1. Different error metrics
   2.1 https://github.com/janivanecky/Depth-Estimation
   2.2 https://fenix.tecnico.ulisboa.pt/downloadFile/1689244997256744/Thesis.pdf

Table 4.2: Comparison of results on the KITTI 2012 dataset.

KITTI 2012	Saxena	Eigen	Liu	Mancini	Our method
mean log	-	-	**0.211**	-	0.372
relative abs	0.280	**0.190**	0.217	-	0.525
relative square	3.012	**1.515**	-	-	2.311
linear RMS	8.734	7.156	**7.046**	7.508	13.093
log RMS	0.361	**0.270**	-	0.524	0.590
scale-invariant	0.327	0.246	-	**0.196**	0.347

Table 4.3: Results on the KITTI 2015 dataset.

KITTI 2015	Our method
mean log	0.291
relative abs	0.322
relative square	1.921
linear RMS	12.498
log RMS	0.452
scale-invariant	0.204

2. Android implementations
3. Segmentation from depth images

Segmenting Unknown 3D Objects from Real Depth Images using Mask R-CNN Trained on Synthetic Data

Michael Danielczuk[1], Matthew Matl[1], Saurabh Gupta[1],
Andrew Li[1], Andrew Lee[1], Jeffrey Mahler[1], Ken Goldberg[1,2]

Abstract— **The ability to segment unknown objects in depth images has potential to enhance robot skills in grasping and object tracking. Recent computer vision research has demonstrated that Mask R-CNN can be trained to segment specific categories of objects in RGB images when massive hand-labeled datasets are available. As generating these datasets is time-consuming, we instead train with synthetic depth images. Many robots now use depth sensors, and recent results suggest training on synthetic depth data can transfer successfully to the real world. We present a method for automated dataset generation and rapidly generate a synthetic training dataset of 50,000 depth images and 320,000 object masks using simulated heaps of 3D CAD models. We train a variant of Mask R-CNN with domain randomization on the generated dataset to perform category-agnostic instance segmentation without any hand-**

Fig. 1: Color image (left) and depth image segmented by SD Mask RCNN (right) for a heap of objects. Despite clutter, occlusions, and complex geometries, SD Mask RCNN is able to correctly mask each of the objects.

4. Generate high-quality depth maps
5. Improve this model by replacing the encoding part with a pretrained network.
6. Loss functions play an important role in solving this problem. Tuning the loss functions may yield significant improvement.
7. FastDepth: Fast Monocular Depth Estimation on Embedded Systems
   Replace the encoder-decoder with a fast efficient and lightweight encoder-decoder network architecture and apply network pruning to further reduce computational complexity and latency.
   Replace the network with fast computing encoder-decoder network and compare the results. Also, work on producing the high definition images using the work of https://paperswithcode.com/paper/high-quality-monocular-depth-estimation-via#code

TABLE I

A SUMMARY OF DEEP LEARNING-BASED MONOCULAR DEPTH ESTIMATION. "MONO." REFERS TO "MONOCULAR", AND "MULTI-TASKS" MEANS THAT IN ADDITION TO POSE AND DEPTH ESTIMATION, THERE ARE OTHER TASKS THAT ARE JOINTLY TRAINED IN THE FRAMEWORK, SUCH AS SEMANTIC SEGMENTATION, MOTION SEGMENTATION, OPTICAL FLOW, ETC.

8.

			Supervised (Sup) manner			
Methods	Years	Training set	Sup	Semi-sup	Unsup	Main contributions
Eigen *et al.* [35]	2014	RGB + Depth	√			CNNs
Li *et al.* [66]	2015	RGB + Depth	√			hierarchical CRFs
Liu *et al.* [69]	2015	RGB + Depth	√			continuous CRF
Wang *et al.* [70]	2015	RGB + Depth	√			Semantic labels, hierarchical CRFs
Shelhamer *et al.* [52]	2015	RGB + Depth	√			Fully CNNs
Eigen *et al.* [51]	2015	RGB + Depth	√			Multi-task
Szegedy *et al.* [59]	2015	RGB + Depth	√			Inception Module
Mousavian *et al.* [71]	2016	RGB + Depth	√			Multi-task
Roy *et al.* [73]	2016	RGB + Depth	√			RFs
Mayer *et al.* [33]	2016	RGB + Disparity	√			Multi-task
Laina *et al.* [54]	2016	RGB + Depth	√			Residual learning
Jung *et al.* [63]	2017	RGB + Depth	√			Adversarial learning
Kendall *et al.* [49]	2017	Stereo images + Disparity	√			Disparity Loss
Zhang *et al.* [56]	2018	RGB + Depth	√			Task-attentional, BerHu loss
Xu *et al.* [72]	2018	RGB + Depth	√			Continuous CRF, structured attention
Gwn *et al.* [79]	2018	RGB + Depth	√			Conditional GAN
Fu *et al.* [60]	2018	RGB + Depth	√			Ordinal regression
Facil *et al.* [27]	2019	RGB + Depth	√			Camera model
Woft *et al.* [61]	2019	RGB + Depth	√			Lightweight network

9. Estimate the pose and depth from incoming data and find the error in the actual pose and predicted pose and introduce penalty to the latent space or gaussian process
10. Functionality testing with different Gaussian process parameters
11. Change the cost volume