Machine learning-based occupancy grid mapping

Peiyuan Zhai

Abstract

In this note, there are contents and comments on ML-based occupancy grid mapping algorithms. This is the public version of a private copy of the note. It might be updated in the future.

I. TABLE OF WORKS

TABLE I: Table of Deep Learning-based OGM methods

Type	Name	Year	Code	Structure	Input	Output	Comments	Pro	Con
Single modal	RadarOcc [1]	2024	✓	transformer	4D radar	3D SSC		• DL structure help 4D radar feature learning (i.e., doppler) • second point	
	Occ3D [2]	2023	✓	transformer	RGB	3D SSC			
	VoxelNet [3]	2017	✓	CNN	LiDAR	3D box			
	SurroundOcc [4]	2023	✓	transformer	RGB	3D SSC			
	DeepISM [5]	2019	✗	VAE	radar	2D OGM	Evidence available		
	Evidential DeepISM [6]	2019	✗	VAE	radar	2D (E)OGM	Evidence available		
	PointPillars [7]	2019	✓	CNN	LiDAR	3D box			
	OccupancyNet [8]	2019	✗	CNN	radar BEV	2D OGM			
	SemanticPON [9]	2020	✓	STPN	RGB	2D SSC			
	MotionNet [10]	2020	✓	STPN	LiDAR	2D OGM			
	EviLOG [11]	2021	✓	CNN	radar	2D EOGM			
	RPFA-Net [12]	2021	✓	CNN	radar	2D OGM			
Single sensor augmented			✗ ✗ ✗ ✗						
Multi-sensors	MVDNet [13] ST-MVDNet [14] ST-MVDNet++ [15]	2021 2022 2023	✓ ✗ ✗	CNN	radar LiDAR	2D box		• Multi-sensor fusion • s	• High computational cost • Complex architecture
	MV3D [16]	2017	✓	CNN	LiDAR camera	3D box			
	DTW [17]	2018	✗	CNN+RNN	LiDAR	2D OGM+box			
	AVOD [18]	2018	✗	CNN	LiDAR camera	3D box			
	FISHING [19]	2020	✗	CNN	radar LiDAR camera	3D box	Limited semantic, some obstacles are not detected		
	RadarNet [20]	2020	✗	CNN	radar LiDAR	3D box			

Note: SSC: Semantic Scene Completion, OGM: Occupancy Grid Map, EOGM: Evidential OGM

REFERENCES

[1] F. Ding, X. Wen, Y. Zhu, Y. Li, and C. X. Lu, "RadarOcc: Robust 3D Occupancy Prediction with 4D Imaging Radar," Oct. 2024. 5 citations (Semantic Scholar/arXiv) [2025-01-07] arXiv:2405.14014 [cs].

[2] X. Tian, T. Jiang, L. Yun, Y. Mao, H. Yang, Y. Wang, Y. Wang, and H. Zhao, "Occ3D: a large-scale 3D occupancy prediction benchmark for autonomous driving," Dec. 2023. arXiv:2304.14365 [cs].

[3] Y. Zhou and O. Tuzel, "VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection," Nov. 2017. arXiv:1711.06396 [cs].

[4] Y. Wei, L. Zhao, W. Zheng, Z. Zhu, J. Zhou, and J. Lu, "SurroundOcc: multi-camera 3D occupancy prediction for autonomous driving," Aug. 2023. 161 citations (Semantic Scholar/arXiv) [2025-02-23] arXiv:2303.09551 [cs].

[5] R. Weston, S. Cen, P. Newman, and I. Posner, "Probably Unknown: Deep Inverse Sensor Modelling In Radar," May 2019. arXiv:1810.08151 [cs].

[6] D. Bauer, L. Kuhnert, and L. Eckstein, "Deep, spatially coherent Inverse Sensor Models with Uncertainty Incorporation using the evidential Framework," Mar. 2019. 12 citations (Semantic Scholar/arXiv) [2025-01-10] arXiv:1904.00842 [cs].

[7] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, "PointPillars: Fast Encoders for Object Detection from Point Clouds," May 2019. arXiv:1812.05784 [cs].

[8] L. Sless, B. E. Shlomo, G. Cohen, and S. Oron, "Road Scene Understanding by Occupancy Grid Learning from Sparse Radar Clusters using Semantic Segmentation," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 867–875, Oct. 2019. 44 citations (Crossref) [2024-08-21] ISSN: 2473-9944.

[9] T. Roddick and R. Cipolla, "Predicting Semantic Map Representations from Images using Pyramid Occupancy Networks," Mar. 2020. arXiv:2003.13402 [cs].

[10] P. Wu, S. Chen, and D. N. Metaxas, "MotionNet: Joint Perception and Motion Prediction for Autonomous Driving Based on Bird's Eye View Maps," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (Seattle, WA, USA), pp. 11382–11392, IEEE, June 2020.

[11] R. Van Kempen, B. Lampe, T. Woopen, and L. Eckstein, "A Simulation-based End-to-End Learning Framework for Evidential Occupancy Grid Mapping," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, pp. 934–939, July 2021. 6 citations (Crossref) [2024-08-21].

[12] B. Xu, X. Zhang, L. Wang, X. Hu, Z. Li, S. Pan, J. Li, and Y. Deng, "RPFA-net: a 4D RaDAR pillar feature attention network for 3D object detection," in *2021 IEEE International Intelligent Transportation Systems Conference (itsc)*, pp. 3061–3066, Sept. 2021.

[13] K. Qian, S. Zhu, X. Zhang, and L. E. Li, "Robust multimodal vehicle detection in foggy weather using complementary lidar and radar signals," in *2021 Ieee/cvf Conference on Computer Vision and Pattern Recognition (cvpr)*, pp. 444–453, June 2021. ISSN: 2575-7075.

[14] Y.-J. Li, J. Park, M. O'Toole, and K. Kitani, "Modality-agnostic learning for radar-lidar fusion in vehicle detection," in *2022 Ieee/cvf Conference on Computer Vision and Pattern Recognition (cvpr)*, (New Orleans, LA, USA), pp. 908–917, IEEE, June 2022.

[15] Y.-J. Li, M. O'Toole, and K. Kitani, "ST-MVDNet++: Improve Vehicle Detection with Lidar-Radar Geometrical Augmentation via Self-Training," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, June 2023. 5 citations (Semantic Scholar/DOI) [2025-02-24] ISSN: 2379-190X.

[16] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, "Multi-view 3D Object Detection Network for Autonomous Driving," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (Honolulu, HI), pp. 6526–6534, IEEE, July 2017.

[17] J. Dequaire, P. Ondrúška, D. Rao, D. Wang, and I. Posner, "Deep tracking in the wild: End-to-end tracking using recurrent neural networks," *The International Journal of Robotics Research*, vol. 37, pp. 492–512, Apr. 2018. 64 citations (Crossref) [2024-08-21] Publisher: SAGE Publications Ltd STM.

[18] J. Ku, M. Mozifian, J. Lee, A. Harakeh, and S. Waslander, "Joint 3D Proposal Generation and Object Detection from View Aggregation," July 2018. arXiv:1712.02294 [cs].

[19] N. Hendy, C. Sloan, F. Tian, P. Duan, N. Charchut, Y. Xie, C. Wang, and J. Philbin, "FISHING Net: Future Inference of Semantic Heatmaps In Grids," June 2020. 73 citations (Semantic Scholar/arXiv) [2025-02-24] arXiv:2006.09917 [cs].

[20] B. Yang, R. Guo, M. Liang, S. Casas, and R. Urtasun, "RadarNet: Exploiting Radar for Robust Perception of Dynamic Objects," in *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII*, (Berlin, Heidelberg), pp. 496–512, Springer-Verlag, Aug. 2020. 101 citations (Semantic Scholar/DOI) [2025-01-09].