

3. Corner Cases Caused By Occlusions
Amodal Video Dataset Generation Using CARLA: Bridging the Domain Gap

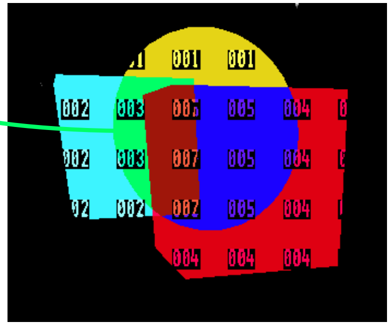
CARLA simulator: generation of data for automated driving with different sensors (e.g., RGB, semantic segmentation, instance segmentation)

Idea: Generate pre-specified occlusion scenarios in CARLA with amodal ground truth

2 problems:
… how to obtain amodal ground truth?
… how to run specific scenarios in CARLA?

amodal ground truth:
- From computer graphics: custom depth stencil (in underlying UnrealEngine)
- Custom depth stencil is also used to obtain other segmentations
- Amodal information is written in the custom depth stencil if instances are occluded
- New: change the custom depth stencil values dynamically after actors have spawned in the world

Specific overlaps are calculated from a graph-like structure to obtain individual unique IDs per overlap



3. Corner Cases Caused By Occlusions

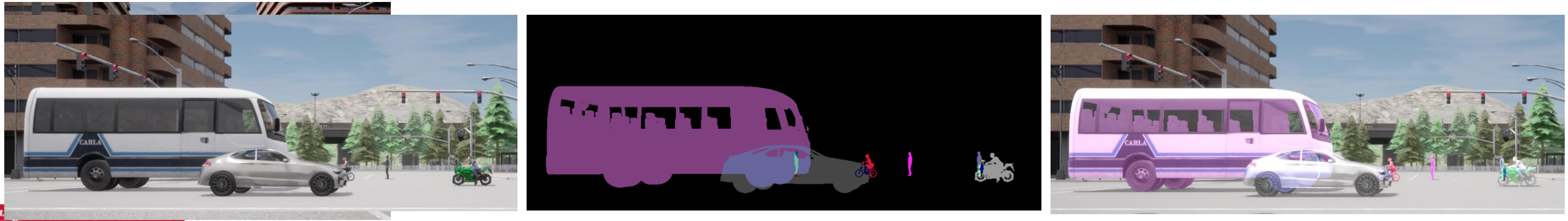
Amodal Video Dataset Generation Using CARLA: Bridging the Domain Gap

CARLA simulator: generation of data for automated driving with different sensors (e.g., RGB, semantic segmentation, instance segmentation)

Idea: Generate pre-specified occlusion scenarios in CARLA with amodal ground truth

2 problems:
… how to obtain amodal ground truth?
… how to run specific scenarios in CARLA?

CARLA scenarios:
- CARLA scenario runner has been used in related work to generate corner case scenarios [Bogdoll et al.,2022]
- Scenario definitions have to be extended to cover certain time spans
- Trajectories of traffic participants were very undeterministic → extension to pre-define trajectories of traffic participants



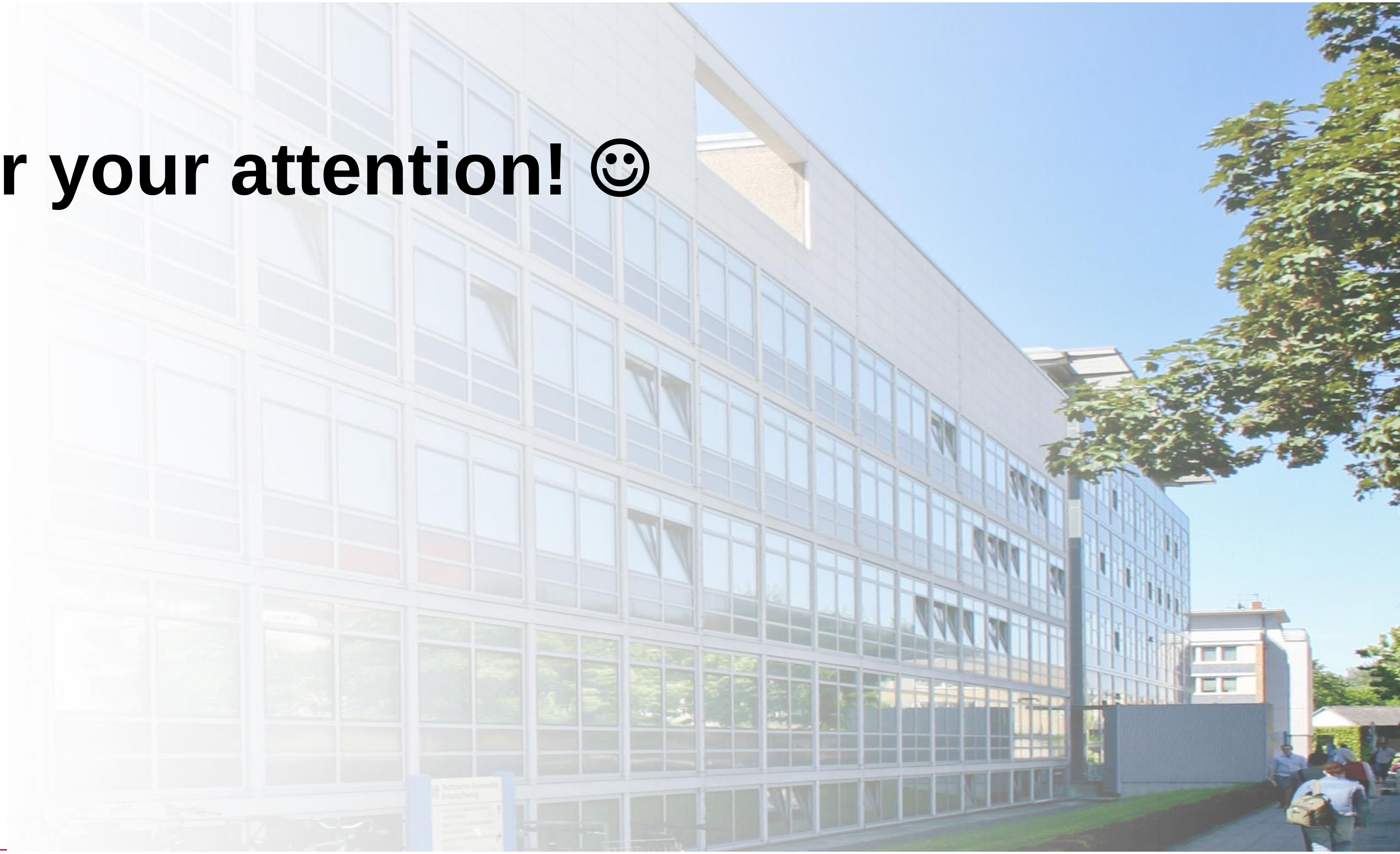
[D. Bogdoll, S. Guneshka, J. M. Zöllner, “One Ontology to Rule Them All: Corner Case Scenarios for Autonomous Driving”, in Proc. of ECCV-Workshops, 2022]

4. Summary and Future Work

Summary:
- Corner case systematization for better understanding and unified definition of corner cases
- Corner cases can be mapped to sensors and perception methods
- Relevance estimation measures the effect corner case types have on a segmentation method
- Amodal perception is an important ability to treat corner cases caused by occlusion
- Amodal segmentation is possible on both images and videos

Future Work:
- Automatization of corner case description to facilitate data generation and model validation
- Deeper investigation of method layer corner cases:
    - Which type of data results on corner cases for perception methods?
    - What are corner cases beside the ones “imaginable” for humans?
- Generation of automotive amodal video data using the CARLA simulator (ongoing)
- Self-supervised amodal video instance segmentation to soften the label requirements

Thank you for your attention! ☺



Jasmin Breitenstein

j.breitenstein@tu-braunschweig.de

2. Corner Case Definitions
Summary and Research Directions

- Corner case definitions are needed for:
  - Development of directed corner case detection methods
  - Scenario descriptions
  - Validation and verification of machine learning methods

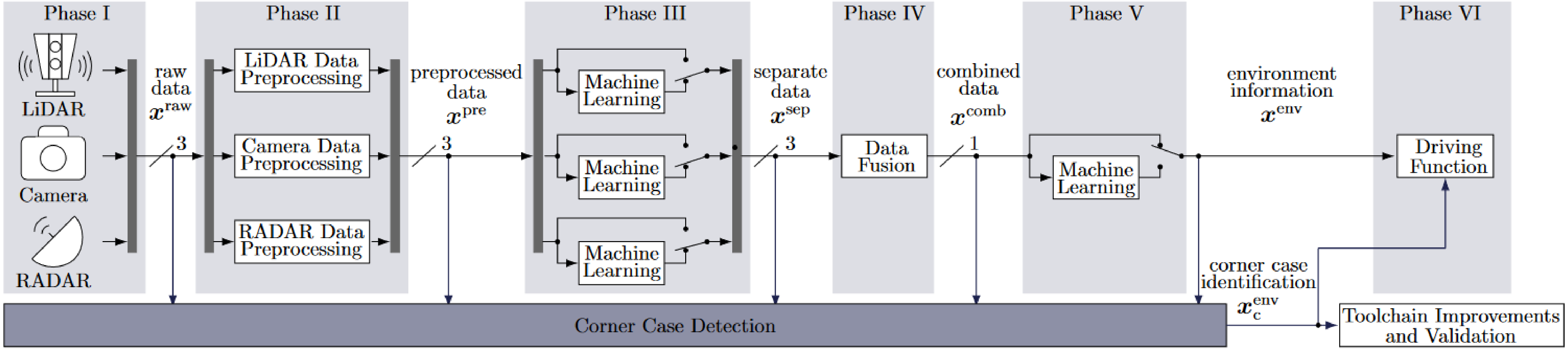
- Systematization defines corner cases based on scenario content on different layers and levels
- Corner case descriptions can be developed based on the systematization

- Relevance estimation measures the effect corner case types have on a segmentation method
  - Provides a mapping between corner case systematization and perception method
  - For the evaluated network collective anomalies are the most relevant corner case type

Research directions:
- Automatic corner case descriptions to generate new data and analyze content of given data
- Deeper investigation of method layer corner cases:
  - Which type of data results on corner cases for perception methods?
  - What are corner cases beside the ones “imaginable” for humans?

2. Corner Case Definitions
Systematization of Corner Cases: Extension to Other Sensors

Toolchain for data processing in automated driving:



- Phases: data acquisition (I), data preprocessing (II), early machine learning (III), data fusion (IV), late machine learning (V), and application (VI)
- Three individual sensors: LiDAR, camera, RADAR

[F. Heidecker, J. Breitenstein, K. Rösch, J. Löhdefink, M. Bieshaar, C. Stiller, T. Fingscheidt, B. Sick, “An Application-Driven Conceptualization of Corner Cases for Perception in Highly Automated Driving”, in Proc. of IV, 2021]

3. Corner Cases Caused By Occlusions
Standard Perception Fails at Occlusions

- Amodal perception is the ability to hallucinate the full shape of (partially) occluded objects
- Automated driving: perceiving occludees behind occlusions is crucial for safe environment perception
- Humans are good at this, perception methods in general not

Testing accuracy of object classification methods under extreme occlusion on the VehicleOcclusion dataset [Zhu et al., 2019]

Humans/Methods	w/o occlusion	w/ occlusion
Humans	-	**93.3%**
AlexNet	89.8%	50.0%
ResNet	90.1%	54.0%
VGG16	**94.7%**	**62.6%**

example images

- Worse perception of occluded objects results in corner cases
- Other types of corner cases can form behind occlusions

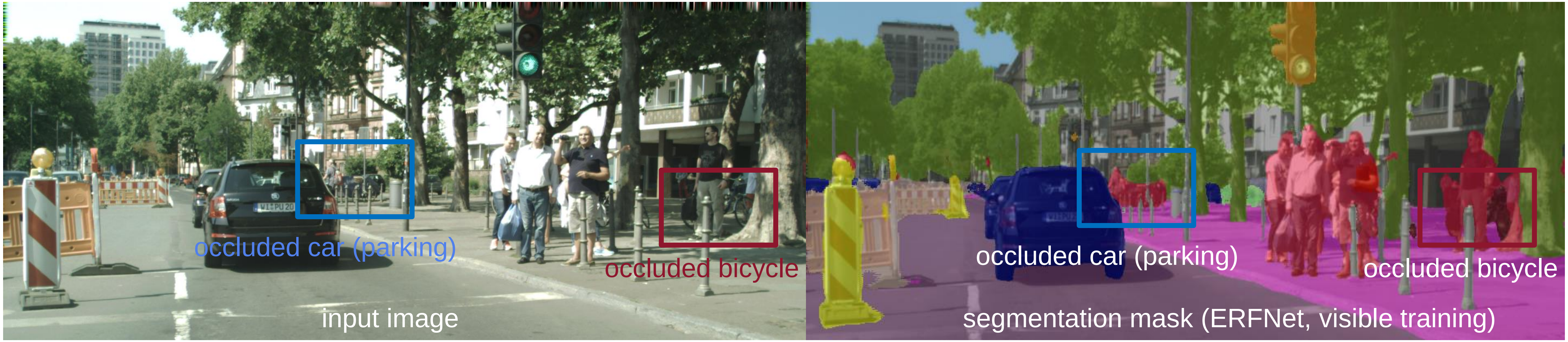
→ Amodal methods learn to anticipate occluded objects

[H. Zhu, P. Tang, J. Park, S. Park, A. Yuille: "Robustness of Object Recognition under Extreme Occlusion in Humans and Computational Models," arXiv:1905.04598, 2019]

3. Corner Cases Caused By Occlusions
Standard Perception Fails at Occlusions

- Amodal perception is the ability to hallucinate the full shape of (partially) occluded objects
- Automated driving: perceiving occludees behind occlusions is crucial for safe environment perception
- Humans are good at this, perception methods in general not

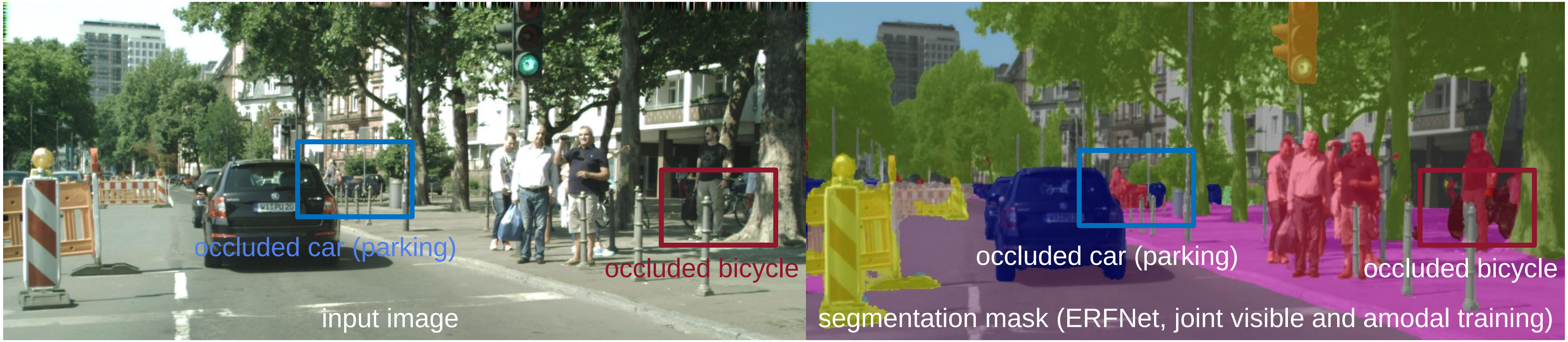
Intuition (qualitative) for semantic segmentation:



3. Corner Cases Caused By Occlusions
Standard Perception Fails at Occlusions

- Amodal perception is the ability to hallucinate the full shape of (partially) occluded objects
- Automated driving: perceiving occludees behind occlusions is crucial for safe environment perception
- Humans are good at this, perception methods in general not

Intuition (qualitative) for semantic segmentation:



- Worse perception of occluded objects results in corner cases
- Other types of corner cases can form behind occlusions, especially: knowledge of VRUs at all times!

→ Amodal segmentation methods learn to anticipate occluded objects

Amodal Segmentation on Images: Amodal Cityscapes

- Training requires an amodal ground truth
- Copy-paste of instances to synthesize Amodal Cityscapes based on Cityscapes

Standard Cityscapes ($\mathcal{D}^{\mathrm{CS}}$) dataset:

Split	Training $\mathcal{D}^{\mathrm{CS}}_{\mathrm{train}}$	Validation $\mathcal{D}^{\mathrm{CS}}_{\mathrm{val}}$	Test $\mathcal{D}^{\mathrm{CS}}_{\mathrm{test}}$
#images	2975	500	1525
#ground truth masks	2975	500	-

Amodal Cityscapes ($\mathcal{D}^{\mathrm{amCS}}$) dataset:

Split	Training $\mathcal{D}^{\mathrm{amCS}}_{\mathrm{train}}$	Validation $\mathcal{D}^{\mathrm{amCS}}_{\mathrm{val}}$	Test $\mathcal{D}^{\mathrm{amCS}}_{\mathrm{test}}$
#images	2900	75	500
corresponding $\mathcal{D}^{\mathrm{CS}}$ split	training	training	validation
#ground truth masks	2900	75	500
#source images	2900-1	75-1	500-1
#occluders for pasting	36303 - N	832- N	6014- N

N number instances per image

[Jasmin Breitenstein, Tim Fingscheidt, "Amodal Cityscapes: A New Dataset, its Generation, and an Amodal Semantic Segmentation Challenge Baseline", In Proc. of IV, Aachen, Germany, Jun. 2022]

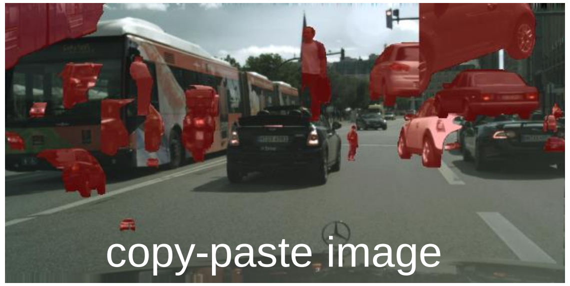
3. Corner Cases Caused By Occlusions

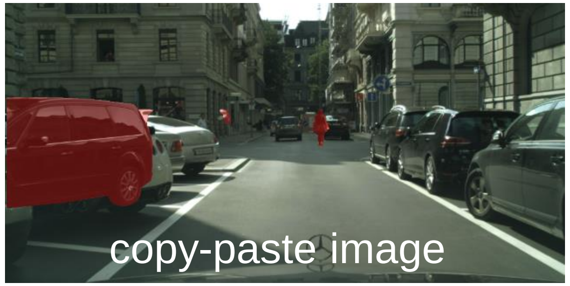
Amodal Cityscapes Dataset: Dataset Generation

randomly inserted instances

instances inserted with horizontal heuristic

…inserted instances shaded in red for visualization


copy-paste image


copy-paste image

…not very plausible images

…harder to say which instances are inserted

Example data from the Amodal Cityscapes dataset


copy-paste image


visible mask (ground truth)


amodal mask (ground truth)

End2end Amodal Video Instance Segmentation
Future Work

- Hard label requirements: annotated amodal masks for entire video sequences are needed
- Can likely only be satisfied for synthetic data

→ SAIL-VOS is not from ego-vehicle perspective for automated driving? Can we bridge the domain gap?
→ How to soften the label requirements? Can we work with real data?

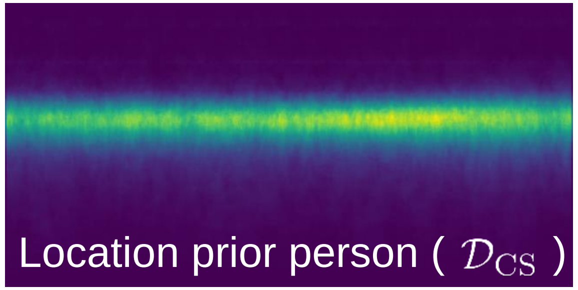
Available in VIS:

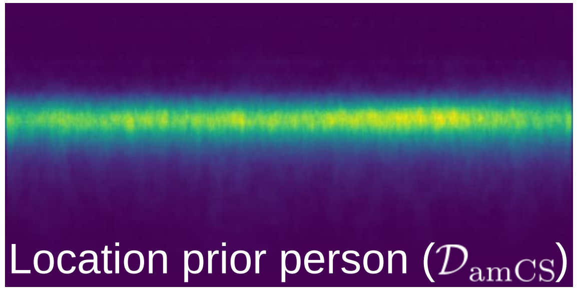
Example sequence of BDD-Mots

Examples with annotated labels for VIS (tracking ID, class, visible instance masks): BDD-Mots, KITTI-360
→ The only requirement missing is the amodal video instance mask annotation

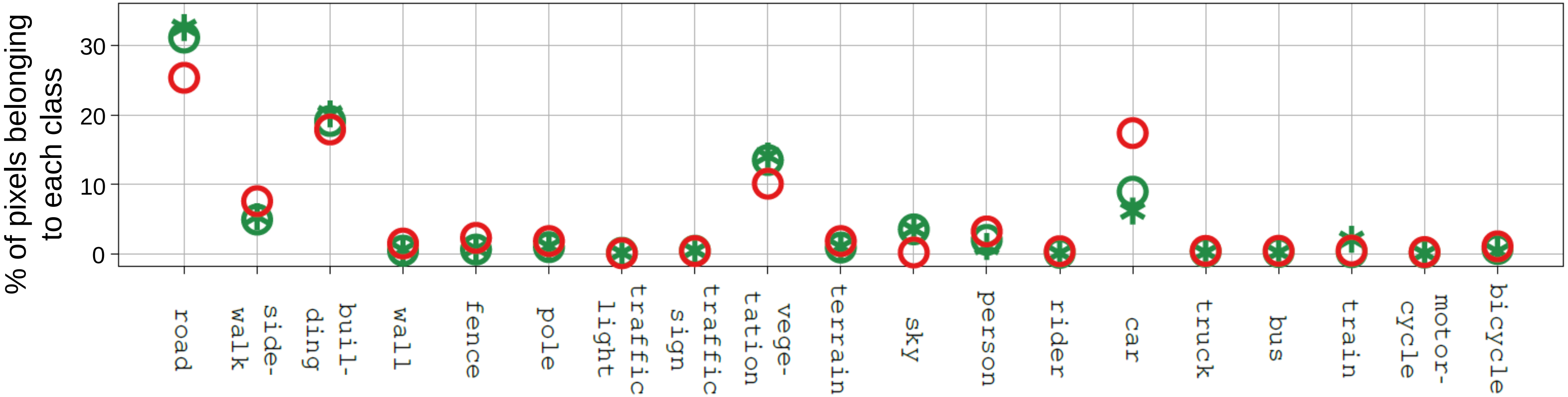
1. Can we bridge the SAIL-VOS domain gap partly and simulate data from automated driving?
2. Can we find a way to train our framework such that amodal annotations are not needed?

We compare the amodal Cityscapes dataset to the Cityscapes dataset:





Comparison of the location priors for class person: No visual difference



→ % of pixels belonging to each class similar for both datasets and visible and occluded areas

% of pixels belonging to each class on Cityscapes (∗) and Amodal Cityscapes (○) for visible and occluded