Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

Bonn-Aachen
International Center for
Information Technology

R&D Project

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

Kevin Patel

Submitted to Hochschule Bonn-Rhein-Sieg,
Department of Computer Science
in partial fullfilment of the requirements for the degree
of Master of Science in Autonomous Systems

Supervised by

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

November 2023

I, the undersigned below, declare that this work has not previously been submitted to this or any other university and that it is, unless otherwise stated, entirely my own work.

Date	Kevin Patel

Abstract

TODO: add abstract

Acknowledgements

TODO: add acknowledgements

Contents

1

Introduction

In the realm of autonomous driving, the ability to detect objects in challenging weather conditions remains a critical area of research. Consider a scenario where a self-driving vehicle navigates a winding mountain road at night amidst fog and rain. The limitations of the vehicle's visual cameras become evident as they struggle to detect objects due to reduced visibility, highlighting the crucial need for advanced object detection methods in adverse weather conditions. This is particularly significant in emergency situations, like when a deer suddenly appears on the road, necessitating quick and accurate object detection to prevent accidents [1] [2] [3].

As depicted in Figure 1.1, the performance of various sensors in automated systems under different conditions has been extensively analyzed. Cameras, for instance, excel in recognizing colors and signs but falter in dark or distant object measurement. Conversely, thermal sensors maintain efficiency in poor weather but lack color detection and texture information. Radar sensors are adept at speed measurement and are less hindered by visual obstructions, though they produce sparse and noisy data. LiDAR sensors, meanwhile, provide excellent object shape and size mapping but underperform in poor weather conditions [4].

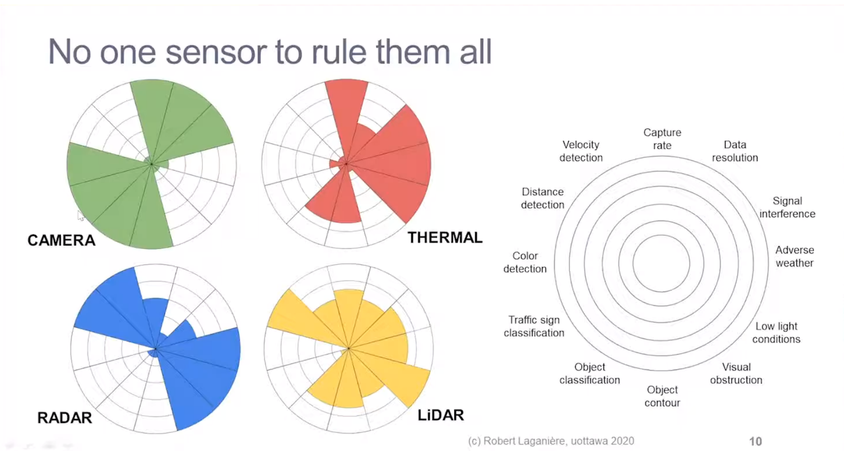


Figure 1.1: Sensors modality characteristics [5]

To overcome these limitations, this project proposes the development of a tightly-coupled multimodal sensor fusion system, as exemplified in Figure 1.2. By integrating cameras, radar, and LiDAR sensors, this approach harnesses the unique strengths of each to create a comprehensive and reliable object detection framework. The fusion of these complementary sensors, coupled with advanced machine learning algorithms, aims to significantly enhance the range, accuracy, and reliability of detection in adverse weather conditions. Effectively synthesizing diverse data sources enables the creation of a robust and responsive system that overcomes the weaknesses of any individual sensor type. A sample from DENSE dataset [6] is shown in Figure 1.3 to illustrate the importance of fusion in adverse weather conditions.

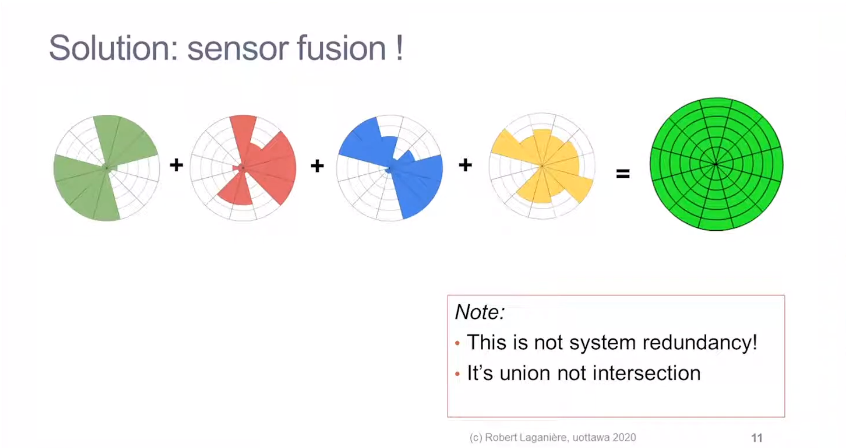


Figure 1.2: Sensors modality characteristics [5]

However, the integration of different sensor types presents its own set of challenges, such as varying resolutions, sampling rates, and the need for sophisticated calibration and alignment techniques. These hurdles necessitate the development of efficient and scalable algorithms and hardware architectures capable of processing large volumes of sensor data with minimal latency.

Despite these challenges, the anticipated outcomes of this project are transformative, with potential applications extending beyond autonomous vehicles to drones, surveillance, and security systems. By enhancing situational awareness through optimized sensor fusion, the project aims to foster safer and more efficient operations across various sectors. The primary objective is to ensure safe and efficient autonomous driving in adverse weather conditions, prioritizing the safety of passengers, other drivers, and pedestrians. This will be achieved through a sensor fusion system designed for minimal latency, enabling the processing of data from multiple sensors in near real-time.

In this research, the focus is specifically on 2D object detection to manage the intricacies of network and computational complexity. This scope simplifies the task of handling high-dimensional data, allowing for a more streamlined implementation while addressing the challenge of object detection in adverse weather conditions. Such conditions, including fog, snow, and rain, significantly hinder the visibility and recognition of various entities like cars, trucks, pedestrians, and cyclists. The 'tightly-coupled' approach involves the intricate integration of various data modalities, combining features at multiple levels for optimal results. The 'data-driven' aspect emphasizes leveraging existing datasets to enhance performance,

Figure 1.3: Importance of sensor fusion in adverse weather conditions, *Fusion of Camera, Lidar, Radar [6]

while 'multimodal' pertains to the utilization of diverse sensor types. Finally, 'sensor fusion' is central to the project, entailing the amalgamation of data from different sensors to improve environmental perception and object detection capabilities.

1.1 Motivation

The motivation for the research topic is deeply rooted in the understanding that weather phenomena significantly impact traffic and transportation safety. Adverse weather conditions, such as rain, snow, sleet, and fog, are major contributors to the high number of vehicle crashes and fatalities worldwide. This problem is particularly acute in regions with severe weather variations.

In the United States, over 30,000 vehicle crashes occur annually on snowy or icy roads, resulting in over 5,000 fatalities and 418,000 injuries, as reported by the Federal Highway Administration [7] and the US Department of Commerce [8]. The Insurance Institute for Highway Safety (IIHS) underscores this issue, noting that the likelihood of fatal crashes increases by 21% during snowy weather conditions, primarily due to reduced visibility, compared to clear roads. It has been proven that the risk of accident in rain conditions is 70% higher than normal [9]. A 2018 report by the National Highway Traffic Safety Administration (NHTSA) further underscores that adverse weather conditions were involved in 4,000 fatal crashes [10].

In Europe, adverse weather conditions are responsible for 25% of all road accidents. Frost, ice, snow, and rain significantly reduce drivers' perceptible range from hundreds of meters to just a few meters, leading to over 12,000 fatalities annually on European roads due to weather-related accidents, as indicated by the European Commission and the European Transport Safety Council (ETSC) [11]. This drastic reduction in visibility poses a major challenge for effective obstacle or object detection, a critical component of road safety in adverse conditions.

In the context of current advancements in autonomous vehicles (AVs), a persistent challenge remains their operation in inclement weather conditions such as heavy rain or snow, which raises safety concerns. Extensive research and trials have been conducted under adverse weather conditions, yet the Mcity shuttle, for instance, ceases operation when continuous use of windshield wipers is necessitated in rain or snow [12]. Tesla's autopilot system exhibits a limited capability to navigate through mild rain or snow, provided road markings are visible, but encounters difficulties in more severe conditions like heavy storms or obscured lane lines [13]. Similarly, General Motors' Super Cruise, another prominent Level 2 autonomous driving system, explicitly restricts the use of its self-driving function in hazardous conditions, including rain, sleet, fog, ice, or snow [14]. These limitations indicate that, despite advancements, autonomous vehicles are not yet reliable enough to operate independently of human drivers in all weather conditions. Consequently, overcoming these environmental constraints is imperative for the advancement of autonomous driving systems (ADS) into a new era of complete autonomy.

The relevance of this research extends beyond road safety, impacting sectors such as autonomous vehicles, healthcare, precision agriculture, environmental monitoring, aerospace, defense, and industrial automation. For instance, the market for sensor fusion in autonomous vehicles, vital for enhancing object detection capabilities under adverse weather conditions, was valued at USD 594.4 million in 2019. It is projected to grow to USD 1563.5 million by 2025, according to MarketInsider [15]. This growth reflects the increasing reliance on advanced sensor technologies, such as radar and vision sensors, for autonomous emergency braking (AEB) systems and other ADAS applications.

Similarly, In the healthcare sector, the market for wearable sensors, which typically integrate multiple sensors for comprehensive health data, is expected to reach over $1.5 billion by 2030 [16]. The precision agriculture and environmental monitoring markets are anticipated to reach $10.5 billion by 2026 [17], while the aerospace and defense sector is projected to reach $23.83 billion by 2027 [18]. Each of these sectors can benefit significantly from advancements in sensor fusion technology, enhancing efficiency and reducing production costs in industrial automation.

The focus of this research on developing and applying data-driven multimodal sensor fusion techniques for object detection in adverse weather conditions is vital not only for improving road safety in autonomous vehicles but also holds the potential to revolutionize various technology-driven sectors. By addressing the challenges of object detection in complex and unpredictable weather conditions, this research aims to contribute significantly to the fields of autonomous vehicles, healthcare, environmental monitoring, and beyond.

1.2 Challenges and Difficulties

The field of multimodal sensor fusion, particularly in the context of adverse weather conditions, presents a series of distinct challenges that are critical to the advancement of reliable object detection systems. These challenges range from the scarcity of specialized datasets to the intricacies of fusion architecture, computational demands, and the search for effective network design and generalization. This section outlines the key challenges in this field.

Scarcity of Adverse Weather-Related Datasets: A fundamental challenge in this field is the lack of datasets that capture adverse weather conditions through multimodal sensors. Most existing datasets are oriented towards clear weather scenarios. For example, popular datasets for multimodal sensor fusion include camera and LiDAR but often lack radar sensors. These include KITTI [19], ApolloScape [20], and Waymo [21]. There are datasets that incorporate radar sensors, such as VoD [22], TJ4DRadSet [23], and PixSet [24], but they do not cover an extensive range of adverse weather conditions. This limitation restricts the development and validation of object detection models in more challenging environmental conditions and hinders the advancement of robust object detection systems capable of performing reliably under diverse weather conditions.

Limitations in Fusion Architecture: The predominant focus in the current landscape of fusion architectures is on middle or feature fusion, with only limited exploration in the area of tightly coupled fusion networks. This trend poses a particular challenge for sensor fusion that involves radar data, which tends to be noisier and sparser compared to the data produced by cameras and LiDAR. The scarcity of research on tightly coupled fusion methods for integrating such diverse data types is a significant obstacle in developing efficient and effective multimodal fusion systems.

Computational Constraints: Another significant barrier is the computational demand required to train large, complex networks. This requirement often exceeds the resources available to many researchers, restricting exploration in this field and slowing the advancement of more sophisticated fusion methods.

General Guidelines for Network Architecture: Currently, there is no standard or widely accepted framework for the design of network architectures in multimodal sensor fusion, leading to several unanswered questions. Key considerations, as highlighted by Feng et al. [25], include:

- ”What to fuse,” such as LiDAR, radar, various camera types (color, thermal, event), or ultrasonic sensors;
- ”How to fuse,” with possibilities including addition or averaging, concatenation;
- ”When to fuse,” which can range from early, mid, late, or a combination of these fusion stages.

This absence of a clear guideline results in uncertainties regarding the optimal choices for integrating different modalities in sensor fusion.

Limited Generalization from Previous Studies: Many existing studies in multimodal sensor fusion have focused on results from their baseline models and custom datasets. This narrow scope limits the generalizability of their findings, as these models may not perform as well under different conditions or with alternative datasets.

Advanced Fusion Methods and Temporal Information: While recent datasets have begun to include baseline models featuring basic fusion methods, there is notable potential for significant performance enhancement. This can be achieved by incorporating advanced transformer-based architectures, known for their superior handling of complex data patterns and scalability. Additionally, employing sophisticated fusion techniques, such as tightly-coupled fusion, which integrates data more closely and efficiently, could further optimize the sensor fusion process. Additionally, the integration of temporal information in sensor fusion is an area that is yet to be extensively explored. Although not a focus of this project, it represents a promising direction for future research.

Addressing these challenges is critical for advancing the field of object detection in adverse weather conditions using multimodal sensor fusion. This research will contribute to this effort by exploring these underdeveloped areas, aiming to find more robust and effective object detection systems.

1.3 Problem Statement

The field of object detection using multiple modalities has become a topic of increasing interest in recent years, reflecting the growing demand for advanced sensing technologies in various applications, especially in autonomous vehicles and robotics. However, the comprehensive analysis and practical application of state-of-the-art multimodal object detection methods under adverse weather conditions remain largely unexplored areas. The crux of this research project is to bridge this gap by providing an in-depth analysis and practical implementation of cutting-edge techniques in multimodal object detection. The project will explore the integration of multiple sensor modalities, including but not limited to cameras, LiDAR, and radar, to enhance detection capabilities in challenging weather scenarios.

One key aspect of this project is its focus on 2D object detection, a deliberate choice to reduce the complexities associated with network and computer processing. While simplifying the technical challenges, this focus does not diminish the project's primary objective: to explore object detection in adverse weather conditions through a multimodal approach and tightly coupled fusion architecture.

A pivotal challenge in this field is exploring an effective fusion strategy that capitalizes on the complementary strengths of different sensors while mitigating their individual limitations. For example, the integration of visual camera data with radar information presents a complex yet crucial research question. This research aims to explore an optimal fusion strategy, a technique that effectively combines the unique strengths of each sensor type, thereby creating a more robust and accurate detection system. The strategic combination of these modalities holds the potential to significantly enhance object detection performance, especially in conditions where traditional single-sensor systems fall short.

To empirically validate the effectiveness of the proposed methods, this project will conduct comprehensive experiments under a variety of adverse weather conditions. The performance of these methods will be rigorously tested and compared using renowned publicly available datasets, such as DENSE [6] and nuScenes [26]. These datasets, known for their suitability in evaluating performance under challenging environmental conditions, will serve as the benchmark for assessing the robustness and reliability of the proposed multimodal object detection techniques.

Additionally, this study will compare the proposed methods to the best existing methods by comparing their outcomes. This comparison will not only highlight the efficacy of the new methods in adverse conditions but will also shed light on the strengths and weaknesses of each approach. Through this evaluation, the project aims to contribute significantly to the body of knowledge in multimodal object detection, providing valuable insights for future research and practical applications in this rapidly evolving field.

2

Related Works

2.1 Adverse Weather Conditions Influence on Sensors

In the evolving landscape of autonomous robotics, particularly in the domains of self-driving vehicles and autonomous drones, object detection stands as a paramount challenge in computer vision. These cutting-edge applications necessitate precise 2D or 3D bounding boxes for objects within complex and often unpredictable real-world environments. These scenarios commonly involve cluttered scenes, varying lighting conditions, and notably, adverse weather conditions. To tackle these multifaceted challenges, state-of-the-art autonomous vehicle systems are increasingly reliant on a suite of redundant sensor modalities. Recent studies, such as those by Caesar et al. [26], Sun et al. [27], and Ziegler et al. [28], highlight this trend. These sensor modalities extend beyond traditional cameras and LiDAR, encompassing radar and emerging technologies like far-infrared (FIR) and near-infrared (NIR) sensors, which are proving instrumental in enabling reliable object detection in adverse conditions [6].

For standard perception systems in autonomous vehicles, the camera remains an indispensable, yet highly vulnerable sensor to adverse weather conditions. Despite its high resolution, a camera's functionality can be severely compromised by a single water drop on the lens or emitter during rain [29], as illustrated in Figure 2.1. In conditions like heavy snow or hail, image intensity can fluctuate, and object edges may become obscured, leading to detection failures [30]. Additionally, cameras are susceptible to strong light interference, either from direct sunlight or artificial sources like urban light pollution, causing significant operational challenges [31].

TODO: refer table and update this content

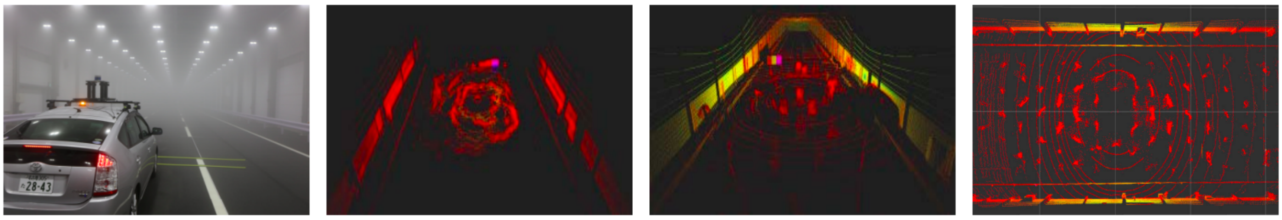
Table 2.1: The influence level of various weather conditions on sensors [32]

Modality	**Light rain** <4mm/hr	**Heavy rain** <25mm/hr	**Dense smoke** vis<0.1km	**Fog** vis<0.5km	**Haze** vis<2km	**Snow**	**Strong light** (over emitter)	**Cost**
LiDAR	2	3	5	4	0	5	2	high
Radar	0	1	2	0	0	2	0	medium
Camera	3	4	5	4	3	3	5	lowest
NIR	2	3	2	1	0	2	4	low
FIR	2	3	1	0	2	4	3	low

Figure 2.1: Van occluded by a water droplet on the lens [33]

LiDAR, the second-most common sensor in autonomous driving systems, exhibits a different response to adverse weather. As Fersch et al. [34] suggest, LiDAR sensors with small apertures are relatively unaffected by moderate rainfall. However, intense and uneven precipitation can generate fog clusters, potentially resulting in false obstacle detection by LiDAR systems. Hasirlioglu et al. [35] demonstrated that rainfall rates exceeding 40 mm/hr significantly reduce signal reflection intensity. Dense fog and smoke, along with strong light, can adversely affect LiDAR sensors in challenging conditions [31, 32]. This is exemplified in Figure 2.2, which showcases an instance of LiDAR's performance in fog, where it erroneously creates small false obstacle clouds. Similarly, Figure 2.3 illustrates how LiDAR's ability to measure distances is compromised in foggy environments. This contrasts with radar technology, whose outputs remain largely unaffected under similar conditions. Such discrepancies highlight the limitations of LiDAR in adverse weather and the need for integrating complementary sensor modalities for enhanced reliability in autonomous driving systems.

Radar, the third critical sensor in autonomous driving systems, is frequently used in mass-produced cars for active safety functions like Automatic Emergency Braking (AEB) and Forward Collision Warning (FCW). Its role in perception tasks for autonomous driving, however, is often undervalued. Unlike cameras operating in the visible light bands (384–769 THz) and LiDARs in the infrared bands (361–331 THz), radar utilizes longer wavelength radio bands (77–81 GHz). This attribute



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

Figure 2.2: LiDAR performance test (a sample from LIBRE [2] dataset)

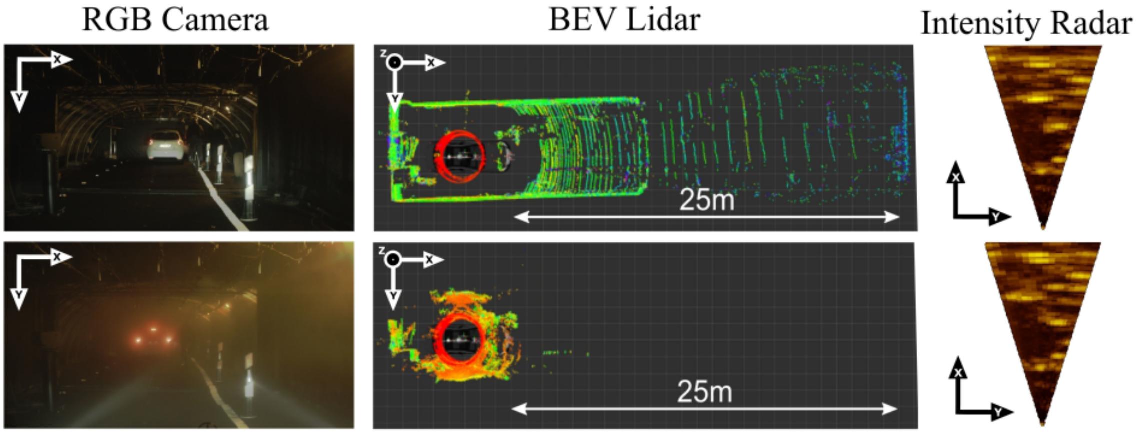
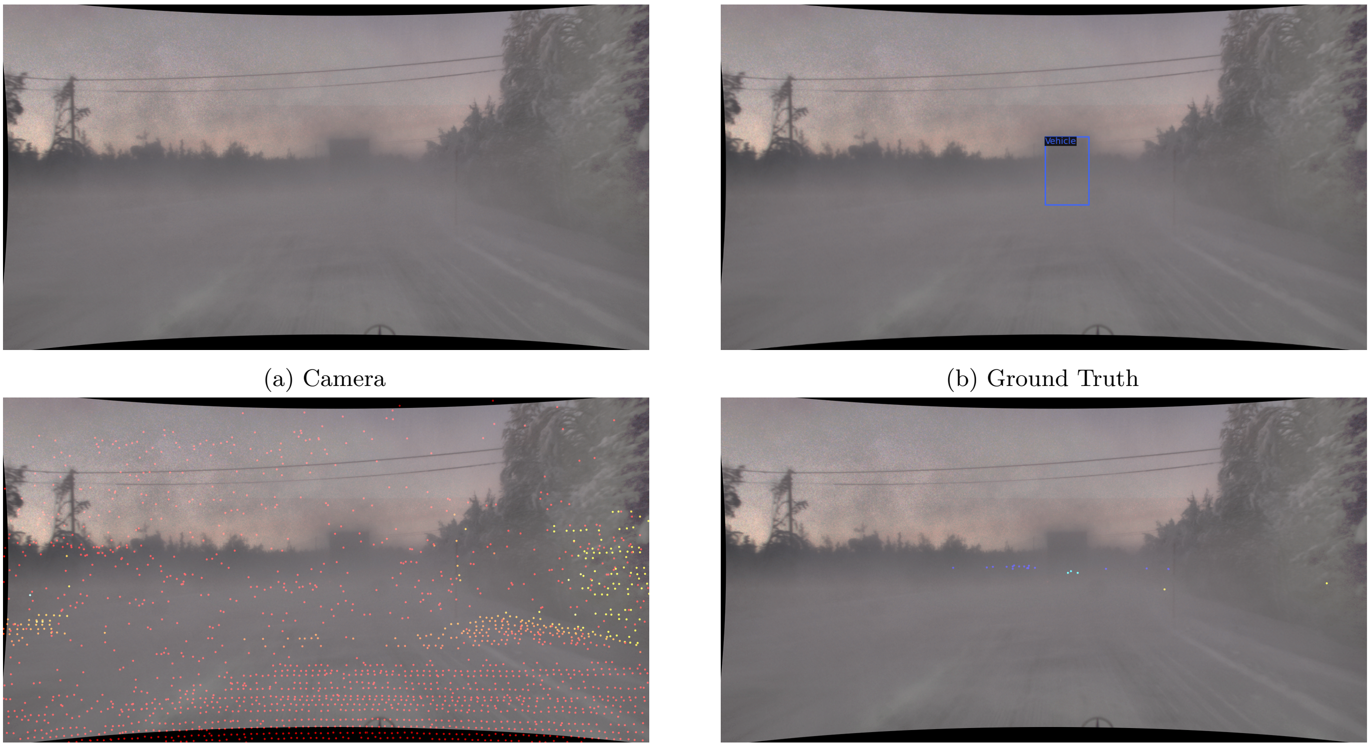


Figure 2.3: 1st row: clear weather condition, 2nd row: with fog. Shows that lidar affects by the fog but radar intensity remains the same [6]

ensures its robust performance in adverse weather conditions [36]. Studies by Ijaz et al. [37] and Ismail [38] indicate radar's lower attenuation in rainy conditions compared to LiDAR. At 77 GHz, radar exhibits approximately 3.5 times less attenuation (10 dB/km) than LiDAR at 905 nm (35 dB/km), showcasing its superior robustness. Various experiments [30, 39–42] have demonstrated that radar's performance is minimally affected by dust, fog, snow, and light rain, although it degrades in heavy rainfall conditions. Figure 2.4 from the DENSE [6] dataset exemplifies this, showing radar's ability to detect vehicles under dense fog conditions, where cameras and LiDAR fail. Nevertheless, radar's low resolution and sparser point clouds compared to LiDAR limit its utility in perception tasks. The emerging 4D radar technology, while promising denser point clouds, lacks public datasets in adverse weather conditions for validation.

The combination of LiDAR and camera technologies alone has proven insufficient for navigating through adverse weather conditions with adequate safety assurance. However, the integration of these with radar, infrared cameras, gated cameras, stereo cameras, weather stations, and other weather-related sensors presents a new paradigm in autonomous vehicle perception. This multimodal sensor fusion, as evidenced in Figures 1.1 and 1.2, offers a composite strength that individual systems lack. Consequently, research groups worldwide are exploring various permutations and combinations of these sensors to enhance the reliability and safety of autonomous driving systems in challenging weather conditions.



(a) Camera (b) Ground Truth

(c) Lidar (d) Radar

Figure 2.4: Dense fog influence on sensors (a sample from DENSE dataset [6])

2.2 Multimodal Sensor Fusion

The concept of multimodal sensor fusion has become increasingly pivotal in the field of object detection, particularly under adverse weather conditions. This approach integrates various sensor inputs to enhance perception and object detection capabilities, addressing the limitations inherent to individual sensors. A notable contribution in this field is the work of Radecki et al. [43], who conducted a thorough review of sensor efficacy across diverse weather conditions, including wet environments, varying light conditions, and dusty atmospheres. They developed a sophisticated system capable of object tracking and classification through a real-time, joint probabilistic perception algorithm. This algorithm dynamically selects the most appropriate sensor subsets based on prevailing weather conditions. By intelligently weighting sensors and accurately quantifying parameters specific to each weather scenario, the system not only improves the baseline of perception ability but also enhances its robustness and reliability. This work underscores the importance of real-time strategy adaptation and smart sensor subset selection in maximizing the accuracy and dependability of multimodal perception systems.

However, Radecki et al.'s study [43] has certain limitations, notably its focus on optimal weather conditions and the omission of heavy traffic and urban environments. Additionally, the study does not explore deep learning-based fusion techniques. Future research could address these gaps by employing similar approach for object detection in challenging weather datasets.

In 2019, FLIR System Inc. [44] and VSI Labs [45] tested the first-ever fused Automated Emergency Braking (AEB) sensor suite, comprising a thermal long-wave infrared (LWIR) camera, a radar, and a visible camera. The LWIR camera, operating in the 8 µm to 14 µm range at ambient temperature, was part of this groundbreaking sensor suite. The suite's performance was evaluated against standard AEB systems using radar and visible cameras under various conditions, including daytime, nighttime, and transitions from tunnel exits into sun glare. The results revealed that while most AEB systems function effectively during the day, they almost invariably failed under adverse conditions, often colliding with mannequins. In stark contrast, the LWIR sensor suite successfully avoided collisions in these challenging scenarios, highlighting the efficacy of fusing camera and radar data in adverse weather conditions.

However, the use of a thermal camera as part of the sensor fusion raises concerns about the durability of such temperature-sensitive devices in real-world settings, necessitating further validation to ascertain their effectiveness in adverse weather [30].

Another significant development in multimodal sensor fusion is the CameraRadarFusionNet (CRF-Net) proposed by Nobis et al. [46]. Inspired by previous works on camera-LiDAR fusion [47, 48], the CRF-Net was designed to determine the most beneficial stage for sensor data fusion within a neural network architecture for detection tasks. Utilizing the nuScenes [26] dataset and their own TUM dataset, they introduced a novel training strategy, BlackIn, focusing on a specific sensor type. The fusion method employed, element-wise addition, demonstrated superior performance over an image-only network on both datasets, underscoring the significance of incorporating radar data into detection tasks.

Nevertheless, Nobis et al.'s approach [46] exhibited only marginal improvements in detection performance over the baseline image network. The absence of an RGB sensor ablation study in their work left questions about the system's robustness in case of camera failure. According to Safa et al. [49], pre-processing the radar data before fusion could further enhance performance, suggesting an area for potential improvement in future research endeavors.

Building on the concept of multimodal sensor fusion, Yang et al. [50] introduced RadarNet, a novel framework for object detection and velocity estimation. RadarNet is distinctive for its dual strategy of leveraging both radar and LiDAR sensors for perception. It employs an early fusion technique to learn joint representations from these sensors, while its late fusion phase incorporates radar's radial velocity evidence to enhance object velocity estimation. This approach was rigorously evaluated using the nuScenes dataset [26]. However, a limitation of RadarNet, as noted by Yang et al. [50], lies in the radar sensor data from the nuScenes dataset, which is characterized by low resolution, thus restricting its effectiveness in object detection. This issue of low resolution and erroneous elevation estimates, as highlighted in studies like Ulrich et al. [51] and Drews et al. [52], suggests a potential improvement for RadarNet: the integration of a higher-resolution radar, such as the one from the K-Radar dataset. This 4D radar, with a significantly wider elevation angle, could enhance the performance of RadarNet, overcoming the limitations of conventional radars.

In a similar vein, Bijelic et al. [6] from Mercedes-Benz AG conducted an extensive study focusing on enhancing detection performance in adverse weather through deep multimodal sensor fusion. Their experimental setup included a diverse array of sensors on their test vehicle, including stereo RGB cameras, a NIR camera, a 77 GHz radar, dual LiDARs, an FIR camera, a weather station, and a road-friction sensor. They introduced an innovative entropy-steered fusion approach, attenuating regions of low entropy while amplifying those with high entropy during feature extraction. This method, trained using clear weather data, showed impressive adaptability to adverse weather conditions. The fusion network was designed to maintain consistency across different scenarios, with all sensor data projected into the camera coordinate system. Their findings demonstrated that this fused approach significantly outperformed LiDAR or image-only methods, particularly under foggy conditions.

Additionally, Bijelic et al. [6] contributed to the research community by providing the SeeingThroughFog or DENSE dataset. This comprehensive dataset, comprising 10,000 km of driving data recorded in Northern Europe, spans a range of weather and illumination conditions. The dataset includes detailed annotations for various weather frames, including clear weather, dense fog, light fog, and snow/rain, making it a valuable resource for future studies on multimodal sensor fusion in challenging weather scenarios.

However, the study by Bijelic et al. [6] also presents certain challenges. The projection transformation technique used in their approach may lead to the loss of crucial radar spatial information. Moreover, the extensive array of sensors required exceeds typical expectations for autonomous driving systems, posing implementation challenges in real-world scenarios. The large volume of data from multiple sensors could potentially impact the algorithm's response and reaction time. Additionally, the radar's performance was constrained by its limited azimuth and elevation resolution [32]. Future research could focus on enhancing the network architecture, possibly through the integration of higher resolution radars and transformer-based approaches, to further refine the performance of sensor fusion in adverse weather conditions.

Liu et al. [53] presented another innovative approach to target recognition and tracking by fusing radar and camera data. In their methodology, radar is the primary sensor, complemented by camera data as secondary information. This fusion was evaluated under challenging weather conditions, including rain, fog, and low visibility nighttime scenarios. The results demonstrated that radar-based detection was highly accurate in detecting moving targets in wet weather, while the camera excelled in target classification. The combined radar and camera data exhibited a superior performance, surpassing LiDAR-based methods by over 33%, highlighting the effectiveness of this fusion approach in challenging weather scenarios.

The exploration of multimodal sensor fusion for enhanced vehicle detection under adverse weather conditions continues with the work of Qian et al. [54], who developed the Multimodal Vehicle Detection Network (MVDNet). MVDNet uniquely incorporates LiDAR and radar data, utilizing a two-stage attention block within its fusion module. The network first applies self-attention to each modality to extract features, followed by the blending of these features with region-wise features through cross-attention mechanisms. This fusion technique has shown to be particularly effective in foggy conditions. The performance of MVDNet was rigorously tested and validated on the DENSE [6] and Oxford Radar Robotcar [55] datasets, where it demonstrated a notably improved performance compared to LiDAR-only systems in foggy environments.

Despite its robust performance, MVDNet's design by Qian et al. [54] has certain limitations. One significant issue is the misalignment between LiDAR and radar data within the dataset, which could potentially compromise the network's effectiveness. Furthermore, the simple label assignment strategy employed in the loss computation and the region-of-interest (ROI) assisted fusion design might limit the model's overall performance. These aspects suggest potential areas for improvement, possibly through the adoption of more advanced fusion techniques and refined label assignment strategies, as indicated in studies like Yang et al. [56].

Continuing in this vein, Rawashdeh et al. [57] proposed a CNN-based sensor fusion approach aimed at detecting drivable paths. This approach integrates data from cameras, LiDAR, and radar and was evaluated using the DENSE dataset [6]. Their multi-stream encoder-decoder network is designed to counter the asymmetric degradation of the input sensors effectively. The depth and number of blocks allocated to each sensor in the architecture were determined by their respective input data densities, with the camera being the most dense, followed by LiDAR, and radar being the least dense. The fully connected network's outputs are transformed into a 2-D array for input into the decoder. The researchers demonstrated that their model could effectively ignore misleading road lines and edges, thereby accurately delineating the general drivable area.

However, Rawashdeh et al.'s approach [57] also presents certain challenges. The method lacks a comparative analysis with other state-of-the-art methods in the field, leaving its relative effectiveness somewhat uncertain. Additionally, the study does not delve into the real-time processing requirements or the computational costs of the proposed algorithm. These factors are crucial for practical applicability, especially in scenarios requiring rapid decision-making and processing, like autonomous driving in adverse weather conditions. Addressing these gaps could significantly enhance the feasibility and implementation potential of their sensor fusion approach in real-world applications.

2.3 Synthetic Data for Adverse Weather Conditions

The use of synthetic data and advanced image processing techniques has been a subject of considerable research focus. One prominent area is the application of de-hazing techniques to mitigate the impacts of adverse weather on visual data. Historically, physical priors have been employed for this purpose [58, 59]. However, with the advent of data-driven approaches, particularly deep learning, new methodologies have emerged. These deep de-hazing models, while innovative, often suffer from high computational complexity, making them less suitable for ultra-high-definition images. Chen et al. [60] highlighted a significant limitation of these models: their training on synthetic images does not effectively generalize to real-world hazy conditions. Conversely, Zhang et al. [61] leveraged temporal redundancy in video de-hazing, assembling a dataset comprising real-world hazy and haze-free videos. The challenge in this domain lies in obtaining paired hazy and haze-free ground-truth images, which is difficult in natural settings. However, this obstacle can be somewhat mitigated through the use of professional haze/fog generators that simulate real-world conditions [62, 63].

Another emerging trend in this field is the exploration of synthetic data generation for adverse weather conditions using Generative Adversarial Networks (GANs). Researchers such as Sun et al. [64], Zheng et al. [65], and Lee et al. [66] have investigated this approach, utilizing clean weather datasets like KITTI [19] and Cityscapes [67] as a basis. These methods predominantly involve the creation of artificial fog or rain images, supplemented by a limited selection of actual images captured under specific fog or rain conditions. However, the efficacy of these algorithms in diverse adverse weather scenarios remains somewhat ambiguous. A critical concern is whether these synthetic data-driven approaches can perform effectively under various real-world adverse weather conditions. Additionally, the methods for evaluating their real-world applicability and effectiveness in such scenarios are not fully established [68]. This gap in the research indicates a need for further investigation and development to enhance the reliability and applicability of synthetic data generation techniques in the context of adverse weather conditions for object detection.

2.4 The Role of Simulation in Autonomous Driving Research

The advent of autonomous driving technology, especially in challenging weather conditions, has significantly benefited from the utilization of simulation platforms and specialized experimental setups. One such noteworthy tool is the CARLA simulator [69], a widely recognized virtual platform. CARLA is particularly advantageous for researchers as it enables the creation of intricate road environments and the simulation of numerous non-ego entities in scenarios that would otherwise be impractical or prohibitively expensive to replicate in real-life experiments. This capability is crucial, considering that specific weather conditions, particularly those related to extreme climates or certain seasons, are not always readily available for testing. For example, tropical regions cannot easily conduct tests in snow conditions, and the unpredictability and brevity of natural rain showers may impede the collection of comprehensive experimental data. Most importantly, conducting tests in actual adverse weather conditions not only presents logistical challenges but also introduces significant safety risks. In contrast, simulators like CARLA offer a completely safe environment, eliminating the dangers associated with real-world testing [32].

However, the effectiveness of virtual datasets and simulation platforms in accurately representing real-world phenomena remains a topic of debate. The extent to which a simulator can truly mirror real-world conditions is an open question. Developing more realistic simulators is a key challenge in this domain. Additionally, determining the most effective methods for integrating real and virtual data is another critical area of ongoing research [25]. These aspects underline the need for continuous improvement in simulation technologies to ensure that they can effectively support the development and testing of autonomous driving systems, especially in the face of adverse weather conditions. The quest for more realistic simulators and the optimal blend of real and virtual data stand as important open questions, driving the future direction of research in autonomous driving simulations. Due to these reasons, this project will focus on the real-world datasets for adverse weather conditions.

3

Methodology

3.1 Available Datasets

The majority of deep multimodal perception approaches rely on supervised learning, which necessitates the use of high-quality, large-scale multimodal datasets with labeled ground truth for training deep neural networks. Several multimodal datasets, such as KITTI [19], ApolloScape [20], and Waymo [21], are prevalent in the domain of LiDAR-camera fusion. However, a significant number of these datasets are collected under clear weather conditions or lack a comprehensive array of sensors, including cameras, LiDAR, and Radar. A notable limitation is the scarcity of multimodal datasets that are collected under adverse weather conditions and incorporate at least all three of these essential sensors. Table 3.1 summarizes some of the available multimodal datasets[1] for evaluating the performance of deep multimodal perception techniques in adverse weather conditions. The dataset are sorted in ascending order with respect to year.

Table 3.1: Multimodal Adverse Weather Conditions Datasets. Sensors†: C-R-L-N-F denote Camera, Radar, LiDAR, Near-infrared, and Far-infrared sensors, respectively. Weather Conditions‡: F-SN-R-O-SL-N denote Fog, Snow, Rain, Overcast, Sleet, and Night conditions, respectively. Note that highlighted datasets are used for the project.

Name	**Sensors†**	**Weather Cond.‡**	**Size (GB)**	**Year**	**Citation Cnt.**	**Ref.**
DENSE	**CRLNF**	**F, SN, R, O, N**	**582**	**2020**	**269**	**[6]**
nuScenes	**CRL**	**R, N**	**400**	**2020**	**3459**	**[26]**
The Oxford RobotCar	CRL	R, SN, F	4700	2020	317	[55]
EU Long-term	CRL	SN, R, O, N	NA	2020	72	[70]
RADIATE	CRL	F, SN, R, O, SL, N	NA	2021	132	[71]
K-Radar	CRL	F, R, SN	13000	2022	15	[36]
Boreas	CRL	SN, R, O, N	4400	2022	38	[72]
aiMotive	CRL	R, O, N	85	2023	3	[73]

The selection of appropriate datasets is crucial. The criteria for selecting datasets cover several key areas, focusing on the availability of diverse sensors, specifically cameras, radar, and lidar, which are crucial for robust object detection in challenging environments. Additionally, the datasets must represent adverse

[1] For all the datasets, formal registration form is required to fill to access the dataset

weather conditions effectively, as this is a critical aspect of the research. Furthermore, the accessibility and thorough documentation of the datasets are considered, ensuring that the data can be easily understood and utilized in the research process. Another vital criterion is the dataset's popularity in existing research, as this allows for comparative analysis with publicly available methods, thereby validating the research findings. Moreover, the specific perception task, in this case, object detection, and the type of radar data, particularly in point cloud format, are essential considerations. The requirement for time-synchronized and calibrated data is also emphasized to ensure accuracy and reliability in sensor fusion and object detection algorithms.

After a comprehensive evaluation of these criteria, two datasets have been chosen for this research: DENSE and nuScenes. The DENSE dataset is particularly suited for this study as it includes data from various sensors under adverse weather conditions, which is crucial for testing the efficacy of multimodal sensor fusion in challenging environments. The nuScenes dataset, on the other hand, is widely used in the field, providing a rich source of data with camera, radar, and lidar sensors. Its extensive use in the community allows for a meaningful comparison with existing methods. Both datasets provide time-synchronized and calibrated data, which is essential for the accuracy of object detection algorithms in adverse weather conditions. The selection of these datasets aligns perfectly with the research objectives, offering a comprehensive platform for exploring and advancing the capabilities of data-driven multimodal sensor fusion in object detection under challenging weather scenarios.

3.1.1 DENSE dataset

The DENSE dataset, as detailed in Bijelic et al. (2020) [6], is a critical asset for evaluating multi-modal fusion algorithms in adverse weather conditions. Its standout feature is the extensive sensor array, including LiDAR, a stereo camera, a frontal long-range radar, a gated camera operating in the NIR band, a FIR camera, and a weather station sensor, as illustrated in Figure 3.1. These sensors allow for detailed data capture under various adverse weather conditions, such as rain, snow, light fog, and dense fog. Notably, the DENSE dataset uniquely offers a split for light and dense fog conditions, essential for assessing the detection performance of Lidar and Radar in varying visibility scenarios. The range of these conditions and their distribution are visually depicted in Figure 3.2. Additionally, the dataset includes data from a controlled lab environment within a fog chamber, offering a distinct view of sensor performance under simulated conditions. However, it's important to note that for the purposes of this project, only real-world data from the DENSE dataset is utilized.

The dataset covers a broad spectrum of environments, encompassing urban cities, suburban areas, highways, and tunnels. Its geographical scope is extensive, with data collection spanning over two months and covering 10,000 km across Germany, Sweden, Denmark, and Finland. This diverse environmental range enhances the dataset's applicability in various real-world scenarios.

Technically, the DENSE dataset offers radar targets in a point cloud format, aligning well with the Lidar data. Given the inherent noise in radar data, preprocessing has been performed to eliminate false points, thereby bolstering the dataset's accuracy and reliability. The radar data includes 3D information - range, azimuth, and velocity. Moreover, it's noteworthy that the latest generation of radar sensors in the dataset provides 4D data, adding elevation to the existing dimensions. There are total 3 classes available in the dataset, including car, pedestrian, and cyclist. Object annotations in the DENSE dataset are provided in the COCO style format [74], with bounding box (bbox) parameters specified as x, y, width, and height. The annotation processed is well described in the supplementary material from the paper [75]. This meticulous approach to data collection, processing, and annotation positions the DENSE dataset as a powerful and adaptable tool for research in adverse weather conditions, especially in the domain of data-driven multimodal sensor fusion.

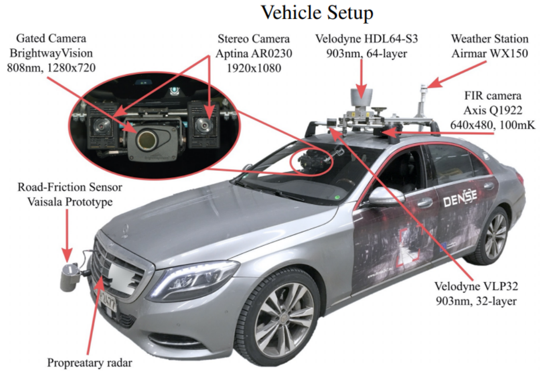


Figure 3.1: Test Vehicle Setup [6]

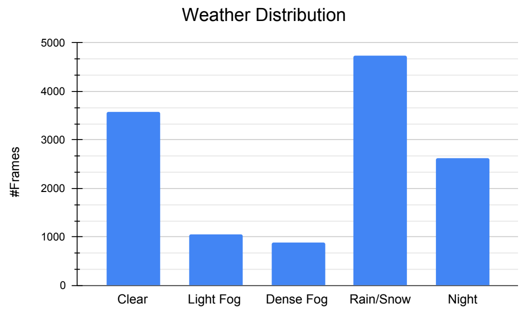


Figure 3.2: Distribution of Weather Conditions [6]

TODO: add Classwise distribution

TODO: put two in 4x4 grid, random samples of the dataset, one for clean and other with adverse weather conditions

3.1.2 nuScenes dataset

In addition to the DENSE dataset, this project also uses the nuScenes dataset [26] as a benchmark. The nuScenes dataset stands out for its large-scale and diverse scenarios. The data collection vehicle for nuScenes, as depicted in Figure 3.3, is equipped with a comprehensive set of sensors, including a 32-beam LiDAR, six cameras, five long-range multi-mode radars, and a GPS/IMU system. This dataset provides 3D annotations for 23 classes of road users across 1,000 scenes, accumulating to a total of 1.3 million frames. Although the radar data in nuScenes is sparse, its extensive documentation makes it a good starting point for research in object detection.

The nuScenes dataset focuses on urban, suburban, and highway areas, but it covers fewer adverse weather conditions compared to the DENSE dataset, primarily rain and night scenarios, as shown in Figure 3.4. Like DENSE, nuScenes also provides radar data in point cloud format. There are total 10 classes in the dataset, including car, truck, trailer, bus, construction vehicle, bicycle, motorcycle, pedestrian, traffic cone, and barrier. However, a notable distinction is that nuScenes does not provide 2D annotations. Researchers using this dataset typically generate their own 2D annotations based on the 3D annotations provided. Once these 2D annotations are created, the data is converted into the COCO style format [74], similar to DENSE, where the bbox format includes x, y, width, and height dimensions.

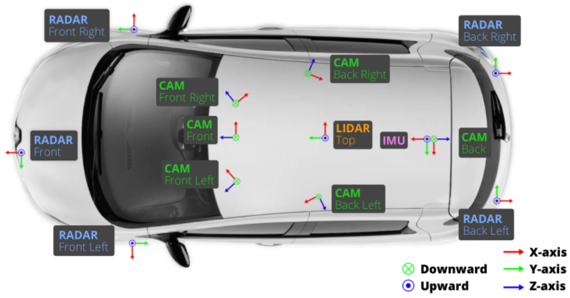


Figure 3.3: Test Vehicle Setup [26]

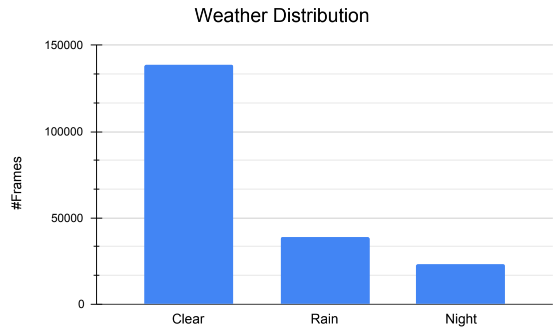


Figure 3.4: Distribution of Weather Conditions [26]

TODO: add Classwise distribution TODO: put two in 4x4 grid, random samples of the dataset, one for clean and other with adverse weather conditions

Table 3.2 highlights the overall comparison of the sensor setup and dataset statistics for datasets used in this project.

Table 3.2: Comparison of Dataset Features

Dataset	**NuScenes [26]**	**DENSE [6]**
RGB Cameras	6	2
RGB Resolution	1600x900	1920x1024
Lidar Sensors	1	2
Lidar Resolution	32	64
Radar Sensor	4	1
Gated Camera	x	1
FIR Camera	x	1
Frame Rate	1 Hz/10 Hz	10 Hz
Dataset Statistics		
Labeled Frames	40K	13.5K
Labels	1.4M	100K
Scene Tags	✓	✓
Night Time	✓	✓
Light Weather	✓	✓
Heavy Weather	x	✓
Fog Chamber	x	✓

3.2 Evaluation Metrics

Metrics provide a standardized scale for comparing various methods in object detection tasks. Among these, Average Precision (AP) and Average Recall (AR) stand out as the most prominent. The understanding and computation of AP and AR are deeply rooted in more fundamental concepts such as Precision and Recall, which form the basis for these advanced metrics.

Average Precision (AP) is a pivotal metric in object detection tasks, offering a comprehensive measure of a model's precision and recall at various thresholds. A common misconception is that AP represents the average of Precision values, which is not accurate. A more precise interpretation of AP is that it reflects the area under the Precision-Recall curve. Precision and recall are fundamental concepts in this context, defined as follows:

Precision (P) is the ratio of correctly predicted positive observations to the total predicted positive observations, formulated as

$$P = \frac{TP}{TP + FP},$$

where TP is the number of true positives and FP is the number of false positives.

Recall (R) is the ratio of correctly predicted positive observations to all observations in the actual class, given by

$$R = \frac{TP}{TP + FN},$$

with FN representing the number of false negatives.

In addition to these metrics, the COCO benchmark categorizes objects based on their size: small, medium, and large. Specifically, an object is considered *small* if its bounding box area is less than 32^2 pixels, allowing for more detailed assessment of model performance across different object scales.

AP can be calculated in numerous ways, but in the realm of object detection, it is typically computed class-wise and then averaged over all classes, yielding the mean Average Precision (mAP). In this project, we adopt the COCO style benchmarking, where mAP is referred to as AP. The COCO AP is calculated over a range of Intersection over Union (IoU) thresholds and is often denoted as $AP@[0.5 : 0.05 : 0.95]$, indicating multiple IoU thresholds from 0.5 to 0.95 with a step size of 0.05. The formula for calculating COCO AP is defined as:

$$AP@[0.5 : 0.05 : 0.95] = \frac{AP_{0.5} + AP_{0.55} + \ldots + AP_{0.95}}{10},$$

where $AP_{0.5}$ is the area under the Precision-Recall curve for IoU ≥ 0.5.

The COCO detection benchmark encompasses 12 distinct metrics, as illustrated in Table 3.3.

Average Recall (AR) offers a vital alternative perspective to AP in the assessment of object detection models. AR evaluates the ability of a model to accurately recognize *all* relevant examples of the specified categories, disregarding the incidence of false positives. In the COCO benchmark, AR is calculated with varying numbers of detections per image, specifically 1, 10, or 100. Notably, AR at 1 considers only the detection with the highest confidence score for each image, focusing on the model's precision in identifying the most probable object. This contrasts with AR at 10 or 100, where multiple detections per image are considered, reflecting the model's capability to identify numerous objects with varying confidence levels. The importance of achieving a high recall in all categories, as indicated by AR, is particularly critical in scenarios where overlooking an object could lead to severe consequences. For instance, in the context of self-driving vehicles operating under adverse weather conditions such as fog or heavy rain, the failure to detect a pedestrian or an approaching vehicle could pose a threat to human life. AR, by focusing

Table 3.3: COCO Metrics [74]

Metric	**Description**
Average Precision (AP):	
$AP^{IoU=.50:.05:.95}$	AP at IoU=.50:.05:.95 (primary COCO metric)
$AP^{IoU=.50}$	AP at IoU=.50 (PASCAL VOC metric)
$AP^{IoU=.75}$	AP at IoU=.75 (strict metric)
AP Across Scales:	
AP^{small}	AP for small objects: area $< 32^2$
AP^{medium}	AP for medium objects: $32^2 <$ area $< 96^2$
AP^{large}	AP for large objects: area $> 96^2$
Average Recall (AR):	
$AR^{max=1}$	AR given 1 detection per image
$AR^{max=10}$	AR given 10 detections per image
$AR^{max=100}$	AR given 100 detections per image
AR Across Scales:	
AR^{small}	AR for small objects: area $< 32^2$
AR^{medium}	AR for medium objects: $32^2 <$ area $< 96^2$
AR^{large}	AR for large objects: area $> 96^2$

exclusively on the rate of detection and excluding considerations of precision, emphasizes these types of missed detections that might be neglected when only considering AP. Consequently, AR provides crucial insights about the reliability and effectiveness of object detection systems, complementing the focus on precision embodied by AP.

Given the project's focus on object detection in adverse weather conditions, we will extend the evaluation of AP and AR to include performance under specific weather scenarios like fog, rain, and snow. This will provide a more comprehensive understanding of the model's robustness and effectiveness in varying environmental conditions.

In addition to these metrics, **Inference Time** and **FLOPs** (Floating Point Operations Per Second) or **GFLOPs** (GigaFLOPs) are crucial for assessing the computational efficiency and performance of the models. Inference time, significantly influenced by the hardware used, serves as a reliable indicator of a model's practical applicability in various scenarios. In this study, the inference time for all models is tested on an NVIDIA V100 GPU, ensuring a consistent and robust basis for comparison. Furthermore, the **Model Parameters** metric is instrumental in understanding the models' complexity, shedding light on their computational requirements and potential scalability.

4

Evaluation and Results

4.1 Experiment Description

Describe the experiments/evaluation you are performing to analyse your method.

4.2 Experimental Setup

Describe your experimental setup in detail.

4.3 Results

Describe the results of your experiments in detail.

5

Conclusions

5.1 Contributions

5.2 Lessons learned

5.3 Future work

Utilization of 4D Imaging Radar in Adverse Weather: There is a notable lack of research utilizing 4D imaging radar sensors, especially in adverse weather conditions [76]. Given the potential of these sensors in challenging environments, further exploration in this area is essential. The K-Radar dataset [36] is a step in the right direction, but yet to be explored.

A

Design Details

Your first appendix

B

Parameters

Your second chapter appendix

References

[1] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies," *IEEE access*, vol. 8, pp. 58 443–58 469, 2020.

[2] A. Carballo, J. Lambert, A. Monrroy, D. Wong, P. Narksri, Y. Kitsukawa, E. Takeuchi, S. Kato, and K. Takeda, "Libre: The multiple 3d lidar dataset," *IEEE Intelligent Vehicles Symposium (IV)*, pp. 1094–1101, 2020.

[3] U. of Michigan, "Getting traction: Tips for traveling in winter weather," 2020. [Online]. Available: https://mcity.umich.edu/wp-content/uploads/2020/10/mcity-driverless-shuttle-whitepaper.pdf

[4] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, "Sensor and sensor fusion technology in autonomous vehicles: A review," *Sensors*, vol. 21, no. 6, p. 2140, 2021.

[5] R. Laganière, "Sensor fusion for autonomous vehicles: Strategies, methods, and tradeoffs," 2022, accessed on 18.12.2022. [Online]. Available: https://youtu.be/2Fcmh7SLPBI

[6] M. Bijelic, T. Gruber, F. Mannan, F. Kraus, W. Ritter, K. Dietmayer, and F. Heide, "Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather," pp. 11 682–11 692, 2020.

[7] Federal-Highway-Administration, "How Do Weather Events Impact Roads? - FHWA Road Weather Management." [Online]. Available: https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm

[8] N. US Department of Commerce, "Getting traction: Tips for traveling in winter weather," Nov 2016. [Online]. Available: https://www.weather.gov/wrn/getting_traction

[9] J. Andrey and S. Yagar, "A temporal analysis of rain-related crash risk," *Accident Analysis & Prevention*, vol. 25, no. 4, pp. 465–472, 1993.

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

[11] G. Cookson, "Weather-Related Road Deaths in Europe £15bn+ Per Year — INRIX," 2 2022. [Online]. Available: https://inrix.com/blog/blog-cost-of-weather/

[12] U. Briefs, "Mcity grand opening," *Research Review*, vol. 46, no. 3, 2015.

[13] F. Lambert, "Watch tesla autopilot go through a snowstorm," https://electrek.co/2019/01/28/tesla-autopilot-snow-storm/, 2019, [Last accessed 10 May 2021].

[14] Cadillac General Motors, "Designed to take your hands and breath away," https://www.cadillac.com/ownership/vehicle-technology/super-cruise, 2021, last accessed 10 May 2021.

[15] "Sensor fusion market in autonomous vehicles - growth, trends, forecasts (2020-2025)," Nov 2020. [Online]. Available: https://markets.businessinsider.com/news/stocks/sensor-fusion-market-in-autonomous-vehicles-growth-trends-forecasts-2020-2025-1029759714

[16] S. Research, "Wearable sensors market," 2021. [Online]. Available: https://straitsresearch.com/report/wearable-sensors-market#:~:text=Market%20Overview,period%20(2022%E2%80%932030)

[17] Mordor intelligence. Precision farming market analysis, industry report, trends, size, and share. [Online]. Available: https://www.mordorintelligence.com/industry-reports/global-precision-farming-market-industry#:~:text=The%20Precision%20Farming%20Market%20was,period%20(2021%2D2026).

[18] Fortune business insights. Aerospace and defense materials market size, share, and report, forecast 2020-2027. [Online]. Available: https://www.fortunebusinessinsights.com/aerospace-defense-materials-market-102980

[19] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2012, pp. 3354–3361.

[20] X. Huang, P. Wang, X. Cheng, D. Zhou, Q. Geng, and R. Yang, "The apolloscape open dataset for autonomous driving and its application," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 10, pp. 2702–2719, 2019.

[21] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. 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*, 2020, pp. 2446–2454.

[22] A. Palffy, E. Pool, S. Baratam, J. F. Kooij, and D. M. Gavrila, "Multi-class road user detection with 3+ 1d radar in the view-of-delft dataset," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4961–4968, 2022.

[23] L. Zheng, Z. Ma, X. Zhu, B. Tan, S. Li, K. Long, W. Sun, S. Chen, L. Zhang, M. Wan *et al.*, "Tj4dradset: A 4d radar dataset for autonomous driving," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2022, pp. 493–498.

[24] J.-L. Déziel, P. Merriaux, F. Tremblay, D. Lessard, D. Plourde, J. Stanguennec, P. Goulet, and P. Olivier, "Pixset: An opportunity for 3d computer vision to go beyond point clouds with a full-waveform lidar dataset," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2021, pp. 2987–2993.

[25] D. Feng, C. Haase-Schütz, L. Rosenbaum, H. Hertlein, C. Glaeser, F. Timm, W. Wiesbeck, and K. Dietmayer, "Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges," pp. 1341–1360, 2020.

[26] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nuscenes: A multimodal dataset for autonomous driving," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 11 621–11 631.

[27] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, Y. Zhang, J. Shlens, Z. Chen, and D. Anguelov, "Scalability in perception for autonomous driving: Waymo open dataset," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[28] J. Ziegler, P. Bender, M. Schreiber, H. Lategahn, T. Strauss, C. Stiller, T. Dang, U. Franke, N. Appenrodt, C. G. Keller *et al.*, "Making bertha drive—an autonomous journey on a historic route," *IEEE Intelligent transportation systems magazine*, vol. 6, no. 2, pp. 8–20, 2014.

[29] R. Mardirosian, "Lidar vs. camera: driving in the rain," [Last accessed 10 April 2023], 2023. [Online]. Available: https://ouster.com/blog/lidar-vs-camera-comparison-in-the-rain/

[30] S. Zang, M. Ding, D. Smith, P. Tyler, T. Rakotoarivelo, and M. A. 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*, vol. 14, no. 2, pp. 103–111, 2019.

[31] A. Carballo, J. Lambert, A. Monrroy, D. Wong, P. Narksri, Y. Kitsukawa, E. Takeuchi, S. Kato, and K. Takeda, "LIBRE: The multiple 3d lidar dataset," *arXiv preprint arXiv:2003.06129*, 2020, (accepted for presentation at IV2020).

[32] Y. Zhang, A. Carballo, H. Yang, and K. Takeda, "Perception and sensing for autonomous vehicles under adverse weather conditions: A survey," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 196, pp. 146–177, 2023.

[33] F. Nobis, M. Geisslinger, M. Weber, J. Betz, and M. Lienkamp, "A Deep Learning-based Radar and Camera Sensor Fusion Architecture for Object Detection," *arXiv*, May 2020. [Online]. Available: https://arxiv.org/abs/2005.07431v1

[34] T. Fersch, A. Buhmann, A. Koelpin, and R. Weigel, "The influence of rain on small aperture lidar sensors," in *German Microwave Conference (GeMiC)*. IEEE, 2016, pp. 84–87.

[35] S. Hasirlioglu, I. Doric, C. Lauerer, and T. Brandmeier, "Modeling and simulation of rain for the test of automotive sensor systems," in *IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2016, pp. 286–291.

[36] D.-H. Paek, S.-H. Kong, and K. T. Wijaya, "K-Radar: 4D Radar Object Detection for Autonomous Driving in Various Weather Conditions," Jun. 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2206.08171

[37] M. Ijaz, Z. Ghassemlooy, H. Le Minh, S. Rajbhandari, and J. Perez, “Analysis of fog and smoke attenuation in a free space optical communication link under controlled laboratory conditions,” in *International Workshop on Optical Wireless Communications (IWOW)*. IEEE, 2012, pp. 1–3.

[38] I. Gultepe, “Measurements of light rain, drizzle and heavy fog,” in *Precipitation: advances in measurement, estimation and prediction*. Springer, 2008, pp. 59–82.

[39] M. Adams, M. D. Adams, and E. Jose, *Robotic navigation and mapping with radar*. Artech House, 2012.

[40] G. Brooker, R. Hennessey, C. Lobsey, M. Bishop, and E. Widzyk-Capehart, “Seeing through dust and water vapor: Millimeter wave radar sensors for mining applications,” *Journal of Field Robotics*, vol. 24, no. 7, pp. 527–557, 2007.

[41] R. Xu, W. Dong, A. Sharma, and M. Kaess, “Learned depth estimation of 3d imaging radar for indoor mapping,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 13 260–13 267.

[42] R. Gourova, O. Krasnov, and A. Yarovoy, “Analysis of rain clutter detections in commercial 77 ghz automotive radar,” in *European Radar Conference (EURAD)*. IEEE, 2017, pp. 25–28.

[43] P. Radecki, M. Campbell, and K. Matzen, “All weather perception: Joint data association, tracking, and classification for autonomous ground vehicles,” *arXiv preprint arXiv:1605.02196*, 2016.

[44] “Flir. fused aeb with thermal can save lives.” [Online]. Available: https://www.flir.com/globalassets/industrial/oem/adas/flir-thermal-aeb-white-paper---final-v1.pdf

[45] “Research testing on adas autonomous vehicle technologies.” [Online]. Available: https://www.vsi-labs.com/

[46] F. Nobis, M. Geisslinger, M. Weber, J. Betz, and M. Lienkamp, “A deep learning-based radar and camera sensor fusion architecture for object detection,” in *Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. IEEE, 2019, pp. 1–7.

[47] D. Yu, H. Xiong, Q. Xu, J. Wang, and K. Li, “Multi-stage residual fusion network for lidar-camera road detection,” in *IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2019, pp. 2323–2328.

[48] L. Caltagirone, M. Bellone, L. Svensson, and M. Wahde, “Lidar-camera fusion for road detection using fully convolutional neural networks,” *Robotics and Autonomous Systems*, vol. 111, pp. 125–131, 2019.

[49] A. Safa, T. Verbelen, I. Ocket, A. Bourdoux, F. Catthoor, and G. G. Gielen, “Fail-safe human detection for drones using a multi-modal curriculum learning approach,” *IEEE Robotics and Automation Letters*, vol. 7, no. 1, pp. 303–310, 2021.

[50] B. Yang, R. Guo, M. Liang, S. Casas, and R. Urtasun, “Radarnet: Exploiting radar for robust perception of dynamic objects,” in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII 16*. Springer, 2020, pp. 496–512.

[51] M. Ulrich, C. Gläser, and F. Timm, “Deepreflecs: Deep learning for automotive object classification with radar reflections,” in *IEEE Radar Conference (RadarConf21)*. IEEE, 2021, pp. 1–6.

[52] F. Drews, D. Feng, F. Faion, L. Rosenbaum, M. Ulrich, and C. Gläser, “Deepfusion: A robust and modular 3d object detector for lidars, cameras and radars,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 560–567.

[53] Z. Liu, Y. Cai, H. Wang, L. Chen, H. Gao, Y. Jia, and Y. 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*, vol. 23, no. 7, pp. 6640–6653, 2021.

[54] K. Qian, S. Zhu, X. Zhang, and L. E. 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 (CVPR)*, 2021, pp. 444–453.

[55] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, “The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset,” in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 6433–6438.

[56] Y. Yang, J. Liu, T. Huang, Q.-L. Han, G. Ma, and B. Zhu, “Ralibev: Radar and lidar bev fusion learning for anchor box free object detection system,” *arXiv preprint arXiv:2211.06108*, 2022.

[57] N. A. Rawashdeh, J. P. Bos, and N. 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*, vol. 11748. SPIE, 2021, pp. 36–45.

[58] R. T. Tan, “Visibility in bad weather from a single image,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2008, pp. 1–8.

[59] J.-P. Tarel and N. Hautiere, “Fast visibility restoration from a single color or gray level image,” in *IEEE 12th international conference on computer vision*. IEEE, 2009, pp. 2201–2208.

[60] Z. Chen, Y. Wang, Y. Yang, and D. Liu, “Psd: Principled synthetic-to-real dehazing guided by physical priors,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 7180–7189.

[61] X. Zhang, H. Dong, J. Pan, C. Zhu, Y. Tai, C. Wang, J. Li, F. Huang, and F. 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 (CVPR)*, 2021, pp. 9239–9248.

[62] V. Mușat, I. Fursa, P. Newman, F. Cuzzolin, and A. Bradley, "Multi-weather city: Adverse weather stacking for autonomous driving," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 2906–2915.

[63] R. Timofte, S. Gu, J. Wu, and L. Van Gool, "Ntire 2018 challenge on single image super-resolution: Methods and results," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 852–863.

[64] T. Sun, J. Chen, and F. Ng, "Multi-target domain adaptation via unsupervised domain classification for weather invariant object detection," *arXiv preprint arXiv:2103.13970*, 2021.

[65] Z. Zheng, Y. Wu, X. Han, and J. 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*. Springer, 2020, pp. 155–170.

[66] Y. Lee, Y. Ko, Y. Kim, and M. Jeon, "Perception-friendly video enhancement for autonomous driving under adverse weather conditions," in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 7760–7767.

[67] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 3213–3223.

[68] M. Hassaballah, M. A. Kenk, K. Muhammad, and S. Minaee, "Vehicle detection and tracking in adverse weather using a deep learning framework," *IEEE transactions on intelligent transportation systems*, vol. 22, no. 7, pp. 4230–4242, 2020.

[69] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," in *Conference on robot learning*. PMLR, 2017, pp. 1–16.

[70] Z. Yan, L. Sun, T. Krajník, and Y. Ruichek, "Eu long-term dataset with multiple sensors for autonomous driving," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 10 697–10 704.

[71] M. Sheeny, E. De Pellegrin, S. Mukherjee, A. Ahrabian, S. Wang, and A. Wallace, "Radiate: A radar dataset for automotive perception in bad weather," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 1–7.

[72] K. Burnett, D. J. Yoon, Y. Wu, A. Z. Li, H. Zhang, S. Lu, J. Qian, W.-K. Tseng, A. Lambert, K. Y. Leung *et al.*, "Boreas: A multi-season autonomous driving dataset," *The International Journal of Robotics Research*, p. 02783649231160195, 2022.

[73] T. Matuszka, I. Barton, Á. Butykai, P. Hajas, D. Kiss, D. Kovács, S. Kunsági-Máté, P. Lengyel, G. Németh, L. Pető *et al.*, "aimotive dataset: A multimodal dataset for robust autonomous driving with long-range perception," *arXiv preprint arXiv:2211.09445*, 2022.

[74] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13.* Springer, 2014, pp. 740–755.

[75] F. Heide, "Adverse weather fusion supplement," https://www.cs.princeton.edu/~fheide/AdverseWeatherFusion/figures/AdverseWeatherFusion_Supplement.pdf, 2023, accessed: 2023-11-24.

[76] Y. Zhou, L. Liu, H. Zhao, M. López-Benítez, L. Yu, and Y. Yue, "Towards Deep Radar Perception for Autonomous Driving: Datasets, Methods, and Challenges," p. 4208, May 2022. [Online]. Available: https://doi.org/10.3390/s22114208