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

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].

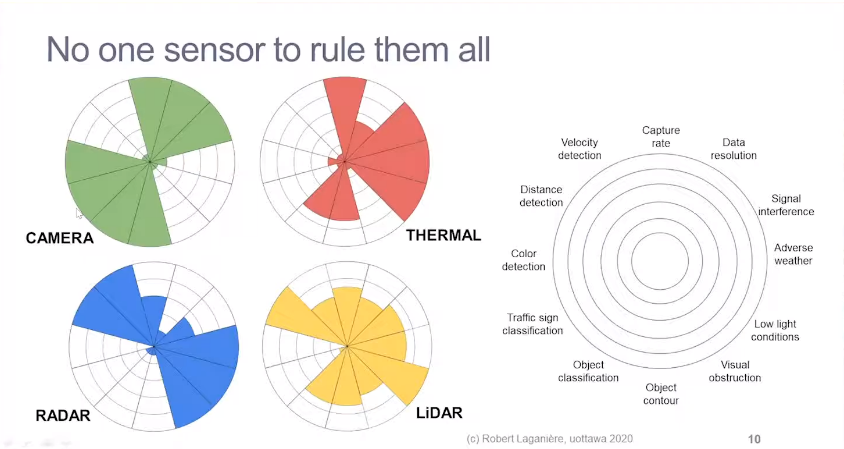


Figure 1.1: Sensors modality characteristics [4]

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.

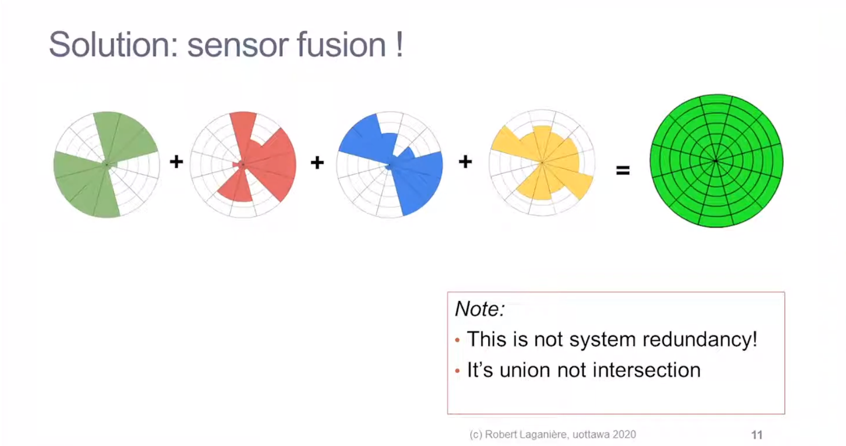


Figure 1.2: Sensors modality characteristics [4]

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.

The project will also delve into specific aspects of object detection, such as identifying various entities like cars, trucks, pedestrians, and cyclists, and will focus on detection in adverse weather conditions like fog, snow, and rain, which pose significant challenges to visibility. The 'tightly-coupled' approach refers to the intricate integration of different data modalities, combining features at various levels for optimal results. The 'data-driven' aspect emphasizes leveraging existing datasets to enhance performance, 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 [5] and the US Department of Commerce [6]. 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. A 2018 report by the National Highway Traffic Safety Administration (NHTSA) further underscores that adverse weather conditions were involved in 4,000 fatal crashes [7].

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) [8]. This drastic reduction in visibility poses a major challenge for effective obstacle or object detection, a critical component of road safety in adverse conditions.

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 [9]. 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 [10]. The precision agriculture and environmental monitoring markets are anticipated to reach $10.5 billion by 2026 [11], while the aerospace and defense sector is projected to reach $23.83 billion by 2027 [12]. 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 [13], ApolloScape [14], and Waymo [15]. There are datasets that incorporate radar sensors, such as VoD [16], TJ4DRadSet [17], and PixSet [18], 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: As identified by Feng et al. [19], there is an absence of a universally accepted framework or guidelines for designing network architectures in multimodal sensor fusion. This leads to unanswered questions regarding the optimal choices for 'what to fuse' (e.g., LiDAR, radar, various camera types), 'how to fuse' (such as through addition, concatenation, or more complex methods), and 'when to fuse' (determining at which stage of the processing to integrate the different modalities).

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 started including baseline models with simple fusion methods, there is a potential for significant performance improvement by employing advanced fusion techniques, such as transformer-based or tightly-coupled fusion. 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.

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 [20]. Given the potential of these sensors in challenging environments, further exploration in this area is essential.

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 develop 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.

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 pinpoint an optimal fusion strategy, a tactic that harmonizes the distinct advantages 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 [21] and nuScenes [22]. 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

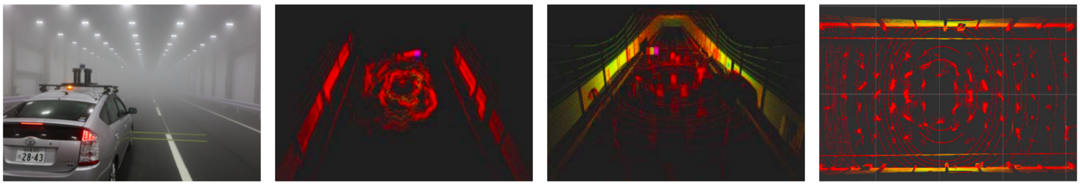
- In the realm of autonomous robots, particularly self-driving vehicles and autonomous drones, object detection has emerged as a critical computer vision problem. These applications demand accurate 2D or 3D bounding boxes for objects in complex real-world scenarios, which often include cluttered scenes, 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 [22–24]. 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 [21].

- 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 [25] as shown in the Figure 2.1. 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 [26]. 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 [27].

- LiDAR is the second most commonly used sensor in autonomous driving systems. Fersch et al. [29] 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. [30] 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 [31] [27]. Figure 2.2 shows an example of LiDAR performance in fog conditions where it creates small false obstacle clouds.

- 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 [28]



(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 [31]

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 [32]. As reported by Ijaz et al. [33] and Ismail [34], 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 [26, 35–38] 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. The same can be seen as discussed before from the Figure 1.1 and 1.2. As a result, groups from all over the world come up with their

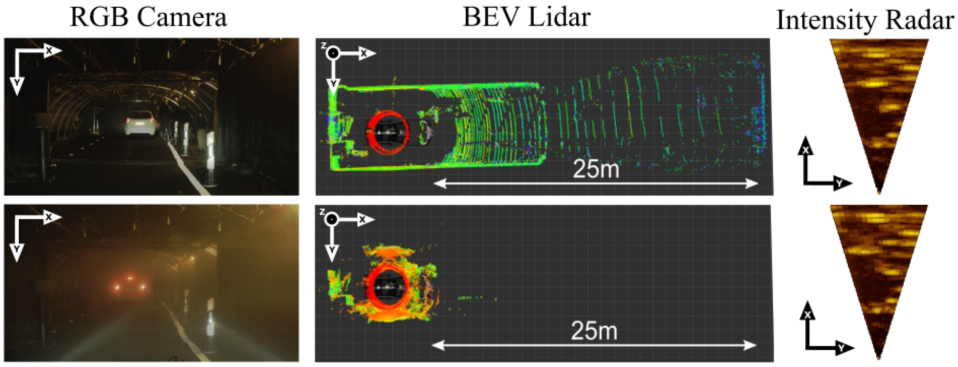


Figure 2.3: 1st row: clear weather condition, 2nd row: with fog. Shows that lidar affects by the fog but radar intensity remains the same [21]

own permutation and combination with camera, LiDAR, radar, infrared camera, gated camera, stereo camera, weather stations and other weather-related sensors.

2.2 Multimodal Sensor Fusion

- Radecki et al. [39] 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.

- Limitations of Radecki et al.'s [39] 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.

- FLIR System Inc. [40] and VSI Labs [41] 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.

- The work done by FLIR System Inc. [40] and the VSI Labs [41] 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 [26].

- To address the problem of when to fuse the data in the neural network architecture, Nobis et al. [42] proposed a CameraRadarFusionNet (CRF-Net), which was inspired from camera-LiDAR fusion [43] and [44], to learn at which level the fusion of the sensor data was the most beneficial for the detection task. They used nuScenes [22] 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.

- Nobis et al. [42] 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 [22] 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. [45], the performance could be further improved by pre-processing the radar data before fusion.

- Yang et al. [46] 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 [22].

- In the RadarNet architecture, proposed by Yang et al. [46], the radar sensor data used from the nuScenes dataset, which has a very low resolution and hence it's not a good choice for studying the role of radar in perception. Object detection using radars is limited by low resolution and erroneous elevation estimates [47] [48]. Therefore, one possible improvement is to include the radar sensor used in K-Radar dataset, which is a 4D radar with elevation angle of 30 degrees compare to 1 or 2 degrees for conventional radars, can be used to improve the performance of the RadarNet architecture.

- Bijelic et al. [21] 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 [21]

- Bijelic et al. [21] 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.

- The study conducted by Bijelic et al. [21] 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 [31]. 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.

- Liu et al. [49] 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. [50] 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 [21] and the Oxford Radar Robotcar [51] datasets. And the evaluation shows much better performance than LiDAR alone in fog conditions.

- Qian et al. [50] 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 [52].

- Rawashdeh et al. [53] developed a CNN-based sensor fusion approach for detecting drivable paths using cameras, LiDAR, and radar, which was evaluated using the DENSE dataset [21]. 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.

- Rawashdeh et al. [53] 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 [53].

- The paper by Chang et al. [54] introduces a novel method for enhancing obstacle detection in autonomous driving systems. This method, called spatial attention fusion (SAF), effectively integrates data from millimeter-wave (mmWave) radar and camera sensors. SAF addresses the sparsity of radar points by generating an attention weight matrix that distinctively fuses vision features, diverging from traditional concatenation or element-wise addition fusion methods. This method can be integrated into the feature-extraction stage of existing deep learning object detection frameworks, facilitating end-to-end training. Additionally, the paper presents a generation model that converts radar points into images for neural network training. This type of image projection called radar imagery. The paper's findings indicate that this fusion approach significantly improves performance on nuScenes [22] dataset.

- Another work by Broedermann et al. [55] presents an extended work on HRNet [56] and HRFormer [57] to integrate multimodal sensors into a single network. It introduces HRFuser, a versatile, multi-resolution, multi-sensor fusion architecture that can efficiently integrate an arbitrary number of sensors like lidar, radar, and gated cameras, alongside standard cameras. HRFuser is built on the HRNet and HRFormer paradigms, preserving high-resolution representations and incorporating a novel multi-window cross-attention (MWCA) block for effective fusion across multiple resolutions. The system's generic design allows for easy scalability with various sensors without the need for specialized components for each sensor. Extensive testing on major autonomous driving datasets, including nuScenes [22], and SeeingThroughFog [21], demonstrates HRFuser's superior performance over existing camera-only networks and sensor fusion methods, proving its efficacy in both standard and adverse weather conditions. HRFuser's adaptability to different sensor sets and its ability to selectively focus on relevant features from high-resolution data of additional sensors mark a significant advancement in the field of 2D object detection.

- Recent work by Chu et al. [58] proposes a novel end-to-end multimodal multistage object detection network called MT-DETR (MulTi-sensor MulTimodal DTtection TRansformer) that leverages data from multiple sensors - camera, lidar, radar and time - to achieve robust detection, especially in adverse weather conditions. Here time modality is an additional binary image input to inform model about day or night. It employs specialized fusion modules - Residual Fusion Module (RFM) and Confidence Fusion Module (CFM) for hierarchical cross-modal feature fusion. The network also uses a Residual Enhancement Module (REM) to strengthen individual sensor branches. A multi-stage loss function further regularizes feature learning across modalities. Extensive experiments on the publicly available SeeingThroughFog [21] dataset demonstrate that MT-DETR significantly outperforms existing unimodal and multimodal detection methods. Notably, when additionally trained on realistic synthetic foggy data generated by a novel camera-lidar synthesis algorithm, the performance boost is even higher.

  TODO: write about my choosen methods limitations

- 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 [19] 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, infrared 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. [19], 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 tightly-coupled fusion, has the potential to improve their performance.
- Temporal information is a crucial aspect of sensor fusion, but very few multimodal fusion algorithms have been developed to handle this type of data [21]. Although this project does not consider the temporal information in the fusion process, it is still a promising area for future research.
- There is a dearth of work in the literature utilizing 4D imaging radar sensor especially for adverse weather conditions, which is a promising area for future research. [20]

TODO: try to accomodate these two sections in the above section

2.3 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 [59,60], 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. [61] found that models trained on synthetic images do not generalize well to real-world hazy images, while Zhang et al. [62] 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 [63, 64].
- Few researchers [65] [66] [67] have also explored synthetic data generation for adverse weather conditions using GAN-based techniques from clean weather dataset eg. KITTI [13], Cityscapes [68], 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 [69].

2.4 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 [70] 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 [31].

2.5 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 [32] 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	[21]	2020
EU Long-term	CRL	[71]	2020
nuScenes	CRL	[22]	2020
The Oxford RobotCar	CRL	[51]	2020
RADIATE	CRL	[72]	2021
K-Radar	CRL	[32]	2022
aiMotive	CRL	[73]	2022
Boreas	CRL	[74]	2022
WADS	CRLNF	[75]	2023

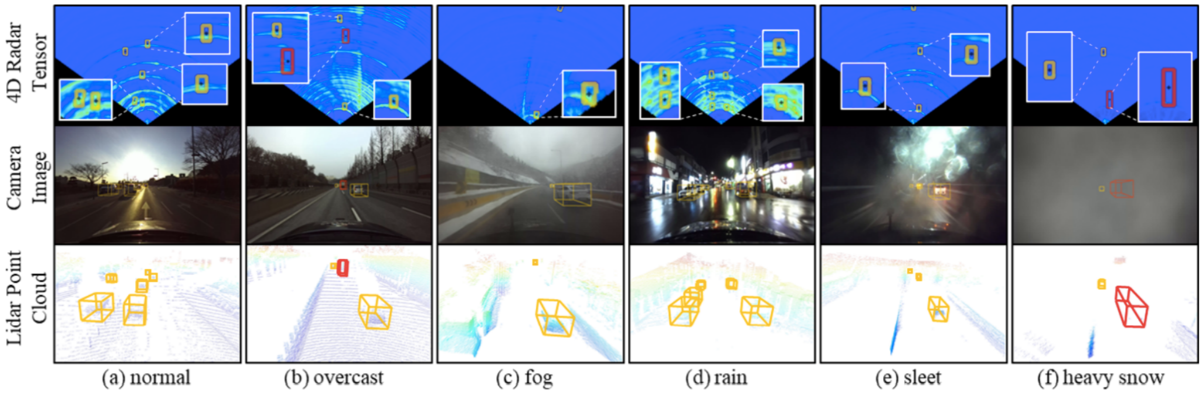


Figure 2.5: Samples of K-Radar datasets for various weather conditions [32]

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

6.1 Contributions

6.2 Lessons learned

6.3 Future work

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] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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/

[26] 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.

[27] 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).

[28] 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

[29] 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.

[30] 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.

[31] 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

[32] 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

[33] 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.

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

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

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] “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

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

[42] 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.

[43] 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.

[44] 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.

[45] 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.

[46] 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.

[47] 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.

[48] 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.

[49] 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.

[50] 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.

[51] 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.

[52] 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.

[53] 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.

[54] S. Chang, Y. Zhang, F. Zhang, X. Zhao, S. Huang, Z. Feng, and Z. Wei, "Spatial attention fusion for obstacle detection using mmwave radar and vision sensor," *Sensors*, vol. 20, no. 4, p. 956, 2020.

[55] T. Broedermann, C. Sakaridis, D. Dai, and L. Van Gool, "Hrfuser: A multi-resolution sensor fusion architecture for 2d object detection," *arXiv preprint arXiv:2206.15157*, 2022.

[56] J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, Y. Zhao, D. Liu, Y. Mu, M. Tan, X. Wang *et al.*, "Deep high-resolution representation learning for visual recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 10, pp. 3349–3364, 2020.

[57] Y. Yuan, R. Fu, L. Huang, W. Lin, C. Zhang, X. Chen, and J. Wang, "Hrformer: High-resolution vision transformer for dense predict," *Advances in Neural Information Processing Systems*, vol. 34, pp. 7281–7293, 2021.

[58] S.-Y. Chu and M.-S. Lee, "Mt-detr: Robust end-to-end multimodal detection with confidence fusion," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 5252–5261.

[59] 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.

[60] 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.

[61] 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.

[62] 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.

[63] 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.

[64] 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.

[65] 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.

[66] 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.

[67] 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.

[68] 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.

[69] 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.

[70] 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.

[71] 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.

[72] 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.

[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] 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.

[75] 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.