

1 Depth Evaluations

	<i>Stereo</i>		<i>Multi-Frame</i>		<i>Add. Sup.</i>		Error metric ($\downarrow$)				Accuracy metric ($\uparrow$)		
	<i>Train</i>	<i>Infer.</i>					Abs Rel	Sq Rel	RMSE	RMSE log	$\delta < 1.25$	$\delta < 1.25^2$	$\delta < 1.25^3$
UnOS	✓	✓					0.049	0.515	3.404	0.121	0.965	0.984	0.992
EffiScene	✓	✓					0.049	0.522	3.461	0.120	0.961	0.984	0.992
Xiang et al.	✓	✓					0.048	0.487	3.447	0.117	0.964	0.985	0.992
Liu et al. (stereo)	✓	✓					0.051	0.532	3.780	0.126	0.957	0.982	0.991
Liu et al. (mono)	✓						0.108	1.020	5.528	0.195	0.863	0.948	0.980
Hur et al. 2020	✓						0.125	0.978	4.877	0.208	0.851	0.950	0.978
PLADE-Net	✓						0.089	0.590	4.008	0.172	0.900	0.967	0.985
DRAFT			✓		S		0.097	0.647	3.991	0.169	0.899	0.968	0.984
ManyDepth			✓				0.090	0.713	4.261	0.170	0.914	0.966	0.983
DepthFormer			✓				0.090	0.661	4.149	0.175	0.905	0.967	0.984
Zhang et al.					I		0.108	0.761	4.608	0.187	0.883	0.962	0.982
Dyna-DM					m		0.115	0.785	4.698	0.192	0.871	0.959	0.982
Boulahbal et al.					m		0.110	0.719	4.486	0.184	0.878	0.964	0.984
SC-DepthV3					D		0.118	0.756	4.709	0.188	0.864	0.960	0.984
GeoNet							0.155	1.296	5.857	0.233	0.793	0.931	0.973
Zhao et al.							0.113	0.704	4.581	0.184	0.871	0.961	0.984
Ours-LM							0.112	0.758	4.505	0.183	0.873	0.959	0.984

Table 8: Additional depth evaluation on the KITTI Dataset. *Stereo Train* and *Stereo Infer.* stand for the use of stereo-view during training and inference, respectively. *Multi-Frame* denotes the use of multiple frames as inputs during inference. *Add. Sup.* stands for additional supervision during training where m, I, S, and D stand for mask-level, IMU-level, synthetic-data, and pretrained depth model supervision, respectively.

2 Odometry Results

	nuScenes	Waymo
Speed (m/s)	5.097 ± 3.786	6.975 ± 6.025
# snippets	23002	19564
Monodepth2*	0.016 ± 0.018	0.017 ± 0.029
LiteMono*	0.017 ± 0.017	0.022 ± 0.043
Ours-MD2 [†]	0.017 ± 0.020	0.016 ± 0.025
Ours-LM [†]	0.018 ± 0.022	0.017 ± 0.037
Ours-LM	0.016 ± 0.019	0.017 ± 0.035

Table 9: Odometry results on the nuScenes and Waymo Open Dataset. Results show the average absolute trajectory error over all overlapping five-frame snippets in the test sequences, and standard deviation, in meters. For consistency, the first 100 test sequences for both datasets are evaluated.