Benchmarking Out-of-Distribution Depth Estimation under Corruptions

Lingdong Kong[1,2] Shaoyuan Xie[3] Hanjiang Hu[4] Lai Xing Ng[5] Benoit Cottereau[6] Wei Tsang Ooi[1]

[1]National University of Singapore, [2]CNRS@CREATE, [3]Huazhong University of Science and Technology, [4]Carnegie Mellon University, [5]A*STAR, [6]CNRS

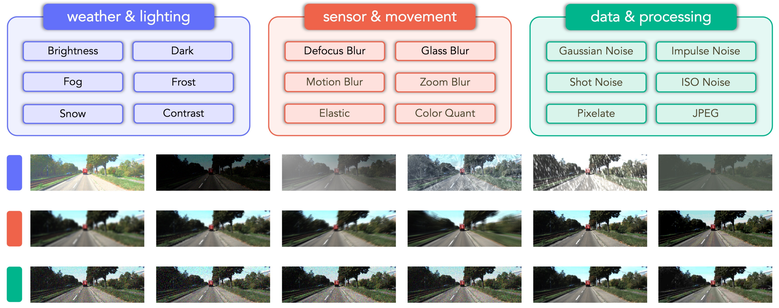
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

RoboDepth

Motivation & Contribution

TL;DR

- **RoboDepth** a comprehensive benchmark for probing the robustness of monocular depth estimation algorithms. It includes **18** common corruption types, ranging from weather and lighting conditions, sensor failure and movement, and noises during data processing.



Motivation

- Existing supervised & self-supervised learning-based depth estimation algorithms use clean video inputs for training. Videos captured by **cameras** in the real world, however, may include distortions, noises, and other artifacts introduced by the environment, sensors, or the data processing. In this project, we ask the following questions:
- How robust are the **existing** monocular depth estimation algorithms to the various corruptions occur in the real world?
- What makes an algorithm more **robust** to certain corruptions?
- Can we **design** novel monocular depth estimation algorithms that are robust against common corruptions?

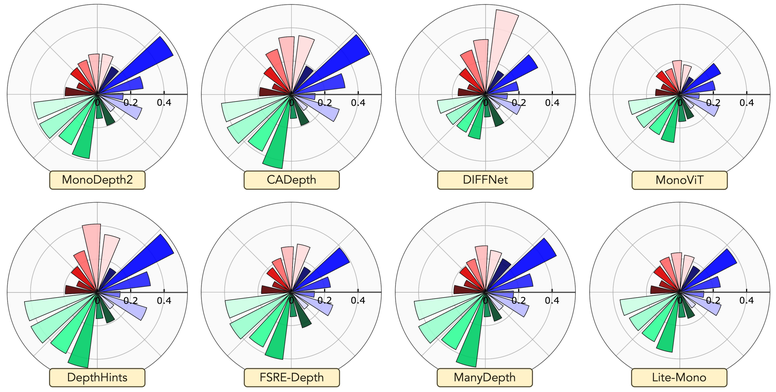
Collaborator

DesCartes · CNRS @CREATE Singapore · NUS National University of Singapore · Agency for Science, Technology and Research SINGAPORE · Carnegie Mellon University · HUST

Dataset & Benchmark

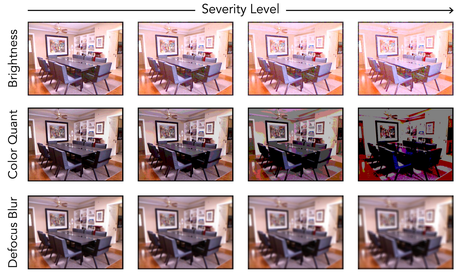
The RoboDepth Benchmark

- We benchmark **42** state-of-the-art depth estimation models from indoor and outdoor scenes, on their robustness against corruptions, via newly established datasets: **KITTI-C**, **NYUDepth2-C**, and **KITTI-S**.



Statistical Analysis

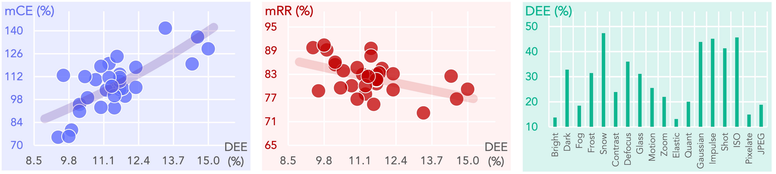
- We design different levels of **severity** for systematic analysis and benchmark.
- We observe that different models exhibit **diverse** strengths and weaknesses.
- Design choices matter for **robustness** to corruptions occur in the real world.



Experiments & Analysis

Benchmarking Results

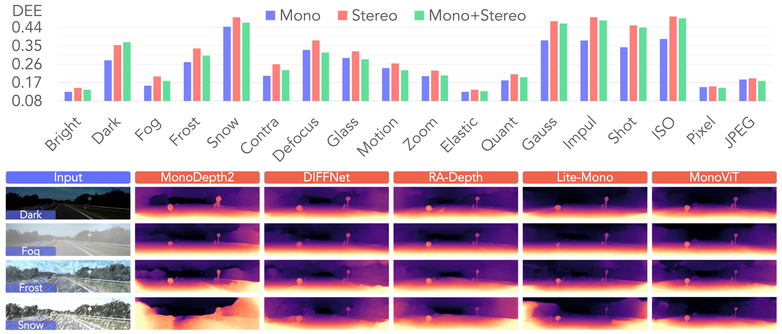
- We use mean Corruption Error **(mCE)** and mean Resilience Rate **(mRR)** to measure the robustness of monocular depth estimation models.



- We find that existing models are **vulnerable** to corruptions, mainly due to the lack of a suitable robustness evaluation suite.

Ablation Study

- We reveal that the factors related to the input modality, resolution, and pretraining strategy are **important** for robust depth estimation.



Summary & Conclusion

- We contribute **RoboDepth**, a new suite to facilitate future research toward robust and reliable monocular depth estimation.
- The code and benchmark toolkit are openly **accessible** at our GitHub repository.

Paper · Code