

Contents

Abstract	v
Acknowledgements	vii
Zusammenfassung	ix
1 Introduction	1
1.1 Summary of contributions	4
1.2 Publications	6
1.3 Industry PhD Program of Daimler AG	7
2 Technical Background	9
2.1 Basic tasks of ADAS and AVs perception appilcation .	9
2.1.1 Object classification	10
2.1.2 Object detection	10
2.1.3 Object segmentation	11
2.1.4 Object Tracking	12
2.2 Automotive Perception Sensor	12
2.2.1 Camera	13
2.2.2 LiDAR	18
2.2.3 Radar	25
2.3 Sensor Fusion Technology	34
2.3.1 Decision level sensor fusion	37
2.3.2 Feature Level Sensor Fusion	47

2.3.3	Data Level Sensor Fusion	48
2.3.4	Multi level sensor fusion	49
3	High Dimensional Frustum PointNet for 3D Object Detection from Camera, LiDAR, and Radar	51
3.1	Introduction	51
3.2	Related work	54
3.2.1	Object detection based on LiDAR	55
3.2.2	Object detection based on radar	57
3.2.3	Object detection based on sensor fusion	58
3.3	High Dimensional Frustum PointNet	60
3.3.1	2D frustum proposal	61
3.3.2	7D PointNet for LiDAR point clouds	61
3.3.3	8D PointNet for radar point clouds	63
3.3.4	Network architecture details	65
3.3.5	Loss	67
3.4	Experimental setup	68
3.4.1	Dataset and training details	68
3.4.2	Comparison between different methods	72
3.5	Conclusions	74
3.6	Acklowdegments	74
4	Sparse PointNet: See further in Autonomous Vehicles	75
4.1	INTRODUCTION	75
4.2	Related Work	80
4.2.1	LiDAR-based 3D object detection	80
4.2.2	Multimodal-based 3D object detection	81
4.3	Sparse-PointNet	82
4.3.1	Preprocessing	84
	LiDAR	84

	Radar	85
4.3.2	Dynamic continuous occupancy heatmap . . .	85
4.3.3	Sampling	87
4.3.4	Proposal generation module	89
4.3.5	Proposal refinement module	91
4.4	Experimental setup	91
4.4.1	Dataset	92
4.4.2	Implementation details	92
4.4.3	Main Results	95
4.4.4	Ablation Study	99
4.5	CONCLUSIONS	101
5	L2R GAN: LiDAR-to-Radar Translation	103
5.1	Introduction	103
5.1.1	Contributions	105
5.2	Background and related work	105
5.2.1	Cross-Domain data translation with conditional GAN	106
5.2.2	Representations of radar data	107
5.3	The pix2pix and pix2pixHD baseline implementations	109
5.3.1	Architecture of pix2pix and pix2pixHD	109
5.3.2	Drawbacks of the baseline approaches	110
5.4	Our L2R-GAN framework	111
5.4.1	Occupancy-grid-mask guidance for the global generator	113
5.4.2	Local region generator and discriminator . . .	114
5.4.3	Objective functions	115
5.5	Experimental results	116
5.5.1	Baseline comparisons	117
	Implementation details.	117

	Training, validation and test datasets.	118
	Evaluation metrics.	120
	Baseline comparisons.	121
5.5.2	Ablation Analysis	121
	Analysis of the framework structure.	121
	Analysis of the loss functions.	122
5.6	Application: data augmentation for ADAS	123
5.7	Conclusion	124
5.8	Supplementary Material	124
5.8.1	Occupancy Grid Mask	124
5.8.2	Emergency Scenario Generator	126
5.9	Acklowdegments	127
6	Radar Ghost Target Detection via Multimodal Transform-	
	ers	131
6.1	INTRODUCTION	131
6.2	Related Work	136
6.2.1	Radar Ghost Target Detection	136
6.2.2	Transformers on Point Clouds	137
6.3	Attention modules	138
6.3.1	Background: Scaled Dot Product Attention	138
6.3.2	Self-Attention Module for Radar Points	140
6.3.3	Multimodal Attention Module	141
6.4	Multimodal transformer network	142
6.4.1	Data Encoding	142
6.4.2	Network Architecture	143
6.5	Experimental setup	145
6.5.1	Dataset	145
6.5.2	Ground Truth Generation	146
6.5.3	Implementation Details	147

6.5.4	Impact on Object Detection	149
6.6	Results	149
6.6.1	Baselines	149
6.6.2	Ablation Studies	150
6.6.3	Qualitative Analysis	151
6.6.4	Analysis of Impact on Object Detection	151
6.7	CONCLUSIONS	152
7	Conclusion, Challenge, and Outlook	159
7.1	Conclusion	159
7.2	Challenges and Outlook	161
7.2.1	Data representation	161
7.2.2	Fusion Strategies	162
7.2.3	Industrial project	163
	Bibliography	167