The KITTI Vision Benchmark Suite

A project of Karlsruhe Institute of Technology and Toyota Technological Institute at Chicago


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

Multi-Object Tracking and Segmentation (MOTS) Evaluation

The Multi-Object and Segmentation (MOTS) benchmark consists of 21 training sequences and 29 test sequences. It is based on the KITTI Tracking Evaluation 2012 and extends the annotations to the Multi-Object and Segmentation (MOTS) task. To this end, we added dense pixelwise segmentation labels for every object. We evaluate submitted results using the common metrics CLEAR MOT and MT/PT/ML (adapted for the segmentation case).

- Project page
- Download (trainset images + annotations / testset images)
- Description of annotation format
- Tools for loading data, evaluation, and visualization

Important Policy Update: As more and more non-published work and re-implementations of existing work is submitted to KITTI, we have established a new policy: from now on, only submissions with significant novelty that are leading to a peer-reviewed paper in a conference or journal are allowed. Minor modifications of existing algorithms or student research projects are not allowed. Such work must be evaluated on a split of the training set. To ensure that our policy is adopted, new users must detail their status, describe their work and specify the targeted venue during registration. Furthermore, we will regularly delete all entries that are 6 months old but are still anonymous or do not have a paper associated with them. For conferences, 6 month is enough to determine if a paper has been accepted and to add the bibliography information. For longer review cycles, you need to resubmit your results.

Additional information used by the methods

CAR

Method	Setting	Code	sMOTSA	MOTSA	MOTSP	MOTSAL	MODSA	MODSP	MT	ML	IDS	Frag	Runtime	Environment	Compare
1 PointTrack			**78.50 %**	**90.90 %**	87.10 %	**91.80 %**	**91.80 %**	89.70 %	**90.80 %**	0.60 %	346	645	0.045 s	GPU @ 2.5 Ghz (Python)	
2 MOTSFusion		code	75.00 %	84.10 %	**89.30 %**	84.70 %	84.70 %	**91.70 %**	66.10 %	6.20 %	201	572	0.44 s	GPU @ 2.5 Ghz (Python)	
J. Luiten, T. Fischer and B. Leibe: Track to Reconstruct and Reconstruct to Track. arXiv preprint arXiv:1910.00130 2019.															
3 PointTrack(MF)			75.00 %	84.20 %	89.30 %	84.70 %	84.70 %	91.70 %	66.10 %	6.20 %	187	547	0.008 s	GPU @ 2.5 Ghz (Python)	
4 MOTSNet			71.00 %	81.70 %	87.30 %	82.80 %	82.80 %	89.90 %	69.40 %	3.80 %	428	875	0.01 s	GPU @ 2.5 Ghz (Python)	
5 TrackR-CNN		code	67.00 %	79.60 %	85.10 %	81.50 %	81.50 %	88.30 %	74.90 %	2.30 %	692	1058	0.5 s	GPU @ 2.5 Ghz (Python)	
P. Voigtlaender, M. Krause, A. O\usep, J. Luiten, B. Sekar, A. Geiger and B. Leibe: MOTS: Multi-Object Tracking and Segmentation. CVPR 2019.															
6 PointTrack(MF)			0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	**0.00 %**	0	0	0.008 s	GPU @ 2.5 Ghz (C/C++)	
7 PointTrack			0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0	0	0.045 s	GPU @ 2.5 Ghz (Python)	
8 MOTSFusion		code	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0	0	0.44 s	1 core @ 2.5 Ghz (C/C++)	
J. Luiten, T. Fischer and B. Leibe: Track to Reconstruct and Reconstruct to Track. arXiv preprint arXiv:1910.00130 2019.															

Table as LaTeX | Only published Methods

PEDESTRIAN

Method	Setting	Code	sMOTSA	MOTSA	MOTSP	MOTSAL	MODSA	MODSP	MT	ML	IDS	Frag	Runtime	Environment	Compare
1 PointTrack			**61.50 %**	**76.50 %**	81.00 %	**77.40 %**	**77.40 %**	93.80 %	**48.90 %**	9.30 %	176	632	0.045 s	GPU @ 2.5 Ghz (Python)	
2 PointTrack(MF)			59.40 %	73.50 %	**81.50 %**	74.20 %	74.20 %	**94.10 %**	47.40 %	15.60 %	150	481	0.008 s	GPU @ 2.5 Ghz (C/C++)	
3 MOTSFusion		code	58.70 %	72.90 %	81.50 %	74.20 %	74.20 %	94.10 %	47.40 %	15.60 %	279	534	0.44 s	1 core @ 2.5 Ghz (C/C++)	
J. Luiten, T. Fischer and B. Leibe: Track to Reconstruct and Reconstruct to Track. arXiv preprint arXiv:1910.00130 2019.															
4 MOTSNet			48.70 %	62.00 %	79.40 %	63.30 %	63.30 %	93.60 %	34.80 %	22.20 %	273	663	0.01 s	GPU @ 2.5 Ghz (Python)	
5 TrackR-CNN		code	47.30 %	66.10 %	74.60 %	68.40 %	68.40 %	91.80 %	45.60 %	13.30 %	481	861	0.5 s	GPU @ 2.5 Ghz (Python)	
P. Voigtlaender, M. Krause, A. O\usep, J. Luiten, B. Sekar, A. Geiger and B. Leibe: MOTS: Multi-Object Tracking and Segmentation. CVPR 2019.															
6 MOTSFusion		code	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	**0.00 %**	0	0	0.44 s	GPU @ 2.5 Ghz (Python)	
J. Luiten, T. Fischer and B. Leibe: Track to Reconstruct and Reconstruct to Track. arXiv preprint arXiv:1910.00130 2019.															
7 PointTrack(MF)			0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0	0	0.008 s	GPU @ 2.5 Ghz (Python)	
8 PointTrack			0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0	0	0.045 s	GPU @ 2.5 Ghz (Python)	

Table as LaTeX | Only published Methods

Citation

When using this dataset in your research, we will be happy if you cite us:

```
@INPROCEEDINGS{Voigtlaender2019CVPR,
  author = {Paul Voigtlaender and Michael Krause and Aljosa Osep and Jonathon Luiten and Berin Balachandar Gnana Sekar and Andreas Geiger and Bastian Leibe},
  title = {MOTS: Multi-Object Tracking and Segmentation},
  booktitle = {Conference on Computer Vision and Pattern Recognition (CVPR)},
  year = {2019}
}
```