

VOLKSWAGEN

AKTIENGESELLSCHAFT

GROUP INNOVATION



VOLKSWAGEN GROUP INNOVATION

INTERNATIONAL INNOVATION ECOSYSTEM

VOLKSWAGEN

AKTIENGESELLSCHAFT

GROUP INNOVATION

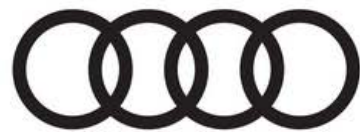


Sensor Data Annotation and Deep Neural Networks for Environment Perception

Dr. Volker Patricio Schomerus, Dr. Thorsten Bagdonat

VOLKSWAGEN

AKTIENGESELLSCHAFT



Audi



ŠKODA



BENTLEY



PORSCHE



**Commercial
Vehicles**



Data-driven Perception Projects

Deep Neural Networks for:

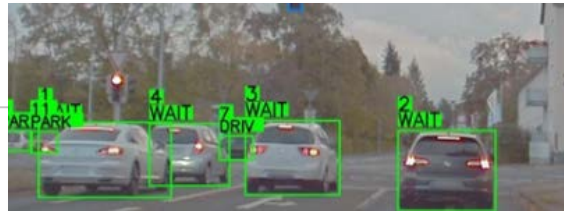
- ▶ Semantic Pixel Labeling



- ▶ Instance Segmentation



- ▶ 2D Action Prediction



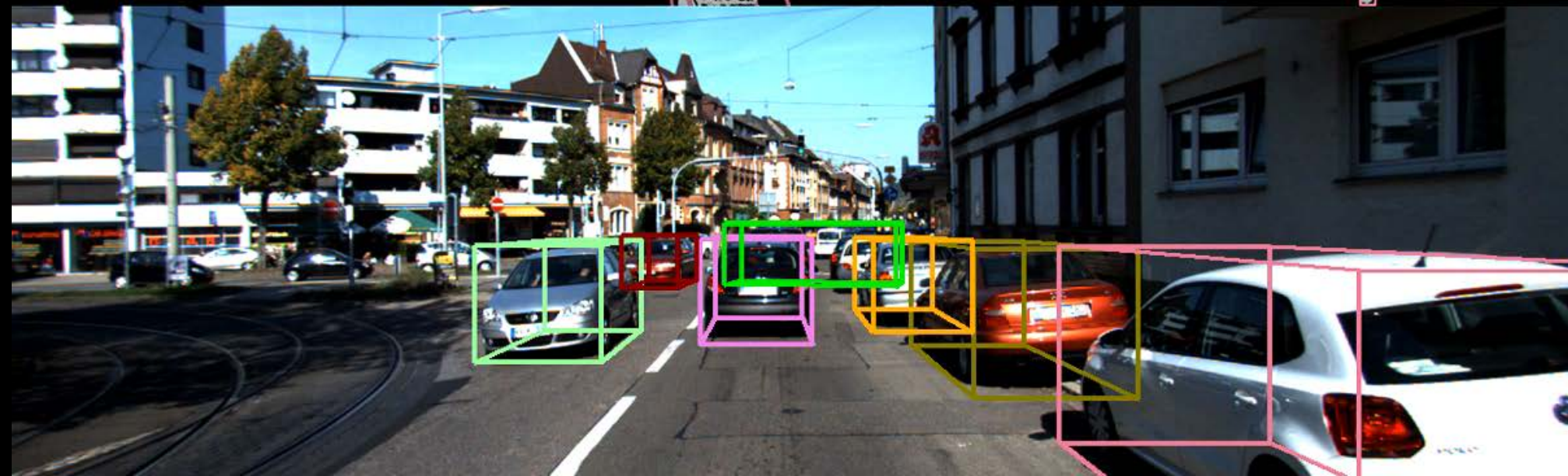
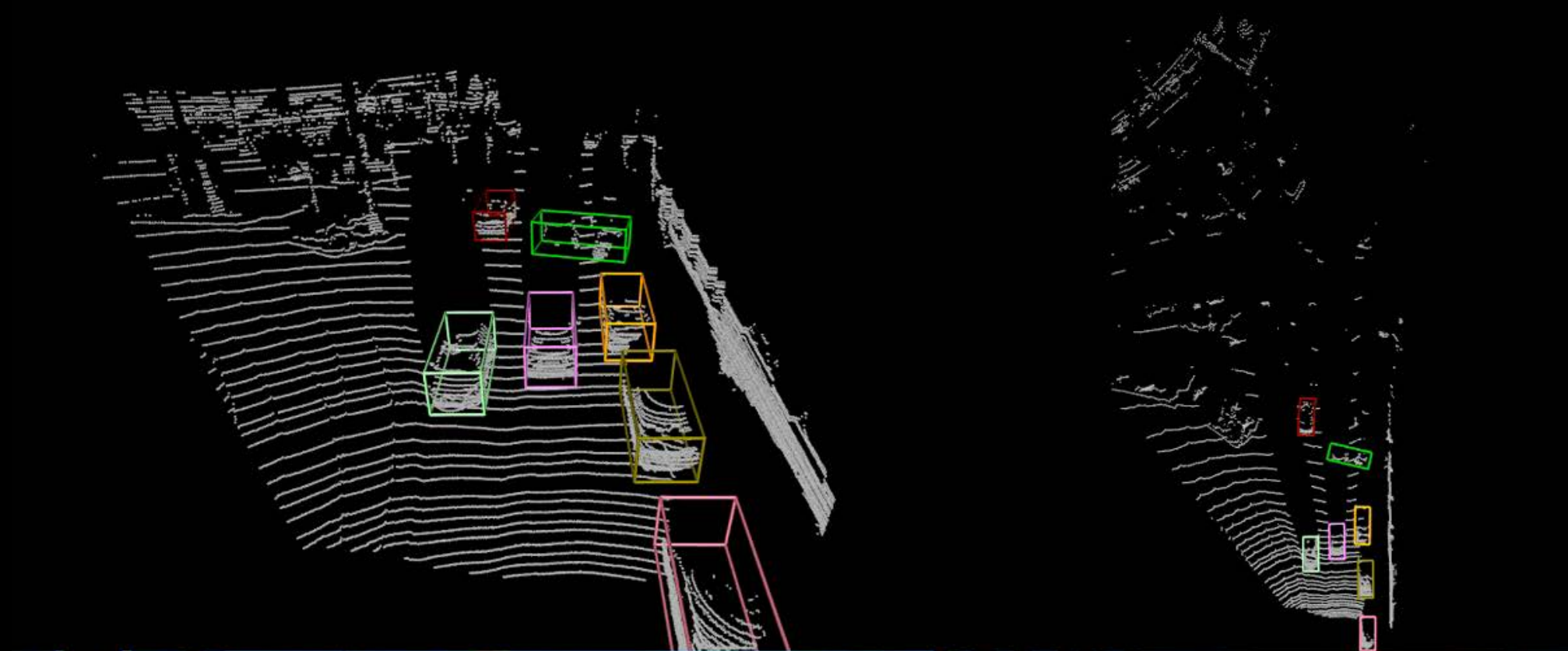
- ▶ 3D Object Detection



- ▶ 3D Object Tracking







Outline

- I. Data-driven Perception Projects
- II. Semantic Segmentation → Multi Class Pixel Labeling
- III. Objects and Attributes



Multi Class Pixel Labeling

What to teach a DNN?

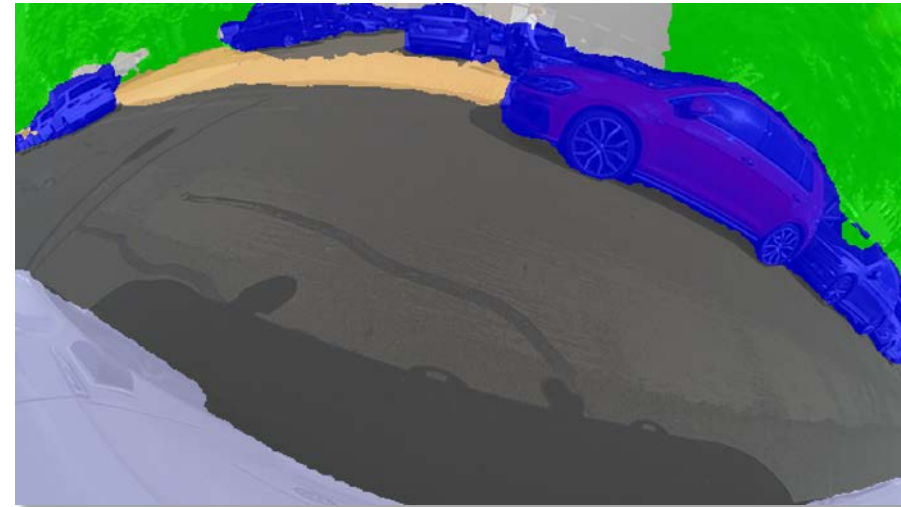
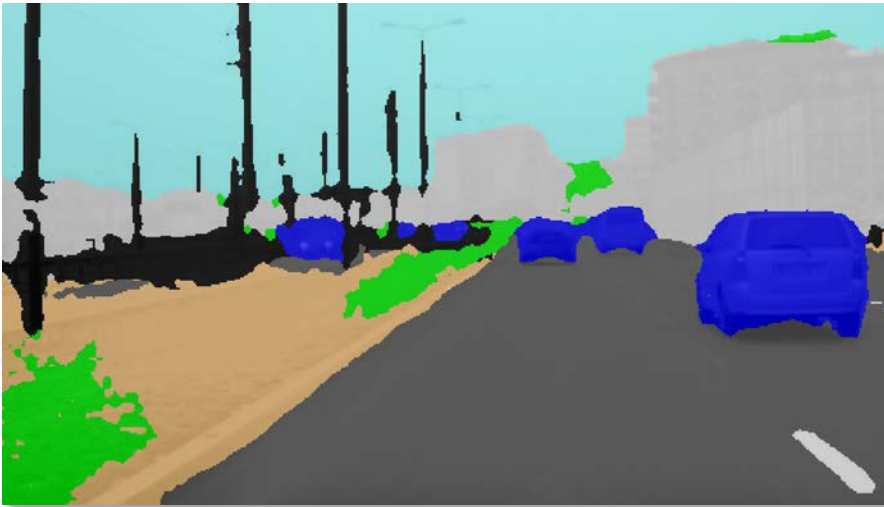
- ▶ *Functional meaning:* drivable / non-drivable / sidewalk / parking spot...
- ▶ *Texture:* asphalt / cobblestone / snow / gravel...
- ▶ *Road features:* dashed marking / solid marking / curbstone / pot hole...



Multi Class Pixel Labeling

What to teach a DNN?

- ▶ DNNs for semantic image segmentation struggle with ground classification



- ▶ Unclear class definition, inconsistent annotation
- ▶ Depending on the situation, a DNN is better or worse in classifying function, texture or features

Multi Class Pixel Labeling

Class Definitions

A, B, C: Ground Surface

D: Dynamic Objects (instances)

E: Static Objects

The image shows a detailed table of class definitions. A red box highlights the 'Surface: Semantic / Functional' section (A) and the 'Surface: Textural / Material' section (B). The table includes columns for 'Category', 'Class No.', and 'Class name'. The 'Surface: Semantic / Functional' section includes classes 1 through 8, and the 'Surface: Textural / Material' section includes classes 9 through 15.

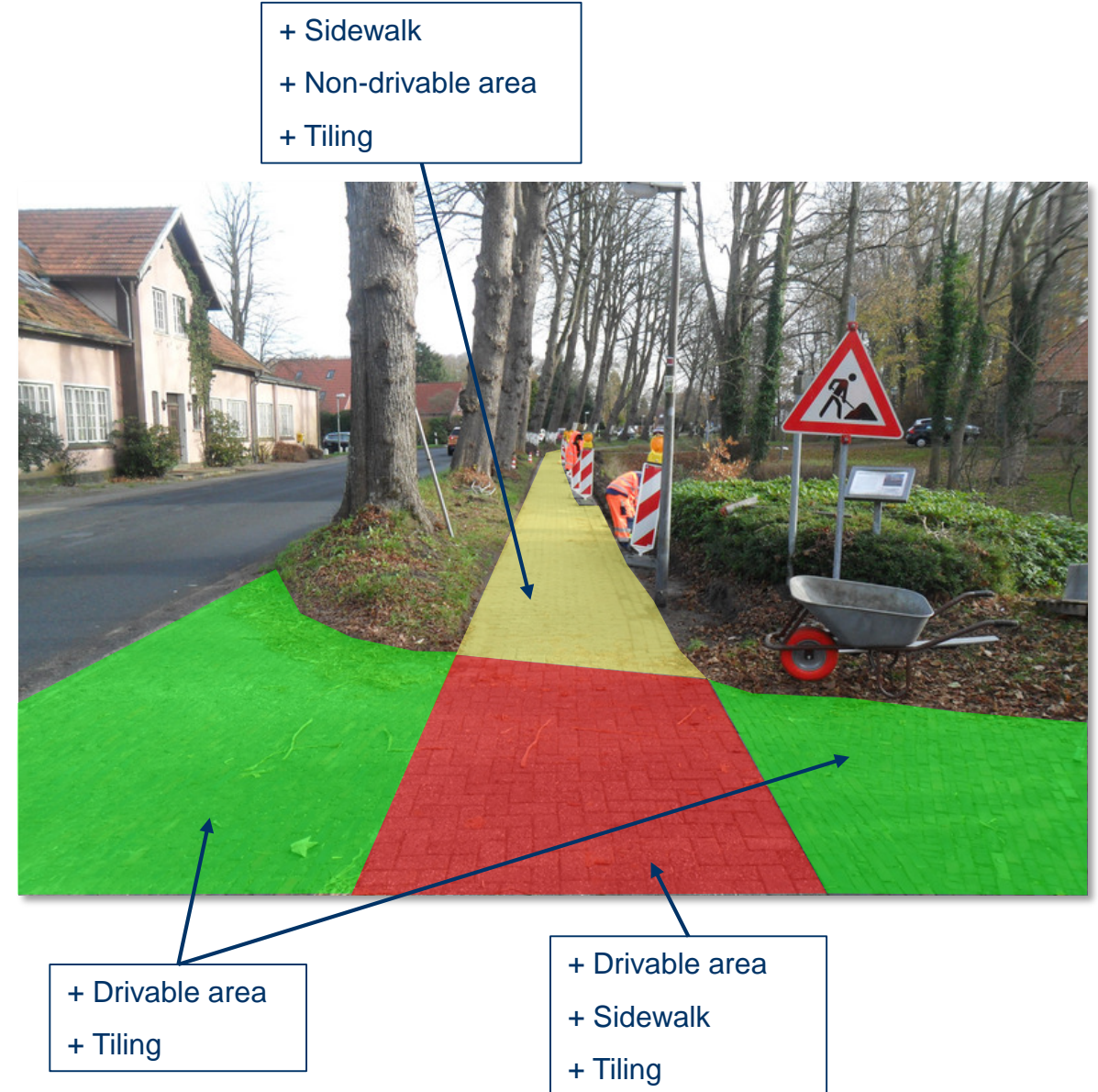
Category	Class No.	Class name
A Surface: Semantic / Functional	1	Drivable road
	2	Street gutter
	3	Parking spot
	4	Drivable area
	5	Sidewalk
	6	Non-drivable area
	7	Zebra crossing (area)
	8	Restricted road area
B Surface: Textural / Material	9	Asphalt
	10	Tiling
	11	Cobblestone
	12	Gravel
	13	Nature surface
	14	Snow
	15	Leaves
C Surface: Road Features	16	Curbstone
	17	Marking: dashed line
	18	Marking: solid line
	19	Marking: painted driving instructions
	20	Marking: restricted road area
	21	Marking: zebra crossing
	22	Yellow marking

Multi Class Pixel Labeling

Ground Plane Annotation

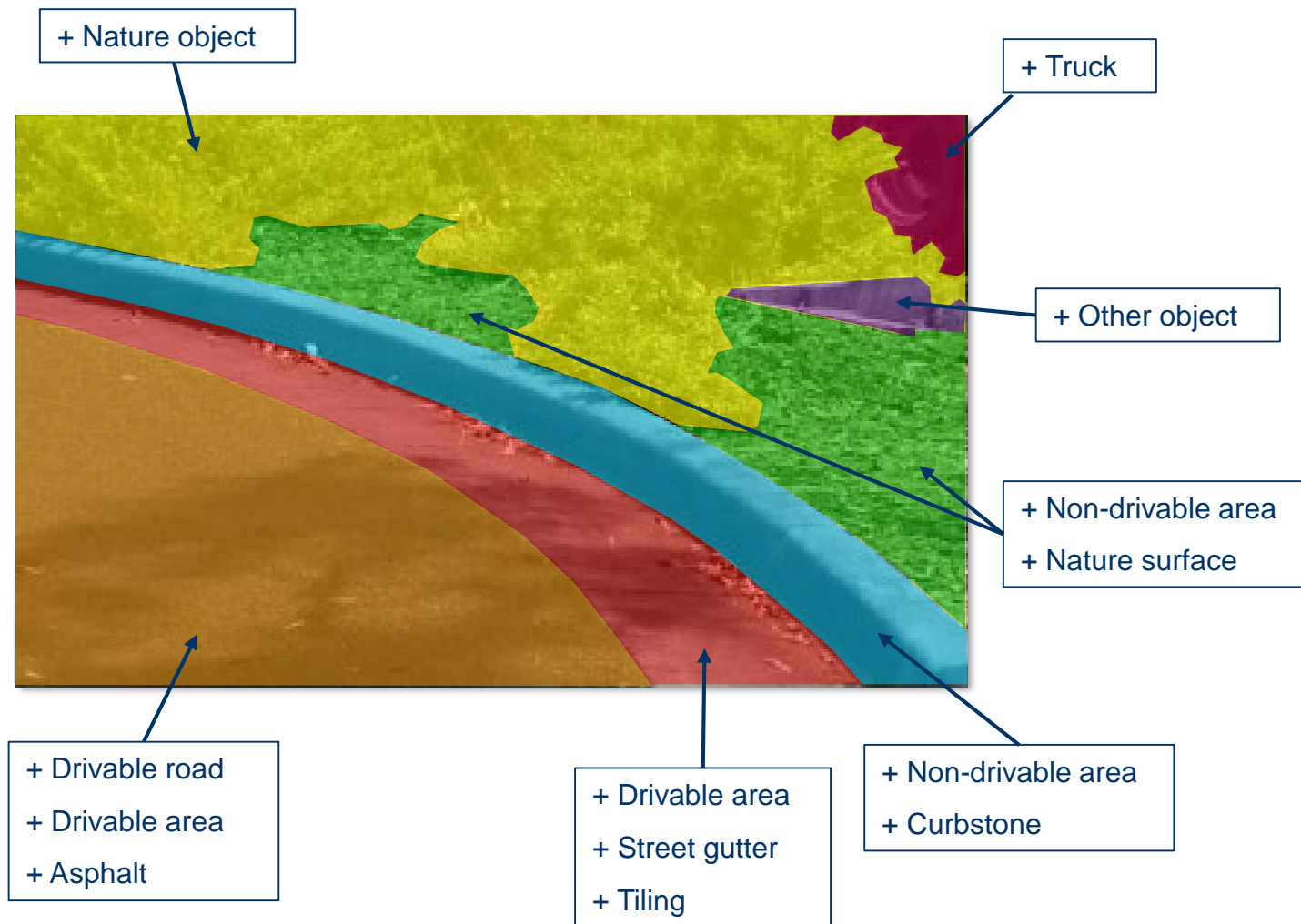
Training data with multiple classes per pixel

- ▶ Select classes for training depending on application
- ▶ Train DNN with multiple outputs per pixel



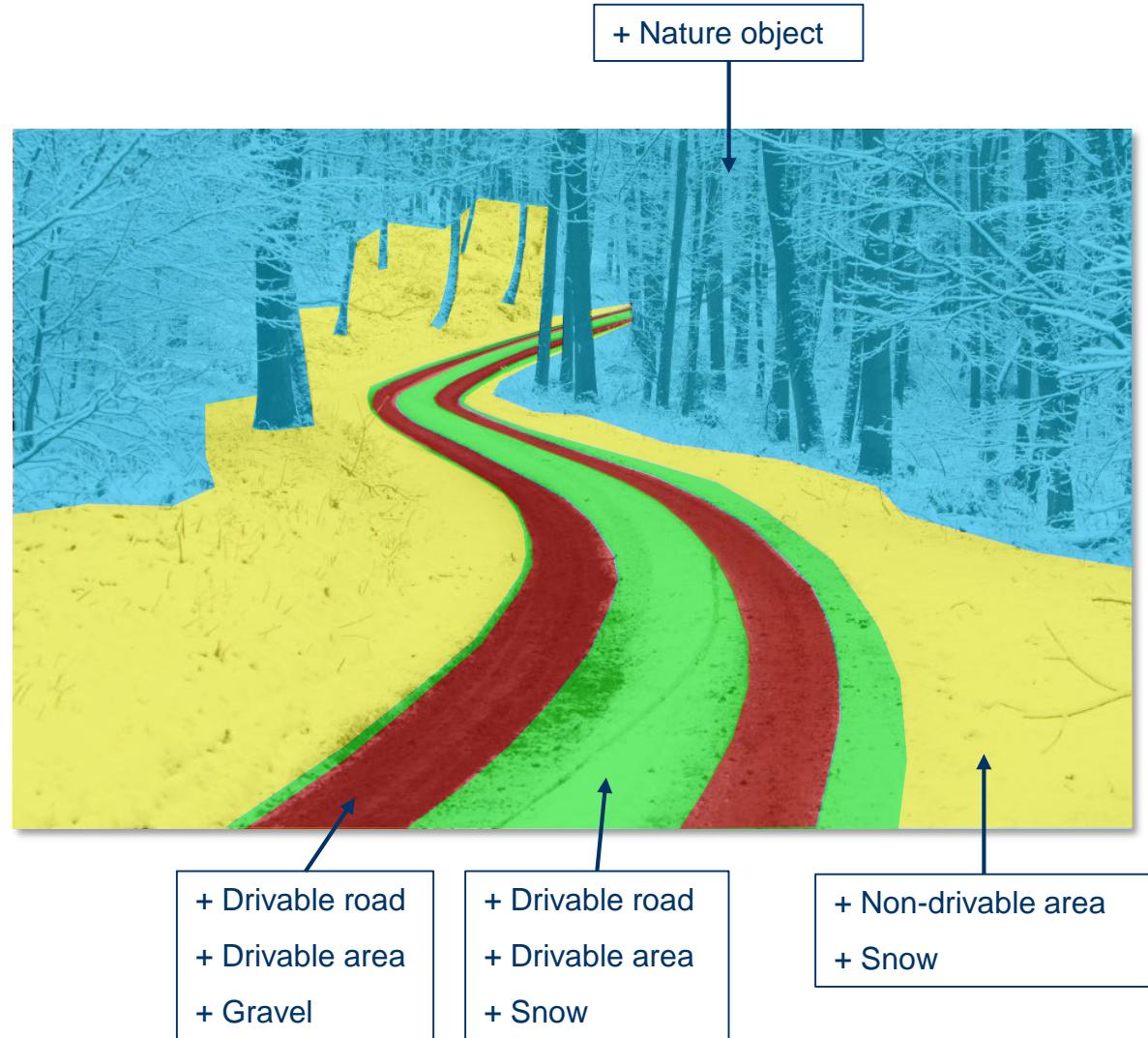
Multi Class Pixel Labeling

Ground Plane Annotation



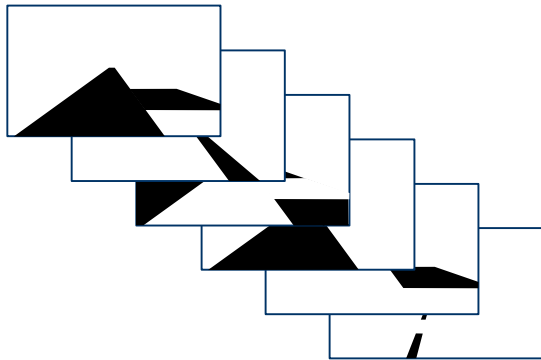
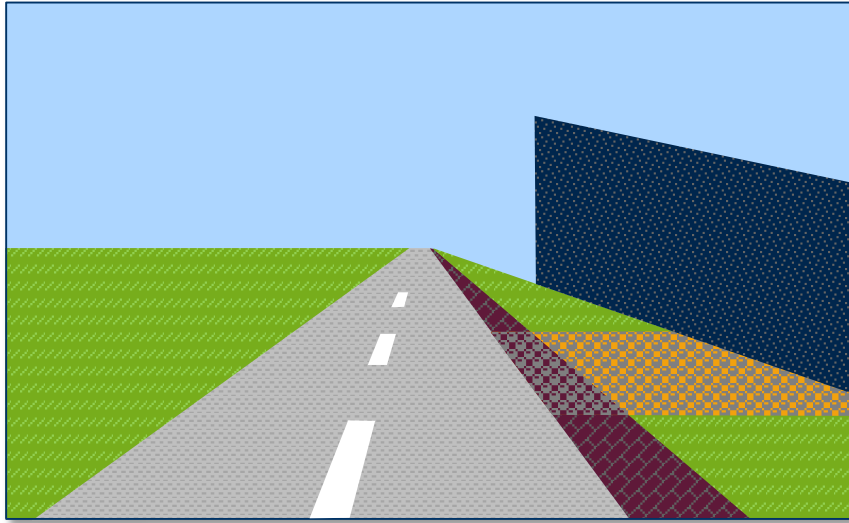
Multi Class Pixel Labeling

Ground Plane Annotation

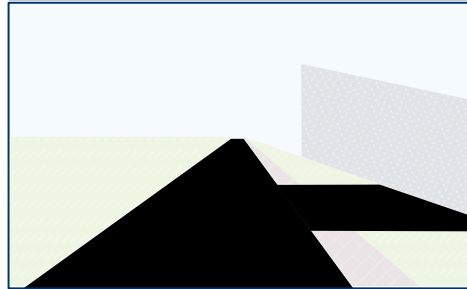


Multi Class Pixel Labeling

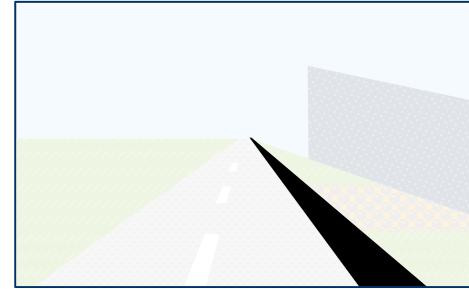
Binary Class Maps



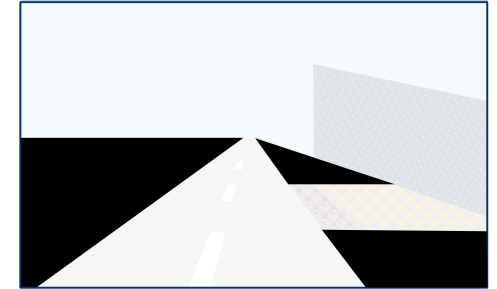
A/4: Drivable area



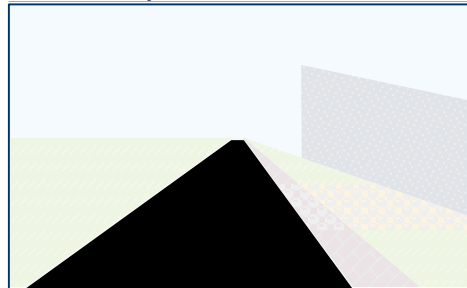
A/5: Sidewalk



A/6: Non-drivable area



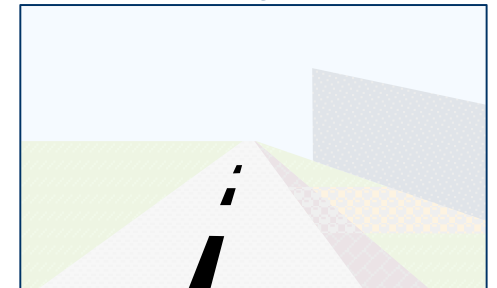
B/9: Asphalt



B/11: Cobblestone

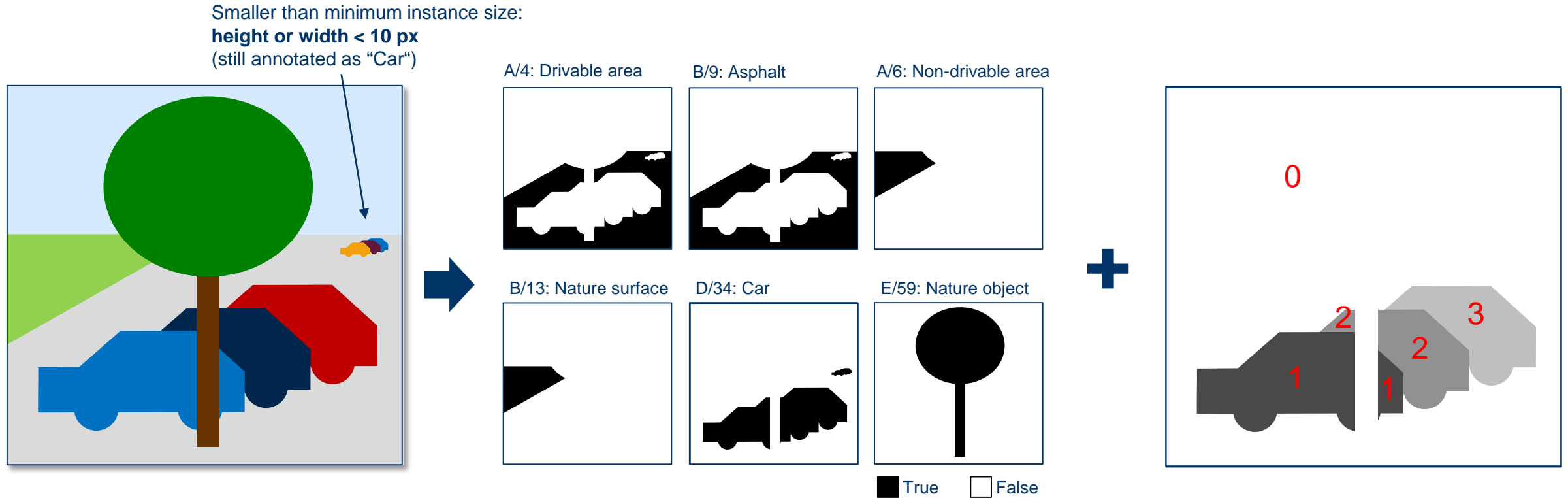


C/17: Marking: Dashed line



■ True □ False

Multi Class Pixel Labeling



Format option A:

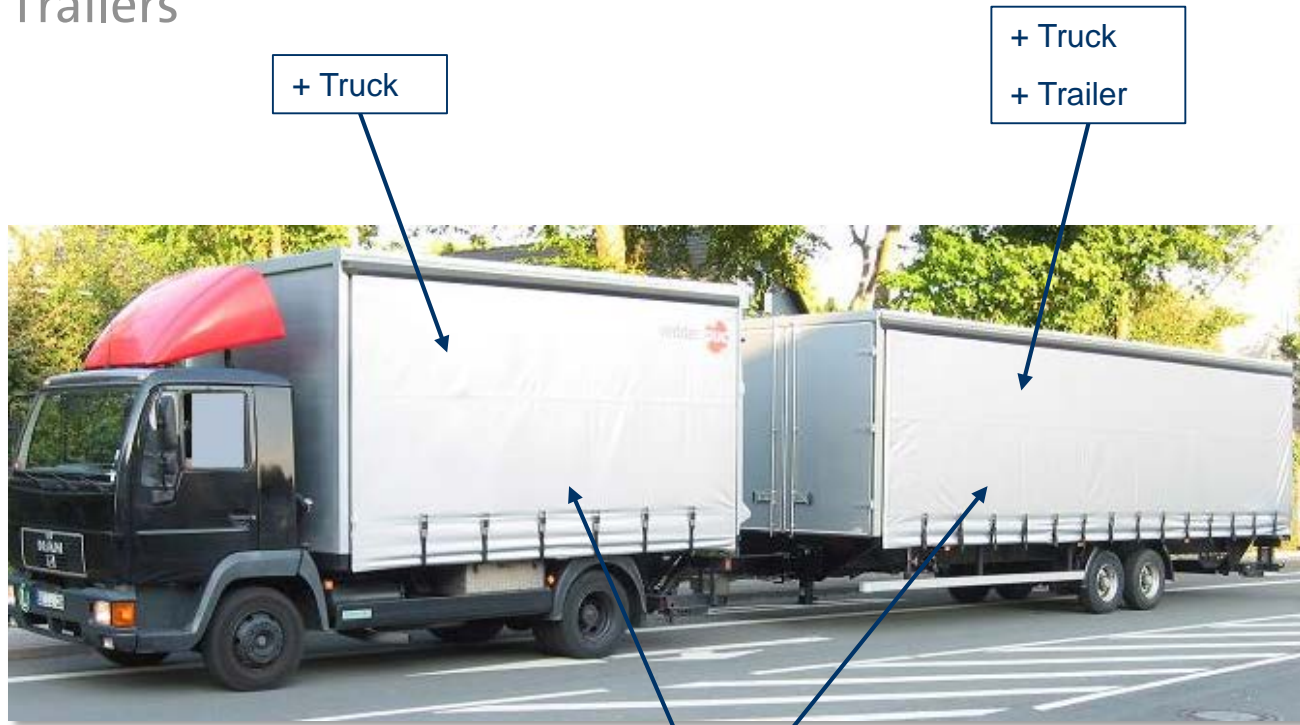
- ▶ Multi-channel (binary) image (e.g. numpy array) for all classes + 1-channel grayscale image for object instances

Format option B

- ▶ json file with polygons assigned to (multiple) classes and to object instance ID (conversion to option A should be possible)

Class descriptions

Trailers



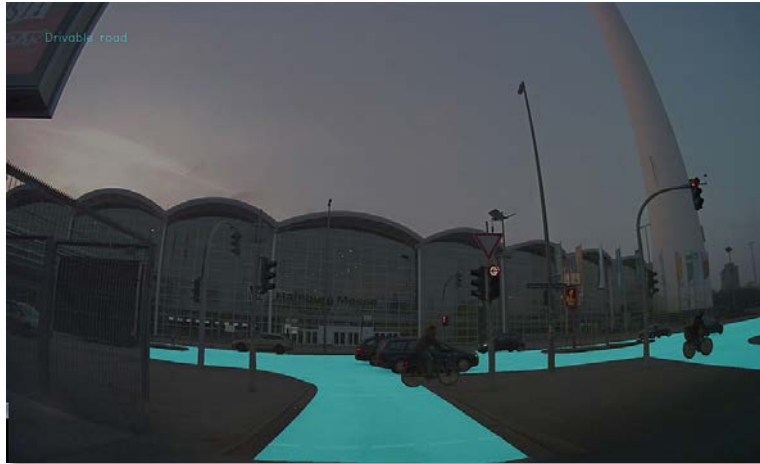
Same instance



Same instance (bicycle + rider + trailer + dog)

Multi Class Pixel Labeling

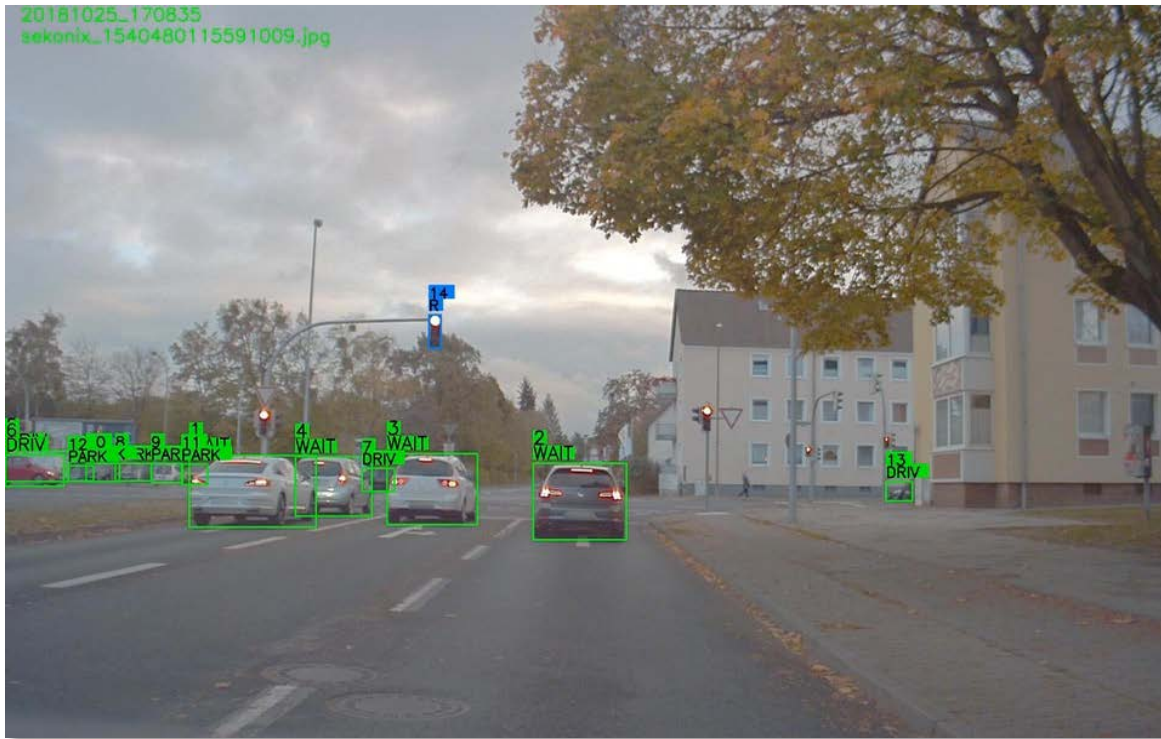
Binary Class Maps



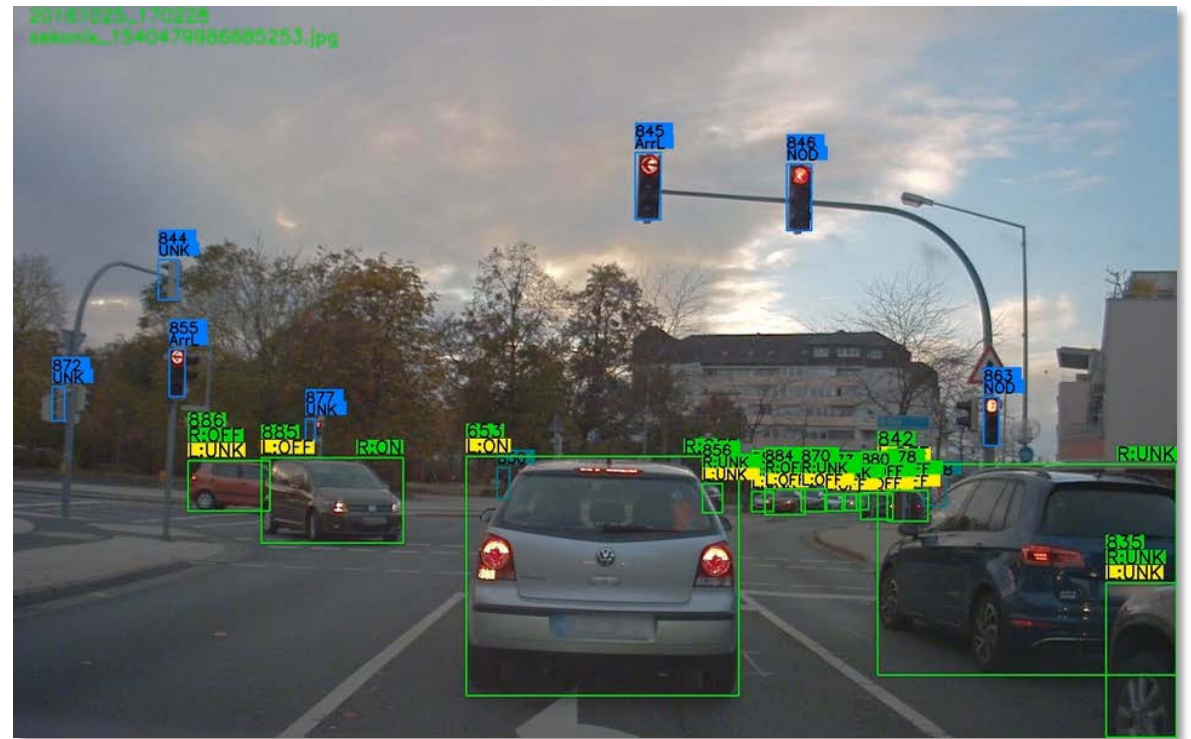
Objects and Attributes

Object Detection, Action Recognition and Prediction

Driving Action



Blinker Status



Objects and Attributes

Action Recognition and Prediction

Attribute	Possible values
action_vehicle	<i>waiting</i>
	<i>driving</i>
	<i>turn_left</i>
	<i>turn_right</i>
	<i>lane_change_left</i>
	<i>lane_change_right</i>
	<i>park_in</i>
	<i>park_out</i>
	<i>parking</i>
	...
driving_direction	<i>oncoming_vehicle</i>
	<i>preceding_vehicle</i>
	<i>cross_traffic_from_right</i>
	<i>cross_traffic_from_left</i>
blinker_right	<i>off</i>
	<i>on</i>
	<i>unknown</i>
...	...

Class	action_vehicle	driving_direction	blinker_right	blinker_left	...
car	x	x	x	x	
bus	x	x	x	x	
motorcycle	x	x	x	x	
special vehicle	x	x	x	x	
truck	x	x	x	x	
bus	x	x	x	x	
pedestrian					
bicycle		x			
traffic light					x
...					

Attributes / Actions: Pedestrian → crossing street...

Sub-objects: Vehicle → left blinker...

Connections: Truck → Trailer...

Dataset Structures

Example

Dataset X

- Folder X_sensordata_cam01 → jpg, bmp ?
- Folder X_sensordata_cam02
- Folder X_sensordata_lidar → txt, bin?
- Folder X_sensordata_radar
- Folder X_labels_semseg_cam01 → json, xml, rgb, binarymaps?
- Folder X_labels_semseg_cam02
- Folder X_labels_2DBB_cam01
- Folder X_labels_2DBB_cam02
- Folder X_labels_3DBB_lidar
- ...
- Folder X_metadata

- data description
- calibration parameters
- labeling specification
- ...

Structure and datatype conversions are easy for small datasets...

VOLKSWAGEN

AKTIENGESELLSCHAFT

GROUP INNOVATION



VOLKSWAGEN GROUP INNOVATION

INTERNATIONAL INNOVATION ECOSYSTEM