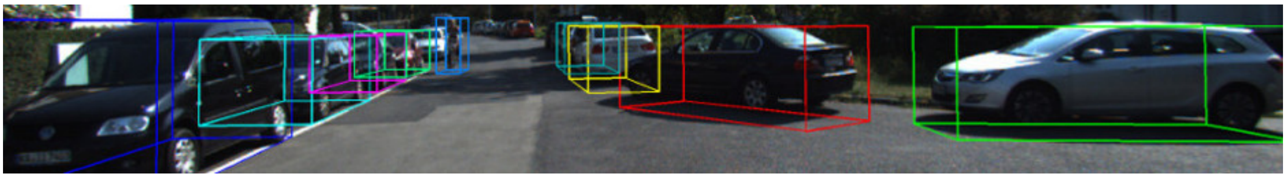
The KITTI Vision Benchmark Suite

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)

3D Object Detection Evaluation 2017



The 3D object detection benchmark consists of 7481 training images and 7518 test images as well as the corresponding point clouds, comprising a total of 80.256 labeled objects. For evaluation, we compute precision-recall curves. To rank the methods we compute average precision. We require that all methods use the same parameter set for all test pairs. Our development kit provides details about the data format as well as MATLAB / C++ utility functions for reading and writing the label files.

- Download left color images of object data set (12 GB)
- Download right color images, if you want to use stereo information (12 GB)
- Download the 3 temporally preceding frames (left color) (36 GB)
- Download the 3 temporally preceding frames (right color) (36 GB)
- Download Velodyne point clouds, if you want to use laser information (29 GB)
- Download camera calibration matrices of object data set (16 MB)
- Download training labels of object data set (5 MB)
- Download object development kit (1 MB) (including 3D object detection and bird's eye view evaluation code)
- Download pre-trained LSVM baseline models (5 MB) used in Joint 3D Estimation of Objects and Scene Layout (NIPS 2011). These models are referred to as LSVM-MDPM-sv (supervised version) and LSVM-MDPM-us (unsupervised version) in the tables below.
- Download reference detections (L-SVM) for training and test set (800 MB)
- Qianli Liao (NYU) has put together code to convert from KITTI to PASCAL VOC file format (documentation included, requires Emacs).
- Karl Rosaen (U.Mich) has released code to convert between KITTI, KITTI tracking, Pascal VOC, Udacity, CrowdAI and AUTTI formats.
- We thank David Stutz and Bo Li for developing the 3D object detection benchmark.

We evaluate 3D object detection performance using the PASCAL criteria also used for 2D object detection. Far objects are thus filtered based on their bounding box height in the image plane. As only objects also appearing on the image plane are labeled, objects in don't car areas do not count as false positives. We note that the evaluation does not take care of ignoring detections that are not visible on the image plane — these detections might give rise to false positives. For **cars** we require an 3D bounding box **overlap of 70%**, while for pedestrians and cyclists we require a 3D bounding box overlap of 50%. Difficulties are defined as follows:

- **Easy:** Min. bounding box height: 40 Px, Max. occlusion level: Fully visible, Max. truncation: 15 %
- **Moderate:** Min. bounding box height: 25 Px, Max. occlusion level: Partly occluded, Max. truncation: 30 %
- **Hard:** Min. bounding box height: 25 Px, Max. occlusion level: Difficult to see, Max. truncation: 50 %

All methods are ranked based on the moderately difficult results.

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

- Stereo: Method uses left and right (stereo) images
- Flow: Method uses optical flow (2 temporally adjacent images)
- Multiview: Method uses more than 2 temporally adjacent images
- Laser Points: Method uses point clouds from Velodyne laser scanner
- Additional training data: Use of additional data sources for training (see details)

Car

	Method	Setting	Code	Moderate	Easy	Hard	Runtime	Environment	Compare
1	MMLab-PartA^2			**77.86 %**	85.94 %	72.00 %	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)	
	S. Shi, Z. Wang, X. Wang and H. Li: Part-A^2 Net: 3D Part-Aware and Aggregation Neural Network for Object Detection from Point Cloud. arXiv preprint arXiv:1907.03670 2019.								
2	HRI-FusionRCNN			77.84 %	86.60 %	69.15 %	0.1 s	1 core @ 2.5 Ghz (C/C++)	
3	STD			77.63 %	86.61 %	**76.06 %**	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)	
	Z. Yang, Y. Sun, S. Liu, X. Shen and J. Jia: STD: Sparse-to-Dense 3D Object Detector for Point Cloud. ICCV 2019.								
4	Patches - EMP			77.20 %	87.85 %	72.78 %	0.5 s	GPU @ 2.5 Ghz (Python)	
5	Patches			77.16 %	**87.87 %**	68.91 %	0.15 s	GPU @ 2.0 Ghz	
6	UberATG-MMF			76.75 %	86.81 %	68.41 %	0.08 s	GPU @ 2.5 Ghz (Python)	
	M. Liang*, B. Yang*, Y. Chen, R. Hu and R. Urtasun: Multi-Task Multi-Sensor Fusion for 3D Object Detection. CVPR 2019.								
7	F-ConvNet			76.51 %	85.88 %	68.08 %	0.47 s	GPU @ 2.5 Ghz (Python + C/C++)	
	Z. Wang and K. Jia: Frustum ConvNet: Sliding Frustums to Aggregate Local Point-Wise Features for Amodal 3D Object Detection. IROS 2019.								
8	3D IoU Loss			76.28 %	84.43 %	68.22 %	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)	
	D. Zhou: IoU Loss for 2D/3D Object Detection. International Conference on 3D Vision (3DV) 2019.								
9	SRF			76.25 %	85.09 %	68.10 %	0.05 s	GPU @ 2.5 Ghz (Python + C/C++)	
10	HRI-VoxelFPN			76.14 %	85.48 %	68.05 %	0.02 s	GPU @ 2.5 Ghz (Python + C/C++)	
	B. Wang, J. An and J. Cao: Voxel-FPN: multi-scale voxel feature aggregation in 3D object detection from point clouds. arXiv preprint arXiv:1907.05286v2 2019.								
11	RGB3D			75.92 %	85.72 %	68.29 %	0.39 s	GPU @ 2.5 Ghz (Python)	
12	SegVoxelNet			75.81 %	84.19 %	67.80 %	0.04 s	1 core @ 2.5 Ghz (Python)	
13	epBRM			75.79 %	83.95 %	67.88 %	0.1 s	GPU @ >3.5 Ghz (Python + C/C++)	
14	MMLab-PointRCNN		code	75.76 %	85.94 %	68.32 %	0.1 s	GPU @ 2.5 Ghz (Python + C/C++)	

S. Shi, X. Wang and H. Li: Pointrcnn: 3d object proposal generation and detection from point cloud. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2019.

15	PFPN		75.74 %	85.66 %	67.68 %	0.02 s	4 cores @ >3.5 Ghz (Python)
16	Fast Point R-CNNv1.1		75.73 %	84.28 %	67.39 %	0.06 s	GPU @ 2.5 Ghz (Python + C/C++)
17	Alibaba-AILabsX		75.69 %	84.59 %	67.80 %	0.05 s	1 core @ >3.5 Ghz (C/C++)
18	PTS	code	75.67 %	83.95 %	67.71 %	0.01 s	1 core @ 2.5 Ghz (C/C++)
19	PointRCNN-deprecated		75.42 %	84.32 %	67.86 %	0.1 s	GPU @ 2.5 Ghz (Python + C/C++)
20	SECOND-V1.5	code	75.38 %	84.04 %	67.36 %	0.04 s	GPU @ 2.0 Ghz (Python + C/C++)
21	MMV		75.35 %	84.21 %	67.36 %	0.4 s	GPU @ 2.5 Ghz (C/C++)
22	TBA		75.18 %	79.90 %	68.48 %	0.07 s	1 core @ 2.5 Ghz (Python)
23	mypointrcnn		75.12 %	85.01 %	68.09 %	1 s	1 core @ 2.5 Ghz (Python)
24	ARPNET		75.03 %	84.44 %	67.37 %	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)
25	PointPillars	code	74.99 %	79.05 %	68.30 %	16 ms	1080ti GPU and Intel i7 CPU

A. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang and O. Beijbom: PointPillars: Fast Encoders for Object Detection from Point Clouds. CVPR 2019.

26	NU-optim		74.94 %	83.92 %	66.88 %	0.04 s	GPU @ >3.5 Ghz (Python)
27	MPNet		74.92 %	83.84 %	67.30 %	0.02 s	GPU @ 2.5 Ghz (Python + C/C++)
28	DH-ARI		74.90 %	80.02 %	68.46 %	0.2 s	1 core @ >3.5 Ghz (Python + C/C++)
29	Alibaba-AILabsX		74.87 %	84.02 %	67.55 %	0.2 s	GPU @ >3.5 Ghz (Python)
30	A-VoxelNet		74.84 %	84.10 %	66.92 %	0.029 s	GPU @ 2.5 Ghz (Python)
31	MVSLN		74.76 %	84.61 %	67.44 %	0.1s s	1 core @ 2.5 Ghz (C/C++)
32	3DBN		74.64 %	83.56 %	66.76 %	0.13s	1080Ti (Python+C/C++)

X. Li, J. Guivant, N. Kwok and Y. Xu: 3D Backbone Network for 3D Object Detection. CoRR 2019.

33	CFR		74.49 %	84.25 %	66.41 %	0.06 s	1 core @ 2.5 Ghz (C/C++)
34	FOFNet		74.45 %	84.15 %	66.97 %	0.04 s	GPU @ 2.5 Ghz (Python)
35	Fast Point R-CNN		74.43 %	83.45 %	66.38 %	0.06 s	GPU @ 2.5 Ghz (Python + C/C++)
36	PC-CNN-V2		73.80 %	84.33 %	64.83 %	0.5 s	GPU @ 2.5 Ghz (Matlab + C/C++)

X. Du, M. Ang, S. Karaman and D. Rus: A General Pipeline for 3D Detection of Vehicles. 2018 IEEE International Conference on Robotics and Automation (ICRA) 2018.

37	AILabs3D		73.70 %	83.32 %	65.77 %	0.6 s	GPU @ >3.5 Ghz (Python)
38	SECOND	code	73.66 %	83.13 %	66.20 %	38 ms	1080Ti

Y. Yan, Y. Mao and B. Li: SECOND: Sparsely Embedded Convolutional Detection. Sensors 2018.

39	Tencent_ADlab_Lidar		73.03 %	79.61 %	65.98 %	0.1 s	GPU @ 2.5 Ghz (Python + C/C++)
40	MVX-Net		72.67 %	83.19 %	65.22 %	0.06 s	GPU @ 3.0 Ghz (Python + C/C++)
41	MDC		72.67 %	82.07 %	64.60 %	0.17 s	GPU @ 2.5 Ghz (Python)
42	IPOD		72.57 %	79.75 %	66.33 %	0.2 s	GPU @ 2.5 Ghz (Python + C/C++)
43	AVOD-FPN	code	71.88 %	81.94 %	66.38 %	0.1 s	Titan X (Pascal)

J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: [Joint 3D Proposal Generation and Object Detection from View Aggregation](). IROS 2018.

44	CONV-BOX			70.47 %	79.98 %	64.49 %	0.2 s	Tesla V100	
45	F-PointNet		code	70.39 %	81.20 %	62.19 %	0.17 s	GPU @ 3.0 Ghz (Python)	

C. Qi, W. Liu, C. Wu, H. Su and L. Guibas: Frustum PointNets for 3D Object Detection from RGB-D Data. arXiv preprint arXiv:1711.08488 2017.

46	PAD			68.33 %	76.72 %	65.49 %	0.15 s	1 core @ 2.5 Ghz (Python)	
47	PP_v1.0		code	68.12 %	77.99 %	65.34 %	0.02s	1 core @ 2.5 Ghz (C/C++)	
48	SeoulRobotics-HFD			66.98 %	76.09 %	64.92 %	0.035 s	GPU (C++)	
49	ELLIOT			66.86 %	76.38 %	64.47 %	0.1 s	1 core @ 2.5 Ghz (C/C++)	
50	DFD			66.56 %	76.36 %	64.11 %	0.05 s	GPU @ 2.0 Ghz (Python + C/C++)	
51	SCANet			66.30 %	76.09 %	58.68 %	0.09s	GPU @ 2.5 Ghz (Python)	
52	UberATG-ContFuse			66.22 %	82.54 %	64.04 %	0.06 s	GPU @ 2.5 Ghz (Python)	

M. Liang, B. Yang, S. Wang and R. Urtasun: Deep Continuous Fusion for Multi-Sensor 3D Object Detection. ECCV 2018.

53	SCANet			65.99 %	75.66 %	63.48 %	0.17 s	>8 cores @ 2.5 Ghz (Python)	
54	AVOD		code	65.78 %	73.59 %	58.38 %	0.08 s	Titan X (pascal)	

J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation. IROS 2018.

55	SECA			65.75 %	75.98 %	58.44 %	1 s	GPU @ 2.5 Ghz (Python)	
56	VSE			65.75 %	75.98 %	58.44 %	0.15 s	GPU @ 2.5 Ghz (Python)	
57	RTL3D			65.72 %	80.42 %	63.50 %	0.02 s	GPU @ 2.5 Ghz (Python)	
58	Multi-3D			65.33 %	76.57 %	56.11 %	0.15 s	1 core @ 2.5 Ghz (C/C++)	
59	FNV1_RPN			65.18 %	74.61 %	57.75 %	0.12 s	1 core @ 2.5 Ghz (Python + C/C++)	
60	FNV1_Fusion			65.07 %	74.78 %	57.74 %	0.11 s	GPU @ 2.5 Ghz (Python)	
61	X_MD			64.82 %	74.06 %	57.49 %	0.2 s	1 core @ 2.5 Ghz (Python + C/C++)	
62	VoxelNet(Unofficial)			64.80 %	74.59 %	57.38 %	0.5 s	GPU @ 2.0 Ghz (Python)	
63	SECA			64.59 %	73.70 %	57.21 %	0.09 s	GPU @ 2.5 Ghz (Python)	
64	FailNet-Fusion			64.36 %	78.54 %	57.21 %	0.1 s	1 core @ 2.5 Ghz (Python)	
65	NLK			63.99 %	73.81 %	60.90 %	0.02 s	1 core @ 2.5 Ghz (Python)	
66	FailNet-LIDAR			63.08 %	73.26 %	56.24 %	0.1 s	1 core @ 2.5 Ghz (Python)	
67	MV3D			62.35 %	71.09 %	55.12 %	0.36 s	GPU @ 2.5 Ghz (Python + C/C++)	

X. Chen, H. Ma, J. Wan, B. Li and T. Xia: Multi-View 3D Object Detection Network for Autonomous Driving. CVPR 2017.

68	FNV1			61.69 %	71.93 %	55.41 %	0.11 s	GPU @ 2.5 Ghz (Python)	
69	FNV2			59.26 %	67.67 %	51.97 %	0.18 s	GPU @ 2.5 Ghz (Python)	
70	CLF3D			58.48 %	65.54 %	46.54 %	0.13 s	GPU @ 2.5 Ghz (Python)	
71	A3DODWTDA		code	56.81 %	59.35 %	50.51 %	0.08 s	GPU @ 3.0 Ghz (Python)	

F. Gustafsson and E. Linder-Norén: Automotive 3D Object Detection Without Target Domain Annotations. 2018.

72	anm		56.76 %	68.02 %	49.39 %	3 s	1 core @ 2.5 Ghz (C/C++)
73	avodC		55.47 %	65.71 %	48.74 %	0.1 s	GPU @ 2.5 Ghz (Python)
74	PL V2 (SDN+GDC)		54.66 %	68.45 %	51.21 %	0.6 s	GPU @ 2.5 Ghz (C/C++)
75	E-VoxelNet		54.33 %	67.73 %	47.70 %	0.1 s	GPU @ 2.5 Ghz (Python)
76	MV3D (LIDAR)		52.73 %	66.77 %	51.31 %	0.24 s	GPU @ 2.5 Ghz (Python + C/C++)

X. Chen, H. Ma, J. Wan, B. Li and T. Xia: Multi-View 3D Object Detection Network for Autonomous Driving. CVPR 2017.

77	Complexer-YOLO		49.44 %	55.63 %	44.13 %	0.06 s	GPU @ 3.5 Ghz (C/C++)

M. Simon, K. Amende, A. Kraus, J. Honer, T. Samann, H. Kaulbersch, S. Milz and H. Michael Gross: Complexer-YOLO: Real-Time 3D Object Detection and Tracking on Semantic Point Clouds. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops 2019.

78	Pseudo-LiDAR V2	code	44.56 %	60.41 %	38.52 %	0.4 s	GPU @ 2.5 Ghz (Python)
79	a		38.80 %	55.11 %	31.86 %	0.35 s	1 core @ 2.5 Ghz (Python + C/C++)
80	Pseudo-LiDAR	code	37.17 %	55.40 %	31.37 %	0.4 s	GPU @ 2.5 Ghz (C/C++)

Y. Wang, W. Chao, D. Garg, B. Hariharan, M. Campbell and K. Weinberger: Pseudo-LiDAR from Visual Depth Estimation: Bridging the Gap in 3D Object Detection for Autonomous Driving. CVPR 2019.

81	Stereo R-CNN	code	34.05 %	49.23 %	28.39 %	0.3 s	GPU @ 2.5 Ghz (Python)

P. Li, X. Chen and S. Shen: Stereo R-CNN based 3D Object Detection for Autonomous Driving. CVPR 2019.

82			34.04 %	49.68 %	28.45 %		
83	SA_3D		31.21 %	41.70 %	25.96 %	0.3 s	GPU @ 2.5 Ghz (Python)
84	RT3DStereo		24.10 %	28.50 %	20.32 %	0.08 s	GPU @ 2.5 Ghz (C/C++)
85	RT3D		21.27 %	23.49 %	19.81 %	0.09 s	GPU @ 1.8Ghz

Y. Zeng, Y. Hu, S. Liu, J. Ye, Y. Han, X. Li and N. Sun: RT3D: Real-Time 3-D Vehicle Detection in LiDAR Point Cloud for Autonomous Driving. IEEE Robotics and Automation Letters 2018.

86	StereoFENet		20.37 %	29.93 %	16.59 %	0.15 s	1 core @ 3.5 Ghz (Python)

W. Bao, B. Xu and Z. Chen: MonoFENet: Monocular 3D Object Detection with Feature Enhancement Networks. 2019.

87	AM3D		16.08 %	21.48 %	15.26 %	0.4 s	GPU @ 2.5 Ghz (Python + C/C++)

X. Ma, Z. Wang, H. Li, P. Zhang, W. Ouyang and X. Fan: Accurate Monocular Object Detection via Color- Embedded 3D Reconstruction for Autonomous Driving. Proceedings of the IEEE international Conference on Computer Vision (ICCV) 2019.

88	M3D-RPN	code	15.70 %	20.65 %	13.32 %	0.16 s	GPU @ 1.5 Ghz (Python)

G. Brazil and X. Liu: M3D-RPN: Monocular 3D Region Proposal Network for Object Detection . ICCV 2019 .

89	MonoDIS		15.12 %	11.81 %	12.71 %	0.1 s	1 core @ 2.5 Ghz (C/C++)
90	DLMB		14.49 %	15.16 %	12.94 %	0.03 s	8 cores @ 3.5 Ghz (C/C++)
91	BirdNet		13.44 %	14.75 %	12.04 %	0.11 s	Titan Xp GPU

J. Beltrán, C. Guindel, F. Moreno, D. Cruzado, F. García and A. Escalera: BirdNet: A 3D Object Detection Framework from LiDAR Information. 2018 21st International Conference on Intelligent Transportation Systems (ITSC) 2018.

92	Mono3D_PLiDAR	code	13.44 %	17.12 %	12.38 %	0.1 s	NVIDIA GeForce 1080 (pytorch)

X. Weng and K. Kitani: Monocular 3D Object Detection with Pseudo-LiDAR Point Cloud. arXiv:1903.09847 2019.

93	MonoGRNet	code	12.90 %	11.29 %	11.34 %	0.04s	NVIDIA P40

Z. Qin, J. Wang and Y. Lu: MonoGRNet: A Geometric Reasoning Network for 3D Object Localization. The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19) 2019.

94	Licar		12.88 %	16.25 %	13.67 %	0.09 s	GPU @ 2.0 Ghz (Python)
95	TopNet-HighRes		12.58 %	15.29 %	12.25 %	101ms	NVIDIA GeForce 1080 Ti (tensorflow-gpu)

S. Wirges, T. Fischer, C. Stiller and J. Frias: Object Detection and Classification in Occupancy Grid Maps Using Deep Convolutional Networks. 2018 21st International Conference on Intelligent Transportation Systems (ITSC) 2018.

96	FailNet-Mono		11.58 %	8.64 %	10.14 %	0.1 s	1 core @ 2.5 Ghz (Python)
97	MVRA + I-FRCNN+		11.01 %	12.92 %	10.45 %	0.18 s	GPU @ 2.5 Ghz (Python)
98	MonoPSR		10.85 %	12.57 %	9.06 %	0.2 s	GPU @ 3.5 Ghz (Python)

J. Ku*, A. Pon* and S. Waslander: Monocular 3D Object Detection Leveraging Accurate Proposals and Shape Reconstruction. CVPR 2019.

99	ROI-10D		10.30 %	12.30 %	9.39 %	0.2 s	GPU @ 3.5 Ghz (Python)

F. Manhardt, W. Kehl and A. Gaidon: ROI-10D: Monocular Lifting of 2D Detection to 6D Pose and Metric Shape. Computer Vision and Pattern Recognition (CVPR) 2019.

100	DT3D		9.92 %	15.37 %	9.26 %	0,21s	GPU @ 2.5 Ghz (Python)
101	SS3D		9.58 %	11.74 %	7.77 %	48 ms	Tesla V100 (Python)

E. Jörgensen, C. Zach and F. Kahl: Monocular 3D Object Detection and Box Fitting Trained End-to-End Using Intersection-over-Union Loss. CoRR 2019.

102	mylsi-faster-rcnn		9.49 %	11.80 %	9.19 %	0.3 s	1 core @ 2.5 Ghz (Python)
103	CSoR		6.79 %	6.76 %	6.14 %	3.5 s	4 cores @ >3.5 Ghz (Python + C/C++)

L. Plotkin: PyDriver: Entwicklung eines Frameworks für räumliche Detektion und Klassifikation von Objekten in Fahrzeugumgebung. 2015.

104	mymask-rcnn		6.65 %	10.90 %	6.34 %	0.3 s	1 core @ 2.5 Ghz (Python)
105	A3DODWTDA (image)	code	6.45 %	6.76 %	4.87 %	0.8 s	GPU @ 3.0 Ghz (Python)

F. Gustafsson and E. Linder-Norén: Automotive 3D Object Detection Without Target Domain Annotations. 2018.

106	MonoFENet		6.36 %	9.31 %	5.61 %	0.15 s	1 core @ 3.5 Ghz (Python)

W. Bao, B. Xu and Z. Chen: MonoFENet: Monocular 3D Object Detection with Feature Enhancement Networks. 2019.

107	GS3D		6.29 %	7.69 %	6.16 %	2 s	1 core @ 2.5 Ghz (C/C++)

B. Li, W. Ouyang, L. Sheng, X. Zeng and X. Wang: GS3D: An Efficient 3D Object Detection Framework for Autonomous Driving. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2019.

108	Shift R-CNN (mono)	code	5.22 %	8.13 %	4.78 %	0.25 s	GPU @ 1.5 Ghz (Python)

A. Naiden, V. Paunescu, G. Kim, B. Jeon and M. Leordeanu: Shift R-CNN: Deep Monocular 3D Object Detection With Closed-form Geometric Constraints. ICIP 2019.

109	RAR-Net		4.22 %	6.55 %	3.26 %	0.5 s	1 core @ 2.5 Ghz (C/C++)
110	TopNet-UncEst		3.93 %	6.21 %	3.78 %	0.09 s	NVIDIA GeForce 1080 Ti (tensorflow-gpu)

S. Wirges, M. Braun, M. Lauer and C. Stiller: Capturing Object Detection Uncertainty in Multi-Layer Grid Maps. 2019.

111	MF3D		3.17 %	3.81 %	3.25 %	0.03 s	GPU @ 2.5 Ghz (C/C++)
112	OFT-Net		2.50 %	3.28 %	2.27 %	0.5 s	8 cores @ 2.5 Ghz (Python + C/C++)
113	FQNet		2.42 %	3.48 %	1.96 %	0.5 s	1 core @ 2.5 Ghz (Python)
114	3D-SSMFCNN	code	2.28 %	2.39 %	1.52 %	0.1 s	GPU @ 1.5 Ghz (C/C++)

L. Novak: Vehicle Detection and Pose Estimation for Autonomous Driving. 2017.

115	3DVSSD		1.14 %	1.38 %	1.27 %	0.06 s	1 core @ 2.5 Ghz (C/C++)
116	monoref3d		0.55 %	0.53 %	0.55 %	0.1 s	1 core @ 2.5 Ghz (Python)
117	ref3D		0.00 %	0.00 %	0.00 %	0.1 s	1 core @ 2.5 Ghz (Python)
118	ref3D		0.00 %	0.00 %	0.00 %	0.1 s	1 core @ 2.5 Ghz (Python + C/C++)
119	mBoW		0.00 %	0.00 %	0.00 %	10 s	1 core @ 2.5 Ghz (C/C++)

J. Behley, V. Steinhage and A. Cremers: Laser-based Segment Classification Using a Mixture of Bag-of-Words. Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2013.

Table as LaTeX | Only published Methods

Pedestrian

	Method	Setting	Code	Moderate	Easy	Hard	Runtime	Environment	Compare
1	A-VoxelNet			**46.64 %**	54.83 %	42.39 %	0.029 s	GPU @ 2.5 Ghz (Python)	
2	F-ConvNet			45.61 %	52.37 %	41.49 %	0.47 s	GPU @ 2.5 Ghz (Python + C/C++)	
	Z. Wang and K. Jia: Frustum ConvNet: Sliding Frustums to Aggregate Local Point-Wise Features for Amodal 3D Object Detection. IROS 2019.								
3	VMVS			45.01 %	53.98 %	41.72 %	0.25 s	GPU @ 2.5 Ghz (Python)	
	J. Ku, A. Pon, S. Walsh and S. Waslander: Improving 3D object detection for pedestrians with virtual multi-view synthesis orientation estimation. IROS 2019.								
4	F-PointNet		code	44.89 %	51.21 %	40.23 %	0.17 s	GPU @ 3.0 Ghz (Python)	
	C. Qi, W. Liu, C. Wu, H. Su and L. Guibas: Frustum PointNets for 3D Object Detection from RGB-D Data. arXiv preprint arXiv:1711.08488 2017.								
5	IPOD			44.68 %	**56.92 %**	**42.39 %**	0.2 s	GPU @ 2.5 Ghz (Python + C/C++)	
6	STD			44.24 %	53.08 %	41.97 %	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)	
	Z. Yang, Y. Sun, S. Liu, X. Shen and J. Jia: STD: Sparse-to-Dense 3D Object Detector for Point Cloud. ICCV 2019.								
7	epBRM			43.90 %	50.38 %	40.91 %	0.10 s	1 core @ 2.5 Ghz (C/C++)	
8	PointPillars		code	43.53 %	52.08 %	41.49 %	16 ms	1080ti GPU and Intel i7 CPU	
	A. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang and O. Beijbom: PointPillars: Fast Encoders for Object Detection from Point Clouds. CVPR 2019.								
9	Multi-3D			42.87 %	51.17 %	38.94 %	0.15 s	1 core @ 2.5 Ghz (C/C++)	
10	AVOD-FPN		code	42.81 %	50.80 %	40.88 %	0.1 s	Titan X (Pascal)	
	J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation. IROS 2018.								
11	SECOND		code	42.56 %	51.07 %	37.29 %	38 ms	1080Ti	
	Y. Yan, Y. Mao and B. Li: SECOND: Sparsely Embedded Convolutional Detection. Sensors 2018.								
12	MDC			42.54 %	50.79 %	36.56 %	0.17 s	GPU @ 2.5 Ghz (Python)	
13	MMLab-PointRCNN		code	41.78 %	49.43 %	38.63 %	0.1 s	GPU @ 2.5 Ghz (Python + C/C++)	
	S. Shi, X. Wang and H. Li: Pointrcnn: 3d object proposal generation and detection from point cloud. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2019.								
14	Tencent_ADlab_Lidar			41.64 %	49.83 %	39.28 %	0.1 s	GPU @ 2.5 Ghz (Python + C/C++)	
15	ARPNET			41.62 %	50.00 %	39.19 %	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)	
16	SCANet			41.44 %	50.66 %	36.60 %	0.17 s	>8 cores @ 2.5 Ghz (Python)	
17	FOFNet			41.21 %	49.44 %	36.42 %	0.04 s	GPU @ 2.5 Ghz (Python)	
18	CONV-BOX			41.01 %	47.74 %	35.98 %	0.2 s	Tesla V100	
19	DSS			40.93 %	47.68 %	38.34 %	0.03 s	GPU @ 2.5 Ghz (C/C++)	
20	PP_v1.0		code	38.86 %	46.28 %	36.25 %	0.02s	1 core @ 2.5 Ghz (C/C++)	
21	ELLIOT			37.78 %	45.94 %	34.94 %	0.1 s	1 core @ 2.5 Ghz (C/C++)	
22	CFR			36.86 %	44.64 %	35.57 %	0.06 s	1 core @ 2.5 Ghz (C/C++)	
23	anm			34.71 %	45.89 %	32.43 %	3 s	1 core @ 2.5 Ghz (C/C++)	

24	anonymous			33.37 %	40.19 %	27.90 %	0.75 s	GPU @ 3.5 Ghz (C/C++)	
25	X_MD			33.23 %	40.34 %	28.19 %	0.2 s	1 core @ 2.5 Ghz (C/C++)	
26	SA_3D			32.58 %	39.20 %	27.65 %	0.3 s	GPU @ 2.5 Ghz (Python)	
27	CLF3D			31.65 %	35.85 %	26.94 %	0.13 s	GPU @ 2.5 Ghz (Python)	
28	AVOD		code	31.51 %	38.28 %	26.98 %	0.08 s	Titan X (pascal)	
	J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation. IROS 2018.								
29				31.30 %	38.00 %	28.77 %			
30	a			21.85 %	28.14 %	20.92 %	0.35 s	1 core @ 2.5 Ghz (Python + C/C++)	
31	Complexer-YOLO			15.32 %	19.45 %	14.80 %	0.06 s	GPU @ 3.5 Ghz (C/C++)	
	M. Simon, K. Amende, A. Kraus, J. Honer, T. Samann, H. Kaulbersch, S. Milz and H. Michael Gross: Complexer-YOLO: Real-Time 3D Object Detection and Tracking on Semantic Point Clouds. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops 2019.								
32	BirdNet			11.80 %	14.31 %	10.55 %	0.11 s	Titan Xp GPU	
	J. Beltrán, C. Guindel, F. Moreno, D. Cruzado, F. García and A. Escalera: BirdNet: A 3D Object Detection Framework from LiDAR Information. 2018 21st International Conference on Intelligent Transportation Systems (ITSC) 2018.								
33	MonoPSR			10.66 %	12.65 %	10.08 %	0.2 s	GPU @ 3.5 Ghz (Python)	
	J. Ku*, A. Pon* and S. Waslander: Monocular 3D Object Detection Leveraging Accurate Proposals and Shape Reconstruction. CVPR 2019.								
34	Shift R-CNN (mono)		code	10.59 %	13.36 %	10.59 %	0.25 s	GPU @ 1.5 Ghz (Python)	
	A. Naiden, V. Paunescu, G. Kim, B. Jeon and M. Leordeanu: Shift R-CNN: Deep Monocular 3D Object Detection With Closed-form Geometric Constraints. ICIP 2019.								
35	M3D-RPN		code	10.54 %	11.82 %	10.29 %	0.16 s	GPU @ 1.5 Ghz (Python)	
	G. Brazil and X. Liu: M3D-RPN: Monocular 3D Region Proposal Network for Object Detection . ICCV 2019 .								
36	TopNet-HighRes			9.66 %	13.45 %	9.64 %	101ms	NVIDIA GeForce 1080 Ti (tensorflow-gpu)	
	S. Wirges, T. Fischer, C. Stiller and J. Frias: Object Detection and Classification in Occupancy Grid Maps Using Deep Convolutional Networks. 2018 21st International Conference on Intelligent Transportation Systems (ITSC) 2018.								
37	TopNet-UncEst			4.55 %	6.19 %	4.55 %	0.09 s	NVIDIA GeForce 1080 Ti (tensorflow-gpu)	
	S. Wirges, M. Braun, M. Lauer and C. Stiller: Capturing Object Detection Uncertainty in Multi-Layer Grid Maps. 2019.								
38	RT3DStereo			4.25 %	4.27 %	4.26 %	0.08 s	GPU @ 2.5 Ghz (C/C++)	
39	SS3D			3.28 %	3.52 %	2.37 %	48 ms	Tesla V100 (Python)	
	E. Jörgensen, C. Zach and F. Kahl: Monocular 3D Object Detection and Box Fitting Trained End-to-End Using Intersection-over-Union Loss. CoRR 2019.								
40	mylsi-faster-rcnn			1.85 %	2.22 %	1.77 %	0.3 s	1 core @ 2.5 Ghz (Python)	
41	mymask-rcnn			1.44 %	1.86 %	1.38 %	0.3 s	1 core @ 2.5 Ghz (Python)	
42	DT3D			1.14 %	1.14 %	1.14 %	0,21s	GPU @ 2.5 Ghz (Python)	
43	OFT-Net			1.11 %	1.06 %	1.06 %	0.5 s	8 cores @ 2.5 Ghz (Python + C/C++)	
44	mBoW			0.00 %	0.00 %	0.00 %	10 s	1 core @ 2.5 Ghz (C/C++)	
	J. Behley, V. Steinhage and A. Cremers: Laser-based Segment Classification Using a Mixture of Bag-of-Words. Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2013.								

Table as LaTeX | Only published Methods

Cyclist

	Method	Setting	Code	Moderate	Easy	Hard	Runtime	Environment	Compare

1	F-ConvNet		**64.68 %**	**79.58 %**	57.03 %	0.47 s	GPU @ 2.5 Ghz (Python + C/C++)

Z. Wang and K. Jia: Frustum ConvNet: Sliding Frustums to Aggregate Local Point-Wise Features for Amodal 3D Object Detection. IROS 2019.

2	MMLab-PartA^2		62.73 %	78.58 %	**57.74 %**	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)

S. Shi, Z. Wang, X. Wang and H. Li: Part-A^2 Net: 3D Part-Aware and Aggregation Neural Network for Object Detection from Point Cloud. arXiv preprint arXiv:1907.03670 2019.

3	STD		62.53 %	78.89 %	55.77 %	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)

Z. Yang, Y. Sun, S. Liu, X. Shen and J. Jia: STD: Sparse-to-Dense 3D Object Detector for Point Cloud. ICCV 2019.

4	FOFNet		59.65 %	75.36 %	53.03 %	0.04 s	GPU @ 2.5 Ghz (Python)
5	MMLab-PointRCNN	code	59.60 %	73.93 %	53.59 %	0.1 s	GPU @ 2.5 Ghz (Python + C/C++)

S. Shi, X. Wang and H. Li: Pointrcnn: 3d object proposal generation and detection from point cloud. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2019.

6	Multi-3D		59.40 %	75.99 %	51.50 %	0.15 s	1 core @ 2.5 Ghz (C/C++)
7	ARPNET		59.12 %	72.29 %	53.35 %	0.08 s	GPU @ 2.5 Ghz (Python + C/C++)
8	PointPillars	code	59.07 %	75.78 %	52.92 %	16 ms	1080ti GPU and Intel i7 CPU

A. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang and O. Beijbom: PointPillars: Fast Encoders for Object Detection from Point Clouds. CVPR 2019.

9	Tencent_ADlab_Lidar		58.19 %	72.12 %	51.54 %	0.1 s	GPU @ 2.5 Ghz (Python + C/C++)
10	MDC		57.27 %	75.27 %	49.75 %	0.17 s	GPU @ 2.5 Ghz (Python)
11	epBRM		56.94 %	70.52 %	51.70 %	0.10 s	1 core @ 2.5 Ghz (C/C++)
12	A-VoxelNet		56.86 %	70.65 %	50.76 %	0.029 s	GPU @ 2.5 Ghz (Python)
13	F-PointNet	code	56.77 %	71.96 %	50.39 %	0.17 s	GPU @ 3.0 Ghz (Python)

C. Qi, W. Liu, C. Wu, H. Su and L. Guibas: Frustum PointNets for 3D Object Detection from RGB-D Data. arXiv preprint arXiv:1711.08488 2017.

14	CONV-BOX		54.45 %	68.27 %	52.26 %	0.2 s	Tesla V100
15	SECOND	code	53.85 %	70.51 %	46.90 %	38 ms	1080Ti

Y. Yan, Y. Mao and B. Li: SECOND: Sparsely Embedded Convolutional Detection. Sensors 2018.

16	IPOD		53.46 %	71.40 %	48.34 %	0.2 s	GPU @ 2.5 Ghz (Python + C/C++)
17	SCANet		53.07 %	67.97 %	50.81 %	0.17 s	>8 cores @ 2.5 Ghz (Python)
18	AVOD-FPN	code	52.18 %	64.00 %	46.61 %	0.1 s	Titan X (Pascal)

J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation. IROS 2018.

19	DSS		51.90 %	65.04 %	46.94 %	0.03 s	GPU @ 2.5 Ghz (C/C++)
20	ELLIOT		51.17 %	68.87 %	46.35 %	0.1 s	1 core @ 2.5 Ghz (C/C++)
21	CFR		50.73 %	65.70 %	44.93 %	0.06 s	1 core @ 2.5 Ghz (C/C++)
22	PP_v1.0	code	50.60 %	66.86 %	44.84 %	0.02s	1 core @ 2.5 Ghz (C/C++)
23	AVOD	code	44.90 %	60.11 %	38.80 %	0.08 s	Titan X (pascal)

J. Ku, M. Mozifian, J. Lee, A. Harakeh and S. Waslander: Joint 3D Proposal Generation and Object Detection from View Aggregation. IROS 2018.

24	X_MD		37.22 %	51.69 %	36.44 %	0.2 s	1 core @ 2.5 Ghz (C/C++)
25	anm		35.86 %	50.06 %	31.11 %	3 s	1 core @ 2.5 Ghz (C/C++)
26	CLF3D		35.39 %	50.58 %	33.55 %	0.13 s	GPU @ 2.5 Ghz (Python)

27	Complexer-YOLO			23.48 %	28.36 %	22.85 %	0.06 s	GPU @ 3.5 Ghz (C/C++)	
	M. Simon, K. Amende, A. Kraus, J. Honer, T. Samann, H. Kaulbersch, S. Milz and H. Michael Gross: Complexer-YOLO: Real-Time 3D Object Detection and Tracking on Semantic Point Clouds. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops 2019.								
28	a			21.25 %	32.66 %	19.77 %	0.35 s	1 core @ 2.5 Ghz (Python + C/C++)	
29	SA_3D			14.15 %	17.99 %	13.52 %	0.3 s	GPU @ 2.5 Ghz (Python)	
30	BirdNet			12.43 %	18.35 %	11.88 %	0.11 s	Titan Xp GPU	
	J. Beltrán, C. Guindel, F. Moreno, D. Cruzado, F. García and A. Escalera: BirdNet: A 3D Object Detection Framework from LiDAR Information. 2018 21st International Conference on Intelligent Transportation Systems (ITSC) 2018.								
31	MonoPSR			11.01 %	13.43 %	9.93 %	0.2 s	GPU @ 3.5 Ghz (Python)	
	J. Ku*, A. Pon* and S. Waslander: Monocular 3D Object Detection Leveraging Accurate Proposals and Shape Reconstruction. CVPR 2019.								
32	SS3D			9.09 %	10.84 %	9.09 %	48 ms	Tesla V100 (Python)	
	E. Jörgensen, C. Zach and F. Kahl: Monocular 3D Object Detection and Box Fitting Trained End-to-End Using Intersection-over-Union Loss. CoRR 2019.								
33	TopNet-UncEst			7.36 %	8.53 %	6.93 %	0.09 s	NVIDIA GeForce 1080 Ti (tensorflow-gpu)	
	S. Wirges, M. Braun, M. Lauer and C. Stiller: Capturing Object Detection Uncertainty in Multi-Layer Grid Maps. 2019.								
34	RT3DStereo			6.63 %	6.62 %	4.03 %	0.08 s	GPU @ 2.5 Ghz (C/C++)	
35	TopNet-HighRes			5.98 %	4.48 %	6.18 %	101ms	NVIDIA GeForce 1080 Ti (tensorflow-gpu)	
	S. Wirges, T. Fischer, C. Stiller and J. Frias: Object Detection and Classification in Occupancy Grid Maps Using Deep Convolutional Networks. 2018 21st International Conference on Intelligent Transportation Systems (ITSC) 2018.								
36	Shift R-CNN (mono)		code	3.03 %	3.03 %	3.03 %	0.25 s	GPU @ 1.5 Ghz (Python)	
	A. Naiden, V. Paunescu, G. Kim, B. Jeon and M. Leordeanu: Shift R-CNN: Deep Monocular 3D Object Detection With Closed-form Geometric Constraints. ICIP 2019.								
37	mylsi-faster-rcnn			1.68 %	2.53 %	1.41 %	0.3 s	1 core @ 2.5 Ghz (Python)	
38	DT3D			1.20 %	1.76 %	1.26 %	0,21s	GPU @ 2.5 Ghz (Python)	
39	M3D-RPN		code	1.03 %	1.72 %	1.05 %	0.16 s	GPU @ 1.5 Ghz (Python)	
	G. Brazil and X. Liu: M3D-RPN: Monocular 3D Region Proposal Network for Object Detection . ICCV 2019 .								
40	mymask-rcnn			0.84 %	1.18 %	0.83 %	0.3 s	1 core @ 2.5 Ghz (Python)	
41	OFT-Net			0.43 %	0.43 %	0.43 %	0.5 s	8 cores @ 2.5 Ghz (Python + C/C++)	
42	mBoW			0.00 %	0.00 %	0.00 %	10 s	1 core @ 2.5 Ghz (C/C++)	
	J. Behley, V. Steinhage and A. Cremers: Laser-based Segment Classification Using a Mixture of Bag-of-Words. Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2013.								

Table as LaTeX | Only published Methods

Related Datasets

- CERV Vehicle Lights Dataset: Annotations of vehicle lights for a subset of the object detection benchmark.
- PASCAL3D+: Augments 12 rigid object classes of PASCAL VOC 2012 with 3D annotations.
- The PASCAL Visual Object Classes Challenges: Dataset and benchmarks for object class recognition.
- TME Motorway Dataset: 28 video sequences with vehicle annotations captured from VisLab's BRAiVE vehicle.
- LabelMe: Online annotation tool to build image databases for computer vision research.
- MIT Street Scenes: Street-side images with labels for 9 object categories (including cars, pedestrians, buildings, trees).
- Daimler Pedestrian Datasets: Datasets focusing on pedestrian detection for autonomous driving.
- Caltech Pedestrian Detection Benchmark: 10 hours of video with 350.000 annotated pedestrian bounding boxes.
- Robust Multi-Person Tracking from Mobile Platforms: Videos with annotated pedestrians captured from a stroller.

Citation

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

```
@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}
}
```