# The KITTI Vision Benchmark Suite

## A project of [Karlsruhe Institute of Technology](#)
## and [Toyota Technological Institute at Chicago](#)



- [home](#)
- [setup](#)
- [stereo](#)
  - [Stereo 2012](#)
  - [Stereo 2015](#)
- [flow](#)
  - [Flow 2012](#)
  - [Flow 2015](#)
- [sceneflow](#)
- [depth](#)
  - [Depth Completion](#)
  - [Depth Prediction](#)
- [odometry](#)
- [object](#)
  - [2d object](#)
  - [3d object](#)
  - [bird's eye view](#)
- [tracking](#)
  - [multi-object tracking](#)
  - [multi-object tracking and segmentation](#)
- [road](#)
- [semantics](#)
  - [pixel-level](#)
  - [instance-level](#)
- [raw data](#)
- [submit results](#)

[Andreas Geiger (MPI Tübingen)](#) | [Philip Lenz (KIT)](#) | [Christoph Stiller (KIT)](#) | [Raquel Urtasun (University of Toronto)](#)



## Welcome to the KITTI Vision Benchmark Suite!

We take advantage of our [autonomous driving platform Annieway](#) to develop novel challenging real-world computer vision benchmarks. Our tasks of interest are: stereo, optical flow, visual odometry, 3D object detection and 3D tracking. For this purpose, we equipped a standard station wagon with two high-resolution color and grayscale video cameras. Accurate ground truth is provided by a Velodyne laser scanner and a GPS localization system. Our datsets are captured by driving around the mid-size city of [Karlsruhe](#), in rural areas and on highways. Up to 15 cars and 30 pedestrians are visible per image. Besides providing all data in raw format, we extract benchmarks for each task. For each of
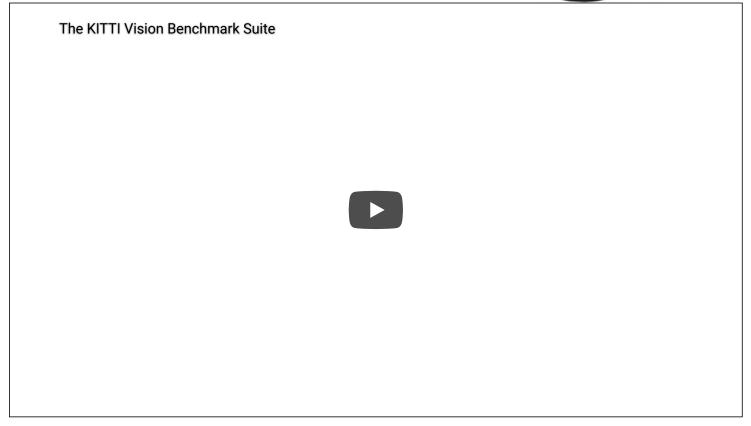
our benchmarks, we also provide an evaluation metric and this evaluation website. Preliminary experiments show that methods ranking high on established benchmarks such as Middlebury perform below average when being moved outside the laboratory to the real world. Our goal is to reduce this bias and complement existing benchmarks by providing real-world benchmarks with novel difficulties to the community.

**Share**

**To get started, grab a cup of your favorite beverage and watch our video trailer (5 minutes):**



The KITTI Vision Benchmark Suite

# Copyright

All datasets and benchmarks on this page are copyright by us and published under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License. This means that you must attribute the work in the manner specified by the authors, you may not use this work for commercial purposes and if you alter, transform, or build upon this work, you may distribute the resulting work only under the same license.

# Citation

When using this dataset in your research, we will be happy if you cite us! (or bring us some self-made cake or ice-cream)
For the **stereo 2012**, **flow 2012**, **odometry**, **object detection** or **tracking benchmarks**, please cite:
@INPROCEEDINGS{Geiger2012CVPR,
  author = {Andreas Geiger and Philip Lenz and Raquel Urtasun},
  title = {Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite},
  booktitle = {Conference on Computer Vision and Pattern Recognition (CVPR)},
  year = {2012}
}
For the **raw dataset**, please cite:
@ARTICLE{Geiger2013IJRR,
  author = {Andreas Geiger and Philip Lenz and Christoph Stiller and Raquel Urtasun},
  title = {Vision meets Robotics: The KITTI Dataset},
  journal = {International Journal of Robotics Research (IJRR)},

```
  year = {2013}
}
```
For the **road benchmark**, please cite:
```
@INPROCEEDINGS{Fritsch2013ITSC,
  author = {Jannik Fritsch and Tobias Kuehnl and Andreas Geiger},
  title = {A New Performance Measure and Evaluation Benchmark for Road Detection Algorithms},
  booktitle = {International Conference on Intelligent Transportation Systems (ITSC)},
  year = {2013}
}
```
For the **stereo 2015**, **flow 2015** and **scene flow 2015 benchmarks**, please cite:
```
@INPROCEEDINGS{Menze2015CVPR,
  author = {Moritz Menze and Andreas Geiger},
  title = {Object Scene Flow for Autonomous Vehicles},
  booktitle = {Conference on Computer Vision and Pattern Recognition (CVPR)},
  year = {2015}
}
```

# Changelog

- 25.2.2021: We have updated the evaluation procedure for Tracking and MOTS. Evaluation now uses the HOTA metrics and is performed with the TrackEval codebase.
- 04.12.2019: We have added a novel benchmark for multi-object tracking and segmentation (MOTS)!
- 18.03.2018: We have added novel benchmarks for semantic segmentation and semantic instance segmentation!
- 11.12.2017: We have added novel benchmarks for depth completion and single image depth prediction!
- 26.07.2017: We have added novel benchmarks for 3D object detection including 3D and bird's eye view evaluation.
- 26.07.2016: For flexibility, we now allow a maximum of 3 submissions per month and count submissions to different benchmarks separately.
- 29.07.2015: We have released our new stereo 2015, flow 2015, and scene flow 2015 benchmarks. In contrast to the stereo 2012 and flow 2012 benchmarks, they provide more difficult sequences as well as ground truth for dynamic objects. We hope for numerous submissions :)
- 09.02.2015: We have fixed some bugs in the ground truth of the road segmentation benchmark and updated the data, devkit and results.
- 11.12.2014: Fixed the bug in the sorting of the object detection benchmark (ordering should be according to moderate level of difficulty).
- 04.09.2014: We are organizing a workshop on reconstruction meets recognition at ECCV 2014!
- 31.07.2014: Added colored versions of the images and ground truth for reflective regions to the stereo/flow dataset.
- 30.06.2014: For detection methods that use flow features, the 3 preceding frames have been made available in the object detection benchmark.
- 04.04.2014: The KITTI road devkit has been updated and some bugs have been fixed in the training ground truth. The server evaluation scripts have been updated to also evaluate the bird's eye view metrics as well as to provide more detailed results for each evaluated method
- 04.11.2013: The ground truth disparity maps and flow fields have been refined/improved. Thanks to Donglai for reporting!
- 31.10.2013: The pose files for the odometry benchmark have been replaced with a properly interpolated (subsampled) version which doesn't exhibit artefacts when computing velocities from the poses.
- 10.10.2013: We are organizing a workshop on reconstruction meets recognition at ICCV 2013!
- 03.10.2013: The evaluation for the odometry benchmark has been modified such that longer sequences are taken into account
- 25.09.2013: The road and lane estimation benchmark has been released!
- 20.06.2013: The tracking benchmark has been released!
- 29.04.2013: A preprint of our IJRR data paper is available for download now!
- 06.03.2013: More complete calibration information (cameras, velodyne, imu) has been added to the object detection benchmark.
- 27.01.2013: We are looking for a PhD student in 3D semantic scene parsing (position available at MPI Tübingen).
- 23.11.2012: The right color images and the Velodyne laser scans have been released for the object detection benchmark.
- 19.11.2012: Added demo code to read and project 3D Velodyne points into images to the raw data development kit.
- 12.11.2012: Added pre-trained LSVM baseline models for download.
- 04.10.2012: Added demo code to read and project tracklets into images to the raw data development kit.
- 01.10.2012: Uploaded the missing oxts file for raw data sequence 2011_09_26_drive_0093.
- 26.09.2012: The velodyne laser scan data has been released for the odometry benchmark.
- 11.09.2012: Added more detailed coordinate transformation descriptions to the raw data development kit.
- 26.08.2012: For transparency and reproducability, we have added the evaluation codes to the development kits.
- 24.08.2012: Fixed an error in the OXTS coordinate system description. Plots and readme have been updated.
- 19.08.2012: The object detection and orientation estimation evaluation goes online!
- 24.07.2012: A section explaining our sensor setup in more details has been added.
- 23.07.2012: The color image data of our object benchmark has been updated, fixing the broken test image 006887.png.
- 04.07.2012: Added error evaluation functions to stereo/flow development kit, which can be used to train model parameters.
- 03.07.2012: Don't care labels for regions with unlabeled objects have been added to the object dataset.
- 02.07.2012: Mechanical Turk occlusion and 2D bounding box corrections have been added to raw data labels.
- 28.06.2012: Minimum time enforced between submission has been increased to 72 hours.
- 27.06.2012: Solved some security issues. Login system now works with cookies.

- 02.06.2012: The training labels and the development kit for the object benchmarks have been released.
- 29.05.2012: The images for the object detection and orientation estimation benchmarks have been released.
- 28.05.2012: We have added the average disparity / optical flow errors as additional error measures.
- 27.05.2012: Large parts of our raw data recordings have been added, including sensor calibration.
- 08.05.2012: Added color sequences to visual odometry benchmark downloads.
- 24.04.2012: Changed colormap of optical flow to a more representative one (new devkit available). Added references to method rankings.
- 23.04.2012: Added paper references and links of all submitted methods to ranking tables. Thanks to Daniel Scharstein for suggesting!
- 05.04.2012: Added links to the most relevant related datasets and benchmarks for each category.
- 04.04.2012: Our CVPR 2012 paper is available for download now!
- 20.03.2012: The KITTI Vision Benchmark Suite goes online, starting with the stereo, flow and odometry benchmarks.

# Privacy

This dataset is made available for academic use only. However, we take your privacy seriously! If you find yourself or personal belongings in this dataset and feel unwell about it, please contact us and we will immediately remove the respective data from our server.

# Credits

We thank Karlsruhe Institute of Technology (KIT) and Toyota Technological Institute at Chicago (TTI-C) for **funding** this project and Jan Cech (CTU) and Pablo Fernandez Alcantarilla (UoA) for **providing initial results**. We further thank our **3D object labeling task force** for doing such a great job: Blasius Forreiter, Michael Ranjbar, Bernhard Schuster, Chen Guo, Arne Dersein, Judith Zinsser, Michael Kroeck, Jasmin Mueller, Bernd Glomb, Jana Scherbarth, Christoph Lohr, Dominik Wewers, Roman Ungefuk, Marvin Lossa, Linda Makni, Hans Christian Mueller, Georgi Kolev, Viet Duc Cao, Bünyamin Sener, Julia Krieg, Mohamed Chanchiri, Anika Stiller. Many thanks also to Qianli Liao (NYU) for helping us in getting the don't care regions of the object detection benchmark correct. Special thanks for providing **the voice** to our video go to Anja Geiger!