Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen
International Center for
Information Technology

Master's Thesis

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

ii

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

x

Contents

<div align="right">

1

</div>

Introduction

- 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].
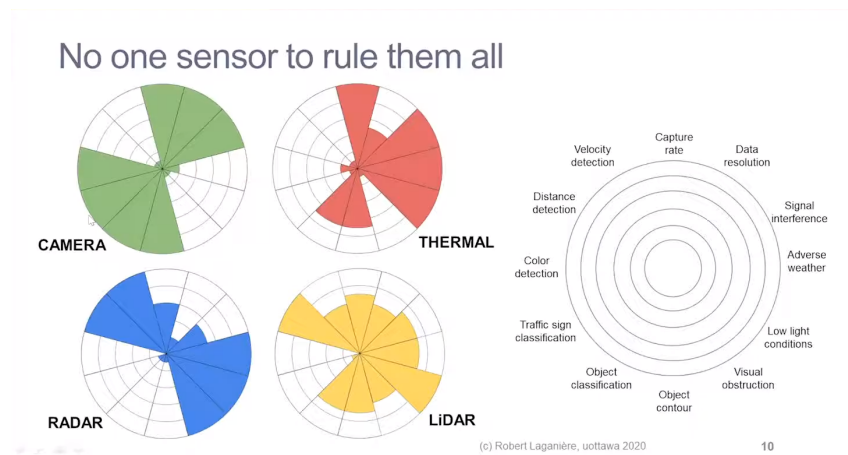

Figure 1.1: Sensors modality characteristics [4]

- 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.

- 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.
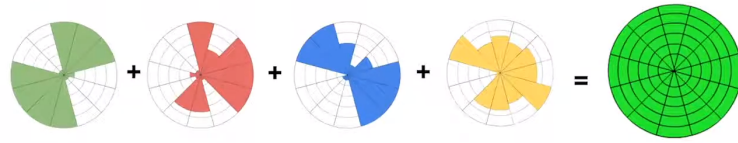
Figure 1.2: Sensors modality characteristics [4]

- 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 objects such as cars, trucks, pedestrians, and cyclists.
    - Adverse weather conditions
        * Refers to conditions such as fog, snow, rain, overcast skies, sleet, and dust.

2

* 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 camera, LIDAR, RADAR, IMU, GPS, and infrared, 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.

1.1 Relevance

* 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.

- In 2019, the global market for sensor fusion used in autonomous vehicles had a value of USD 594.4 million. According to estimates, this market is expected to grow significantly and reach a value of USD 1563.5 million by 2025, reflecting a Compound Annual Growth Rate (CAGR) of 19.6% during the period between 2020 and 2025. According to MarketInsider, autonomous emergency braking (AEB) is expected to drive significant growth in the use of sensor fusion, as AEB systems rely on radar and vision sensors to identify potential collision partners ahead of the ego vehicle. North America is predicted to be one of the largest markets for ADAS-enabled vehicles and self-driven transport solutions [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.

1.2 Motivation

1.2.1 ...

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

1.2.2 ...

1.3 Challenges and Difficulties

1.3.1 ...

Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo.

Aliquam lectus. Vivamus leo. Quisque ornare tellus ullamcorper nulla. Mauris porttitor pharetra tortor. Sed fringilla justo sed mauris. Mauris tellus. Sed non leo. Nullam elementum, magna in cursus

sodales, augue est scelerisque sapien, venenatis congue nulla arcu et pede. Ut suscipit enim vel sapien. Donec congue. Maecenas urna mi, suscipit in, placerat ut, vestibulum ut, massa. Fusce ultrices nulla et nisl.

Etiam ac leo a risus tristique nonummy. Donec dignissim tincidunt nulla. Vestibulum rhoncus molestie odio. Sed lobortis, justo et pretium lobortis, mauris turpis condimentum augue, nec ultricies nibh arcu pretium enim. Nunc purus neque, placerat id, imperdiet sed, pellentesque nec, nisl. Vestibulum imperdiet neque non sem accumsan laoreet. In hac habitasse platea dictumst. Etiam condimentum facilisis libero. Suspendisse in elit quis nisl aliquam dapibus. Pellentesque auctor sapien. Sed egestas sapien nec lectus. Pellentesque vel dui vel neque bibendum viverra. Aliquam porttitor nisl nec pede. Proin mattis libero vel turpis. Donec rutrum mauris et libero. Proin euismod porta felis. Nam lobortis, metus quis elementum commodo, nunc lectus elementum mauris, eget vulputate ligula tellus eu neque. Vivamus eu dolor.

Nulla in ipsum. Praesent eros nulla, congue vitae, euismod ut, commodo a, wisi. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Aenean nonummy magna non leo. Sed felis erat, ullamcorper in, dictum non, ultricies ut, lectus. Proin vel arcu a odio lobortis euismod. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Proin ut est. Aliquam odio. Pellentesque massa turpis, cursus eu, euismod nec, tempor congue, nulla. Duis viverra gravida mauris. Cras tincidunt. Curabitur eros ligula, varius ut, pulvinar in, cursus faucibus, augue.

1.3.2 ...

1.3.3 ...

1.4 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 under 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.

- 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 [13], DENSE [14], and aiMotive [15].

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.

State of the Art

TODO: include MY 3 chosen Methods

2.1 Adverse Weather Conditions Influence on Sensors

- 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 [16–18]. 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 [14].

- 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 [19]. 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 [20]. 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 [21].

- LiDAR is the second most commonly used sensor in autonomous driving systems. Fersch et al. [23] 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. [24] 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 [25] [21].

- 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

Figure 2.1: Van occluded by a water droplet on the lens [22]

(a) Fog chamber

(b) Point clouds in fog

(c) Point clouds in rain

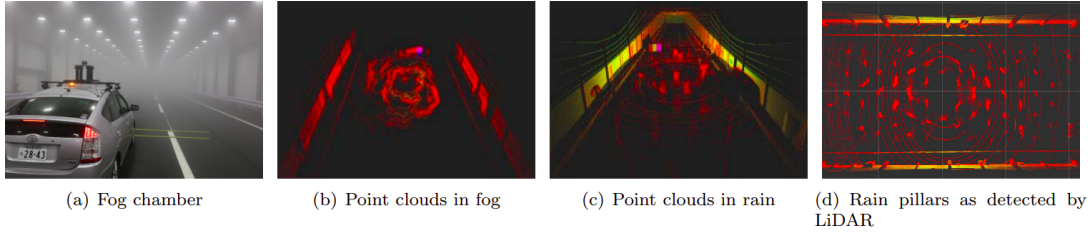(d) Rain pillars as detected by LiDAR

Figure 2.2: LiDAR performance test [25]

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 [13]. As reported by Ijaz et al. [26] and Ismail [27], 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 [20, 28–31] 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.

- 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
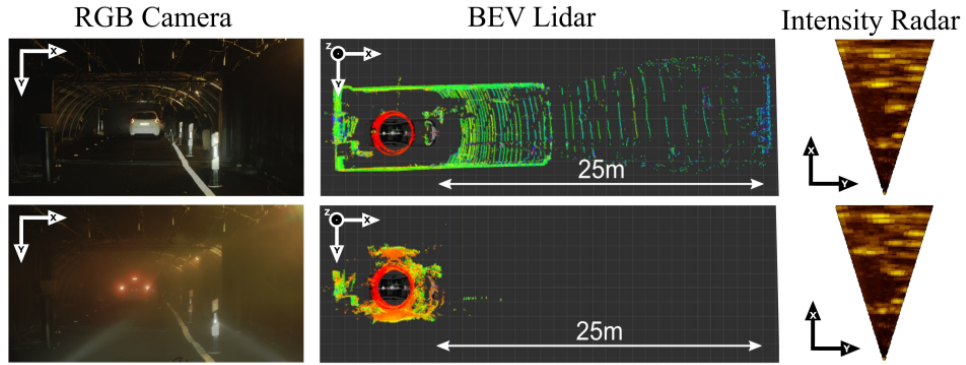
Figure 2.3: 1st row: clear weather condition, 2nd row: with fog. Shows that lidar affects by the fog but radar intensity remains the same [14]

camera, gated camera, stereo camera, weather stations and other weather-related sensors.

- Radecki et al. [32] conducted a comprehensive review, investigating the efficacy of various sensors in a range of weather conditions including wet conditions, day and night, cloudy skies, glare, and dust. They engineered a robust system capable of tracking and classifying objects through a real-time, joint probabilistic perception algorithm. This advanced algorithm employs a dynamic selection of sensor subsets that are particularly suited to the prevailing weather conditions. The system not only elevates the general lower bound of perception ability but also optimizes its robustness and reliability through intelligent weighting of sensors and precise quantification of parameters based on the specific weather context. Such a design emphasizes the crucial role of real-time strategy shifts and intelligent sensor subset selection in maximizing the accuracy and reliability of multimodal perception systems.

- FLIR System Inc. [33] and VSI Labs [34] 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.

- To address the problem of when to fuse the data in the neural network architecture, Nobis et al. [35] proposed a CameraRadarFusionNet (CRF-Net), which was inspired from camera-LiDAR fusion [36] and [37], to learn at which level the fusion of the sensor data was the most beneficial for the detection task. They used nuScenes [16] dataset and released their own TUM dataset. Furthermore, they

introduced a new training strategy to focus the learning on a specific sensor type, which was called BlackIn. For feature fusion, the element-wise addition was adopted as the fusion operation. Their fusion method outperformed the image-only network on both datasets, which again shows the importance of fusing radar data into the detection task.

- Yang et al. [38] 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 [16].

- Bijelic et al. [14] 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.

Figure 2.4: Highlighting the significance of fusing multimodal sensor data [14]

- Bijelic et al. [14] 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 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.

- Liu et al. [39] 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. [40] introduced a Multimodal Vehicle Detection Network (MVDNet) featuring LiDAR and radar. In the network architecture, MVDNet has a two-stage attention block in the fusion module. It first applied self-attention to each modality to extract features and then mixed them with region-wise features through cross attentions. Experiments showed that the fusion mechanism performs robustly in foggy weather. The authors trained and evaluated the model performance on SeeingThroughFog [14] and the Oxford Radar Robotcar [41] datasets. And the evaluation shows much better performance than LiDAR alone in fog conditions.

- Rawashdeh et al. [42] developed a CNN-based sensor fusion approach for detecting drivable paths using cameras, LiDAR, and radar, which was evaluated using the DENSE dataset [14]. 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.

2.2 Synthetic Data for Adverse Weather Conditions

- There are studies out there that use de-hazing techniques to remove the bad effects from adverse weather. While physical priors were previously used [43, 44], 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. [45] found that models trained on synthetic images do not generalize well to real-world hazy images, while Zhang et al. [46] 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 [47, 48].

- Few researchers [49] [50] [51] have also explored synthetic data generation for adverse weather conditions using GAN-based techniques from clean weather dataset eg. KITTI [52], Cityscapes [53], 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 [54].

2.3 Simulation

- 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 [55] 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 [25].

2.4 Available Dataset

TODO: remove this from here

- 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 [13] 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 2.1: List multimodal datasets with adverse weather conditions

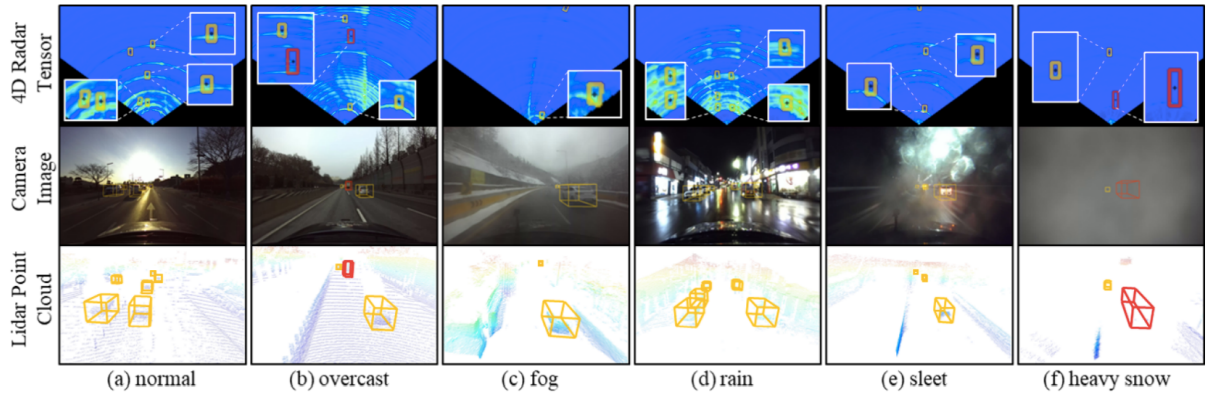| Name	Sensors	Reference	Year
DENSE	CRLNF	[14]	2020
EU Long-term	CRL	[56]	2020
nuScenes	CRL	[16]	2020
The Oxford RobotCar	CRL	[41]	2020
RADIATE	CRL	[57]	2021
K-Radar	CRL	[13]	2022
aiMotive	CRL	[58]	2022
Boreas	CRL	[59]	2022
WADS	CRLNF	[60]	2023

Figure 2.5: Samples of K-Radar datasets for various weather conditions [13]

2.5 Limitations of previous work

- Limitations of Radecki et al.'s [32] study include an emphasis on optimal weather conditions and a lack of investigation into heavy traffic or urban areas, as well as deep learning-based fusion techniques. Future research should consider these limitations and explore the use of proposed Similarly-based methods for object detection in adverse weather datasets.

- The work done by FLIR System Inc. [33] and the VSI Labs [34] 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 [20].

- Nobis et al. [35] 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 [16] 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.

- The study conducted by Bijelic et al. [14] 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 [25]. 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. [40] 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 [62].

- Rawashdeh et al. [42] proposed a multimodal fusion approach for drivable path detection in poor weather conditions. However, the proposed method lacks a comparison with other state-of-the-art methods and does not investigate the real-time processing requirements and computational cost of the proposed algorithm, which could limit its practical applicability [42].

- 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 [63] 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. [63], 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.

16

- Temporal information is a crucial aspect of sensor fusion, but very few multimodal fusion algorithms have been developed to handle this type of data. [14]

- There is a dearth of work in the literature utilizing 4D imaging radar sensors, which is a promising area for future research. [64]

3

Methodology

How you are planning to test/compare/evaluate your research. Criteria used.

3.1 Setup

3.2 Experimental Design

3.3 Evaluation Metrics

4

Solution

Your main contributions go here

4.1 Proposed algorithm

4.2 Implementation details

5

Evaluation and Results

5.1 Experiment Description

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

5.2 Experimental Setup

Describe your experimental setup in detail.

5.3 Results

Describe the results of your experiments in detail.

6
Conclusions

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

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

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

[7] 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.

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

[9] "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

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

[11] 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).

[12] 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

[13] 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

[14] 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.

[15] T. Matuszka, I. Barton, Á. Butykai, P. Hajas, D. Kiss, D. Kovács, S. Kunsági-Máté, P. Lengyel, G. Németh, L. Pető, D. Ribli, D. Szeghy, S. Vajna, and B. Varga, "aiMotive Dataset: A Multimodal Dataset for Robust Autonomous Driving with Long-Range Perception," Nov. 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2211.09445

[16] 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.

[17] 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.

[18] 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.

[19] 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/

[20] 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.

[21] 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).

[22] 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

[23] 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.

[24] 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.

[25] Y. Zhang, A. Carballo, H. Yang, and K. Takeda, "Autonomous Driving in Adverse Weather Conditions: A Survey," Dec. 2021. [Online]. Available: https://doi.org/10.48550/arXiv.2112.08936

[26] 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.

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

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

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] "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

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

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] 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.

[44] 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.

[45] 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.

[46] 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.

[47] 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.

[48] 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.

[49] 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.

[50] 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.

[51] 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.

[52] 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.

[53] 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.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] A. Kurup and J. 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*, vol. 12115. SPIE, 2022, pp. 146–152.

[61] 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.

[62] 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.

[63] 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.

[64] 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