

WoodScape: A multi-task, multi-camera fisheye dataset for autonomous driving

- Supplementary Material

Senthil Yogamani, Ciarán Hughes, Jonathan Horgan, Ganesh Sistu, Padraig Varley, Derek O’Dea, Michal Uříčář, Stefan Milz, Martin Simon, Karl Amende, Christian Witt, Hazem Rashed, Sumanth Chennupati, Sanjaya Nayak, Saquib Mansoor, Xavier Perroton, Patrick Pérez

<https://github.com/valeoai/WoodScape>

firstname.lastname@valeo.com

1. Accuracy breakdown of semantic tasks

Table 1 summarizes the statistics of semantic segmentation and bounding box detection tasks in our 10k dataset. The samples are carefully chosen to ensure a good distribution of object. We refined the instance level annotations and generated bounding box annotations in PASCAL VOC style. It is be noted that these are preliminary results and final results may vary due to sampling variations and training configuration. When the baseline network is released in our github, corresponding accuracies will be provided too.

Table 1: Summary of results of baseline experiments. From top to bottom: (1) Distribution of frames across different camera positions, (2) Number of 2D box samples for each class and its corresponding mAP score using FasterRCNN and (3) Class-wise accuracies of semantic segmentation network.

Camera Distributions			
Front	Rear	Left	Right
3500	3500	750	750

Bounding Box Object Distributions			
4-Wheeler	2-Wheeler	Person	Animals
56.5k	8.7k	54k	0.5k
FasterRCNN mAP@0.5IOU			
0.63	0.17	0.44	0.01k

Semantic Segmentation (ENet) IOU scores				
4-Wheeler	2-Wheeler	Person	Road	Void
0.55	0.22	0.43	0.95	0.42

2. Annotated class hierarchy

Here we provide the complete class hierarchy for the annotation of the WoodScape dataset:

- sidewalk

- curb
- free space
- sidewalk
- road
 - road
 - road surface
- four wheelers
 - dynamic car
 - car
 - dynamic van
 - van
 - caravan
- four wheelers heavy
 - dynamic bus
 - dynamic truck
 - dynamic rv
 - dynamic trailer
 - dynamic on rails
 - bus
 - truck
 - trailer
 - train/tram
- four wheelers groups
 - dynamic vehicle group
 - grouped vehicles
- two wheelers
 - dynamic bicycle

- dynamic motorcycle
 - bicycle
 - motorcycle
- two wheelers groups
 - dynamic cycle group
- person
 - dynamic person
 - person
- rider
 - dynamic rider
 - rider
- person group
 - dynamic person group
- lane marks
 - road lane line
 - road parking line
 - lane marking
 - parking marking
 - zebra crossing
 - parking line
- ground markings
 - road reserved parking
 - empty reserved
 - other ground marking
 - zebra crossing
- catseye bottsdots
 - road botts dots
 - road group botts dots
- animals
 - dynamic animal
 - dynamic animal group
 - animal
 - grouped animals
- structure
 - fence guardrail
 - structure
 - construction
 - fence
- object

- dynamic object
 - pole-post
 - pole
- traffic signs
 - other traffic sign
 - other traffic sign indistinguishable
 - red traffic light
 - yellow traffic light
 - green traffic light
 - unknown traffic light
 - unknown traffic light
 - traffic light red
 - traffic sign indistinguishable
 - traffic light green
 - traffic sign
 - traffic light yellow
- nature
 - terrain
 - nature
- sky
 - sky
- void
 - other nosight
 - other blur
 - ego vehicle
 - green strip
 - dynamic
 - background
 - void
 - cats eyes and botts dots
 - movable object
 - grouped botts dots
 - other wheeled transport
 - grouped pedestrian and animals

3. Synthetic Data

The synthetic data is generated using a commercial physically accurate simulation engine **ANYVERSE**. Our camera model including intrinsics and extrinsics were incorporated as illustrated in Figure 1. Annotation for instance segmentation and depth estimation will be provided.

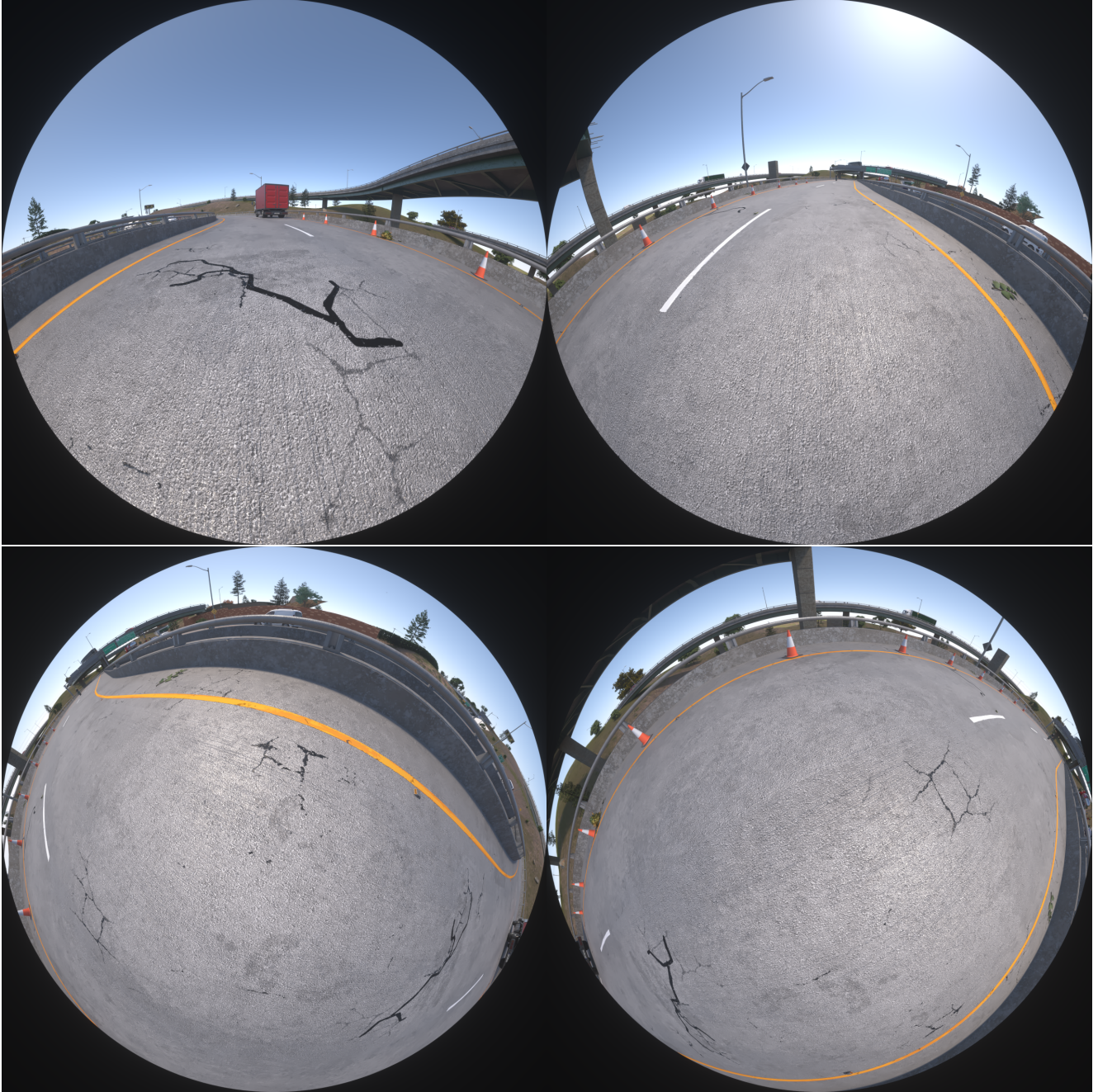


Figure 1: Illustration of synthetic images generated using our fisheye camera model.