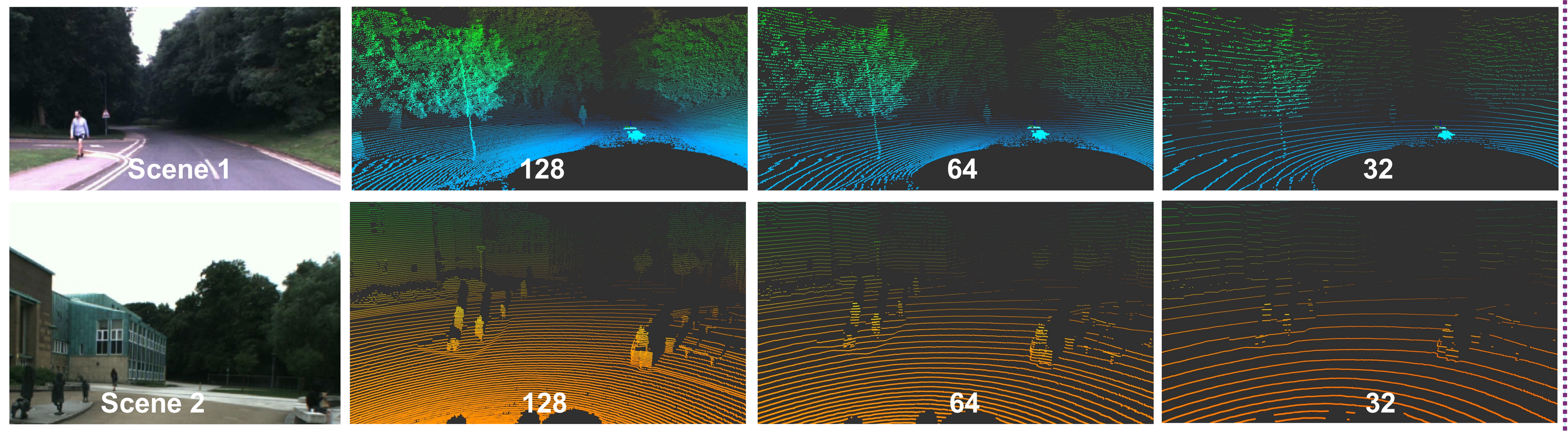
DurLAR: A High-Fidelity 128-Channel LiDAR Dataset with Panoramic Ambient and Reflectivity Imagery for Multi-Modal Autonomous Driving Applications

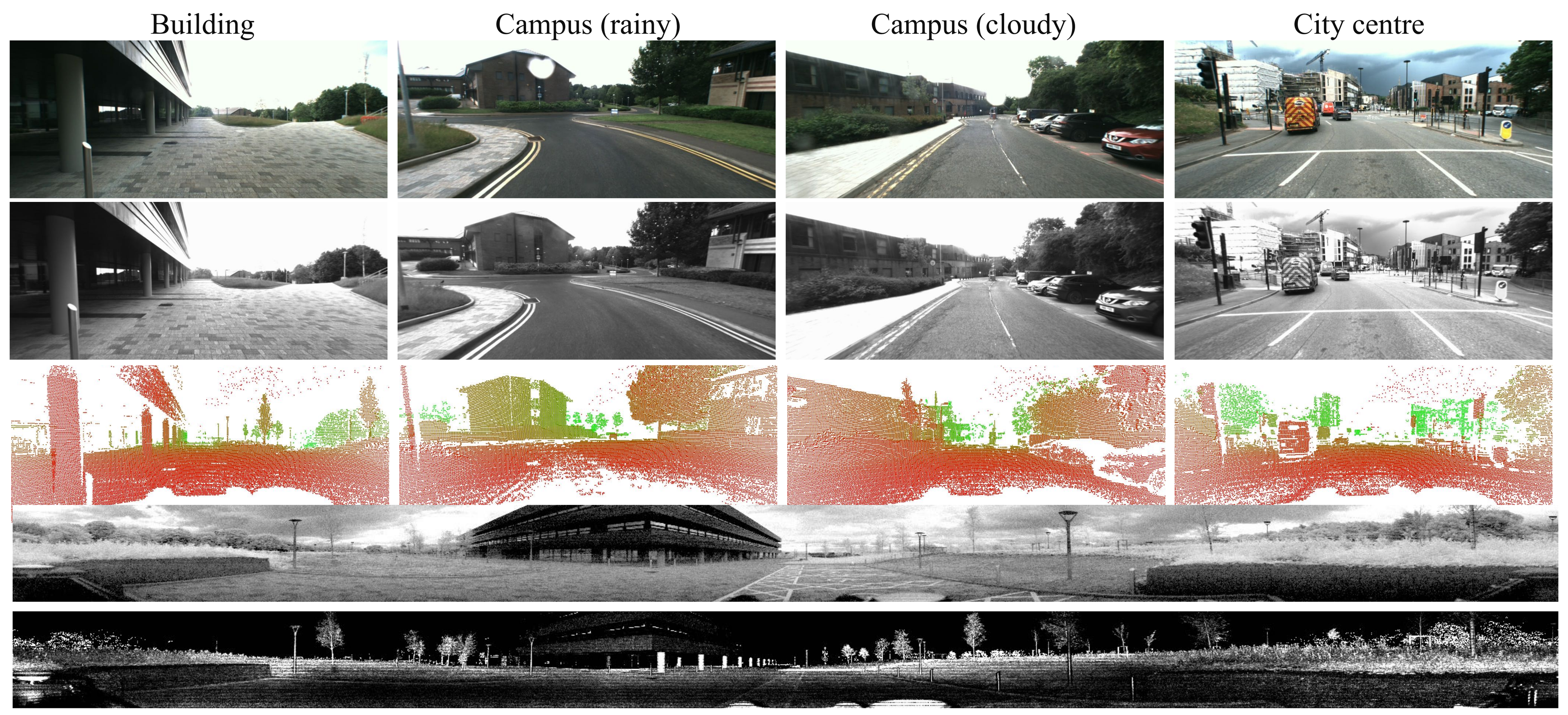
Li Li, Khalid N. Ismail, Hubert P. H. Shum, Toby P. Breckon, Durham University, UK

Main Contributions

A novel **large-scale dataset** comprising contemporary high-fidelity 3D **LiDAR** (**128** channels), **stereo/ambient/reflectivity imagery**, **GNSS/INS** and **illumination** information under repeated route, variable environment conditions. The first autonomous driving task dataset to additionally comprise usable **ambience and reflectivity** LiDAR obtained imagery.



LiDAR point clouds from two exemplar scenes with differing vertical LiDAR resolution (left to right: colour RGB images, [32 → 64 → 128] LiDAR vertical channels).



Examples from DurLAR dataset. From top to bottom, RGB left camera images (top), grayscale right camera images (2nd row), LiDAR point cloud (3rd row), ambient (near infrared, 4th row) and reflectivity (bottom) panoramic images.

Dataset	Res	Range/m	Diversity	Image	#Frames	Other sensors
DENSE	64	120	E/W/T	I	1M	D/M/F/T/B
H3D	64	120	E	I	28k	G/M
KITTI | SemanticKITTI	64	120	E	I	93k	N/S/G/M/B
KITTI-360	64	120	E	I	320k	N/S/G/M/B
LiVi-Set	32	100	E	I	10k	
Lyft Level 5	64	200	E/W/T	I	170k	D/B
nuScenes	32	100	E/W/T	I	1M	M/D/B
Oxford RobotCar	4	50	E/W/T	I	3M	N/S/G/M/B
Stanford Track	64	120	E	I	14k	M
Sydney Urban Objects	64	120	E	I	0.6k	
➡ **DurLAR (ours)**	**128**	**120**	**E/W/T/L**	**I/A/R**	**100k**	**U/N/S/G/M/B**

An exemplar **monocular depth estimation benchmark** to compare the performance of supervised/self-supervised variants of three leading approaches [1, 2, 3] when trained and evaluated on low resolution (KITTI [4]), high resolution (DurLAR) ground truth LiDAR depth, or our novel KITTI/DurLAR dataset partition, with the observation that increased resolution enables superior monocular depth estimation performance via our joint supervised/self-supervised loss formulation.

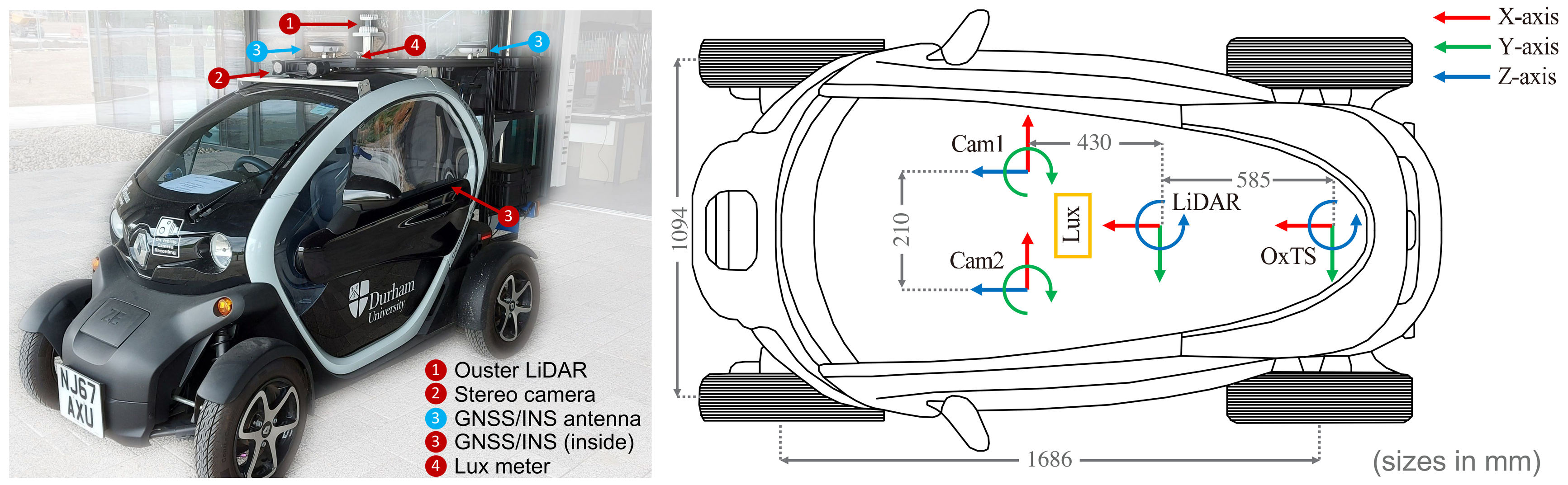
Sensor Setup

LiDAR: Ouster OS1-128 LiDAR sensor with 128 channels vertical resolution, 2048 @10 Hz rotation rate mode).

Stereo Camera: Carnegie Robotics MultiSense S21 stereo camera with grayscale, colour, and IR enhanced imagers.

GNSS/INS: OxTS RT3000v3 global navigation satellite and inertial navigation system, 0.03º pitch/roll accuracy, 0.1-1.5m position accuracy.

Lux Meter: Yocto Light V3, a USB ambient light sensor (lux meter).

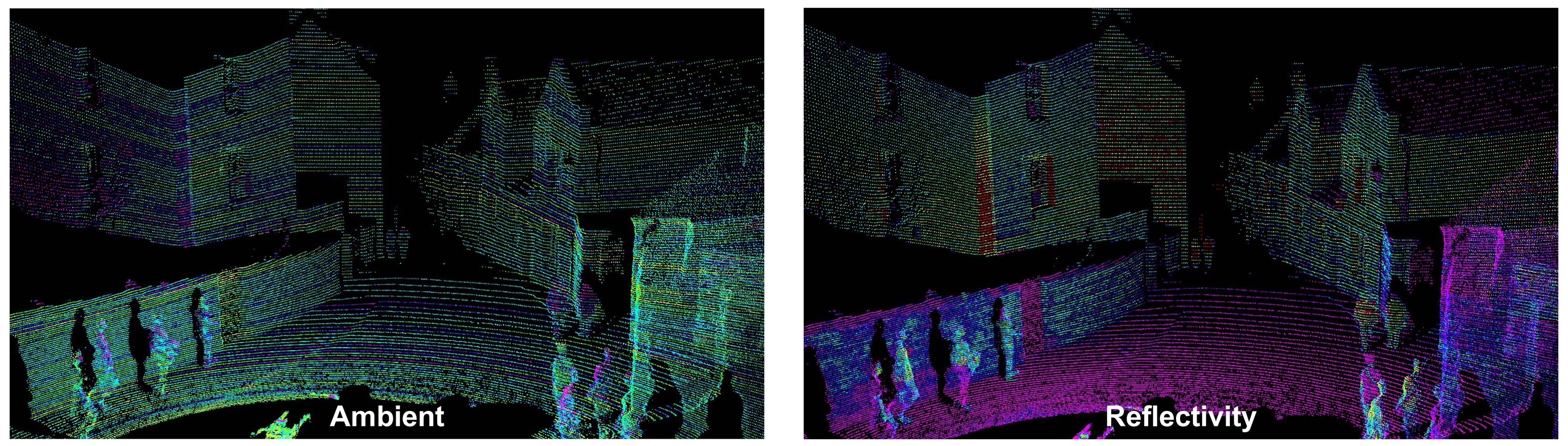


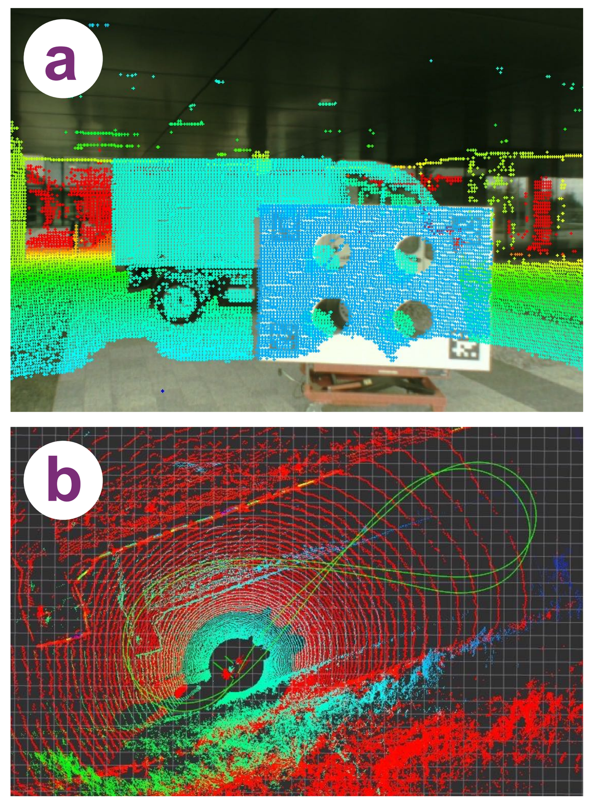
Test vehicle (Renault Twizy) and **sensor placements**. Coordinate axes follow the right-hand rule.

The DurLAR Dataset

Ambient images: day/night scene visibility in the near-IR spectrum.

Reflectivity images: information indicative of the material properties of the object itself and offer good consistency across illumination conditions and range.





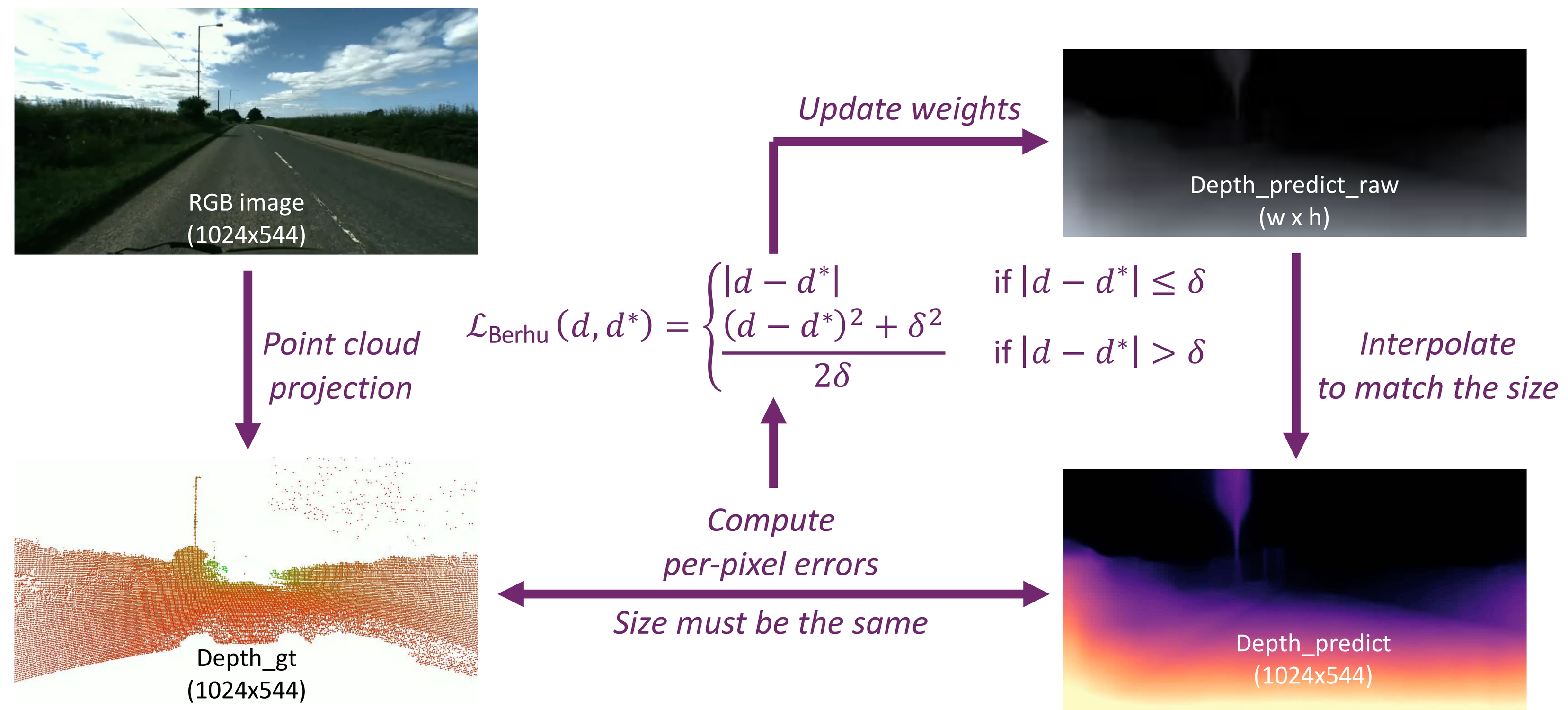
Calibration and Synchronisation

All sensor synchronisation is performed at a rate of **10 Hz**, using ROS (version Noetic) timestamps operating over a Gigabit Ethernet backbone to a common host (Intel Core i5, 16 GB RAM).

Sensor	Collecting rate	External calibration
LiDAR	10Hz	(a) Stereo; (b) GNSS/INS
GNSS/INS	100Hz	(b) LiDAR; stereo
Stereo	30Hz	(a) LiDAR; GNSS/INS
Lux meter	30Hz	

Monocular Depth Estimation

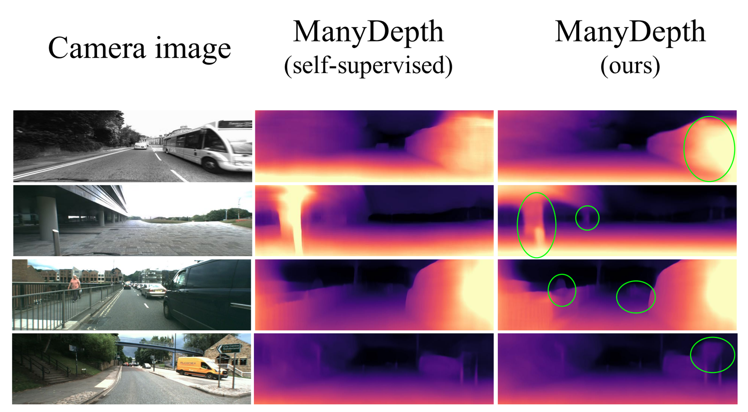
Leading approach for monocular depth estimation: ManyDepth [1].



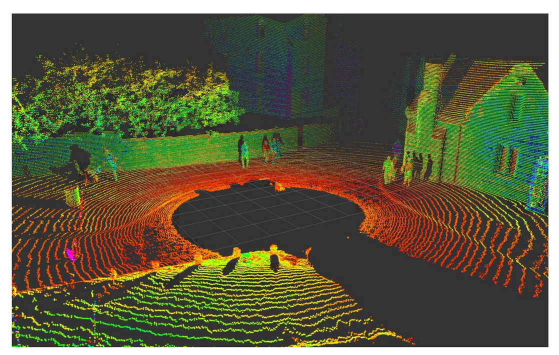
Dataset	Method	+S	Abs Rel	Sq Rel	RMSE	RMSE log	$\delta < 1.25$	$\delta < 1.25^2$	$\delta < 1.25^3$
KITTI	ManyDepth (MR)	x	0.098	0.770	4.459	0.176	0.900	0.965	0.983
	ManyDepth (MR)	x	0.093	0.715	4.245	0.172	0.909	0.966	0.983
Cityscapes	ManyDepth	x	0.114	1.193	6.223	0.170	0.875	0.967	0.989
DurLAR	Depth-hints	x	0.122	1.070	4.148	0.211	0.870	0.946	0.972
	Depth-hints	√	0.121	1.109	4.121	0.210	0.874	0.946	0.972
	MonoDepth2	x	0.111	1.114	4.002	0.187	0.895	0.960	0.981
	MonoDepth2	√	<u>0.108</u>	<u>1.010</u>	<u>3.804</u>	0.185	0.898	0.963	0.982
	ManyDepth (MR)	x	0.115	1.227	4.116	0.186	0.892	0.962	0.982
	ManyDepth (MR)	√	0.109	0.936	3.711	<u>0.176</u>	0.895	0.964	<u>0.984</u>
	ManyDepth (HR)	x	0.109	1.111	3.875	0.177	<u>0.901</u>	<u>0.966</u>	<u>0.984</u>
	ManyDepth (HR)	√	0.104	0.936	3.639	0.171	0.906	0.969	0.986

Performance comparison over the KITTI Eigen split [4], Cityscapes [5] and DurLAR datasets.

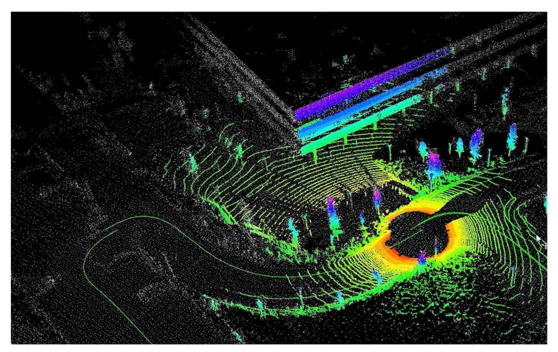
Config.		Abs Rel	Sq Rel	RMSE	RMSE log	$\delta < 1.25$	$\delta < 1.25^2$	$\delta < 1.25^3$
Cross-datasets	K	0.159	1.536	5.101	0.244	0.798	0.923	0.963
	D	0.189	1.764	5.580	0.264	0.758	0.908	0.959
	K+D	0.188	1.941	5.182	0.262	0.769	0.912	0.958
	D+K	0.151	1.123	4.744	0.233	0.805	0.927	0.967
Ablation - Channels	32/+S	0.115	0.908	3.677	0.179	0.888	0.966	0.985
	64/+S	0.107	0.918	3.735	0.175	0.895	0.967	0.986
	128/-S	0.109	1.111	3.875	0.177	0.901	0.966	0.984
	128/+S	0.104	0.936	3.639	0.171	0.906	0.969	0.986



Future Applications



reflectivity/ambient



SLAM

Driver Attention Monitoring

[1] Watson et al., The Temporal Opportunist: Self-Supervised Multi-Frame Monocular Depth. CVPR, 2021.
[2] Godard et al., Digging Into Self-Supervised Monocular Depth Estimation. CVPR, 2019.
[3] Zhou et al., Unsupervised learning of depth and ego-motion from video. CVPR, 2017.
[4] Geiger et al., Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite. CVPR, 2012.
[5] Cordts et al., The Cityscapes Dataset for Semantic Urban Scene Understanding. CVPR, 2016.

Online access for the dataset, github.com/l1997i/DurLAR.