Monocular depth estimation

Stefan Kapunac, Marija Milićević

Faculty of Mathematics, University of Belgrade

7.9.2020.

Overview

Depth estimation from a single image

- Ill posed problem
- Supervised learning works
- But requires ground truth depth
- Lidar sensors - expensive...

Solution

- Self-supervised learning
- Image reconstruction is the supervisory signal
- Stereo
- Monocular video
  - We have to estimate camera pose

In general

- No labels
- Supervision comes from data

Example

ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness (2018.)

(a) Texture image

81.4%	**Indian elephant**
10.3%	indri
8.2%	black swan

(b) Content image

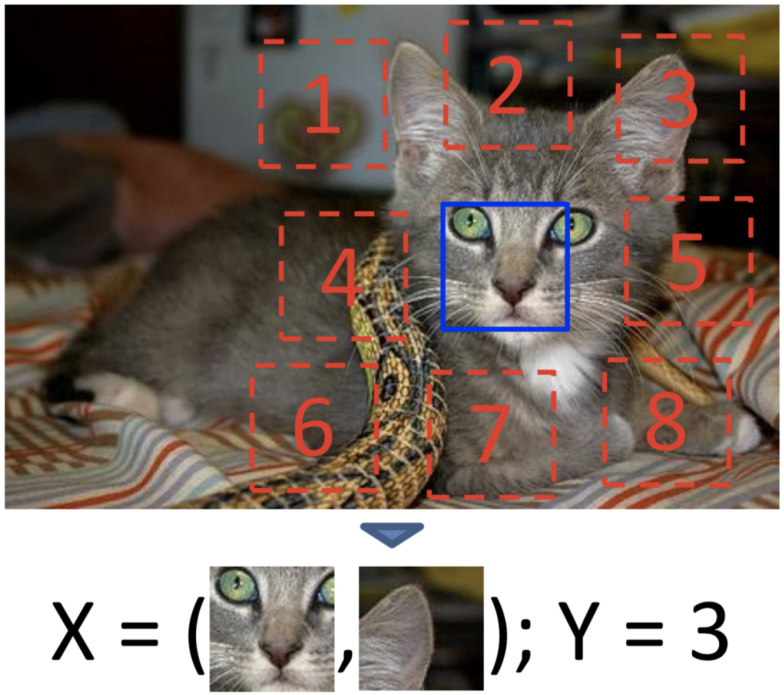
71.1%	**tabby cat**
17.3%	grey fox
3.3%	Siamese cat

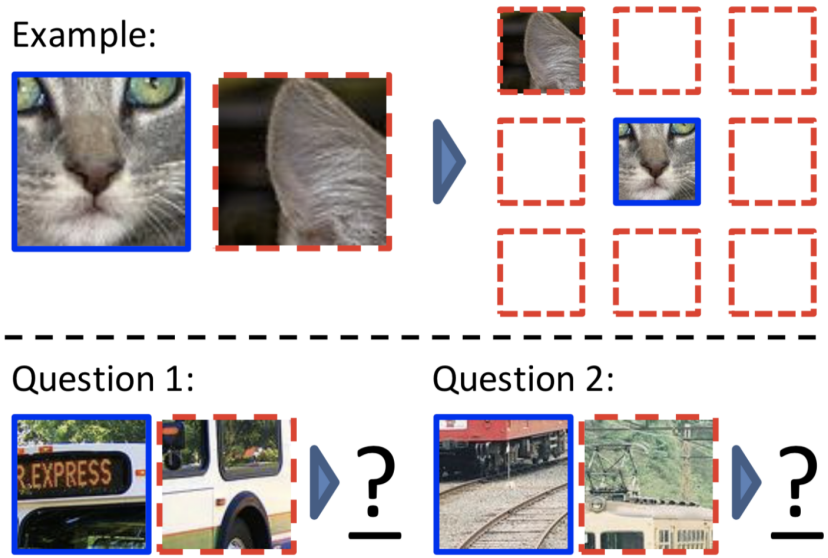
(c) Texture-shape cue conflict

63.9%	**Indian elephant**
26.4%	indri
9.6%	black swan

Example

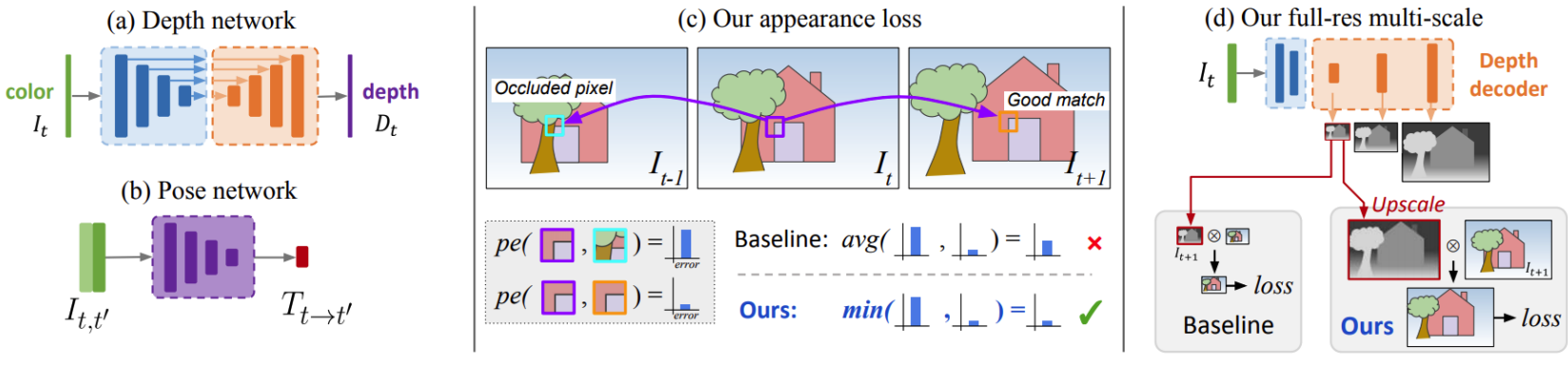
Unsupervised Visual Representation Learning by Context Prediction (2015.)





Monodepth2

Digging Into Self-Supervised Monocular Depth Estimation (2018.)

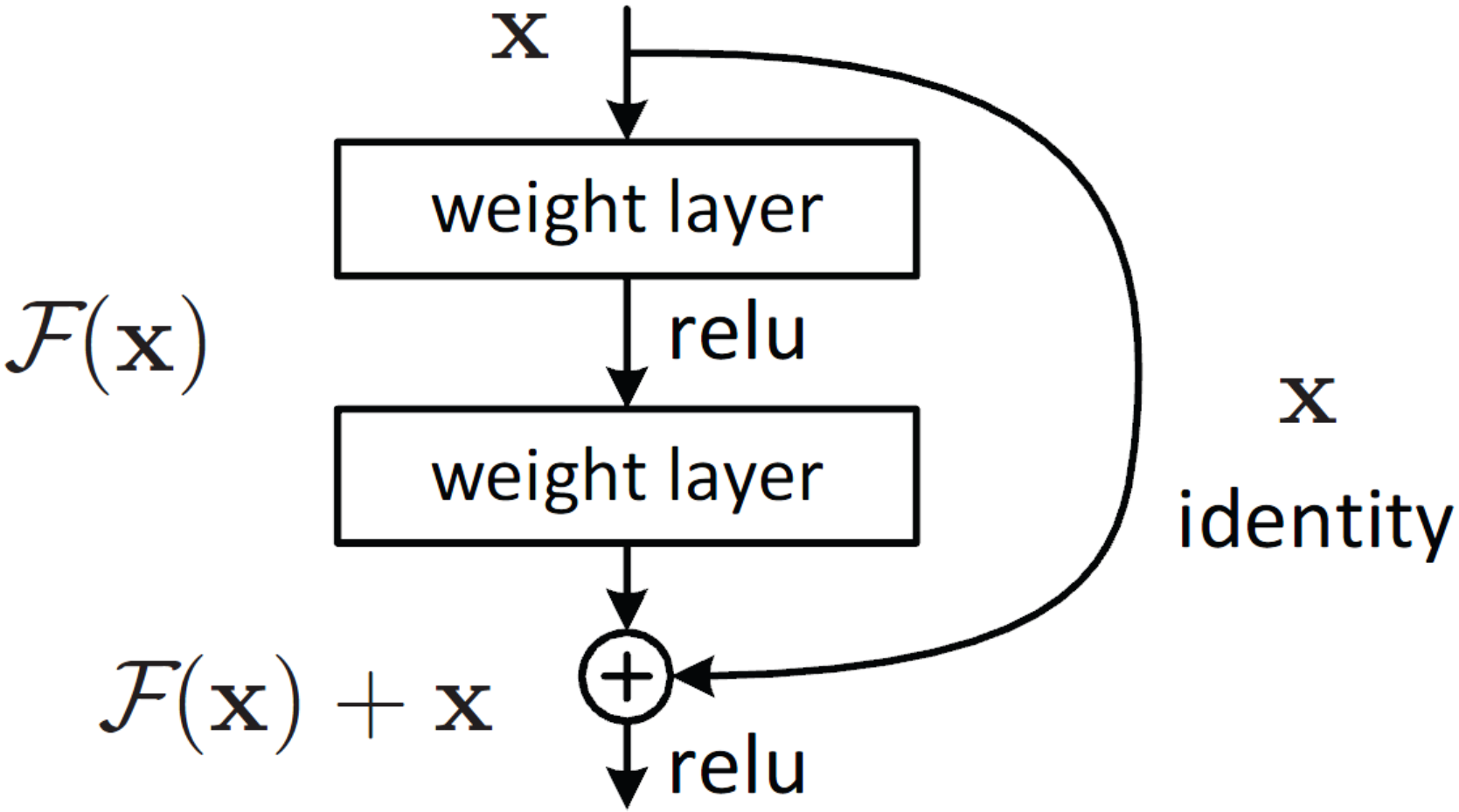


Improvements

- Minimum reprojection loss
- Auto-masking stationary pixels
- Multi-scale estimation

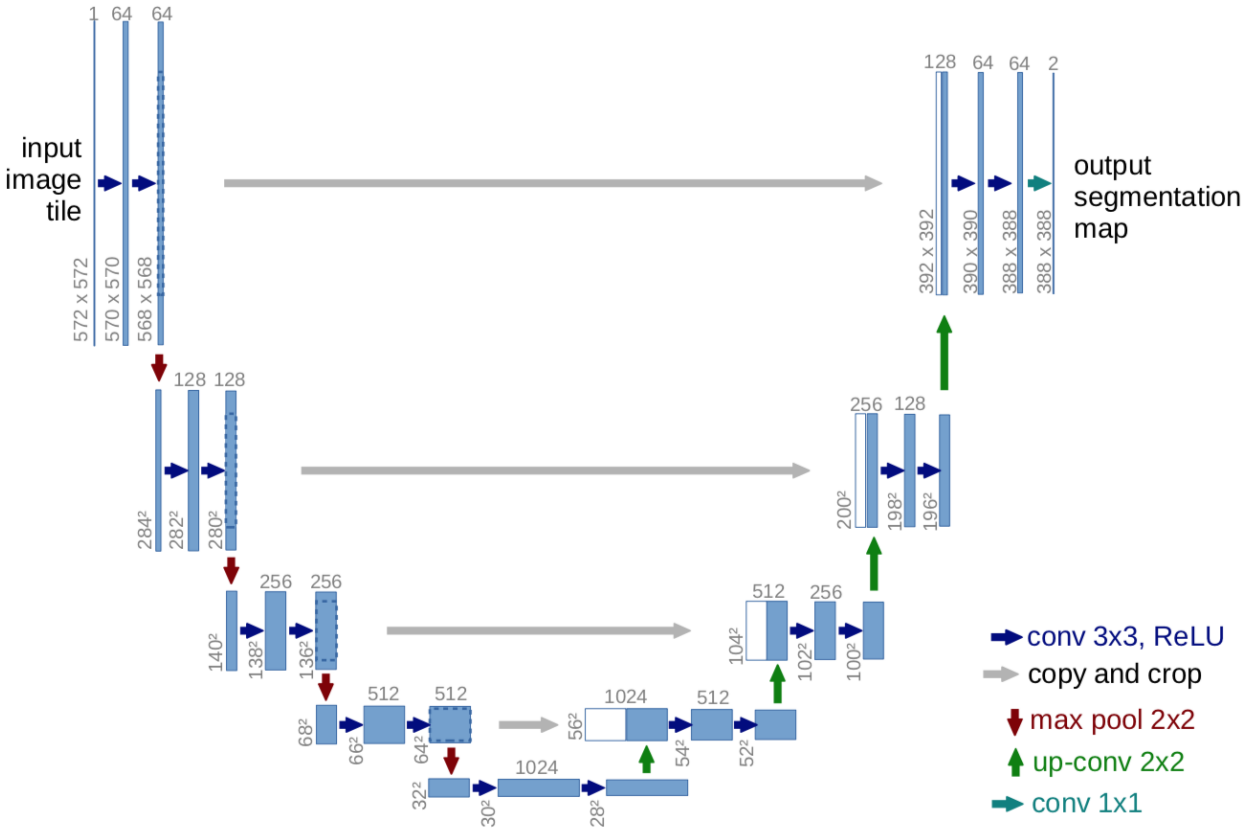
Resnet

Deep Residual Learning for Image Recognition (2015.)



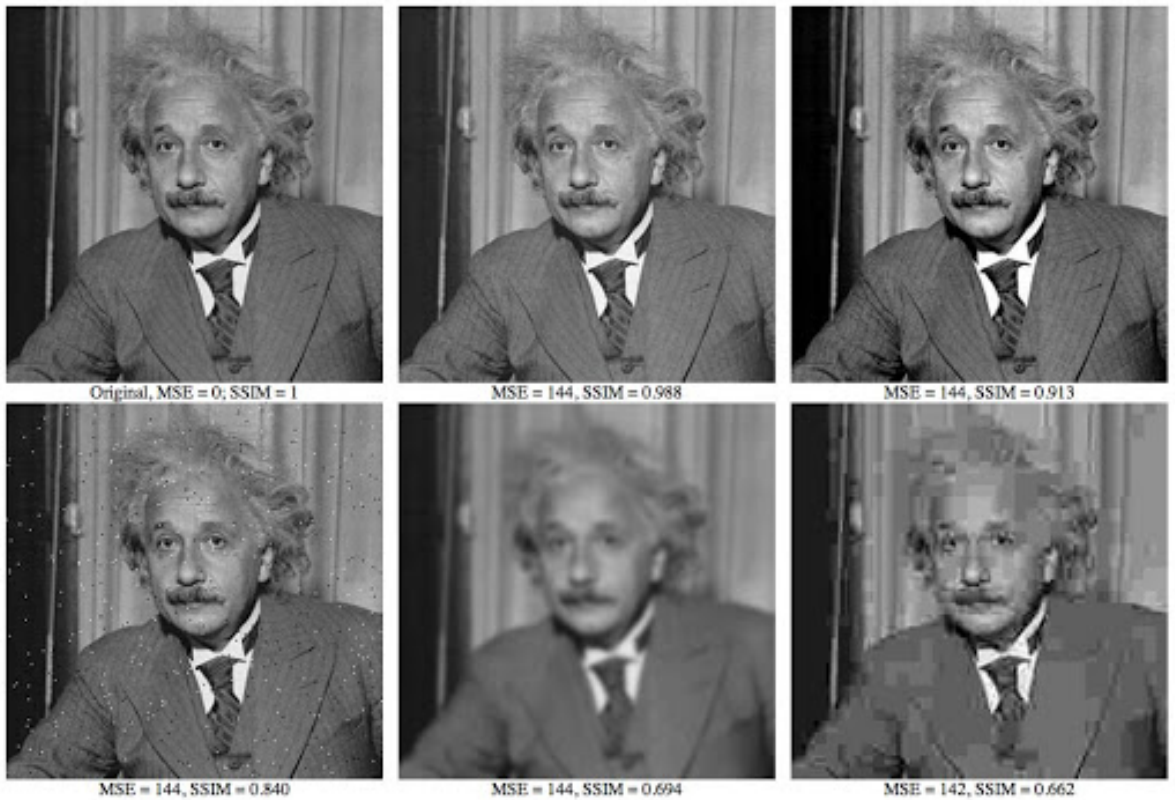
U-Net

U-Net: Convolutional Networks for Biomedical Image Segmentation (2015.)



SSIM

Image Quality Assessment: From Error Visibility to Structural Similarity (2004.)



Bilinear sampling

Spatial Transformer Networks (2015.)

$$V_i^c = \sum_n^H \sum_m^W U_{nm}^c k(x_i^s - m; \Phi_x) k(y_i^s - n; \Phi_y) \quad \forall i \in [1 \dots H'W'] \quad \forall c \in [1 \dots C]$$

KITTI

- Raw dataset 175 GB
- Odometry dataset 65 GB
- ...

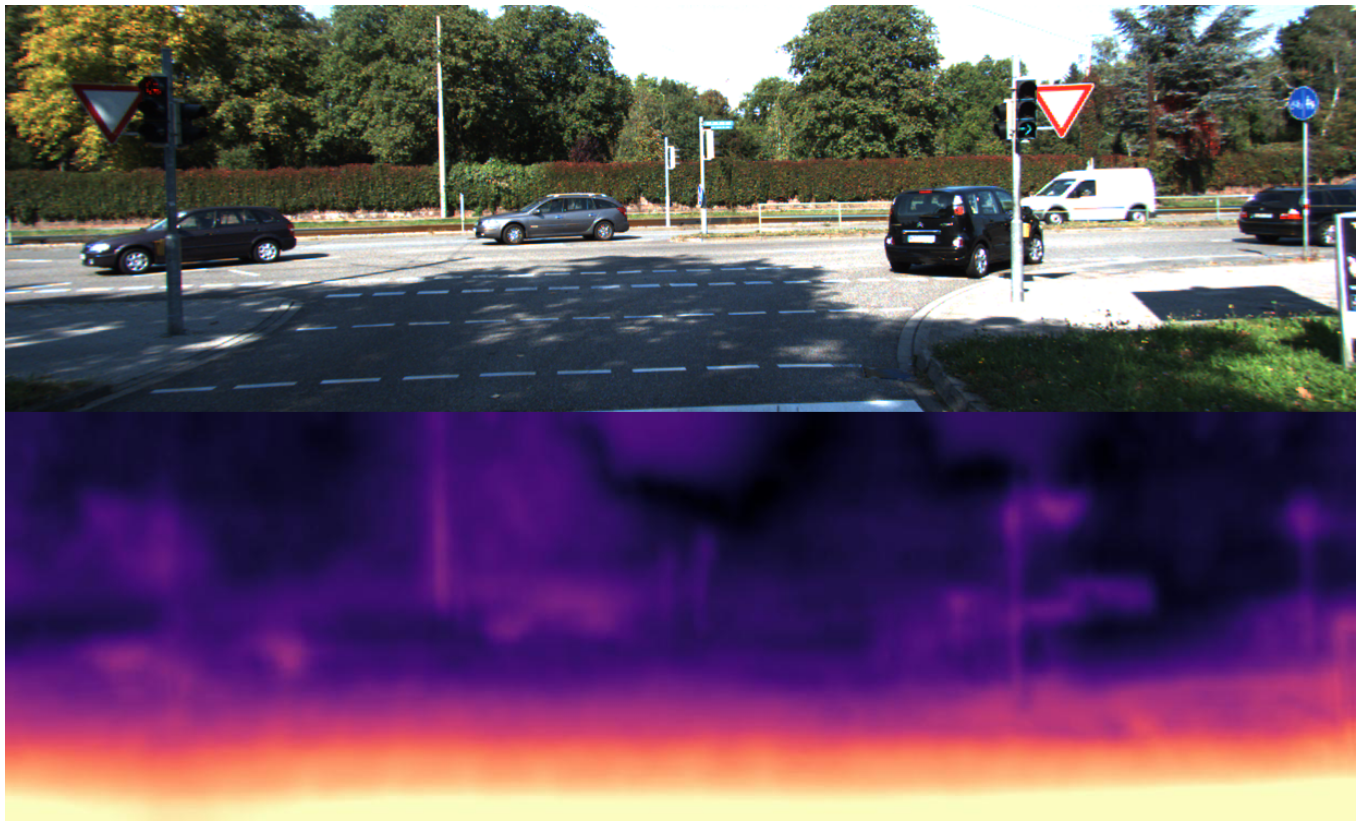
Keras problems

- No ResNet 18 (github?)
- U-Net architecture
- ...

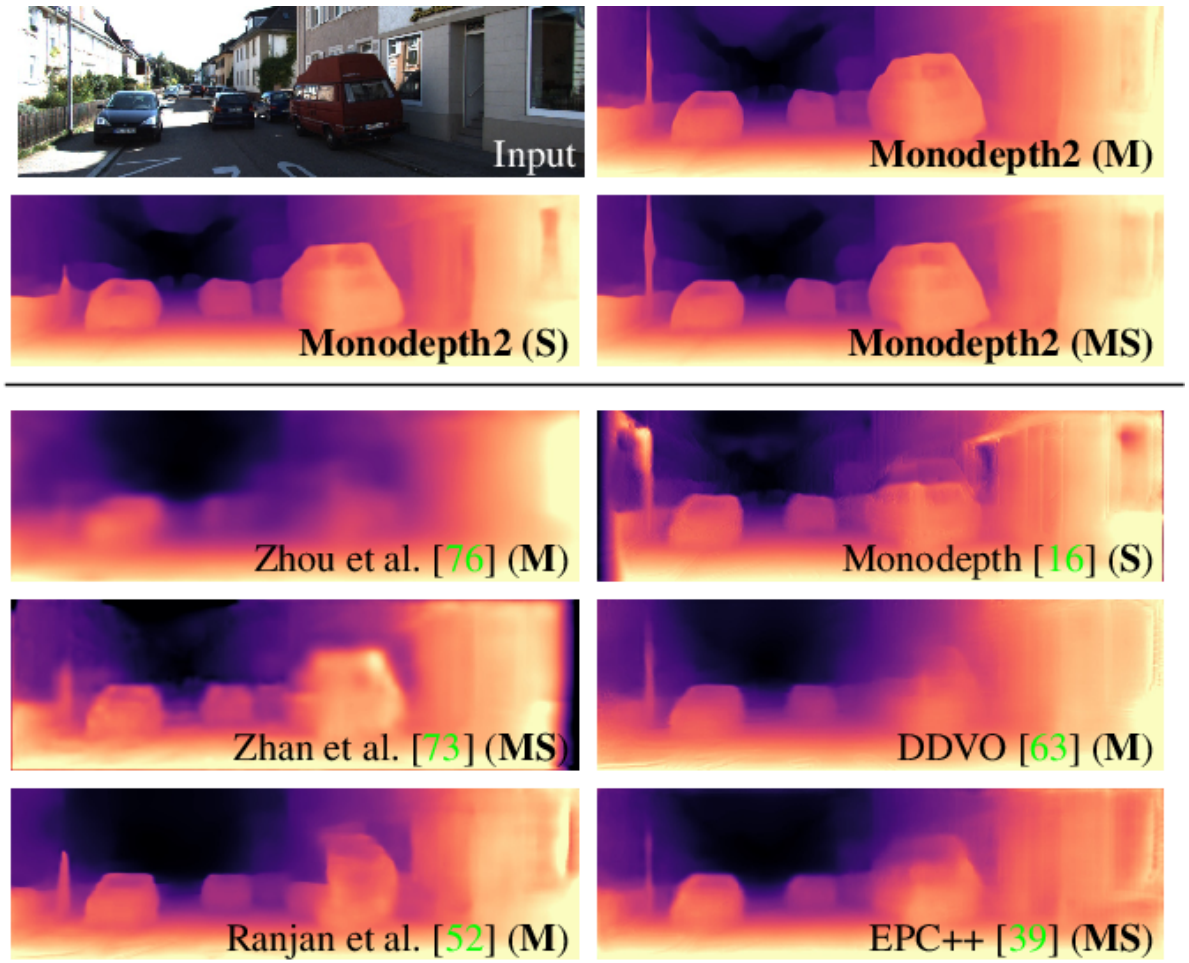
Pytorch

- Original implementation
- Our adaptation
- Split

Example



Comparison


Input
Monodepth2 (M)
Monodepth2 (S)
Monodepth2 (MS)
Zhou et al. [76] (M)
Monodepth [16] (S)
Zhan et al. [73] (MS)
DDVO [63] (M)
Ranjan et al. [52] (M)
EPC++ [39] (MS)