

**“A Study on Advancements in Camera-Radar
Sensor Fusion For Accurate Radar Key-Point
Predictions”**



OLLSCOIL NA GAILLIMHE

UNIVERSITY OF GALWAY

Aditi Vilas Sonkusare
School of Computer Science
University of Galway

Supervisor(s)

Dr. Takfarinas Saber

In partial fulfillment of the requirements for the degree of
MSc in Computer Science (Data Analytics)

Abstract

In recent years, autonomous vehicles and Advanced driver-assistance systems (ADAS) have received a lot of attention. These systems rely significantly on sensor data to detect and comprehend their surroundings. Sensor fusion is essential for improving perception, object detection, tracking, and prediction capabilities by merging data from cameras, and radar sensors. The advancements in camera and radar sensor fusion technologies for forecasting radar key points are the emphasis of this thesis.

Keywords: Camera, Radar, Perception, Object Detection, Early Sensor Fusion, Object Tracking, Synergy, Environmental Awareness, Sensors, Pedestrians, Car, Bus, Truck, Barrier, Motorcycle, Training models, Deep Learning Algorithm

Contents

1	Introduction	1
1.1	Motivation	2
1.2	Purpose	3
2	Background Sensors in Autonomous Vehicles	4
2.1	Sensors	4
2.2	Sensor Fusion	5
3	Related Work	6
3.1	Sensor Data Fusion in Autonomous Vehicles	6
3.2	Sensor Data Fusion in Related Domain	11
3.3	Algorithms and Mechanism Used	12
3.4	Evaluation metrics	14
	References	18

Chapter 1

Introduction

The safe navigation of autonomous vehicles in the upcoming era depends on their ability to accurately perceive the environment. This is accomplished by fusing various sensors using camera radar to create a comprehensive picture of the environment. In order to improve accuracy and dependability, sensor fusion involves combining data from various sensors. This fusion is essential for autonomous vehicles because decisions are made in nanoseconds based on sensor input. In order for these vehicles to perceive and make decisions in real-time, sensor fusion is essential. The main sensors are cameras and radar, and their combination provides increased dependability, accuracy, and a thorough understanding of the environment. Radar provides distance, velocity, and angle information for obstacle detection and motion prediction, while cameras provide high-resolution images and color information to aid in object recognition and road sign identification. From a distance, radar can detect an obstacle and its key points, and a camera can further compare those key points with those discovered by radar. Sensor fusion enables autonomous vehicles to obtain a complete and accurate view of their surroundings, even in difficult circumstances, by making up for the shortcomings of individual sensors.

1.1 Motivation

Reliability and redundancy:

The combination of camera and radar sensors gives redundant and complementary information. The system can increase dependability by fusing data from both sensors, especially in cases where one sensor may have limits or fail.

Improved object identification and tracking:

When it comes to object recognition and tracking, cameras, and radars have various strengths. By combining the data from both sensors, it is able to take advantage of each sensor's strengths and create a more precise and robust understanding of the surrounding environment, resulting in increased object detection and tracking capabilities.

Improved perception in difficult settings:

Cameras can struggle in severe weather conditions such as heavy rain, fog, or glare, which can impair image quality and visibility. Radars, on the other hand, are unaffected by these factors and can offer accurate distance and velocity measurements. The fusion system can compensate for the limitations of each sensor and retain perception capabilities even under difficult settings by integrating camera and radar data.

Improved security and taking decisions:

By combining camera and radar data, the system may gain a more extensive and accurate picture of the environment, boosting the accuracy and reliability of item recognition, categorization, and prediction. As a result, safer and more informed judgments about vehicle control, accident avoidance, and path planning are made.

1.2 Purpose

Camera-radar sensor fusion has numerous practical uses, including the development of autonomous vehicles and enhanced driver assistance systems. Choosing this topic may suggest a desire to learn more about the underlying technology and its potential impact on enhancing safety, dependability, and performance in autonomous vehicles.

Relevance to industry:

The concept of camera-radar sensor fusion has received considerable attention in businesses involved in autonomous systems, such as automotive and technology corporations.

Advancements:

Camera-radar sensor fusion is a new technology that has the potential to improve perception capabilities in applications other than autonomous cars, such as surveillance systems, drones, and robotics. Exploring this topic helps to stay up to date on the most recent advances and anticipated future breakthroughs in this field.

Industrial interest:

The automotive and technology industries are investing heavily in the development of autonomous vehicles and ADAS technologies. The selection of camera-radar sensor fusion as a topic corresponds with industry expectations and the need for novel techniques to improve the perception capabilities of autonomous systems. Having a keen interest in this field, to make a contribution to ongoing breakthroughs and address the issues involved with sensor fusion for autonomous driving applications.

Chapter 2

Background Sensors in Autonomous Vehicles

Sensors in autonomous vehicles are crucial for perceiving the environment and gathering information needed for safe and efficient navigation. As the center of this project's autonomous cars, cameras and radar are two critical sensors. This section will provide a detailed discussion of the primary sensors used as well as the sensor fusion employing the primary sensors.

2.1 Sensors

Camera:

Like the human eye, cameras in autonomous vehicles capture visual information from their surroundings. They use image sensors to transform light into digital data, which the vehicle's perception system may then analyze and interpret. Cameras can help with activities including object detection, lane detection, traffic sign recognition, and pedestrian detection. They produce high-resolution images that allow for detailed item analysis and recognition.

Radar:

Radar (Radio Detection and Ranging) sensors detect and measure the distance, velocity, and direction of objects in the vehicle’s surroundings using radio waves. They send out radio waves and examine the reflections that bounce back from nearby objects. Radar sensors are especially useful in bad weather and low-light situations, where cameras may struggle. They provide vital information regarding an object’s position and speed, making them essential for collision avoidance and adaptive cruise control systems.

2.2 Sensor Fusion

Camera-Radar Sensor fusion is the process of merging data from various sensors to provide a more complete and accurate representation of the environment. Camera and radar sensor fusion is a common strategy in the area of autonomous cars. By combining camera and radar data, autonomous cars can boost perception capabilities by leveraging the strengths of both sensors. Cameras, for example, give rich visual information that allows for accurate object recognition and identification. Radar, on the other hand, can offer accurate distance and velocity information even under adverse weather conditions. The vehicle’s perception system may build a more robust picture of the surroundings by merging input from both sensors, improving object detection, tracking, and prediction. Sensor fusion algorithms process and combine data from cameras and radar to produce a unified sensory output. Other components within the autonomous vehicle, use this output to plan suitable actions and maneuvers. Overall, the combination of cameras and radar, as well as sensor fusion algorithms, enables autonomous vehicles to precisely perceive their surroundings, make intelligent decisions, and navigate safely in a variety of driving circumstances.

Chapter 3

Related Work

This section is divided into four sections, the first of which gives a brief overview of research papers in the autonomous vehicle domain, the second of which gives a brief overview of different domains that are part of related work, the third of which tells about the main algorithms and mechanisms that are widely used, and the fourth of which gives information about the various evaluation metrics used to predict the accuracy and precision of the model.

3.1 Sensor Data Fusion in Autonomous Vehicles

Authors Shenglin Li and Hwan-Sik Yoon [1] offer a technique for improving vehicle detection and tracking accuracy and reliability at a traffic intersection by combining data from a single camera and radar sensor. Sensor fusion is the merging of data from various sensors to gain a more complete and accurate understanding of the environment. The suggested method intends to address the constraints and challenges associated with employing a single sensor in complicated traffic scenarios, such as occlusions or problematic lighting conditions, by combining data from both sensors. The combination of camera and radar data allows for

3.1 Sensor Data Fusion in Autonomous Vehicles

more robust and accurate vehicle detection and tracking, which is critical for a variety of applications in intelligent transportation systems.

Authors Keni Bernardin and Rainer Stiefelhagen [2] tackles the need for objective methods to evaluate the accuracy and efficacy of algorithms that track many objects in films or other types of visual data. The MOT measures are introduced by the authors as a thorough evaluation methodology for multiple object tracking. MOT measurements contain a variety of metrics that quantify the quality of multiple objects tracking outcomes. These metrics analyze tracking accuracy, consistency, false positives, false negatives, and other indicators linked to the algorithm’s capacity to correctly identify and track objects over time.

Authors Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, Oscar Beijbom [3] states the relevance of large-scale, diversified datasets for training and evaluating autonomous driving systems. Existing datasets, they explain, frequently lack the complexity and diversity required to effectively depict real-world driving circumstances. To solve this restriction, the authors provide nuScenes, a comprehensive data set that includes a variety of sensor modalities such as LIDAR, RADAR, and cameras. This paper also explains the evaluation metrics; MOTA (Multiple Object Tracking Accuracy) and MOTP (Multiple Object Tracking Precision) are two extensively used evaluation metrics in computer vision and object tracking. These metrics are used to evaluate the effectiveness of object-tracking algorithms.

Authors Jelena Kocić, Nenad Jovičić, and Vujo Drndarević [4] discuss the significance of sensors in allowing autonomous cars to detect and interpret their surroundings. It emphasizes that autonomous cars collect data about their environment using various sensors such as cameras, RADAR (Radio Detection and Ranging), and ultrasonic sensors. Sensor fusion combines the strengths of vari-

3.1 Sensor Data Fusion in Autonomous Vehicles

ous sensors to acquire a complete and more accurate picture of the environment, overcoming the limitations and uncertainties that individual sensors may have. The paper will most likely examine various sensor fusion techniques and algorithms used in autonomous vehicles, such as Kalman filters, particle filters, and probabilistic approaches. It may also investigate the difficulties and concerns related to sensor fusion, such as data synchronization, sensor calibration, and data fusion methods. The paper delves into the role of sensors and the significance of sensor fusion in enabling autonomous cars to perceive and navigate their environment safely and effectively.

Authors De Jong Yeong, Gustavo Velasco-Hernandez, John Barry, and Joseph Walsh [5] discuss the relevance of sensors in enabling autonomous cars to sense their environment and make informed judgments is discussed in the study. Cameras, LIDAR, RADAR, ultrasonic sensors, and IMUs are among the sensors widely utilized in autonomous vehicles. This paper analyses the concepts, capabilities, and limitations of these sensors. The authors explain various sensor fusion approaches and algorithms that are used in autonomous vehicles to combine input from various sensors and improve perception capabilities. They include Kalman filters, particle filters, or machine learning methods. The study provides a thorough review of sensor technology and sensor fusion in autonomous cars, providing useful insights into their operation.

Authors Federico Castanedo, D. Ursino, Y. Takama [6] focus on combining multiple data sources to achieve more accurate and dependable information. The authors explain numerous data fusion techniques utilized in diverse fields, such as sensor fusion, image fusion, and information fusion. They also investigate various data fusion approaches, such as statistical methods, mathematical models, machine learning algorithms, and expert systems. Topics such as data preprocessing, feature extraction, decision fusion, and uncertainty handling in the context

3.1 Sensor Data Fusion in Autonomous Vehicles

of data fusion may be covered in the article. The writers may also explore data fusion difficulties such as data quality, sensor calibration, and the fusing of diverse data sources.

Authors Joseph Redmon, and Ali Farhadi [7] propose several enhancements, including a larger network design, the use of various scales for detection, and the inclusion of a technique known as "darknet-53" for feature extraction. These improvements aim to solve earlier versions' shortcomings and produce better detection performance across a wide range of item sizes and classes. The study is likely to include thorough information regarding YOLOv3's technical features, such as the network design, training methods, and particular improvements made to boost performance. The study proposes YOLOv3, an improved version of the YOLO object detection system, emphasizing incremental gains in accuracy and speed for real-time object recognition workloads.

Authors Arindam Sengupta; Feng Jin; Siyang Cao [8] addresses the topic of using many sensors to track targets such as vehicles or pedestrians. It combines data from an mm-Wave radar sensor, which provides range and velocity information, with data from a camera sensor, which gathers visual information. The authors suggest a fusion method based on a DNN-LSTM model. The DNN model analyses the visual data from the camera sensor, extracting key features, while the LSTM model handles the tracking task's temporal elements. The suggested approach intends to increase the accuracy and robustness of target tracking by merging information from both sensors. Overall, the research presents a sensor fusion technique that uses mm-Wave radar and camera sensors, as well as DNN and LSTM models, to improve target tracking capabilities.

Authors Martin Simon, Stefan Milz, Karl Amende, and Horst-Michael Gross [9] address the problem of recognizing objects in 3D point clouds generated by sensors. Complex-YOLO seeks to detect objects accurately and efficiently in real-

3.1 Sensor Data Fusion in Autonomous Vehicles

time circumstances. The authors suggest a method for generating probable object areas in a point cloud using Euler Region Proposals (ERPs). These ERPs are then analyzed for object detection and localization using a deep neural network inspired by the YOLO (You Only Look Once) framework. It may also explain the evaluation of Complex-YOLO on benchmark data sets, emphasizing its accuracy and speed in comparison to other existing approaches. Complex-YOLO is introduced in the research as a real-time 3D object recognition system that combines Euler Region Proposals and deep neural networks to detect objects accurately and efficiently.

Authors Teck-Yian Lim, Amin Ansari, Bence Major, Daniel Fontijne, Michael Hamilton, Radhika Gowaikar, Sundar Subramanian [10] The paper discusses the importance of accurate and dependable vehicle identification in ADAS applications to improve safety and enable intelligent driving features. It suggests a fusion strategy that integrates data from radar sensors and camera sensors early in the processing process. The authors discuss the architecture and technique of their suggested early fusion system. They might explain how radar and camera data are synchronized, calibrated, and fused to gain a comprehensive understanding of the cars in the area. This fusion strategy attempts to increase vehicle identification performance by leveraging the complementary strengths of radar (such as accurate range and velocity measurements) and cameras (such as rich visual information). The authors compare the performance of their method to that of other existing systems, demonstrating the benefits of early fusion for vehicle identification in ADAS. This study offers a radar and camera early fusion technique for vehicle recognition in ADAS, emphasizing the benefits of merging data from these two sensors at an early stage to increase accuracy and dependability.

3.2 Sensor Data Fusion in Related Domain

Object recognition using cameras is widely employed in a variety of fields, including autonomous driving, transportation, robotics, and even medicine. Using an object detection algorithm, numerous features of objects can be detected based on photos taken by the camera. You only look once (YOLO) is a well-known object detection method. YOLO [7,9] detects objects as a regression issue using a unique convolutional neural network, and darknet framework and delivers class probabilities for the observed objects in a single run. Detected objects are thus displayed in a rectangular box to indicate the kind, size, and location of the objects. YOLO has great accuracy and rapid processing speed when compared to other methods such as single shot multi-box detector (SSD), Faster-RCNN, Mask-RCNN, and CenterNet. YOLO can construct a bounding box for each detected car in a picture, allowing the position of several vehicles in the scene to be estimated [1]. Sensor fusion of camera and radar can take place in various domains, including:

Autonomous Vehicles:

Camera and radar sensor fusion are widely used in autonomous vehicles for perception and object detection. Cameras provide high-resolution visual information, while radars offer accurate distance and velocity measurements. Combining data from both sensors enhances the robustness and reliability of object detection and tracking systems.

Robotics:

Camera and radar fusion is valuable in robotics applications for environment perception, obstacle avoidance, and mapping. By combining the strengths of camera vision (object recognition, color detection) and radar (distance measurement, robustness in adverse weather conditions), robots can navigate and interact with their surroundings more effectively.

3.3 Algorithms and Mechanism Used

Surveillance and Security:

Integrating camera and radar sensors in surveillance systems enables comprehensive monitoring of large areas. Cameras provide visual identification and recognition, while radars can detect and track objects even in low visibility conditions. Sensor fusion enhances the accuracy and reliability of intrusion detection, object tracking, and perimeter security.

Industrial Automation:

In industrial settings, sensor fusion of cameras and radars can be used for object detection, tracking, and monitoring in manufacturing processes. For example, it can assist in quality control, detecting defects, and ensuring safe operation of machinery.

Traffic Management:

Integrating camera and radar data in traffic management systems improves road safety, congestion detection, and traffic flow optimization. Cameras capture visual information such as vehicle types, license plates, and traffic signs, while radars provide real-time data on vehicle speed, distance, and presence. Sensor fusion enables more accurate analysis and decision-making for traffic control.

Augmented Reality (AR) and Virtual Reality (VR):

Combining camera and radar data can enhance the immersive experience in AR and VR applications. Cameras capture the visual surroundings, while radars can provide depth information, enabling realistic virtual object placement and interaction with the physical environment.

3.3 Algorithms and Mechanism Used

1. Sensor Fusion Algorithm

Kalman Filter: The Kalman filter goes through two basic steps: prediction

3.3 Algorithms and Mechanism Used

and update, which are executed as fresh measurements become available. Cameras and radar sensors have various strengths and drawbacks in terms of detection accuracy, detection range, and environmental robustness. Vision-based sensors, such as cameras, are extremely successful at detecting and recognizing objects from up to 100 meters away. However, beyond this range, camera performance steadily diminishes, and accurate object detection becomes almost impossible at distances greater than 150 m. Radar sensors, on the other hand, can offer longer-distance measurements and are well-known for their resilience in low light and poor weather circumstances. While radar sensors may be influenced by uncommon environmental events such as lightning, their performance under difficult conditions is often more reliable than cameras. To take advantage of the advantages of both types of sensors, the camera and radar sensors execute object identification and classification independently, and the results are integrated using a fusion method.

In ADAS, traditional sensor fusion techniques like Kalman filtering and Bayesian methods are widely employed. Kalman filtering estimates the status of things and predicts their future behavior by combining sensor measurements with a dynamic model. Bayesian approaches employ probabilistic models to combine sensor data and more correctly determine object states. These strategies lay the groundwork for sensor fusion, although they may have limits in more complex settings.

2. Deep Learning Fusion:

Deep learning techniques like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have shown significant promise for sensor fusion. These models can learn to extract useful elements from camera and radar data and combine them for better object detection, tracking, and prediction. Deep learning-based fusion techniques take advantage of neural network strengths to handle complicated, high-dimensional sensor data with greater accuracy.

3. Calibration and Synchronization:

Accurate sensor calibration is critical for effective sensor fusion. Intrinsic calibration entails determining each sensor’s inherent properties, such as focal length and distortion coefficients. Extrinsic calibration determines the sensors’ relative position and orientation in a shared coordinate system. Temporal alignment and synchronization guarantee that data from many sensors are properly aligned in time, allowing for accurate fusion.

4. Object Detection and Tracking:

Sensor Fusion allows for robust object detection and tracking by merging information from cameras, radar, and lidar sensors. Cameras give detailed visual information, allowing for exact object recognition and classification. Radar sensors measure distance, velocity, and angle, which aids in long-range detection and tracking, particularly in inclement weather.

5. Radar Key Point Prediction:

Radar key point prediction refers to the estimation and prediction of specific qualities or attributes that are required for radar-based prediction activities. Sensor fusion improves radar critical point prediction by combining the complementary skills of cameras, and radar. The system may acquire a more thorough picture of the environment by fusing data from many sensors, allowing for more accurate predictions.

3.4 Evaluation metrics

1. The Multiple Object tracking precision MOPT:

MOTP assesses an object tracking algorithm’s localization precision. It computes the average distance between the tracked items’ expected positions and their corresponding ground truth positions. A lower MOTP score indicates higher pre-

cision because the algorithm’s projected positions are closer to the ground truth positions. A perfect MOTP score of 0 shows that the projected and ground truth positions are identical. In short, MOTA assesses total tracking accuracy, taking into account false positives, false negatives, and identity swaps, whereas MOTP focuses primarily on object localization precision. A greater MOTA score denotes improved tracking accuracy. A flawless MOTA score of 1 implies that the algorithm tracked all objects correctly and without errors. These metrics assist researchers and developers in evaluating and comparing the performance of various object-tracking methods. The following formula is used to determine MOTP: $\text{MOTP} = (\text{sum of anticipated and ground truth position distances}) / (\text{total number of accurately tracked objects})$

2. The Multiple Object tracking accuracy MOTA:

MOTA assesses an object tracking algorithm’s overall accuracy by taking into account several criteria such as false positives, false negatives, and identity swaps. It considers how well an algorithm finds and tracks objects in a video clip. Summing up the different error ratios gives us the total error rate. The following formula is used to determine MOTA: $\text{MOTA} = 1 - (\text{total number of ground truth objects minus the sum of false positives, false negatives, and identity switches})$

3. Average Multi-Object Tracking Precision AMOTP:

Like AMOTA, AMOTP computes the average tracking precision across numerous video sequences. It considers an algorithm’s tracking precision consistency across diverse contexts. AMOTP is calculated by first calculating MOTP for each video sequence, and then taking the average of these MOTP values across all sequences. AMOTP is a measure of an algorithm’s average localization precision across multiple sequences. A lower AMOTP score indicates more precision in many settings. $\text{AMOTP} = (\text{sum of all video sequence MOTP scores}) / \text{total number of video sequences}$

4. Average Multi-Object Tracking Accuracy AMOTA:

AMOTA is a MOTA extension that computes the average tracking accuracy over many video sequences. It provides a more comprehensive evaluation of the performance of an object-tracking method by taking into account its consistency across different contexts. AMOTA is produced by first calculating MOTA for each video sequence and then taking the average of these MOTA values across all sequences. AMOTA aids in evaluating an algorithm’s overall performance across various tracking settings. A greater AMOTA number suggests more accurate tracking across numerous sequences. $(\text{Sum of MOTA scores for all video sequences}) / \text{total number of video sequences} = \text{AMOTA}$

5. Smoothed Average Multi-Object Tracking Accuracy SMOTA:

Smoothed MOTA (sMOTA) is an upgraded variant of MOTA that takes into account the temporal continuity of object tracking. It addresses MOTA’s sensitivity to rapid changes in tracking performance between consecutive frames. sMOTA computes a smoother version of MOTA by considering prior frame tracking performance. It punishes unexpected changes in false positives, false negatives, and identity shifts. sMOTA contributes to a more steady and consistent assessment of an algorithm’s tracking accuracy over time. It is especially beneficial for examining algorithm performance in videos with demanding circumstances or fast-changing surroundings.

6. Track initialization duration TID:

TID is when an algorithm establishes a track after an object enters the scene or reappears after being temporarily occluded. It measures the efficiency of the track initialization procedure, which is critical for tracking objects accurately from the start. TID is commonly expressed as the number of frames required for an algorithm to successfully initialize a track. A shorter TID indicates that track initiation will be faster and more efficient. A longer TID, on the other hand,

indicates a slower or less effective track initialization process. TID reduction is significant because it reduces the time it takes to recognize and track items entering or returning to the scene. Fast-track initialization allows the algorithm to begin tracking objects quickly, resulting in provide more accurate and trustworthy tracking outcomes.

7. Longest gap duration LGD:

The longest temporal interval between two consecutive detections of the same item in a video clip is measured by LGD. It measures an object tracking algorithm’s ability to keep tracking during extended periods of occlusion or when objects temporarily disappear from the scene. LGD is calculated by counting the number of frames that indicate the length of the longest gap. A lower LGD value implies higher temporal continuity, implying that the system can efficiently sustain tracking even when objects are occluded or briefly not visible. An approach can assure robust tracking performance in circumstances with occlusions, object interactions, or other problems that cause occasional disappearance by minimizing LGD. The capacity to bridge vast temporal gaps helps to produce trustworthy and accurate tracking results.

References

- [1] Shenglin Li and Hwan-Sik Yoon, “Sensor Fusion-Based Vehicle Detection and Tracking Using a Single Camera and Radar at a Traffic Intersection” The University of Alabama, Tuscaloosa, AL 35487, USA, 2023
- [2] Keni Bernardin and Rainer Stiefelhagen, “Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics” at Karlsruhe, 76131 Karlsruhe, Germany, EURASIP Journal on Image and Video Processing Volume 2008,
- [3] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, Oscar Beijbom, “nuScenes: A multimodal dataset for autonomous driving”, nuTonomy: an AP-TIV company, March 2019
- [4] Jelena Kocić, Nenad Jovičić, and Vujo Drndarević, “Sensors and Sensor Fusion in Autonomous Vehicles”, 26th Telecommunications forum TELFOR 2018
- [5] De Jong Yeong, Gustavo Velasco-Hernandez, John Barry and Joseph Walsh, “Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review”, IMaR Research Centre, Munster Technological University, 18 March 2021
- [6] Federico Castanedo, D. Ursino, Y. Takama, “A Review of Data Fusion Techniques”, The Scientific World Journal Volume 2013
- [7] Joseph Redmon, Ali Farhadi; University of Washington, “YOLOv3: An Incremental Improvement”, University of Washington, 2018
- [8] Arindam Sengupta; Feng Jin; Siyang Cao, “A DNN-LSTM based Target Tracking Approach using mmWave Radar and Camera Sensor Fusion”, University of Arizona, Tucson, AZ USA, 2019 IEEE

REFERENCES

[9] Martin Simon, Stefan Milz, Karl Amende, Horst-Michael Gross Valeo Schalter und Sensoren GmbH, Ilmenau University of Technology, “Complex-YOLO: An Euler-Region-Proposal for Real-time 3D Object Detection on Point Clouds”, Ilmenau University of Technology martin.simon,stefan.milz,karl.amende@valeo.com

[10] Teck-Yian Lim: University of Illinois at Urbana-Champaign Urbana, Illinois; Amin Ansari: Qualcomm Technologies Inc. San Diego, California; Bence Major: Qualcomm AI Research San Diego, California; Daniel Fontijne: Qualcomm AI Research San Diego, California; Daniel Fontijne Qualcomm AI Research San Diego, California; Radhika Gowaikar Qualcomm Technologies Inc. San Diego, California; Sundar Subramanian Qualcomm Technologies Inc. San Diego, California, “Radar and Camera Early Fusion for Vehicle Detection in Advanced Driver Assistance Systems”, Machine Learning for Autonomous Driving Workshop at the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada