Rapid Depth from Single Image

Nazgul Mamasheva
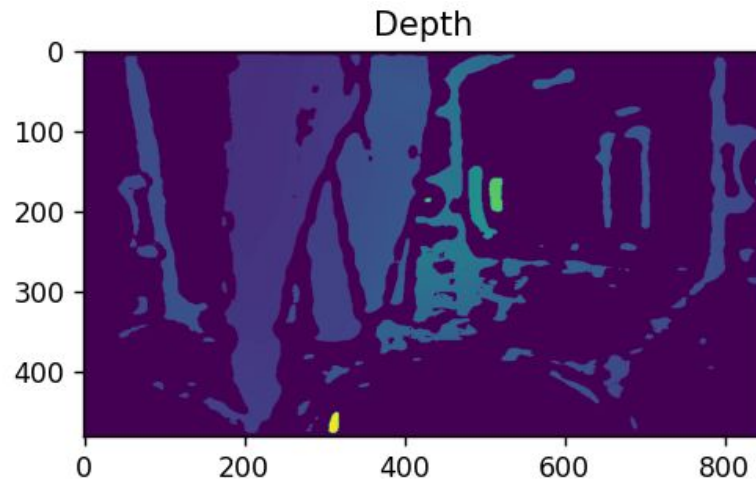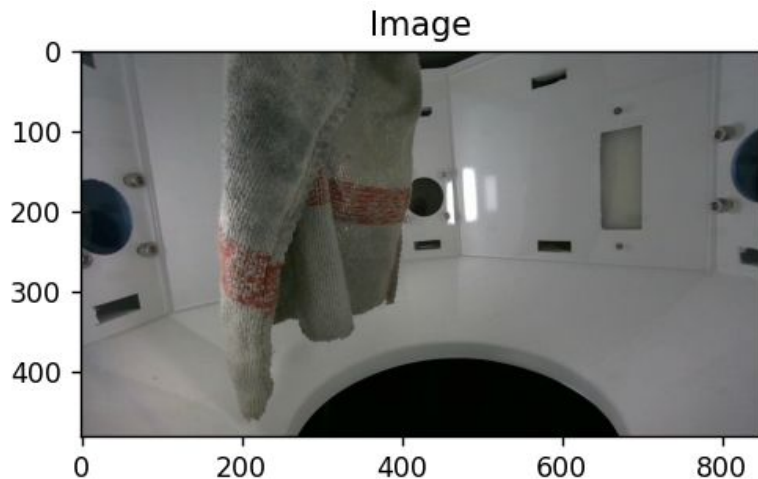
Project task

In this project we want to investigate
if depth information can also be generated
from a single image

Dataset

PTP data consists of:

- images
- depth maps (ground truth)

set of pairs of sample.jpg - sample.npy (depth map)

Dataset

70 pairs of PTP images with depth maps were given as a dataset.

The given depth maps are with missing values.

As in the examples below we can see images and mask images. The black area on the mask images represents the values "0" in depth maps, which are missing values.
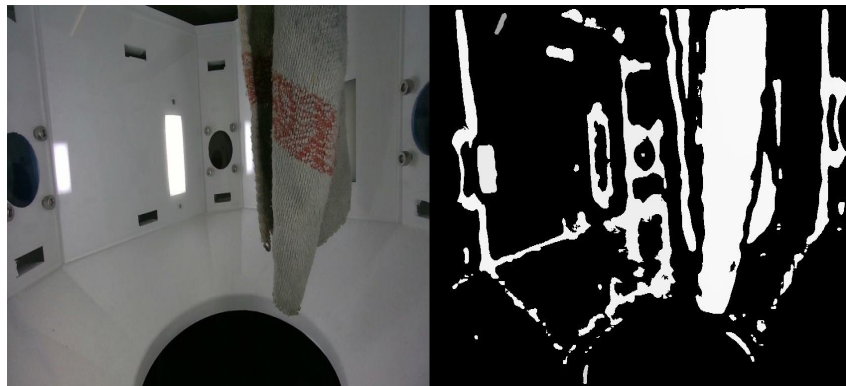
Solution for missing data on depth maps

Missing depth values are filled in using the inpainting method of Colorization using Optimization.

colorization scheme was used in:

- NYU Depth Dataset V2

- KITTI dataset

Missing values examples in NYU Depth Dataset V2

NYU Depth v2 is a dataset that provides images and depth maps for different indoor scenes captured at a resolution of 640 × 480. The dataset contains 120K training samples and 654 testing samples.
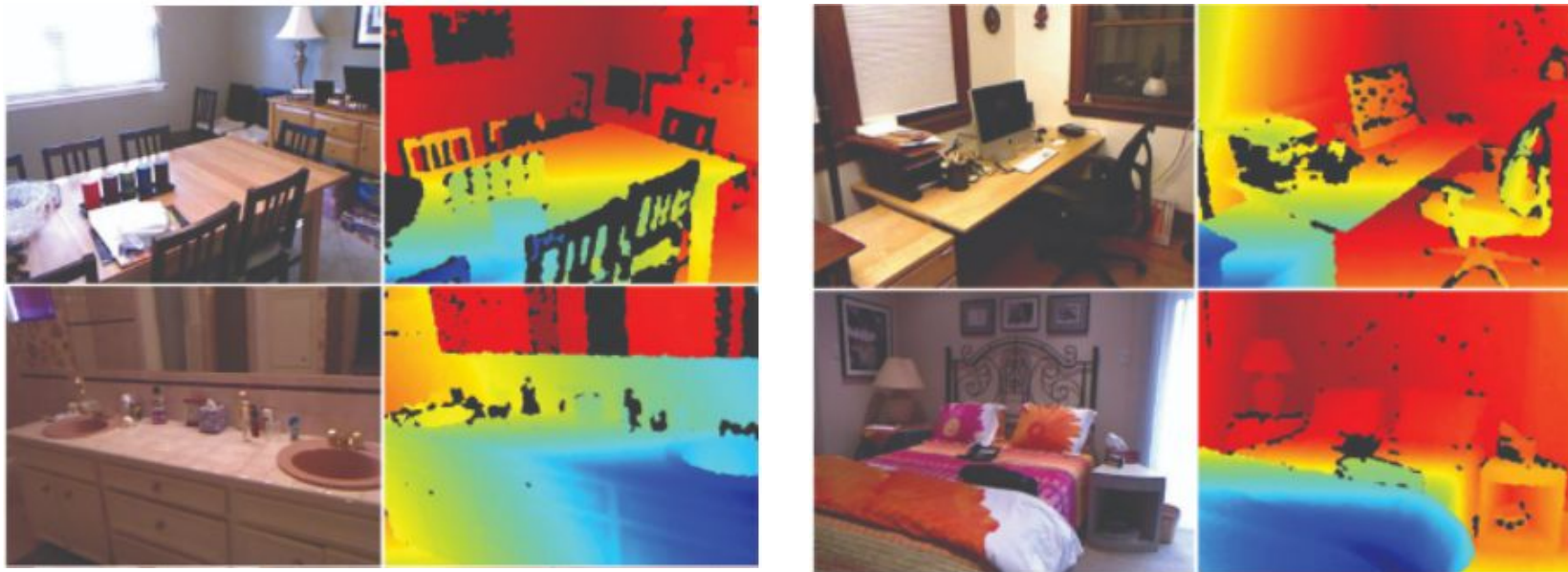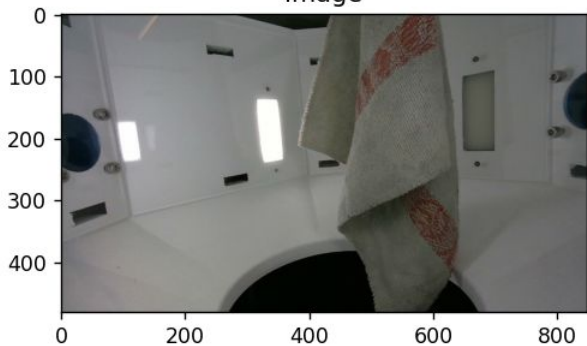
Missing values example in KITTI

KITTI is a dataset that provides stereo images and corresponding 3D laser scans of outdoor scenes captured using equipment mounted on a moving vehicle. The RGB images have a resolution of around 1241 × 376 while the corresponding depth maps are of very low density with lots of missing data.
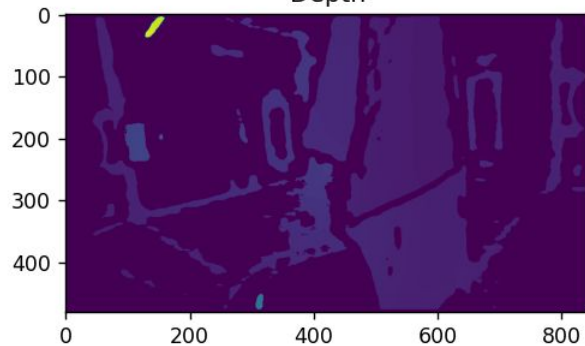
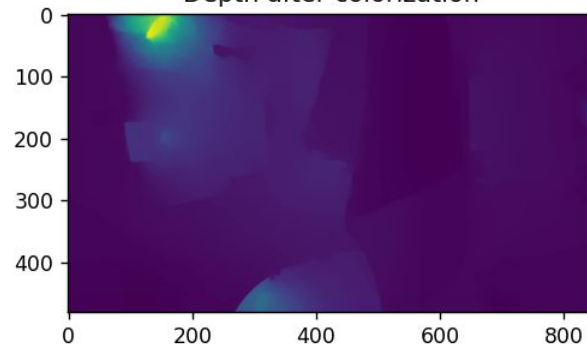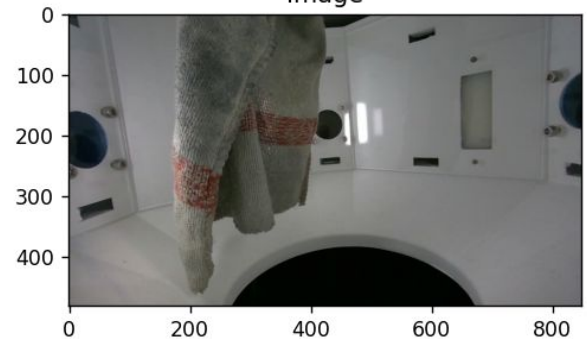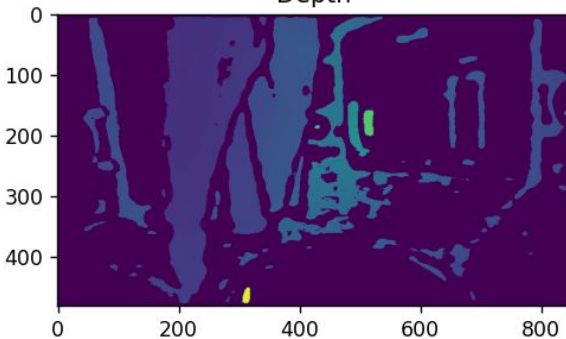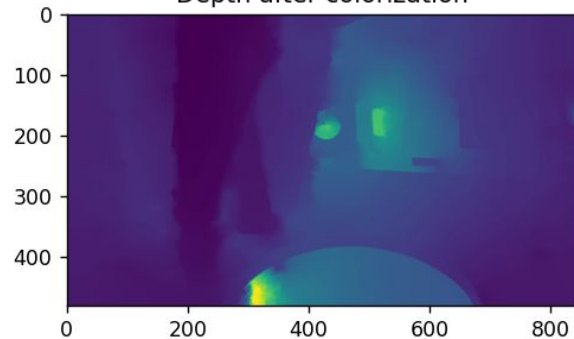# PTP dataset after colorization scheme

MIM Depth Estimation

Masked Image Modeling Depth Estimation

Ranking on KITTI dataset - Monocular-depth-estimation-on-kitti-eigen

Ranking on NYU-Depth V2 - Monocular-depth-estimation-on-nyu-depth-v2

Revealing the Dark Secrets of Masked Image Modeling (Depth Estimation) [Paper]

Ranked #5 Monocular Depth Estimation on NYU-Depth V2   State of the Art Monocular Depth Estimation on KITTI Eigen split

Main results

Results on PTP

Backbone	d1	d2	d3	abs_rel	rmse	rmse_log
Swin-v2-Base	0.5813	0.9641	0.9869	0.2617	0.0011	0.2639
Swin-v2-Large	0.8587	0.9614	0.9837	0.1305	0.0010	0.2024

Results on NYUv2

Backbone	d1	d2	d3	abs_rel	rmse	rmse_log
Swin-v2-Base	0.935	0.991	0.998	0.044	0.304	0.109
Swin-v2-Large	0.949	0.994	0.999	0.036	0.287	0.102

Results on KITTI

Backbone	d1	d2	d3	abs_rel	rmse	rmse_log
Swin-v2-Base	0.976	0.998	0.999	0.052	2.050	0.078
Swin-v2-Large	0.977	0.998	1.000	0.050	1.966	0.075

Visual results
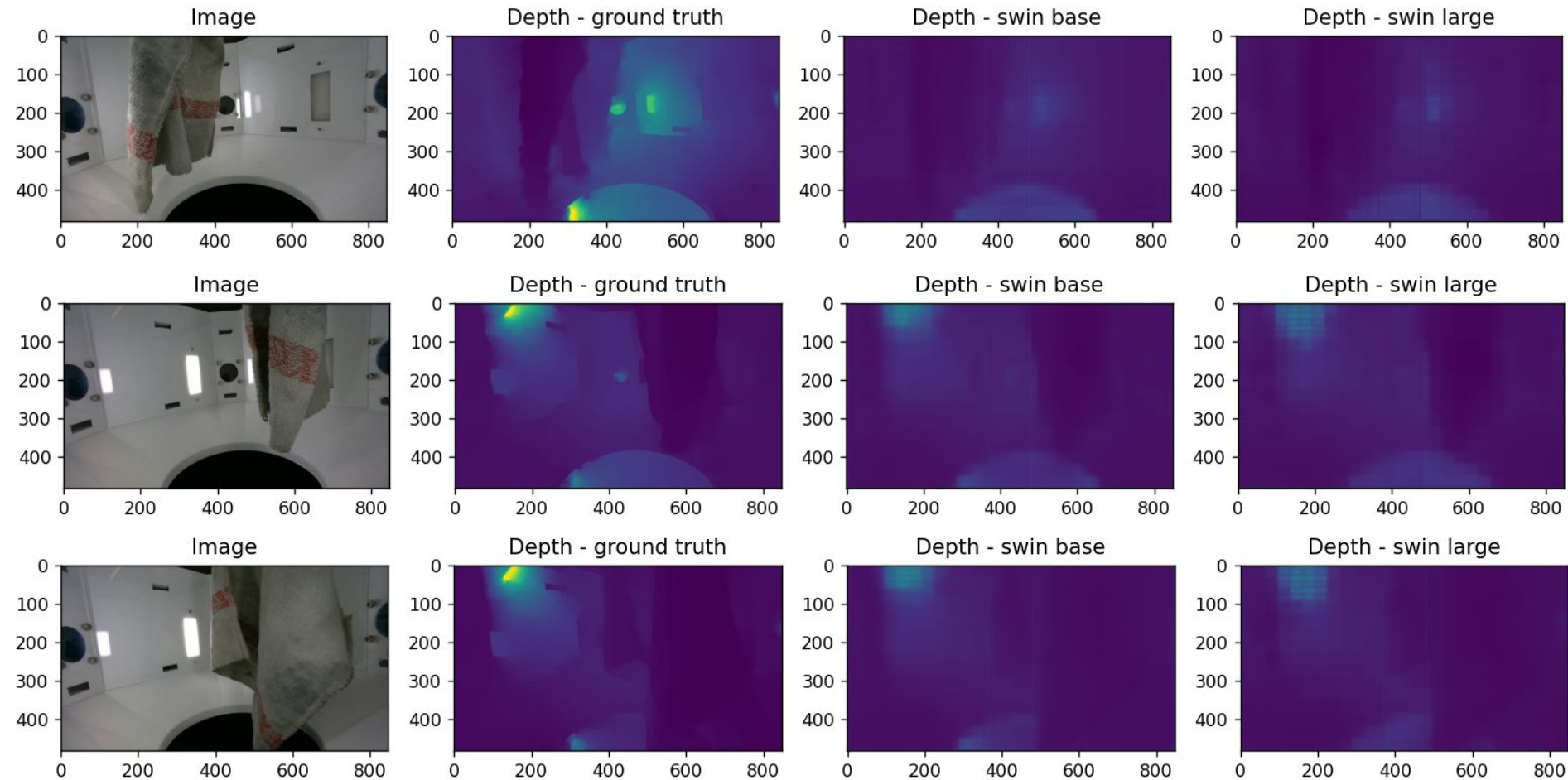
Thank you for your attention!