Pan Meng
School of Computer Science and Engineering
Sun-Yat Sen University
Feb 11, 2022

FINAL PROJECT — **Documentation**

Problem 1.

As one of the most crucial tasks of scene perception, monocular depth estimation (MDE) has made considerable development in recent years. Current MDE researchers are interested in the precision and speed of the estimation, but ignore the generalization ability across scenes for quite a long time. For instance, a MDE network trained on outdoor scenes achieve impressive performance on outdoor scenes but poor performance on indoor scenes, vice versa. In this paper, we propose a self-distillation MDE network to improve the generalization ability across different scenes. Specifically, we design a student encoder that extracts features from two datasets of indoor and outdoor scenes, respectively. After that, we introduce a dissimilarity loss to pull apart encoded features of different scenes in the feature space. Finally, a decoder is adopted to estimate the final depth from encoded features. In doing so, our self-distillation MDE network can learn the depth estimation of two different datasets. To our best knowledge, we are the first one to tackle the generalization problem across datasets of different scenes in the MDE field. Experiments demonstrate that our method achieves competitive estimation performance, compared with state-of-the-art MDE methods. Note that evaluating on two datasets by a single network is more challenging than evaluating on two datasets by two different networks.

$\longrightarrow$ *Answer*

Submitted by Pan Meng on Feb 11, 2022.