Poster: AutoSense: Reliable 3D Bounding Box Prediction for Vehicles

Hem Regmi; Reza Tavasoli; Joseph Telaak; Sanjib Sur; Srihari Nelakuditi
Department of Computer Science and Engineering; University of South Carolina, Columbia, USA
hregmi@email.sc.edu; tavasoli@email.sc.edu; jtelaak@email.sc.edu; sur@cse.sc.edu; srihari@sc.edu

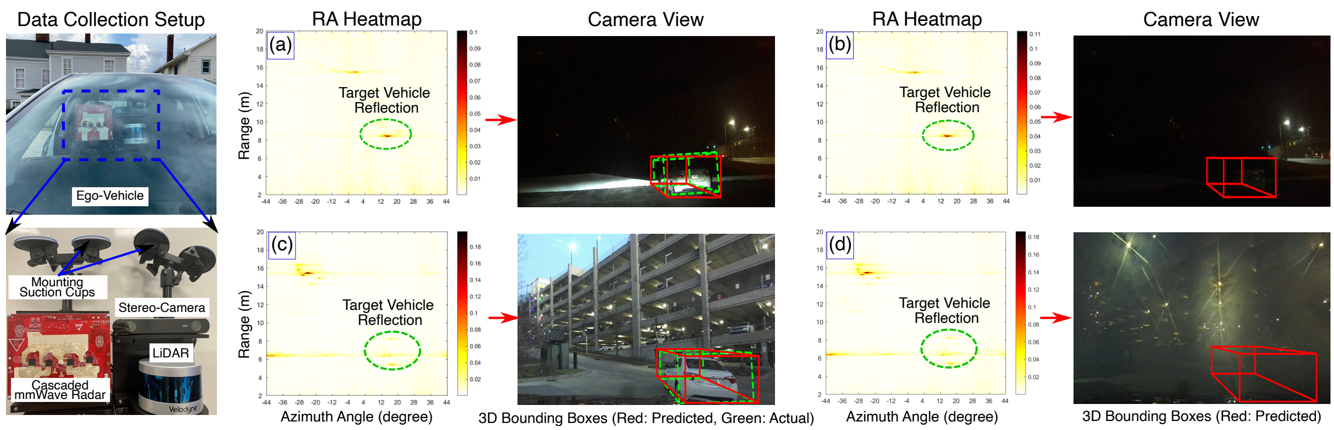

Figure 1: Performance comparison of *AutoSense* and camera-based approach under varying weather and lighting conditions. The figures depict camera view overlaid with predicted (red) and actual (green) 3D bounding boxes, alongside corresponding mmWave radar Range-Azimuth (RA) heatmaps showing the target vehicle's reflection. Nighttime scenarios (a, b): In complete darkness (b), the camera fails while *AutoSense* accurately predicts the vehicle's 3D bounding box. Daytime scenarios (c, d): In clear conditions (c), both approaches perform well; in foggy weather (d), the camera fails while *AutoSense* remains operational.

ABSTRACT

We propose *AutoSense*, a millimeter-wave (mmWave) wireless signal-based system for predicting 3D bounding boxes of vehicles. While cameras and LiDAR can be adversely affected by challenging weather conditions such as heavy rain, fog, or snow, mmWave signals are less susceptible to these environmental factors, making them more resilient. As a result, *AutoSense* can complement other sensors for accurate 3D bounding box predictions in all weather conditions.

CCS CONCEPTS

• **Computing methodologies** → **Object detection**.

KEYWORDS

Object Detection; Deep Learning; Millimeter-wave Radars

ACM Reference Format:
Hem Regmi; Reza Tavasoli; Joseph Telaak; Sanjib Sur; Srihari Nelakuditi, Department of Computer Science and Engineering; University of South Carolina, Columbia, USA, hregmi@email.sc.edu; tavasoli@email.sc.edu; jtelaak@email.sc.edu; sur@cse.sc.edu; srihari@sc.edu. 2024. Poster: AutoSense: Reliable 3D Bounding Box Prediction for Vehicles. In *ACM International Conference on Mobile Systems, Applications, and Services (Mobisys '24), June 3–7, 2024, Minato-ku, Tokyo, Japan.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3643832.3661416

MOBISYS '24, June 3–7, 2024, Minato-ku, Tokyo, Japan

ACM ISBN 979-8-4007-0581-6/24/06
https://doi.org/10.1145/3643832.3661416

1 INTRODUCTION

Every year, millions of vehicle crashes occur, and a considerable number of these accidents are caused by hazardous weather conditions, leading to thousands of fatalities [2]. To improve road safety, Advanced Driver Assistance Systems (ADAS) have been developed, that rely on cameras and LiDAR to provide accurate 3D bounding box predictions of surrounding vehicles on the road. However, the effectiveness of these sensors diminishes in harsh weather, such as fog, rain, or snow, where visibility is severely compromised [4]. Millimeter-wave (mmWave) radars offer a potential solution to the limitations of cameras and LiDAR, as they can penetrate rain and fog, enabling them to operate effectively in adverse weather conditions. However, estimating accurate 3D bounding boxes of vehicles using mmWave radars is difficult due to the motion of vehicles and the specular nature of mmWave reflections.

We introduce AutoSense, *a system that utilizes cascaded mmWave radar to predict 3D bounding boxes of vehicles in all weather conditions (see Figure 1). AutoSense* applies motion-error correction, processes mmWave reflections to generate heatmaps, and employs a deep learning approach to predict accurate 3D bounding boxes. Once trained on mmWave reflections and corresponding ground-truth

3D bounding boxes from camera in clear weather, *AutoSense* relies solely on mmWave radar for prediction.

2 *AUTOSENSE* SYSTEM DESIGN

AutoSense predicts the 3D bounding boxes of vehicles using a cascaded mmWave radar inside the ego-vehicle, providing complementary information to the camera and LiDAR in harsh weather conditions. The mmWave radar employs Frequency Modulated Continuous Wave (FMCW) [3] to accurately detect objects. The radar continuously transmits FMCW signals and receives reflections from objects in the scene. These reflections arrive at the receiver with different frequencies and phases, depending on the target object's distance. By mixing the transmitted and received signals, we can determine the objects' distance from the radar.

AutoSense faces two primary challenges: (1) motion errors due to time division multiplexing of cascaded mmWave radar, and (2) specularity of mmWave and variable reflectivity of vehicles. Motion errors occur due to sequential signal transmission, causing small timing differences that can lead to object displacement on the radar heatmap. Additionally, the specular nature of mmWave reflections results in only a few high-energy points being visible on the Range-Azimuth (RA) heatmap, with most parts of the vehicle missing. To compensate for motion errors, we take advantage of *AutoSense*'s mmWave radar's antenna configuration. The mmwave radar is equipped with 12 transmitters and 16 receivers, forming 192 virtual antennas of which 32 are co-located and overlapping. We leverage the phase differences between these overlapping virtual antennas to resolve motion error.

For 3D bounding box prediction, *AutoSense* employs Convolutional Neural Networks (CNNs) to extract object features from mmWave heatmaps. The system utilizes YOLO-based object detection blocks [1] to extract these features, and head detection modules to predict the center location, length, width, and height of vehicles on the road. During the training phase, *AutoSense* uses 3D bounding boxes obtained from a stereo camera as ground truth labels to supervise the learning process and optimize the network parameters. By using the camera-based 3D bounding boxes as a reference, the system learns to accurately predict the 3D bounding boxes from the mmWave heatmaps alone. Once trained, *AutoSense* only requires mmWave heatmaps as input to predict 3D bounding boxes during the testing or deployment phase.

3 PRELIMINARY RESULTS

We evaluate *AutoSense* using three standard metrics for 3D bounding box prediction: Intersection-over-Union (IoU), mean Average Precision (mAP), and Mean Absolute Error (MAE). IoU indicates the volume overlap between the actual and predicted 3D bounding boxes, with a value between 0 and 1. mAP accounts for False Positives (FP) and False Negatives (FN) and is defined as the area under the precision-recall curve. MAE represents the absolute difference between the predicted and actual values for length, width, height, and center distance of vehicles. Figure 2(a) shows the CDF of IoU. We observe a median IoU of 0.75, indicating a high degree of overlap between the predicted and ground-truth bounding boxes. *AutoSense* achieves an mAP of 0.64 for vehicles at an IoU threshold of 0.5, which drops to 0.39 at an IoU threshold of 0.7 (Figure 2[b]),

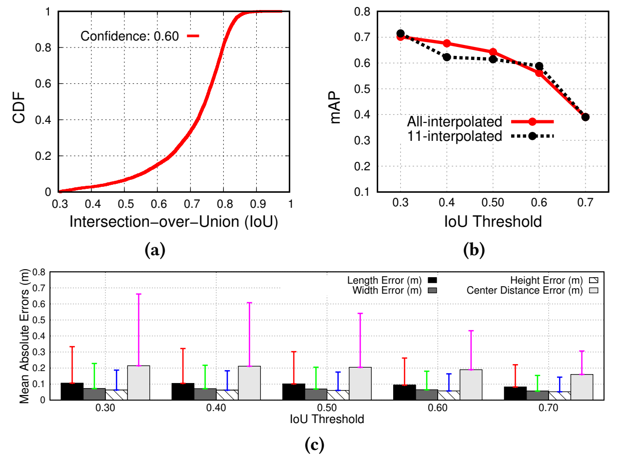

Figure 2: (a) CDF of IoU at IoU threshold of 0.3 and 60% prediction confidence of the vehicle; (b) mAP of vehicles with all-interpolated and 11-interpolated methods; (c) Median and 90^{th} percentile errors of vehicle length, width, height, and 3D center distance at different IoU thresholds.

but it can be improved using multiple cascaded mmWave radars. Figure 2(c) shows the median and 90^{th} percentile plots of length, width, height, and center distance at different IoU thresholds. The average vehicle length, width, and height in our test data samples are 4.22, 2.31, and 1.69 meters, respectively. *AutoSense* estimates vehicles' length, width, and height with less than 7.5%, 9.0%, and 10.5% errors for the 90^{th} percentile of test samples.

4 CONCLUSION AND FUTURE WORKS

In conclusion, *AutoSense* demonstrates the feasibility of using mmWave radar to predict 3D bounding boxes of vehicles, even in challenging weather conditions. We plan to extend *AutoSense* to predict 3D bounding boxes for pedestrians and other objects. We will also explore integrating *AutoSense* with other sensor modalities to create a more robust system.

ACKNOWLEDGMENTS

This work is partially supported by the NSF under grants CAREER-2144505 and MRI-2018966.

REFERENCES

[1] Ultralytics. 2024. Neural Network Blocks. https://github.com/ultralytics/ultralytics/tree/main/ultralytics/nn
[2] U.S. Department of Transportation. 2024. Low Visibility. https://ops.fhwa.dot.gov/weather/weather_events/low_visibility.htm
[3] Wang, Guochao and Munoz-Ferreras, Jose-Maria and Gu, Changzhan and Li, Changzhi and Gomez-Garcia, Roberto. 2014. Application of linear-frequency-modulated continuous-wave (LFMCW) radars for tracking of vital signs. *IEEE Transactions on Microwave Theory and Techniques* 62, 6 (2014), 1387–1399.
[4] Yaqoob, Ibrar and Khan, Latif U and Kazmi, SM Ahsan and Imran, Muhammad and Guizani, Nadra and Hong, Choong Seon. 2019. Autonomous driving cars in smart cities: Recent advances, requirements, and challenges. *IEEE Network* 34, 1 (2019), 174–181.