Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it
Bonn-Aachen
International Center for
Information Technology

R&D Project Proposal

Object detection in adverse weather conditions using tightly-coupled data-driven multimodal sensor fusion

Kevin Patel

Supervised by

Prof. Dr.-Ing. Sebastian Houben
M.Sc. Santosh Thoduka

April 2023

1 Introduction

1.1 Topic of This R&D Project

- Imagine driving on a winding mountain road at night, with fog and rain obscuring your view, your vehicle's self-driving system struggles to detect objects ahead due to the challenging weather conditions. Suddenly, a deer jumps out in front of your car, causing the system to issue an alert and apply the brakes in time to avoid a collision.

- This scenario highlights the importance of object detection in adverse weather conditions for self-driving cars. Visual cameras, which are commonly used for object detection, may be distorted or obscured by rain, fog, snow, or low light, making it difficult to accurately detect objects on the road [1] [2] [3].

- To address these challenges, this project aims to implement a multimodal sensor fusion system that combines cameras, radar, and LiDAR sensors. By fusing data from multiple sensors and leveraging advanced machine learning algorithms, the goal is to enhance object detection's range, accuracy, and reliability in adverse weather conditions.

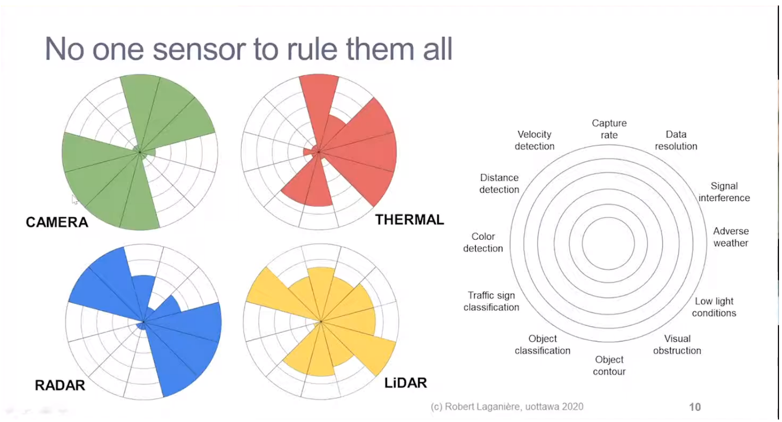


Figure 1: Sensors modality characteristics [4]

- The focus will also be on synchronizing multimodal data, processing dense and sparse resolution sensor data, and using a data-driven approach to optimize object detection performance.

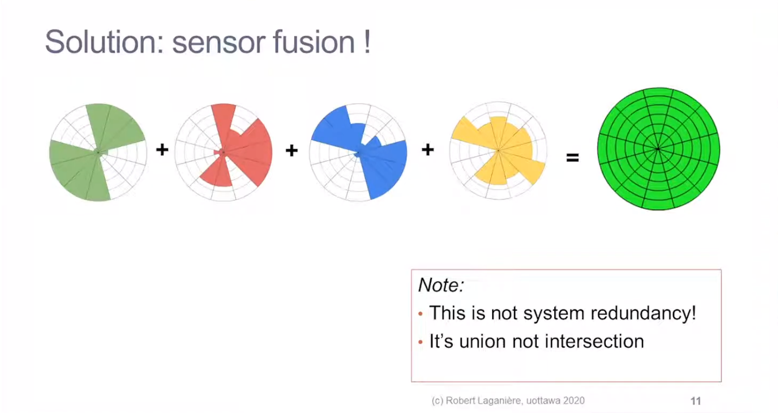


Figure 2: Sensors modality characteristics [4]

- However, this project also faces several challenges. For example, different sensors may have different resolutions, sampling rates, and may require sophisticated calibration and alignment techniques to ensure the accurate fusion of their data. Furthermore, processing large volumes of sensor data with minimal latency requires efficient and scalable algorithms and hardware architectures.

- The proposed system will be trained on a diverse dataset to ensure robustness and adaptability in different weather and lighting conditions. The system's effectiveness will be evaluated by extensive experiments and by comparing existing state-of-the-art methods.

- Despite the challenges, the project has the potential to revolutionize object detection in adverse weather conditions, with applications ranging from self-driving cars, drones to surveillance, and security systems. By fusing multiple sensor data sources and optimizing their fusion, situational awareness can be enhanced, enabling safer and more efficient operations in various domains.

- This research aims to facilitate safe and efficient self-driving in adverse weather conditions, prioritizing the safety of passengers, other drivers, and pedestrians on the road. To accomplish this, the proposed approach is to develop a sensor fusion system that operates with minimal latency, enabling data processing from multiple sensors in near real-time.

- Topic naming convention:
  - Object detection
    - Refers to the task of detecting objects within an image or video stream.
    - In this project, the focus is on detecting 2D objects such as cars, trucks, pedestrians, and cyclists.
  - Adverse weather conditions
    - Refers to conditions such as fog, snow, rain, overcast skies, sleet, and dust.
    - These conditions can make object detection more challenging due to reduced visibility or other environmental factors.
  - Tightly-coupled
    - Refers to how different modalities of data are combined and integrated at different levels.
    - Rather than relying solely on early, mid, or late fusion techniques, a combination of features at different levels is employed to achieve optimal fusion results.
  - Data-driven
    - Refers to the use of previously collected data or publicly available datasets to improve object detection performance.
  - Multimodal
    - Refers to the use of different data modalities to improve object detection performance.
    - Examples include sensors such as LiDAR, camera, IMU, GPS, infrared, and radar, with different datatypes such as point clouds, images, and time series data.
  - Sensor fusion
    - Refers to the process of fusing data from different sensors to get a better estimation of an environment and improve object detection performance.

1.2 Relevance of This R&D Project

- The relevance of the research project lies in the fact that weather phenomena have a significant negative influence on traffic and transportation, which can lead to accidents, injuries, and fatalities.
- The statistics show that adverse weather conditions, such as rain, snow, sleet, and fog, contribute to a high number of vehicle crashes and fatalities worldwide.
- For example, in the United States, over 30,000 vehicle crashes occur on snowy or icy roads each year, causing over 5,000 fatalities and 418,000 injuries due to adverse weather-related crashes, according to the Federal Highway Administration (FHA) [5] [6].
- According to a study by the Insurance Institute for Highway Safety (IIHS), the likelihood of fatal crashes increases significantly during snowy weather conditions, with a 21% higher rate compared to clear roads. The rate is even higher at 37% during sleet and freezing rain. Furthermore, a report by the National Highway Traffic Safety Administration (NHTSA) in 2018 revealed that adverse weather conditions were involved in 4,000 fatal crashes [7].
- In Europe, adverse weather conditions cause 25% of all road accidents, with frost and ice, snow, and rain being the highest contributing factors, according to the European Commission and the European Transport Safety Council (ETSC). Over 12,000 people die on European roads each year in weather-related accidents [8].
- Furthermore, the project's results will benefit various sectors, including autonomous vehicles, healthcare, precision agriculture, environmental monitoring, aerospace and defense, and industrial automation.
- The sensor fusion market for autonomous vehicles is expected to reach $22.2 billion by 2030 at a CAGR of 25.4%, according to Marketsandmarkets [9].
- In the healthcare sector, wearable sensors are estimated to reach over $1.5 billion in revenue by 2030, growing at a CAGR of over 18.3% [10].

- For precision agriculture and environmental monitoring, the market is expected to reach \$10.5 billion by 2026, growing at a CAGR of over 12.6% [11].

- The aerospace and defense sector, including aircraft navigation and control, missile guidance, and military logistics, is expected to reach \$23.83 billion by 2027, at a CAGR of 4.21% [12].

- Even the industrial automation sector benefits from the sensor fusion technology as it can improve the efficiency of the production process and reduce the cost of production.

2 Related Work

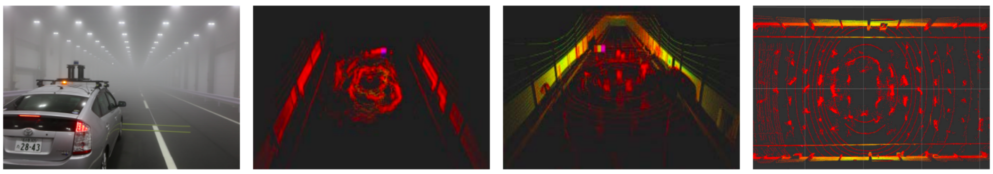
2.1 Survey of Related Work

- In the realm of autonomous robots, particularly self-driving vehicles and autonomous drones, object detection has emerged as a critical computer vision problem. These applications demand accurate 2D or 3D bounding boxes for objects in complex real-world scenarios, which often include cluttered scenes, unpredictable lighting, and adverse weather conditions. To address these challenges, the most promising autonomous vehicle systems rely on input from redundant sensor modalities, as documented by several recent studies [13–15]. These sensor modalities include cameras, LiDAR, radar, and emerging sensors like far-infrared (FIR) and near-infrared (NIR) sensors, which hold great potential for enabling reliable object detection in adverse environments [16].

- For a typical perception system, the most common sensor is camera, and it's actually the one element that is absolutely not replaceable in autonomous driving systems. But it's also one of the most vulnerable sensors to adverse weather conditions. A camera in rain, regardless of however high resolution, can be easily incapacitated by a single water drop on the emitter or lens [17]. Heavy snow or hail could fluctuate the image intensity and obscure the edges of the pattern of a certain object in the image or video which leads to detection failure [18]. A particular weather phenomenon, strong light, directly from the sun or artificial light source like light pollution from a skyscraper may also cause severe trouble to cameras [19].

Figure 3: Van occluded by a water droplet on the lens [20]

- LiDAR is the second most commonly used sensor in autonomous driving systems. Fersch et al. [21] suggest that for moderate levels of rainfall, LiDAR sensors with small apertures are not significantly affected. However, heavy and non-uniform precipitation rates can create clusters of fog that can lead to erroneous obstacle detection by the LiDARs. Hasirlioglu et al. [22] demonstrated that a rainfall rate exceeding 40 mm/hr leads to a significant drop in signal reflection intensity. Dense fog and smoke, as well as strong light, can also affect LiDAR sensors in adverse conditions [23] [19].



(a) Fog chamber (b) Point clouds in fog (c) Point clouds in rain (d) Rain pillars as detected by LiDAR

Figure 4: LiDAR performance test [23]

- Radar is the third most crucial sensor in autonomous driving systems and is widely used in mass-produced cars for active safety functions, such as automatic emergency braking (AEB) and forward collision warning (FCW). However, its significance is often overlooked from the perspective of perception tasks in autonomous driving. Unlike RGB cameras that use visible light bands (384∼769 THz) and LiDARs that use infrared bands (361∼331 THz), Radars use relatively longer wavelength radio bands (77∼81 GHz), resulting in robust measurements in adverse weathers [24]. As reported by Ijaz et al. [25] and Ismail [26], radar exhibits lower attenuation in rainy conditions than LiDAR. The attenuation of radar at 77 GHz is approximately 3.5 times lower (10 dB/km) than that of LiDAR at 905 nm (35 dB/km), demonstrating better robustness. Multiple experiments [18, 27–30] have revealed that attenuation and backscattering under dust, fog, snow, and light rain are negligible for radar, while its performance degrades under heavy rainfall. However, one of the significant drawbacks of radar is its low resolution, which makes it difficult to use in perception tasks. The radar point cloud is much sparser than LiDAR, limiting its usability. Recently, the next generation of 4D radar has emerged, which can provide denser points compared to conventional radar sensors.

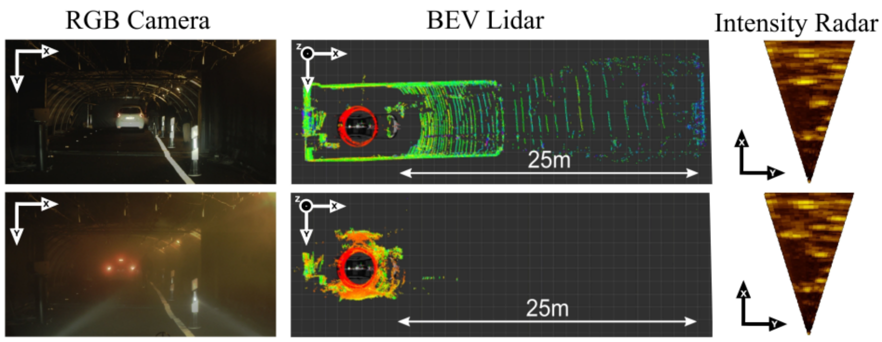


Figure 5: 1st row: clear weather condition, 2nd row: with fog. Shows that lidar affects by the fog but radar intensity remains the same [16]

- By now, it's almost well established that the LiDAR or Camera architecture alone is not going to navigate through adverse weather conditions with enough safety assurance. But two forces combining together would be a different story with the additional strength. As a result, groups from all over the world come up with their own permutation and combination with camera, LiDAR, radar, infrared camera, gated camera, stereo camera, weather stations and other weather-related sensors.

- Radecki et al. [31] provided a comprehensive review of the performance of various sensors in diverse weather conditions such as wet conditions, day and night, cloudy, glare, and dust. They developed a system that could track and classify objects using a joint probabilistic perception algorithm that used sensor subsets appropriate for different weather conditions. The integration of sensor data in real-time improved the general perception ability of the system. However, the optimal sensor weighting and quantified parameters for a given weather condition determined the reliability and robustness of the system. The authors emphasized the need for intelligent selection of sensor subsets to improve the accuracy of perception systems in different weather conditions.

- FLIR System Inc. [32] and VSI Labs [33] conducted a test on the first ever fused automated emergency braking (AEB) sensor suite in 2019, consisting of a thermal long-wave infrared (LWIR) camera, a radar, and a visible camera. The LWIR camera captures wavelengths ranging from 8 µm to 14 µm and operates under ambient temperature, known as the uncooled thermal camera. The sensor suite was evaluated alongside several cars equipped with AEB systems employing radar and visible cameras under various conditions including day-time, nighttime, and tunnel exit into sun glare. The comparison results indicate that while most AEB systems perform adequately during the day, the standard AEB almost collided with every mannequin under adverse conditions, whereas the LWIR sensor suite avoided any collision. This study underscores the potential of fusing camera and radar in challenging weather situations.

- Nobis et al. [34] proposed a solution to the problem of determining when to fuse data in a neural network architecture. Their proposed CameraRadar-

FusionNet (CRF-Net) was inspired by the camera-LiDAR fusion approach [35, 36] and aimed to identify the level at which fusing sensor data was most beneficial for object detection tasks. The authors used nuScenes [13] and their own TUM dataset and introduced a new training strategy called BlackIn to focus on specific sensor types. They used element-wise addition as the fusion operation and found that their fusion method outperformed the image-only network on both datasets, highlighting the significance of fusing multimodal sensor data for effective object detection.

- Yang et al. [37] proposed RadarNet for object detection and velocity estimation, that leverages radar and LiDAR sensors for perception. RadarNet employs early fusion to learn joint representations from both sensors and late fusion to incorporate the radial velocity evidence of radar and enhance the estimated object velocity. The authors evaluated their approach on the nuScenes dataset [13].

- Bijelic et al. [16] from Mercedes-Benz AG conducted a study on improving detection performance in adverse weather conditions using a deep multimodal sensor fusion approach. The authors equipped their test vehicle with various sensors, including stereo RGB cameras, a NIR camera, a 77 GHz radar, two LiDARs, an FIR camera, a weather station, and a road-friction sensor. They proposed an entropy-steered fusion approach where regions with low entropy were attenuated while entropy-rich regions were amplified during feature extraction. The exteroceptive sensor data were concatenated and trained using clear weather data, demonstrating strong adaptation to unseen adverse weather data. The fusion network was designed to generalize across different scenarios, and all the sensor data were projected into the camera coordinate system to ensure consistency. The fused detection performance outperformed LiDAR or image-only approaches under fog conditions.

- Bijelic et al. [16] also provided the SeeingThroughFog or DENSE dataset for further research on multimodal sensor fusion in adverse weather conditions. This dataset comprises 10,000 km of driving data in Northern Europe, recorded during February and December 2019, under varying weather and illumination

Figure 6: Highlighting the significance of fusing multimodal sensor data [16]

conditions. The dataset includes annotations for 5.5 k clear weather frames, 1 k dense fog frames, 1 k light fog frames, and 4 k frames captured in snow/rain.

- There are studies out there that use de-hazing techniques to remove the bad effects from adverse weather. While physical priors were previously used [38, 39], data-driven methods using deep learning have been introduced. However, deep de-hazing models have high computational complexity and are unsuitable for ultra-high-definition images. Chen et al. [40] found that models trained on synthetic images do not generalize well to real-world hazy images, while Zhang et al. [41] used temporal redundancy to perform video de-hazing and collected a dataset of real-world hazy and haze-free videos. Although collecting pairs of hazy and haze-free ground-truth images is challenging, professional haze/fog generators exist to simulate real-world conditions [42, 43].

- Few researchers [44] [45] [46] have also explored synthetic data generation for adverse weather conditions using GAN-based techniques from clean weather dataset eg. KITTI [47], Cityscapes [48], etc. However, the current methods are predominantly assessed on artificially created fog or rain images, along with a limited number of actual images under specific fog or rain models. Consequently, the capability of these algorithms to perform effectively under various adverse weather conditions and how their progress can be assessed in real-world scenarios remain unclear [49].

- Liu et al. [50] presented a novel approach to enhance the target recognition and tracking by fusing radar and camera data. In this approach, radar is considered as the primary sensor, and camera data is used as secondary information to complement the radar measurements. The authors evaluated the performance of their approach in challenging weather conditions, including rain and fog, as well as low visibility scenarios during nighttime. The experimental results revealed that radar-based detection exhibited high accuracy in detecting moving targets in wet weather, while the camera was more effective in target classification. Furthermore, the fusion of radar and camera data showed superior performance compared to LiDAR-based detection methods by over 33

- Qian et al. [51] proposed a Multimodal Vehicle Detection Network (MVDNet) that integrates LiDAR and radar for robust object detection. The network comprises a fusion module with a two-stage attention block that applies self-attention to each modality to extract features and then combines them with region-wise features through cross attentions. The study evaluated the proposed mechanism on SeeingThroughFog [16] and the Oxford Radar Robotcar [52] datasets, showing that the fusion mechanism performs robustly in foggy weather and outperforms LiDAR alone in such conditions.

- Rawashdeh et al. [53] developed a CNN-based sensor fusion approach for detecting drivable paths using cameras, LiDAR, and radar, which was evaluated using the DENSE dataset [16]. Their multi-stream encoder-decoder network was designed to compensate for the asymmetric degradation of the input sensors at the highest level. The depth and number of blocks for each sensor in the architecture were determined by their respective input data densities, with the camera having the highest density, followed by LiDAR, and radar having the lowest. The fully connected network's outputs were reshaped into a 2-D array that was input to the decoder. The researchers showed that their model could effectively disregard road lines and edges that might otherwise cause false interpretations and accurately delineate the general drivable area.

- The emergence of autonomous driving, particularly in harsh weather conditions, has benefited greatly from simulation platforms and experimental facilities such as fog chambers or test roads. CARLA simulator [54] is a popular virtual platform that allows researchers to create complex road environments and non-ego participants in infinite scenarios, which would be difficult and costly to replicate in real-world experiments. Furthermore, weather conditions, especially season-related or extreme climates-related, may not always be available for testing purposes. For instance, tropical regions cannot conduct snow tests, and natural rain showers may not last long enough to collect adequate experimental data. Most importantly, adverse weather conditions pose a danger to driving, and real-world tests always carry the risk of safety hazards, while simulators can provide an environment with zero risks [23].

- The majority of deep multimodal perception approaches rely on supervised learning, and therefore necessitate multimodal datasets with labeled ground truth for training deep neural networks. While several multimodal datasets are available, many of these datasets are collected under clear weather conditions or do not include all sensors, such as cameras, LiDAR, and radar. Unfortunately, the availability of multimodal datasets collected under adverse weather conditions with all three sensors is limited. Table 1 summarizes some of the available multimodal datasets for evaluating the performance of deep multimodal perception techniques in adverse weather conditions. Of these datasets, only the recently released K-Radar [24] incorporates a high-resolution 4D-radar sensor. In the table, C-R-L-N-F denotes the Camera, Radar, LiDAR, Near-infrared, and Far-infrared sensors, respectively.

Table 1: List multimodal datasets with adverse weather conditions

Name	Sensors	Reference	Year
DENSE	CRLNF	[16]	2020
EU Long-term	CRL	[55]	2020
nuScenes	CRL	[13]	2020
The Oxford RobotCar	CRL	[52]	2020
RADIATE	CRL	[56]	2021
K-Radar	CRL	[24]	2022
aiMotive	CRL	[57]	2022
Boreas	CRL	[58]	2022
WADS	CRLNF	[59]	2023

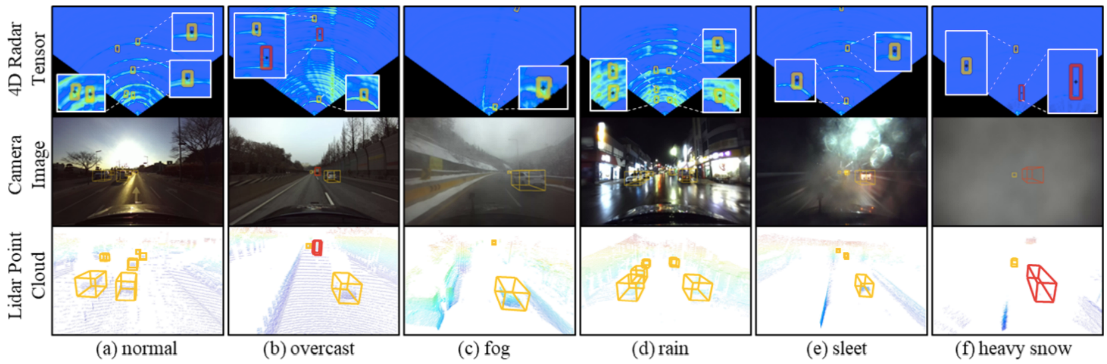


Figure 7: Samples of K-Radar datasets for various weather conditions [24]

2.2 Limitation and Deficits in the State of the Art

- While Radecki et al. [31] provided a comprehensive summary of the performance of various sensors in adverse weather conditions, their study has certain limitations. The majority of the studies considered were conducted under optimal weather conditions, which may not guarantee robustness under harsh weather conditions [60]. Moreover, the study did not explore the performance of the sensors in heavy-traffic or urban areas, which may present different challenges to the perception system. Additionally, the study did not investigate the performance of deep learning-based fusion techniques. To address these limitations, future work could explore the use of the proposed Similarly-based method to detect objects in recent adverse weather datasets.

- The work done by FLIR System Inc. [32] and the VSI Labs [33] by using a thermal camera for fusion as one of the sensors, which raises the concerns of the durability of such temperature-sensitive devices in real environments. This needs further validation in real environments in the future to ensure their usefulness in adverse weather conditions [18].

- Nobis et al. [34] proposed the CameraRadarFusionNet (CRF-Net) to learn the optimal level for sensor data fusion in the neural network architecture for detection tasks. While their fusion method outperformed the image-only network on nuScenes [13] and their own TUM dataset, the improvement in detection performance was limited. The element-wise addition used for feature fusion was a simple method, and the baseline image network's performance was only slightly lower than the CRF-Net's performance. Additionally, the study did not provide an RGB sensor ablation study, so it was unclear whether their system was robust in the case of camera failure. According to Safa et al. [61], the performance could be further improved by pre-processing the radar data before fusion.

- In the RadarNet architecture, proposed by Yang et al. [37], the radar sensor data used from the nuScenes dataset, which has a very low resolution and hence it's not a good choice for studying the role of radar in perception. Object detection using radars is limited by low resolution and erroneous elevation estimates [62] [63]. Therefore, one possible improvement is to include the radar sensor used in K-Radar dataset, which is a 4D radar with elevation angle of 30 degrees compare to 1 or 2 degrees for conventional radars, can be used to improve the performance of the RadarNet architecture.

- The study conducted by Bijelic et al. [16] presents a multimodal sensor fusion approach that outperforms LiDAR or image-only approaches under fog conditions. However, one issue is that some essential radar information may be lost during the projection transformation used by the proposed method, leading to a loss of spatial information. Additionally, the large number of sensors required by the approach exceeds the typical expectations for an autonomous driving system, making it challenging to implement in real-world scenarios. The response and reaction time of the algorithm may also become a concern due to the bulk amount of data from multiple sensors. While the study demonstrated strong adaptation to adverse weather data, the performance of the radar used was limited by its low azimuth and elevation resolution [23]. To address these limitations, future work could focus on improving the network architecture by using a higher resolution radar and using a transformer-based approaches to improve the performance of the sensor fusion approach in adverse weather conditions.

- Qian et al. [51] proposed a Multimodal Vehicle Detection Network (MVDNet) that combines LiDAR and radar data using a two-stage attention block in the fusion module. Despite demonstrating robust performance in foggy weather conditions, the study has some limitations. Firstly, the misalignment between the LiDAR and radar data in the dataset is not corrected, which can affect the MVDNet's performance. Secondly, the simple label assignment strategy used in the loss computation procedure and the region-of-interest (ROI) assisted fusion design limits the model's performance. These factors suggest that there is room for improvement in the model's design, which could potentially be addressed by more advanced fusion techniques and better label assignment strategies [64].

- The multimodal fusion approach proposed by Rawashdeh et al. [53] has the potential to address not only snow conditions but also other scenarios of low visibility, including fog, rain, and dust., and it lacks the comparison with other state-of-the-art methods for drivable path detection in poor weather conditions. Additionally, the proposed multi-stream encoder-decoder architecture is designed based on the input data density of each sensor, which may not be optimal for other types of sensors or scenarios. Finally, the study does not investigate the real-time processing requirements and the computational cost of the proposed algorithm, which could limit its practical applicability [53].

- Since most of the advance multimodal sensor fusion techniques, including state-of-the-art ones, are developed on clear weather datasets without paying special attention to the adverse weather conditions. And many of them use LiDAR and camera [65] as the primary sensors. This could be because of the availability of the datasets. Hence, there is a need to study the performance of these advance deep learning techniques in adverse weather conditions with the combination of other sensors like radar, thermal camera, etc.

- Currently, there is no general guideline available for the design of network architecture in multimodal sensor fusion, and several questions remain unanswered. According to Feng et al. [65], these include
  - "what to fuse," such as LiDAR, radar, color camera, thermal camera, event camera, or ultrasonic sensors;
  - "how to fuse," which can include addition or mean, concatenation, ensemble methods, or mixture of experts;
  - "when to fuse," which can involve early, mid, late, or a combination of all fusion methods.

- The lack of comparison with alternative models or datasets is a limitation of previous studies on multimodal sensor fusion. Many studies have only shown results for their own baseline models and custom datasets, which limits the generalization of their findings.

- While recent multimodal datasets have released baseline models with simple fusion methods, the use of more advanced fusion methods, such as transformer-based or gated fusion, has the potential to improve their performance.
- Temporal information is a crucial aspect of sensor fusion, but very few multimodal fusion algorithms have been developed to handle this type of data. [16]
- There is a dearth of work in the literature utilizing 4D imaging radar sensors, which is a promising area for future research. [66]

3 Problem Statement

- Object detection using multiple modalities has become a topic of increasing interest in recent years. However, despite the wealth of research in this area, there is still a lack of comprehensive analysis and practical implementation of state-of-the-art methods udner adverse weather conditions. Therefore, the primary objective of this research project is to provide a thorough analysis and practical implementation of state-of-the-art methods for object detection using multiple modalities, including but not limited to camera, LiDAR, and radar.
- One of the significant challenges in multimodal object detection is determining an appropriate fusion strategy to exploit the complementary characteristics of various sensors. For instance, fusing camera and 4D radar data is a critical research question that needs to be addressed. This research project will focus on identifying an optimal fusion strategy that leverages the strengths of each sensor while minimizing their weaknesses.
- (tentative) Fusing spatial and temporal information from multimodal sensors is another important aspect that will be investigated in this research project. Most existing multimodal fusion algorithms focus on fusing spatial information, but temporal information is equally important. Therefore, this project aims to develop methods that can fuse both spatial and temporal information effectively.

- (tentative) To validate the performance of a model, simulators such as CARLA may be utilized if required. Simulations provide a controlled environment and enable the evaluation of the system's performance under diverse and adverse weather conditions, which is essential for testing the robustness of the model.
- This research project aims to conduct experiments to evaluate the performance of the proposed methods under adverse weather conditions. To achieve this, the outcomes will be compared with various models on recently released datasets such as K-radar [24], DENSE [16], and aiMotive [67]. These datasets are deemed suitable for the purpose of analyzing the performance of the proposed methods under challenging conditions.
- The comparative analysis of the results will include an assessment of the strengths and weaknesses of each approach with respect to existing state-of-the-art methods.

4 Project Plan

4.1 Work Packages

WP1 Literature Study

- WP1.1 Conduct a comprehensive literature review of state-of-the-art methods for object detection under adverse weather conditions using multiple modalities
- WP1.2 Analyze and compare various fusion strategies for exploiting the complementary characteristics of different sensors
- WP1.3 Search for suitable public datasets with adverse weather conditions and multimodal sensors data
- WP1.4 (tentative) Investigate the fusion of spatial and temporal information from multimodal sensors
- WP1.5 (tentative) Survey the use of simulators to validate the performance of multimodal object detection models under adverse weather conditions

WP2 Data Collection and Preparation

- WP2.1 Acquire the necessary datasets for multimodal object detection under adverse weather conditions, such as K-radar, DENSE, and aiMotive
- WP2.2 Develop tools for pre-processing and augmenting the datasets to enhance the performance of the models
- WP2.3 Perform statistical analysis to identify the main characteristics and challenges of the datasets, including data imbalance and class imbalance

WP3 Model Design and Implementation

- WP3.1 Design and implement a multimodal object detection model that integrates camera, LiDAR, and radar data
- WP3.2 Investigate various fusion strategies, such as concatenation, mixture of experts, attention-based fusion etc, to determine the most effective approach
- WP3.3 Explore deep learning architectures, such as CNNs, RNNs, and Transformers, to improve the performance of the multimodal model
- WP3.4 Optimize the model's hyperparameters and train the model on the acquired datasets

WP4 Model Evaluation and Validation

- WP4.1 Evaluate the performance of the developed multimodal object detection model on the acquired datasets under adverse weather conditions
- WP4.2 (tentative) Validate the performance of the model using simulators, such as CARLA, to generate various scenarios and test the robustness of the model
- WP4.3 Compare the proposed model's results to existing state-of-the-art or baseline methods and analyze the strengths and weaknesses of each approach
- WP4.4 Identify the limitations of the proposed model and suggest possible future improvements

WP5 Project Report

- WP5.1 Write a detailed report that includes the research problem, objectives, methodology, results, and conclusion
- WP5.2 Present the research findings in a clear and concise manner, highlighting the contributions and limitations of the proposed multimodal object detection model
- WP5.3 Discuss possible future research directions based on the outcomes of the study

4.2 Milestones

- M1 Literature review completed and best practice identified
- M2 Data collection and preprocessing completed, including cleaning and augmentation
- M3 Initial model development and testing completed
- M4 Evaluation and optimization of the model completed
- M5 Final model development and testing completed, including comparison with existing state-of-the-art methods and analysis of strengths and weaknesses of each approach.
- M6 Project report completed

4.3 Project Schedule

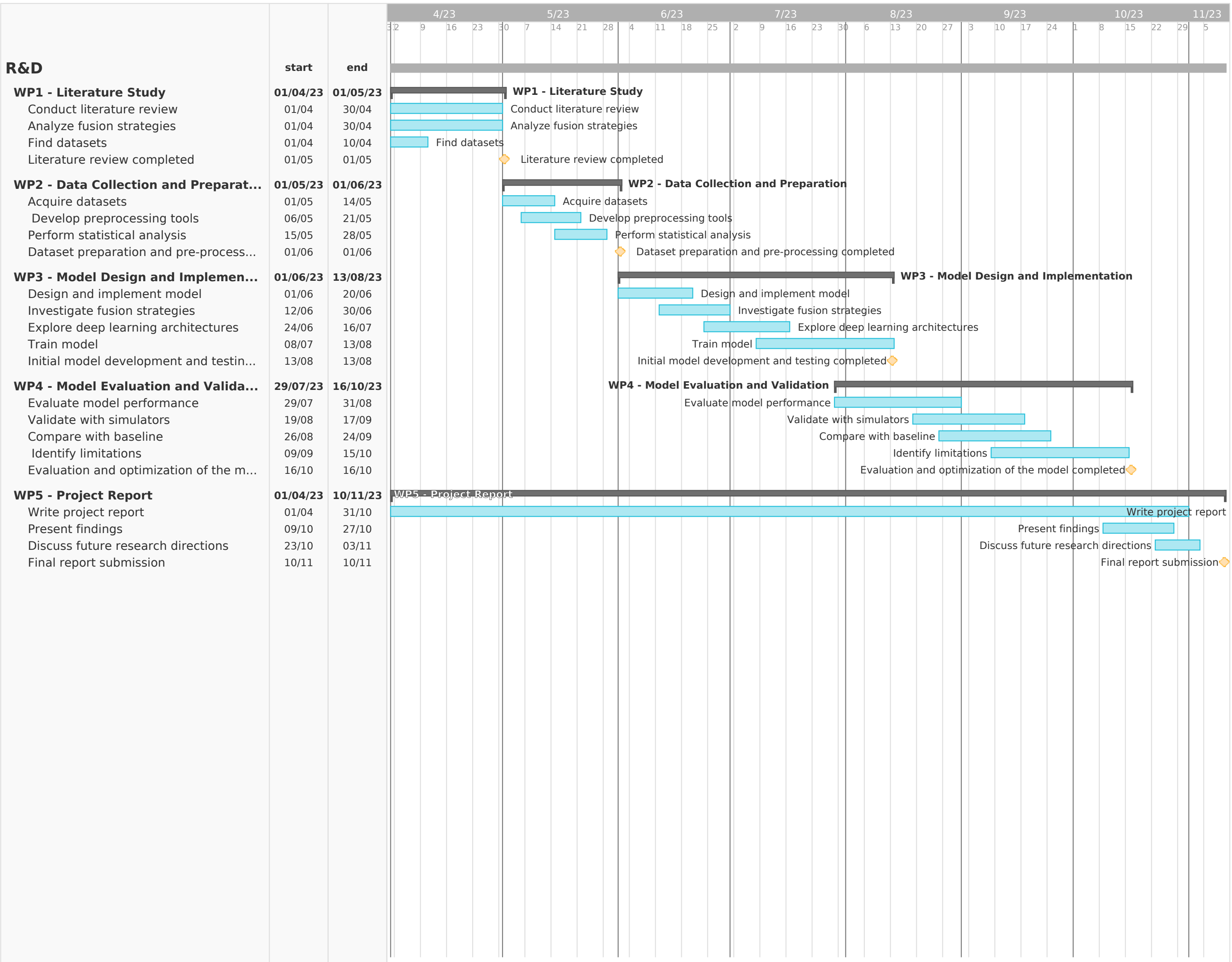


Figure 8: Gantt chart of the project schedule

4.4 Deliverables

Minimum Viable

- Conduct a comprehensive literature review on state-of-the-art multimodal object detection methods and their fusion strategies
- Develop and test an initial model for object detection
- Perform a comparative analysis of at least two methods on two datasets
- Produce a project report that summarizes the work done and the results obtained

Expected

- Compare with more advance methods with baseline methods on different datasets
- Complete the final development and testing of the model, including comparison with existing state-of-the-art methods and analysis of the strengths and weaknesses of each approach.
- Produce a more extensive project report that details the methodology, experimental setup, results, and analysis.

Desired

- Conduct experiments to validate the model's performance by implementing and testing it in CARLA or other simulators
  - Note: CARLA simulator doesn't support 4D radar sensor
- Utilize spatial and temporal information from multimodal sensors in the object detection process

References

[1] Ekim Yurtsever, Jacob Lambert, Alexander Carballo, and Kazuya Takeda. A survey of autonomous driving: Common practices and emerging technologies. *IEEE access*, 8:58443–58469, 2020.

[2] Alexander Carballo, Ekim Yurtsever, and Kazuya Takeda. Libre: The multiple 3d lidar dataset. *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 1094–1101, 2020.

[3] University of Michigan. Getting traction: Tips for traveling in winter weather, 2020.

[4] robert laganière. Sensor fusion for autonomous vehicles: Strategies, methods, and tradeoffs, 2022. Accessed on 18.12.2022.

[5] Federal-Highway-Administration. How Do Weather Events Impact Roads? - FHWA Road Weather Management.

[6] NOAA US Department of Commerce. Getting traction: Tips for traveling in winter weather, Nov 2016.

[7] Matthew L Brumbelow. Light where it matters: Iihs headlight ratings are correlated with nighttime crash rates. *Journal of safety research*, 83:379–387, 2022.

[8] Graham Cookson. Weather-Related Road Deaths in Europe £15bn+ Per Year — INRIX, 2 2022.

[9] MarketsandMarkets. Sensor fusion market by technology - 2022 — marketsandmarkets.

[10] Straits Research. Wearable sensors market, 2021.

[11] Mordor intelligence. Precision farming market analysis, industry report, trends, size, and share.

[12] Fortune business insights. Aerospace and defense materials market size, share, and report, 2027.

[13] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11621–11631, 2020.

[14] Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, et al. Scalability in perception for autonomous driving: Waymo open dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2446–2454, 2020.

[15] Julius Ziegler, Philipp Bender, Markus Schreiber, Henning Lategahn, Tobias Strauss, Christoph Stiller, Thao Dang, Uwe Franke, Nils Appenrodt, Christoph G Keller, et al. Making bertha drive—an autonomous journey on a historic route. *IEEE Intelligent transportation systems magazine*, 6(2):8–20, 2014.

[16] Mario Bijelic, Tobias Gruber, Fahim Mannan, Florian Kraus, Werner Ritter, Klaus Dietmayer, and Felix Heide. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather, 2020.

[17] Raffi Mardirosian. Lidar vs. camera: driving in the rain. [Last accessed 10 April 2023], 2023.

[18] Shizhe Zang, Ming Ding, David Smith, Paul Tyler, Thierry Rakotoarivelo, and Mohamed Ali Kaafar. The impact of adverse weather conditions on autonomous vehicles: how rain, snow, fog, and hail affect the performance of a self-driving car. *IEEE vehicular technology magazine*, 14(2):103–111, 2019.

[19] Alexander Carballo, Jacob Lambert, Abraham Monrroy, David Wong, Patiphon Narksri, Yuki Kitsukawa, Eijiro Takeuchi, Shinpei Kato, and Kazuya Takeda. LIBRE: The multiple 3d lidar dataset. *arXiv preprint arXiv:2003.06129*, 2020. (accepted for presentation at IV2020).

[20] Felix Nobis, Maximilian Geisslinger, Markus Weber, Johannes Betz, and Markus Lienkamp. A Deep Learning-based Radar and Camera Sensor Fusion Architecture for Object Detection. *arXiv*, May 2020.

[21] Thomas Fersch, Alexander Buhmann, Alexander Koelpin, and Robert Weigel. The influence of rain on small aperture lidar sensors. In *2016 German Microwave Conference (GeMiC)*, pages 84–87. IEEE, 2016.

[22] Sinan Hasirlioglu, Igor Doric, Christian Lauerer, and Thomas Brandmeier. Modeling and simulation of rain for the test of automotive sensor systems. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, pages 286–291. IEEE, 2016.

[23] Yuxiao Zhang, Alexander Carballo, Hanting Yang, and Kazuya Takeda. Autonomous Driving in Adverse Weather Conditions: A Survey, December 2021.

[24] Dong-Hee Paek, Seung-Hyun Kong, and Kevin Tirta Wijaya. K-Radar: 4D Radar Object Detection for Autonomous Driving in Various Weather Conditions, June 2022.

[25] Muhammad Ijaz, Zabih Ghassemlooy, Hoa Le Minh, Sujan Rajbhandari, and J Perez. Analysis of fog and smoke attenuation in a free space optical communication link under controlled laboratory conditions. In *2012 International Workshop on Optical Wireless Communications (IWOW)*, pages 1–3. IEEE, 2012.

[26] Ismail Gultepe. Measurements of light rain, drizzle and heavy fog. In *Precipitation: advances in measurement, estimation and prediction*, pages 59–82. Springer, 2008.

[27] Martin Adams, Martin David Adams, and Ebi Jose. *Robotic navigation and mapping with radar*. Artech House, 2012.

[28] Graham Brooker, Ross Hennessey, Craig Lobsey, Mark Bishop, and Eleonora Widzyk-Capehart. Seeing through dust and water vapor: Millimeter wave radar sensors for mining applications. *Journal of Field Robotics*, 24(7):527–557, 2007.

[29] Ruoyang Xu, Wei Dong, Akash Sharma, and Michael Kaess. Learned depth estimation of 3d imaging radar for indoor mapping. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 13260–13267. IEEE, 2022.

[30] Rossiza Gourova, Oleg Krasnov, and Alexander Yarovoy. Analysis of rain clutter detections in commercial 77 ghz automotive radar. In *2017 European Radar Conference (EURAD)*, pages 25–28. IEEE, 2017.

[31] Peter Radecki, Mark Campbell, and Kevin Matzen. All weather perception: Joint data association, tracking, and classification for autonomous ground vehicles. *arXiv preprint arXiv:1605.02196*, 2016.

[32] Flir. fused aeb with thermal can save lives.

[33] Research testing on adas autonomous vehicle technologies.

[34] Felix Nobis, Maximilian Geisslinger, Markus Weber, Johannes Betz, and Markus Lienkamp. A deep learning-based radar and camera sensor fusion architecture for object detection. In *2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–7. IEEE, 2019.

[35] Dameng Yu, Hui Xiong, Qing Xu, Jianqiang Wang, and Keqiang Li. Multistage residual fusion network for lidar-camera road detection. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 2323–2328. IEEE, 2019.

[36] Luca Caltagirone, Mauro Bellone, Lennart Svensson, and Mattias Wahde. Lidar-camera fusion for road detection using fully convolutional neural networks. *Robotics and Autonomous Systems*, 111:125–131, 2019.

[37] Bin Yang, Runsheng Guo, Ming Liang, Sergio Casas, and Raquel 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 16*, pages 496–512. Springer, 2020.

[38] Robby T Tan. Visibility in bad weather from a single image. In *2008 IEEE conference on computer vision and pattern recognition*, pages 1–8. IEEE, 2008.

[39] Jean-Philippe Tarel and Nicolas Hautiere. Fast visibility restoration from a single color or gray level image. In *2009 IEEE 12th international conference on computer vision*, pages 2201–2208. IEEE, 2009.

[40] Zeyuan Chen, Yangchao Wang, Yang Yang, and Dong Liu. Psd: Principled synthetic-to-real dehazing guided by physical priors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7180–7189, 2021.

[41] Xinyi Zhang, Hang Dong, Jinshan Pan, Chao Zhu, Ying Tai, Chengjie Wang, Jilin Li, Feiyue Huang, and Fei Wang. Learning to restore hazy video: A new real-world dataset and a new method. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9239–9248, 2021.

[42] Valentina Mușat, Ivan Fursa, Paul Newman, Fabio Cuzzolin, and Andrew Bradley. Multi-weather city: Adverse weather stacking for autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2906–2915, 2021.

[43] Radu Timofte, Shuhang Gu, Jiqing Wu, and Luc Van Gool. Ntire 2018 challenge on single image super-resolution: Methods and results. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 852–863, 2018.

[44] Ting Sun, Jinlin Chen, and Francis Ng. Multi-target domain adaptation via unsupervised domain classification for weather invariant object detection. *arXiv preprint arXiv:2103.13970*, 2021.

[45] Ziqiang Zheng, Yang Wu, Xinran Han, and Jianbo Shi. Forkgan: Seeing into the rainy night. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*, pages 155–170. Springer, 2020.

[46] Younkwan Lee, Yeongmin Ko, Yechan Kim, and Moongu Jeon. Perception-friendly video enhancement for autonomous driving under adverse weather

conditions. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 7760–7767. IEEE, 2022.

[47] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3354–3361. IEEE, 2012.

[48] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3213–3223, 2016.

[49] Mahmoud Hassaballah, Mourad A Kenk, Khan Muhammad, and Shervin Minaee. Vehicle detection and tracking in adverse weather using a deep learning framework. *IEEE transactions on intelligent transportation systems*, 22(7):4230–4242, 2020.

[50] Ze Liu, Yingfeng Cai, Hai Wang, Long Chen, Hongbo Gao, Yunyi Jia, and Yicheng Li. Robust target recognition and tracking of self-driving cars with radar and camera information fusion under severe weather conditions. *IEEE Transactions on Intelligent Transportation Systems*, 23(7):6640–6653, 2021.

[51] Kun Qian, Shilin Zhu, Xinyu Zhang, and Li Erran Li. Robust multimodal vehicle detection in foggy weather using complementary lidar and radar signals. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 444–453, 2021.

[52] Dan Barnes, Matthew Gadd, Paul Murcutt, Paul Newman, and Ingmar Posner. The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6433–6438. IEEE, 2020.

[53] Nathir A Rawashdeh, Jeremy P Bos, and Nader J Abu-Alrub. Drivable path detection using cnn sensor fusion for autonomous driving in the snow. In *Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure 2021*, volume 11748, pages 36–45. SPIE, 2021.

[54] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. Carla: An open urban driving simulator. In *Conference on robot learning*, pages 1–16. PMLR, 2017.

[55] Zhi Yan, Li Sun, Tomáš Krajník, and Yassine Ruichek. Eu long-term dataset with multiple sensors for autonomous driving. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10697–10704. IEEE, 2020.

[56] Marcel Sheeny, Emanuele De Pellegrin, Saptarshi Mukherjee, Alireza Ahrabian, Sen Wang, and Andrew Wallace. Radiate: A radar dataset for automotive perception in bad weather. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–7. IEEE, 2021.

[57] Tamás Matuszka, Iván Barton, Ádám Butykai, Péter Hajas, Dávid Kiss, Domonkos Kovács, Sándor Kunsági-Máté, Péter Lengyel, Gábor Németh, Levente Pető, et al. aimotive dataset: A multimodal dataset for robust autonomous driving with long-range perception. *arXiv preprint arXiv:2211.09445*, 2022.

[58] Keenan Burnett, David J Yoon, Yuchen Wu, Andrew Z Li, Haowei Zhang, Shichen Lu, Jingxing Qian, Wei-Kang Tseng, Andrew Lambert, Keith YK Leung, et al. Boreas: A multi-season autonomous driving dataset. *The International Journal of Robotics Research*, page 02783649231160195, 2022.

[59] Akhil Kurup and Jeremy Bos. Winter adverse driving dataset (wads): year three. In *Autonomous Systems: Sensors, Processing and Security for Ground, Air, Sea and Space Vehicles and Infrastructure 2022*, volume 12115, pages 146–152. SPIE, 2022.

[60] Yrvann Emzivat, Javier Ibanez-Guzman, Hervé Illy, Philippe Martinet, and Olivier H Roux. A formal approach for the design of a dependable perception system for autonomous vehicles. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2452–2459. IEEE, 2018.

[61] Ali Safa, Tim Verbelen, Ilja Ocket, André Bourdoux, Francky Catthoor, and Georges GE Gielen. Fail-safe human detection for drones using a multi-

modal curriculum learning approach. *IEEE Robotics and Automation Letters*, 7(1):303–310, 2021.

[62] Michael Ulrich, Claudius Gläser, and Fabian Timm. Deepreflecs: Deep learning for automotive object classification with radar reflections. In *2021 IEEE Radar Conference (RadarConf21)*, pages 1–6. IEEE, 2021.

[63] Florian Drews, Di Feng, Florian Faion, Lars Rosenbaum, Michael Ulrich, and Claudius Gläser. Deepfusion: A robust and modular 3d object detector for lidars, cameras and radars. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 560–567. IEEE, 2022.

[64] Yanlong Yang, Jianan Liu, Tao Huang, Qing-Long Han, Gang Ma, and Bing Zhu. Ralibev: Radar and lidar bev fusion learning for anchor box free object detection system. *arXiv preprint arXiv:2211.06108*, 2022.

[65] Di Feng, Christian Haase-Schütz, Lars Rosenbaum, Heinz Hertlein, Claudius Glaeser, Fabian Timm, Werner Wiesbeck, and Klaus Dietmayer. Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges, 2020.

[66] Yi Zhou, Lulu Liu, Haocheng Zhao, Miguel López-Benítez, Limin Yu, and Yutao Yue. Towards Deep Radar Perception for Autonomous Driving: Datasets, Methods, and Challenges, May 2022.

[67] Tamás Matuszka, Iván Barton, Ádám Butykai, Péter Hajas, Dávid Kiss, Domonkos Kovács, Sándor Kunsági-Máté, Péter Lengyel, Gábor Németh, Levente Pető, Dezső Ribli, Dávid Szeghy, Szabolcs Vajna, and Bálint Varga. aiMotive Dataset: A Multimodal Dataset for Robust Autonomous Driving with Long-Range Perception, November 2022.

[68] Author Name. Book title. *Lecture Notes in Autonomous System*, 1001:900–921, 2003.

[69] Cristian Tangemann. Sensor fusion, 2022. Accessed on 29.09.2022.

[70] Markus Willems. A dsp for implementing high-performance sensor fusion on an embedded budget, Nov 2021.

[71] Dan Leibholz. Real-time sensor fusion challenge, 2022. Accessed on 29.09.2022.

[72] Ann Steffora Mutschler. Sensor fusion challenges, 2022. Accessed on 29.09.2022.

[73] Gregory-Koshmak et al. Challenges and issues in multisensor fusion approach, 2022. Accessed on 29.09.2022.

[74] Simon Chadwick, Will Maddern, and Paul Newman. Distant vehicle detection using radar and vision. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 8311–8317. IEEE, 2019.

[75] Shuo Chang, Yifan Zhang, Fan Zhang, Xiaotong Zhao, Sai Huang, Zhiyong Feng, and Zhiqing Wei. Spatial attention fusion for obstacle detection using mmwave radar and vision sensor. *Sensors*, 20(4):956, 2020.

[76] robert laganière. Sensor fusion for autonomous vehicles: Strategies, methods, and tradeoffs, 2022. Accessed on 18.12.2022.