D2.1

Report on Noise Factor Analysis

June 2025

Document Information

Name	Organisation	Contribution
Milan Lovric, Daniel Gummadi, Qasim Sadiq, Pradyot Pendyala, Rufus Eade, Hetian Wang, Valentina Donzella	WMG, University of Warwick	Report layout, sensor testing standards report, noise factors workshop summary report, Winter Testing Campaign data analysis, methodology for sensor noise factor testing in the lab
Imran Mohamed, Hsun Yang, Fengping Li	National Physical Laboratory (NPL)	Winter Testing Campaign
Mike Dempsey	Claytex	Winter Testing Campaign

Abbreviations and Acronyms

ACC	Adaptive Cruise Control
ADAS	Advanced Driver Assistance System
ADS	Automated Driving System
AEB	Automatic Emergency Braking
AV	Automated Vehicle
CFA	Colour Filter Array
CMS	Camera Monitor Systems
CSAC	Compound Semiconductor Applications Catapult
DUT	Device Under Test
FCW	Forward Collision Warning
FOV	Field of View
HDR	High Dynamic Range
IQ	Image Quality
IQA	Image Quality Assessment
ISP	Image Signal Processor
LDW	Lane Departure Warning
MTF	Modulation Transfer Function
NPL	National Physical Laboratory
OECF	Opto-Electronic Conversion Function
OEM	Original Equipment Manufacturer
PCW	Pedestrian Collision Warning
RGB	Red, Green, Blue
RMS	Root Mean Square
SNR	Signal-to-Noise Ratio
WMG	Warwick Manufacturing Group

Contents

Document Information 2

Abbreviations and Acronyms 3

Contents 4

1. Sensor Noise Factors Testing Report 8

1.1. Introduction 8

2. Camera Testing Methods 9

2.1. Requirements and Use Cases in Automotive 9

2.2. Camera Testing Standards 10

- ISO 12233:2023(en): 10
- ISO 12232:2019(en): 10
- ISO 15739:2023(en): 11
- ISO 17850:2015(en): 11
- ISO 17957:2015: 12
- ISO 18844:2017: 12
- ISO 19084:2015: 13
- ISO/TS 19567-1:2016: 13
- ISO 16505:2019: 14
- IEEE P2020 Automotive System Image Quality 14

2.3. Camera Performance Testing 16

2.4. Camera Benchmarking Testing 16

2.5. Current Limitations in Automotive Camera Testing Standards 17

3. LiDAR Testing Methods 19

3.1. Requirements and Use Cases in Automotive 19

3.2. LiDAR Sensor Performance Tests 20

- FoV and Angular Resolution Tests 20
- Various Environmental Conditions Test 22
- Adverse Weather Tests 23
- LiDAR Sensor Benchmark Tests 24

3.3. LiDAR Sensor Testing Standards 25

4. Radar Testing Methods 27

4.1. Radar Testing Standards - Use Case 27
- EN 303 396 - V1.1.1 27
- EN 302 858 V2.1.1 27
- EN 301 091-1 V2.1.1 28
- ITU-R M.2057-1....28

4.2. Radar Testing Standards - System Characteristics 28
- ETSI EN 303 396 V1.1.1....28
- ETSI EN 301 091-1 V2.1.1....29
- ETSI EN 302 858 V2.1.1....30
- ITU-R M.2057-1....31

4.3. Radar Testing Standards - Methods of Measurement 31
- ETSI EN 303 396 V1.1.1....31

5. Sensor Fusion Testing 33

6. Perception Sensor Standards for Object Detection Applications 35

6.1. BS EN ISO 16001:2017 Earth-moving machinery. Object detection systems and visibility aids. Performance requirements and tests 35

6.2. ISO 17386:2023 Intelligent transport systems. Manoeuvring aids for low-speed operation (MALSO) . Performance requirements and test procedures. 36

7. Euro NCAP Protocols....38

7.1. NCAP AEB VRU Systems Test Protocol 38

7.2. NCAP VRU Test Protocol....39

7.3. NCAP AEB Car-to-Car Systems Test Protocol....39

7.4. NCAP Lane Support Systems Test Protocol 40

8. Sim4CAMSens Workshop Summary 41

Participants and Advisory Board....41

8.1. Summary of the Workshop....42

8.2. Outcomes of the Workshop 48
- All Sensors (Summary of Most Important Noise Factors) 49
- Noise factors that could be tested by the Consortium (Most Important Noise Factors) 50
- What noise factors can be simulated? (Most Important Noise Factors) 51

9. Winter Testing Campaign....52

9.1. Site Setup....52

9.2. Weather Stations....54

9.3. Data Logging 55
9.4. Robustness Analysis by Quantifying Sensor Failures 57
- Automotive LiDAR A 57
- Automotive Visible Camera 58
9.5. Qualitative Analysis of Inclement Weather and Other Noise Factors 59
- Automotive Visible Camera 59
- Automotive LiDAR B 64
10. Quantitative Analysis of Winter Testing Campaign Data 66
10.1. Image Quality Assessment (IQA) 67
10.2. Automotive Visible Camera Analysis 68
- Average Visuals (FR & NR IQA Metrics) 68
- Cluster Visuals (FR & NR IQA Metrics) 71
10.3. LiDAR Point Cloud Analysis 74
- Point Cloud Quality Assessment (PCQA) 74
11. Insights from the 2nd Winter Testing Campaign 77
11.1. Image Quality Assessment for Visible Camera 77
11.2. Traffic Sign Recognition from Visible Camera 79
11.3. LiDAR Blooming Effect 85
12. Sensor Noise Factor Testing in the Lab 86
12.1. Development of Novel Methodologies 86
12.1. Equipment and Experimental Setup 87
12.1. Generating Water Droplets on LiDAR Sensor 89
12.2. Generating Mud Occlusion on LiDAR Sensor 90
13. Conclusions 92
References 95
Appendix I – Image Datasets for IQA 101
13.1. Automotive Visible Camera 101
- Visible Camera - Dataset A 101
- Visible Camera - Dataset B 102
- Visible Camera – Dataset C 103
- Visible Camera – Dataset D 104
- Visible Camera – Dataset E 105
- Visible Camera – Dataset F 106

1. Sensor Noise Factors Testing Report

1.1. Introduction

Sensors including cameras, LiDARs (Light Detection and Ranging), and RADARs (Radio Detection and Ranging) are crucial for automated vehicles (AVs) and advanced driver assistance systems (ADAS). They act as the vehicle's eyes and ears, providing essential data for perception, navigation, and decision-making. **Cameras** capture visual data through images and videos, offering colour and high-resolution imagery. They help AVs recognise and understand the environment by identifying pedestrians, vehicles, traffic signs, and lane markings. Cameras also monitor lane markings for lane-keeping and alert drivers if they drift unintentionally. However, they can struggle in adverse weather and low-light situations. **LiDARs** emit laser pulses to measure the time it takes for light to return, providing detailed 3D point cloud data for precise distance and depth information. Effective in low-light conditions, LiDARs excel at object detection, aiding collision avoidance and precise localisation. Yet, they can be costly and sensitive to environmental factors like rain or snow. **Radars** use radio waves to detect objects, measuring speed and distance reliably in various weather conditions. They help maintain safe following distances, detect blind spot vehicles, and assist in collision mitigation. RADARs are less affected by adverse weather and can operate in fog, rain, and snow, though they offer less detailed spatial information compared to LiDARs and cameras.

These sensors are often combined/fused to form a comprehensive perception system for AVs, along with ultrasonic sensors and GPS. This fusion allows vehicles to understand their surroundings, make informed decisions, and navigate safely. In ADAS applications, these sensors enable features such as adaptive cruise control, lane-keeping assistance, and emergency braking, enhancing both safety and convenience.

The Importance of Sensor Testing. Testing cameras, LiDARs, and RADARs is crucial for safe and effective AVs and ADAS. Sensors are the backbone of AV decision-making, and any errors can lead to accidents. Rigorous testing ensures sensors work in all weather and lighting conditions, accurately detect pedestrians and objects, and create precise maps. Testing also verifies the reliability of redundant sensor systems and data fusion for features like adaptive cruise control. Compliance with safety standards through thorough testing builds consumer trust and paves the way for widespread AV adoption.

The Scope of this Report. This report provides a review of existing automotive sensor testing methods and standards from the literature. More specifically, testing standards and methods for automotive cameras (Section 2), LiDARs (Section 3), and RADARs (Section 4). Furthermore, the report covers testing for sensor fusion suites comprised of the aforementioned sensors (Section 5), as well as other testing standards related to perception systems, such as Object Detection (Section 6), and Euro NCAP testing protocols (Section 7). Section 8 includes the summary of the Sim4CAMSens Workshop, held at WMG, University of Warwick. Sections 9-11 present the analysis of different types of sensors data collected during the Winter Testing Campaigns. Section 12 presents the development of novel methodologies for lab testing of sensor noise factors, and Section 13 concludes this report with a summary of key takeaways.

2. Camera Testing Methods

Camera technologies play a pivotal role in modern automotive applications, particularly in the context of automated vehicles (AVs). These technologies have evolved significantly over the years and are now integral components for enhancing vehicle safety, navigation, and decision-making capabilities. In this introduction, we will explore the key aspects of camera technologies and their suitability for automotive, specifically AV, applications.

Automotive cameras come in various types, including monocular (single-lens) and stereo (multi-lens) cameras. Monocular cameras are commonly used for tasks like lane detection and traffic sign recognition, while stereo cameras enable depth perception and object tracking. High-resolution cameras are crucial for capturing detailed information from the environment. These cameras provide clearer images, which are essential for object recognition, lane-keeping, and obstacle avoidance. Features such as frame rates are important aspects of camera technology for automotive use cases. Fast frame rates are essential for capturing real-time information. AVs require cameras with high frame rates to process data quickly and make split-second decisions. Hide Dynamic Range cameras can capture details in both bright and dark areas of the scene simultaneously. This is critical for driving conditions with varying lighting, such as transitioning from tunnels to bright daylight. Lenses and optics are camera accessories that can minimise distortion, increase clarity, and reduce image aberrations. Wide-angle and fish-eye lenses are used for panoramic views and monitoring blind spots. However, the effectiveness and safety of these camera systems rely heavily on rigorous testing methods. Camera testing methods for automotive use cases have become indispensable in ensuring the reliability and performance of these systems. The stakes are high, as cameras play a pivotal role in enhancing road safety, reducing accidents, and advancing the goal of fully autonomous vehicles. As a result, manufacturers, regulators, and consumers alike are increasingly recognising the need for comprehensive testing procedures to assess the functionality and accuracy of automotive cameras.

2.1. Requirements and Use Cases in Automotive

Since driving is largely a visual task, the sensor that has become ubiquitous within vehicles with Advanced Driver Assistance Systems (ADAS) systems are cameras. The use-cases for cameras in the automotive context are very broad. Punke et al. in their book provide several use-cases for automotive grade cameras [1]. They mention that cameras are used both within the vehicle's interior and exterior. For the interior, use-cases of cameras include: Driver state monitoring, Gaze control, Hand Gesture Recognition and Driver and Interior detection. External use-cases for cameras include: high beam detection, traffic sign recognition, lane detection, object recognition, parking assist, etc.

Punke et al. also provide some design requirements/criteria for automotive grade cameras. They mention that the camera FoV plays an important role in the various use-cases for cameras and is defined by the lens and the image sensor. They mention that for front-view cameras, a horizontal field of view (FoV) greater than 40° is preferred as this FoV is optimal for tasks such

as object detection and lane detection. In general, they mention that specific FoVs are dependent on the use-case that the camera is being applied to. Another design requirement they mention is the resolution of the cameras. The authors state that, in general, usual values for the resolution are >15 pixels per degree in the case of driver assistance functions in the field of environmental detection. This is crucial as often objects need to be resolved at distances at great as 100 m away from the ego-vehicle. Another crucial design requirement mentioned by Punke et al. is the accuracy of the colour reproduction of the camera. They state that the colour reproduction is necessary for automotive grade cameras to be used in use-cases such as lane tracking where the camera will need to be able to distinguish between different lane markings which can be coloured differently. Lastly, the authors mention requirements for the cameras dynamic range. They specify that the dynamic range need to be large to prevent adverse noise artifacts, such as glare and ghosting from appearing on an image.

2.2. Camera Testing Standards

▪ ISO 12233:2023(en):

An essential trait of a digital camera is its spatial resolution capability, which is the ability of a camera to optically capture finely spaced detail. To enable the validation and comparison of spatial resolution measurements, there exist standards for measuring resolution, namely ISO 12233:2023(en) [2]. Its primary objective is to provide a clear definition and clarify the significance of terminology within the context of spatial resolution, the creation of test charts, and the test methods involved in executing precise resolution assessments for both analogue and digital cameras. The authors of this standard provide a basic pipeline for conducting tests to determine a camera systems visual resolution. The general process of this test pipeline consists of obtaining images of a specially designed test chart, and then calculating metrics that measure the spatial resolution of the camera by analysing these images. Some metrics outlined by the standard include: SFR, MTF, PSF, SFR10 Frequency, etc. The Standard also specifies what types of test charts are to be used and for what purpose. The authors of the standard state that a camera's resolution is affected by several factors. These include, but are not limited to, the performance of the camera lens, the number of addressable photoelements in the optical imaging device, and the camera image processing, which can include image sharpening, image compression and gamma correction functions. Due to this, the standard includes sections detailing the setup and conditions of the test, in consideration of these factors. The factors they specifically consider during the test setup are: camera framing and lens focal length setting, camera focusing, camera ISP settings, white balance, luminance, and colour measurements as well as gamma correction. This standard also provides methodologies for calculating both edge based and sine wave based frequency responses to measure camera resolution.

▪ ISO 12232:2019(en):

Another standard for digital camera tests is the standard called ISO 12232:2019(en) [3]. The Standard notes that an image from a digital camera is obtained using an insufficient exposure

will contain an unacceptable amount of noise. Conversely if the exposure is increased excessively the image will result in image highlights being clipped to form a uniformly bright area which occludes the objects of interest in the image. Hence, the authors of the standard note that obtaining a measure of the cameras' best exposure settings is necessary to obtain a high quality image. This standard has devised test for four specific metrics/measures that can be used to determine if a camera system has been setup correctly for the capture conditions required in obtaining well-exposed images. The standard defines the first metric, which is called the exposure index. It details what this metric is and how to calculate it. Next the standard specifies how an exposure test needs to be setup. It details the general test conditions, the level of illumination required for a valid test, the required ambient temperature and humidity as well as other camera settings that need to be controlled such as white balance and compression. The standard describes methodologies involving the last three metrics, namely ISO speed, standard output sensitivity and the recommended exposure index. For each metric the standard outlines methods to calculate them as well as tests that can be conducted using these metrics to determine the required exposure settings in a camera for capturing high quality images.

- ISO 15739:2023(en):

There are also standards that have been developed for measuring and testing for noise within digital cameras. One such standard is ISO 15739:2023(en) [4]. This standard was developed to measure and test for the following primary sources of noise, namely: photon shot noise, dark current shot noise, analogue processing readout noise of image sensors, and quantization noise of A/D converters. The standard stresses that if noticeable levels of noise are present in an image captured by a digital camera, then the necessary detail of objects are lost, resulting in the degradation of the image. Therefore, measurement methods for noise are very important. The authors of the standard specify that certain factors affect the level of noise present during a test, and hence they need to be controlled when performing a test. The factors mentioned in the standard are: illumination, temperature, humidity, white balance, compression and the photosite integration time. To measure the different forms of noise, the standard details several metrics to be used, namely different types of signal-to-noise ratios and dynamic ranges. The Standard also mentions different methods of performing a noise measurement test. These include measurements using a test chart, measurements using a camera with exposure control and measurements using a camera with a removable lens.

- ISO 17850:2015(en):

This is another noise based testing standard for digital cameras, but this standard focuses on extrinsic sources of noise, specifically geometric distortion [5]. The standard notes that geometric distortion is caused by the by the variation of magnification in the image field of the camera lens, which causes the proportion of objects in the resulting image to no be preserved and be representative of the real scene. This is a major noise factor that adversely affects the performance of connected and automated vehicles, as noted by Li et al., specifically in tasks such as object detection [6]. ISO 17850:2015(en) itself provides and explains how to conduct measurement methods for both local and line distortion measures. Moreover, the standard

details the testing setup required to perform distortion measurements, which includes the apparatus and hardware requirements, lighting requirements, the required charts (dot and grid), as well as the camera settings. Lastly, standard details the test procedure under the specified test setup to obtain both local and line geometric distortion measurements. It is worth noting that the standard does not mention what level of geometric distortion is acceptable for automotive applications, rather it simply provides a standardised process to measure geometric distortion.

- ISO 17957:2015:

This standard was developed by the International Organization for Standardisation (ISO) under the Technical Committee 42 (TC 42) for Photography [7]. It focuses on digital cameras and specifically deals with shading measurements. Shading, in the context of photography, refers to variations in image brightness or illumination across the frame, which can result in uneven or non-uniform exposure. ISO 17957 establishes a standardised method for measuring and assessing shading in digital camera images. It is intended to provide a consistent and repeatable method for evaluating how evenly light is distributed across the image sensor and, consequently, across the final image. The purpose of this standard is to ensure that digital cameras produce images with minimal shading artifacts. Shading artifacts can be problematic in photography, especially when capturing scenes with uniform lighting or when post-processing requires even illumination. By establishing a standardised testing procedure, the standard aims to help manufacturers assess and improve their camera systems. ISO 17957 specifies a detailed procedure for measuring shading in digital camera images. This procedure involves capturing a test image with the camera in question, which typically includes a uniform, evenly illuminated target. The captured image is then analysed to quantify any deviations in brightness across the frame. The standard outlines how to report the results of the shading measurements, including the format for expressing shading data. This information can be useful for camera manufacturers to assess and compare the performance of different camera models. Manufacturers can use this standard to ensure that their cameras meet certain shading performance criteria before they are released to the market. This standard can be used for calibration and could be used to profile cameras to correct for shading artifacts during post-processing. Shading measurements can also be used to develop software algorithms that correct for uneven illumination in images, improving overall image quality.

- ISO 18844:2017:

Developed by ISO TC 42, it focuses on digital cameras and specifically deals with image flare measurement [8]. Image flare refers to unwanted light scattering within the camera's optical system, which can lead to reduced image contrast and the appearance of artifacts in the final photograph. ISO 18844 provides a standardised method for quantifying and evaluating image flare in digital camera images. ISO 18844 addresses the measurement and characterization of image flare in digital camera images. It aims to establish a standardised method for assessing the level of image flare and its impact on image quality. The purpose of this standard is to ensure that digital cameras produce images with minimal image flare. Image flare can occur when stray light enters the camera's optical system and interacts with the image sensor.

Excessive image flare can lead to reduced image quality and the appearance of unwanted artifacts in photos. ISO 18844 also specifies a detailed procedure for measuring image flare in digital camera images. This involves capturing test images with the camera, typically using controlled lighting conditions. The captured images are then analysed to quantify the presence and extent of image flare. The standard outlines how to report the results of the image flare measurements, including the format for expressing image flare data. This information can be useful for camera manufacturers to assess and improve the performance of their camera systems. As well as manufacturers using this standard to meet image flare performance criteria, Camera lens designers and optical engineers can use image flare measurements to evaluate the impact of lens coatings and optical design on reducing unwanted light scattering.

- ## ISO 19084:2015:

The standard focuses on digital cameras and deals specifically with chromatic displacement measurements [9]. Chromatic displacement refers to the phenomenon where different colours of light are refracted by a camera's lens to varying degrees, causing colour fringing or distortion in the final image. ISO 19084 provides a standardised method for measuring and evaluating chromatic displacement in digital camera images. ISO 19084 addresses the measurement and characterization of chromatic displacement in digital camera images. The standard aims to establish a consistent and repeatable method for assessing the level of chromatic aberration or colour fringing in images produced by digital cameras. The purpose of this standard is to ensure that digital cameras produce images with minimal chromatic displacement. Chromatic aberration can be an undesirable optical artifact that affects image quality, especially in high-contrast scenes. By providing a standardised testing procedure, the standard helps camera manufacturers evaluate and improve their camera systems. ISO 19084 specifies a detailed procedure for measuring chromatic displacement in digital camera images. This typically involves capturing test images of controlled subjects that exhibit colour fringing or chromatic aberration, such as high-contrast edges or fine details. The captured images are then analysed to quantify the extent of chromatic displacement. The standard outlines how to report the results of chromatic displacement measurements, including the format for expressing chromatic aberration data. This information can be useful for camera manufacturers to assess and compare the performance of different camera models. Chromatic displacement measurements are important for various applications, like Quality Assurance for manufacturers, and Camera lens designers to evaluate the performance of camera lenses and coatings in reducing chromatic aberration.

- ## ISO/TS 19567-1:2016:

Developed by ISO TC 42, this standard pertains to digital cameras and focuses specifically on measuring texture reproduction characteristics using cyclic patterns [10]. Methods and procedures are defined for measuring the frequency characteristics of texture reproduction in digital camera images. Texture reproduction refers to how well a camera captures and reproduces fine patterns and textures present in a scene. ISO/TS 19567-1 emphasises the assessment of a camera's ability to capture high-frequency details and patterns. High-frequency details often include fine textures, lines, and small repetitive patterns in images. The use of cyclic patterns

is specified as test targets. These patterns are designed with known spatial frequencies and are used to assess how well a camera reproduces these patterns in images. It also provides guidelines for setting up the test environment, including lighting conditions, camera settings, and the positioning of the cyclic pattern test target further ensuring standardised conditions is crucial for obtaining consistent and meaningful results. The standard outlines the process for analysing images captured with the digital camera under test. This analysis involves quantifying how accurately the camera reproduces the known cyclic patterns, including measuring deviations in frequency and amplitude. ISO/TS 19567-1 offers guidance on reporting the results of frequency characteristics measurements. This typically includes providing data on the camera's performance in reproducing cyclic patterns and may involve presenting frequency response curves or other relevant metrics. This standard assists camera manufacturers, testing laboratories, and users in evaluating and comparing how different camera models reproduce fine patterns and textures in images. This standard provides guidelines for test setups, data analysis, and reporting to ensure consistent and reliable assessments of camera performance in reproducing fine textures and patterns.

- ## ISO 16505:2019:

This standard from ISO TC 42, focuses on camera monitor systems used in road vehicles, and specialises on ergonomic and performance aspects, as well as specifying requirements and test procedures [11]. The standard defines the requirements and test procedures for Camera Monitor Systems (CMS) used in road vehicles. CMS replaces traditional rear-view mirrors with cameras and displays to enhance driver visibility and safety. ISO 16505 addresses the ergonomic design of CMS to ensure they are user-friendly and effective for vehicle operators. It includes considerations for factors like field of view, image quality, display layout, and user interfaces to optimise driver comfort and safety. The standard outlines performance requirements for CMS, which includes criteria for image quality, resolution, contrast, brightness, and colour accuracy to ensure that the system provides clear and reliable visual information to the driver. It also provides specific test methods and procedures to evaluate the performance of CMS. These tests may involve measuring parameters such as image quality, response time, and system reliability under various conditions, including environmental factors. The standard includes safety-related requirements, such as minimizing image distortion and latency to prevent potential hazards while driving. Ensuring that CMS provides real-time and distortion-free images is critical for driver safety. Compatibility issues are addressed within the standard with other vehicle systems, ensuring that CMS can integrate seamlessly with existing vehicle technologies and components. The standard emphasises the importance of providing comprehensive documentation and user information for CMS, including user manuals and instructions for system operation and maintenance. Compliance with this standard is essential for ensuring the usability and safety of these systems in road vehicles.

- ## IEEE P2020 Automotive System Image Quality

The P2020 Automotive System Image Quality is an automotive specific camera and imaging standard that is currently being developed by the IEEE-SA P2020 working group [12]. The IEEE-

SA P2020 working group was established in order to develop a comprehensive automotive imaging standards aims to tackle the significant uncertainty in assessing image quality for automotive imaging systems, encompassing both human and computer vision approaches. The standard's scope includes:

- Covering key attributes impacting image quality in automotive ADAS applications.
- Identifying and incorporating relevant metrics and information about these attributes.
- Establishing a set of consistent objective and subjective testing techniques for assessing automotive camera image quality attributes.
- Defining tools and methodologies to enable standardised communication and comparison among automotive system integrators, vendors, and manufacturers (original equipment manufacturers (OEM's) and Tier 1) regarding ADAS image quality.

A preliminary white paper was published by the IEEE-SA P2020 working group, as a preview into the contents of the final standard [13]. The white paper addresses three initial areas that are currently being considered by the IEEE-SA P2020 working group. The first is the effect of LED flicker on the imaging quality of images produced for automotive applications. They note that, "LED flicker is an artifact observed in digital imaging where a light source or a region of an imaged scene appears to flicker. The light as observed on the vision system display may appear to switch on and off or modulate in terms of brightness or colour, even though the light source itself appears constant to a human observer". The white paper addresses the various causes of LED flicker as well as what needs to be done to address the issue. The authors of the white paper state that the P2020 standard will aim to determine the root cause of LED flicker in automotive systems, assess the various impacts of LED flicker in the ADAS application, define standardised test methodologies and metrics to quantify the flicker effect and finally determine the relationship between the define metrics and flicker observed visually and through computer vision algorithms

The second area addressed by the white paper is automotive image quality for viewing applications (i.e. cases where the camera output is not an input to a computer vision system). It states that the standard will aim to develop metrics to characterise image quality for automotive cameras, such as: rear-view cameras, camera monitor systems, and surround view systems and their components such as lens, colour filter array (CFA), sensor, image signal processors (ISP), and displays. The metrics will aim to address two key aspects for visual image quality, the usefulness of the image (i.e. the level of detail in the image) as well as the level of noise.

Finally, the white paper discusses the effect of various factors and noise in the imaging chain on the performance of computer vision algorithms found in ADAS systems. The authors note that while individual metrics (such as MTF, signal-to-noise ratio (SNR), etc) to quantify various factors and noise types in the imaging chain already exist, there is no metric available that can measure the combined effect of all these factors on a computer vision system. Furthermore, the paper notes that determining the requirements for an ADAS system is a complex procedure. Therefore, the group aims to analyse the imaging chain from the optical level down to electronic signal level, and this must be done considering the use cases and bounds in which computer vision systems are expected to operate. To provide these bounds, the group aims

to develop metrics that both describe various forms of image degradation and give bounds on their confidence (i.e. up to what levels of degradation can my computer vision system be expected to perform well).

2.3. Camera Performance Testing

Testing for optimising camera performance in the automotive context is crucial for enabling better performance in ADAS systems. One example of this within the literature is the paper from Roos et al. [14]. In their paper Roos et al. present a framework for evaluating and optimising cameras under various traffic scenarios in a simulation, based on the camera's raw data quality.

The framework that the authors have developed runs on CARLA [15] and it specifically evaluates the field of view (FoV) of a camera to determine the optimal FoV for the current scenario, by using a novel image quality (IQ) metric that the authors have developed. The IQ metric developed by the authors was analysed using correlation analysis. Their analysis showed that the proposed IQ metric had a strong positive correlation with the performance of state-of-the-art object detector models (evaluated using the recall metric), that were applied to frames taken from cameras at different FoVs. Using this information the authors directly apply their proposed IQ metric to images obtained from a camera at different FoVs. Since they know that the IQ metric correlates with object detector performance, they were able to determine which FoV will maximise object detector performance by checking which FoV produces images that result in the best IQ metric score. They also showed that they can apply this framework in simulation under various traffic scenarios to determine the optimal FoV for each scenario.

2.4. Camera Benchmarking Testing

As camera technology for the automotive industry use cases saturates, the development of smart camera technology for the intelligent vehicle use case is becoming more significant. However, the standards to which these devices are evaluated against is not yet mature. Research from Zhao et al [16] proposes benchmarking criteria made up of 57 possible scenarios to evaluate and compare camera performance. The scenarios are designed for PCW (Pedestrian Collision Warning), FCW (Forward Collision Warning), and LDW (Lane Departure Warning). The scenarios also cover different target conditions for static and dynamic road obstacles, straight and curved roads, different environmental conditions, and varying vehicular velocity, thus attempting to the function of smart cameras in every possible situation. More specifically, researchers developed a smart camera benchmarking based on HiL (Hardware-in-the-Loop) using Carmaker as a virtual testing environment. They tested this using two smart cameras, Mobileye and Maxieye. Through this, it was concluded that one camera, Mobileye, performed better for FCW and PCW, whereas performance in LDW was similar. They showed that the virtual environment developed for benchmarking was appropriate in differentiating the ability between smart cameras but stress that the virtual results should not be compared with real world data, unless comparing the whole suite.

2.5. Current Limitations in Automotive Camera Testing Standards

Since current camera standards were designed for broader use cases, there is generally a gap between what the standard considers versus what is needed for automotive applications. The authors of the P2020 white paper have provided a set of gaps that they believe exist within current camera testing standards [13]. The white paper addresses ISO 15739 as one standard that has limitations. Specifically, the white paper notes that this standard lacks the following definitions of sensor dynamic range for the automotive context: Dynamic range definition on multi-exposure type of sensor, dynamic range of displayed imaged with adaptive tone mapping and dynamic range considering a minimum SNR required over its defined range.

Another standard that the white paper identifies as being limited, is ISO 12232 [3]. This standard considers camera exposure and speed tests, however what this standard fails to consider is the application of such tests in the infra-red (IR) spectrum. The point made in the P2020 white paper is that tests considering the IR spectrum are necessary as many automotive grade cameras are thermal cameras and hence a test method for such exposure and speed in such cameras is needed.

Considering tonal response, one area of camera systems to consider, the white paper lists the following limitations seen in current standards:

- **ISO 19093 - Low Light Performance:** The current standard lacks information on low light performance on actual use-case conditions. For automotive applications a definition considering the trade-off operation according to application is necessary [17].
- **ISO 17957 - Tonal and Chroma Shading:** The current standard lacks adaptation to wide view angle applications [7].
- **ISO 14524 - Methods for measuring Opto-Electronic Conversion Functions (OECFs):** The current standard only evaluates a monotonic response of an input image; it does not consider evaluations on adaptive tone mapping operations [18].

ISO 15739 - Noise Measurements, is a standard for noise measurement tests in digital camera systems, with a focus on photography [4]. The P2020 white paper notes that this standard has the following limitations regarding automotive specific needs:

- **Lack of Definition for Measuring Operation Conditions:** ISO 15739 does not provide clear guidelines or definitions for the measuring operation conditions of noise in camera systems used in automotive applications. It is mentioned that noise reduction is often implemented in these systems, but the standard does not specify how this reduction should be considered in measurements. This lack of guidance can result in inconsistencies when evaluating the noise performance of different systems.
- **No Guidance on Handling Spatial Resolution Degradation:** The standard does not address the issue of spatial resolution degradation that can occur when noise reduction operations are applied. Noise reduction techniques often involve smoothing or averaging pixels, which can lead to a loss of fine details in the captured

images. In the context of automotive applications, where image clarity and detail can be critical for safety and functionality, the absence of any guidance on how to assess and mitigate spatial resolution degradation is a major limitation.

- **Limited Consideration for High Dynamic Range (HDR) Operation:** HDR operation is an important aspect of camera systems used in automotive applications. HDR cameras are designed to capture a wide range of brightness levels in a single image, which is crucial for maintaining visibility in various lighting conditions, such as transitioning from a dark tunnel to bright daylight. The standard lacks specific provisions or recommendations for evaluating the performance of camera systems when HDR operation is applied. Moreover, the P2020 white paper notes that HDR operation can suffer from Signal-to-Noise Ratio (SNR) degradation in intermediate brightness conditions. ISO 15739 does not offer guidance on how to address or assess this degradation, which is essential for ensuring the reliability of automotive camera systems.

Considering testing standards for optical artifacts, the authors of the P2020 white paper mention certain limitations in current camera standards considering Flare, namely ISO 18844 [8]and IEC 62676-5 [19]. They note that in automotive use cases, a vehicle's headlamps direct light and/or direct sun light often enters the FoV or hits the optics of the camera system. Therefore, standards for testing optical artifacts need to evaluate stray light which is incident onto an optical system in terms of the veiling effect that deteriorates image visual or post processing performance. ISO 18844 and IEC 62676-5 currently do not evaluate stray light in such a manner.

Lastly, the P2020 white paper considers ISO 12233, which is a standard for digital camera systems to test camera resolution and spatial frequency responses [2]. The authors of the P2020 white paper note that this standard has the following limitations for automotive applications. First this standard lacks a definition of resolution according to the intended FoV of a specific use case. Second from a visual perspective, this standard lacks a metric to evaluate scene distinguishability performance of camera systems with different resolutions. Lastly, this standard lacks a metric to quantify the partial/converted image resolution from a fisheye lens camera.

3. LiDAR Testing Methods

In this section we will be discussing state-of-the-art LiDAR testing methods. First the use-cases and requirements of automotive LiDAR will be covered to provide an overview of the importance of LiDARs in many automotive applications. Next performance tests for LiDAR sensors will be discussed. This part will cover tests that evaluate specific LiDAR sensor parameters as well as tests that examine LiDAR performance under various noise factors. Following this part, the state-of-the-art of LiDAR benchmarking will be discussed. This section will then conclude with a discussion of Automotive LiDAR testing standards.

3.1. Requirements and Use Cases in Automotive

LiDAR sensors have become an indispensable cutting-edge technology that is essential within any perception system, especially since these are one of the only sensors that facilitate a real-time 3D understanding of the environment [20]. The precise and accurate environmental measurements of LiDAR sensors enable them to have many use-cases within perception systems in automotive applications.

For instance, LiDARs are used extensively to within perception systems to perform object detection. Shi et al. developed a point-cloud based neural network model called PointRCNN that is able to perform 3D object detection using LiDAR point-cloud data [21]. Arnold et al. also conducted a survey, where the Authors showed that LiDAR data is used considerably for both 2D and 3D object detection [22]. LiDAR point-cloud data is also used for object detection under noisy scenarios, for example Wu et al. provided a framework to make object detection robust under adverse weather conditions using LiDAR point-cloud data [23]. LiDARs are also used for object tracking within perception systems. Wang et al. developed a framework to perform pedestrian tracking using LiDAR point-cloud data [24]. This is a very important use-case for LiDARs as, tracking pedestrians is an important safety step which is necessary for automated vehicles to prevent collisions with pedestrians. Peng et al. also provide a detection and tracking framework that specifically uses point-clouds generated from Doppler LiDARs [25]. Some other use-cases of LiDARs within perception systems include road segmentation. For instance, Huang et al. developed a ground segmentation method, which can be used in perception systems to determine the boundaries of roads with respect to the ego vehicle [26]. LiDARs are also used within assisted and automated driving (AAD) to estimate the level of noise present within the environment surrounding the ego vehicle. Karlsson et al. developed a model capable of using LiDAR point-cloud data to estimate levels of rainfall within the ego vehicle's environment [27].

In order for LiDAR measurements to remain precise and accurate, certain parameters that describe the functional limits of the LiDAR need to be within specific bounds for LiDARs to be used within the automotive context. Moreover, LiDARs should also have certain design requirements to be used for automotive applications. Bastos et al. provides a list of such requirements, which is summarised below [28].

- **Range:** A maximum range between 200 and 300 meters is necessary to detect objects when moving at high speeds.
- **Precision:** Sub-centimetre precision is expected for low-range scenarios (up to 50 meters), with a precision of at least 5cm to be guaranteed.
- **Cost:** The cost of a LiDAR sensor should be low in comparison to the cost of the vehicle. A rough cost would be around £1000.
- **Acquisition Rate:** A frame rate of at least 20Hz is necessary for LiDAR data to be used for real-time applications.
- **Angular Resolution:** An angular resolution of 0.1° for both horizontal and vertical planes should be guaranteed.
- **Field of View (FoV):** Horizontal FoV should be as high as possible (360°). The vertical FoV needs to be high enough (at least 20°) to reliably observe the surrounding scene.
- **Power Consumption:** Should be lower than 100 watts, due to the limited available power that vehicle batteries offer.
- **Size and Weight:** Both Compact and light enough to be mounted on the top of the vehicle. Specific weight requirements will depend on the load that a specific vehicle is able to bare.
- **Robustness to Background Light:** The LiDAR receiver must be capable of operating properly when exposed to back light and other external sources of light.
- **Robustness to Interference:** The LiDAR system must be robust to either interference from other LiDAR sources and self-interference.
- **Dynamic Range:** A 300-meter maximum target range corresponds to a receiver dynamic range of approximately 70 dB.
- **Eye Safety:** Automotive LiDAR systems are required to adhere to Class-1 laser regulations for both indoor and outdoor settings [29], [30]. For 1550 nm wavelength LiDARs, the emitted optical signal's average output power should be restricted to around 10 milliwatts.

3.2. LiDAR Sensor Performance Tests

FoV and Angular Resolution Tests

Within academic literature, a range of tests has been developed to assess the performance of LiDAR systems within the automotive domain. In their paper, Gomes et al. showcase an indoor laboratory setup, which can be used to test the accuracy of FoV and angular resolution for a given LiDAR under test [31]. The LiDAR under test is mounted on top of a device that the Authors have called 'RotGon'. The RotGon enables tilting/rotating of LiDAR sensors in three distinct angles. The rotation stage permits continuous 360° motion with a maximum speed of 2°/s and a minimum incremental motion of 0.01°. The RotGon is placed at the end of a rail system. The rail system was designed to enable a base moving that can handle weights of up to 30 kg and can be programmed through external communication. In turn, the base can

support several targets with different reflectivity values. In their setup, Gomes et al. utilise a target with 95% reflectivity mounted on the rail system (opposite to the RotGon).

The test described by Gomes et al. for determining the FoV of the LiDAR involves employing a well-defined target of known dimensions and reflectivity, positioned at a predetermined distance on top of the target holder of the rail system. Given the rail system's variable range capabilities, the RotGon is used to allow the LiDAR to move both horizontally and vertically. This movement is accompanied by monitoring the target's displacement beyond the boundaries of the LiDAR's FoV. Using the position data from the RotGon, it is possible to find the LiDAR's FoV. To begin the test, an initial procedure is executed to identify the target within the LiDAR's point cloud data. Upon successful target detection, the algorithm initiates the FoV measurement process. Initially, it identifies the extreme angles in the vertical plane to establish the vertical FoV. Subsequently, this approach is mirrored along the horizontal axis to obtain the horizontal FoV.

These outcomes serve as baseline parameters for assessing the consistency of both vertical and horizontal FoVs across all quadrants of the target. Throughout these procedures, the system iteratively adjusts RotGon angles to achieve the following four distinct target positions:

- Last position where the target is completely outside of the FoV
- First position where the target is completely inside of the FoV
- Last position where the target is completely inside of the FoV
- First position where the target is completely outside of the FoV

After obtaining these measurements, the FoV test ends by using these measurements to calculate the minimum detection angle and maximum detection angle and FoV are calculated for each axis.

The angular resolution test described by the Authors is simpler compared to the FoV test. In essence, the angular resolution of the LiDAR is calculated by first counting the number of points present in the point cloud obtained from a high-reflectivity target with a known size T_{width} and T_{height} and placed at a known distance T_{dist} with respect to the LiDAR. Next the horizontal and vertical angular resolutions can be calculated using Equations 1 and 2, where $H_{number\_of\_points}$ is the number of points in the horizontal plane of the target and $V_{number\_of\_points}$ is the number of points in the vertical plane of the target.

$$AR_h = \frac{2arcsin(\frac{T_{width}}{2T_{dist}})}{H_{number\_of\_points}} \quad (1)$$

$$AR_v = \frac{2arcsin(\frac{T_{height}}{2T_{dist}})}{V_{number\_of\_points}} \quad (2)$$

Various Environmental Conditions Test

The literature also includes performance tests for LiDARs under various environmental conditions. In their study, Park et al. describe such tests [32]. The tests that they conducted were split into two types of test scenarios. In the study the Authors refer to the LiDAR being tested, as the Device Under Test (DUT). The first type involved simulating harsh driving conditions in a laboratory setting. These simulations considered the effect of environmental factors such as low/high temperatures, direct sunlight on the cover of the DUT and DUT cover contamination. The second type of test scenario involved conducting driving tests for an automotive graded LiDAR sensor (the DUT) under various environmental conditions. The conditions examined include: the position and intensity variations of a lead car as it moved on a country road (i.e. the effect of induced vibrations caused by the roughness of country roads on DUT measurements), interference noise of the DUT caused by an external LiDAR mounted on a vehicle approaching the ego vehicle, signal variations of DUT during day and night transitions, and signal variations of the DUT based on changes in lead vehicle (Sensor Target) colours.

The test metrics that the Authors measured from the DUT include:

- **Number of Points:** This is the number of points seen in the point cloud within a region of interest (RoI)
- **Intensity:** Intensity is proportional to the signal strength of received laser power reflected from a surrounding object. In their study, Park et al. measured the intensity of points.
- **Scan Frequency:** For an automotive LiDAR, the scan frequency is the time gap between two consecutive scan frames.
- **FoV and Angular Resolution:** The FoV of a LiDAR sensor is defined as the angle between two scan points at the end of the scan range. Angular resolution is the minimum deviation of two distinguishable objects in radial coordinates
- **Number of Noise Points:** The number of noise points as the amount of external noise in point clouds from a pre-described interference source.
- **Range Accuracy and Precision:** Range accuracy and precision measures the quality of range detection in scanning of LiDAR sensors. Accuracy is calculated as a difference between a reference measurement of the distance of a specified object and the distance measured by the LiDAR. Range precision is a deviation of the measured distance in LiDAR scanning.

For the **high temperature test**, the Authors first established a realistic the temperature gradient for high temperature by measuring the temperature of the LiDAR (DUT) in a hot environment (which Nevada, USA in the summer). Then, within the Laboratory setting, they placed the DUT mounted on a vehicle within a chamber and measured the test metrics by varying the temperature of the chamber between the limits of the temperature gradient obtained previously (limits were 50°C to 80°C).

The **low temperature test** was similar to the high temperature test, except the realistic temperature gradient for low temperatures was obtained by measuring the temperature of

the DUT when placed outdoors during the winter in Alaska, USA. The limits of the low temperature gradient measured were 0°C to -30°C.

For the **sunlight test**, the Authors established the maximum sunlight value (measured in lux) by measuring the level of sunlight in Nevada, USA during the summer. The maximum level of sunlight they derived from this was 80000 lux. To conduct this test in the lab environment, the Authors analysed the signal strength reduction rate based on the brightness of sunlight using a solar light emulator. In the experiment, the solar emulator was installed in front of the DUT, and the DUT captured points from a reference object attached to the bottom of the emulator.

For the **cover contamination test**, DUT visibility levels based on levels of contamination (i.e. DUT covered by mud) were measured by off-road driving in the Nevada desert, with measurements captured every day while driving on a muddy road (measurements were taken for a total of 2 days). Within the laboratory setting, the authors recreated the levels of visibility on the DUT (by covering it with mud) based on the measurements from off-road drive, and measured the intensity and number of points of a 1.5 m × 1.5 m square target with 50% reflectivity.

In the **vibration test**, the ego vehicle with the DUT mounted on top of it was driven on a bumpy country road in Nevada. The target of the DUT was a reference car driven in front of the ego vehicle at a constant speed. To analyse the impact of the vibrations induced by the bumpy road on the DUT, the standard deviation of the centre position of the reference car in all coordinate axes (x, y, and z), as well as the variation of the DUT's point cloud intensity during the drive was measured.

In the **interference test**, a LiDAR-equipped vehicle (reference vehicle) moved from 0 to 20 m towards (in front of) the ego vehicle. Then the number of noise points generated by the LiDAR sensor mounted on the reference vehicle from the DUT was measured. To isolate the noise, a reference scene was captured with no interference in advance and point clouds from the DUT were captured while driving towards the reference vehicle. To analyse the effect of interference, the intensity and number of noise points in the DUT point cloud was compared against the distance between the reference vehicle and the ego vehicle.

For the **colour change test**, two identical model cars (reference vehicle) (one black and one white) were driven on a straight road in Nevada, USA, at night. Points clouds from both the reference cars as they moved from 0 m to 100 m in front of the ego vehicle mounted with the DUT at a constant speed of 20 km per hour were then obtained. From these point-cloud measurements, the intensity and number of points were compared to the distances between the reference and ego vehicle for both reference vehicles.

Finally, the **Day vs Nighttime test** was conducted in a similar manner to the Colour change test. The difference was that the single reference vehicle was used and measurements were taken during the day and during night.

- ## Adverse Weather Tests

LiDAR performance tests under adverse weather conditions were also found. In their study, Jokela et al. developed and performed tests on LiDARs under Fog and Heavy Snow conditions

in indoor and outdoor settings [33]. The Fog conditions test was conducted in an indoor facility. For this test, the authors placed two pairs of optically calibrated plates as targets with reflectivities of 90% and 5% (white and black, respectively). One target in the pair was larger than the other. In the experiment the targets were placed at 10 m, 15 m and 20 m away from the LiDAR under test. As a reference, measurements from the LiDAR were taken at each target distance when there was no fog. Subsequently, the same measurements were taken for fog visibility distances (**in descending order of fog density**) of 10 m, 15 m, 20 m, 25 m, 30 m, 40 m and 50 m. The resulting raw data for all experiment variations (all combinations of known target distances and fog densities) were processed to obtain the estimated target distances and target region of interest. The heavy snow test was conducted outdoors in northern Finland during winter. Tests were performed at the airport runway in Sodankylä. There was 20 cm of sub-zero snow on the road surface and temperature was at around -4°C. To conduct the test, a reference vehicle was used as the target that the LiDAR under test (which was mounted on the ego vehicle) would track. The ego vehicle would follow the reference vehicle which was leading it in a straight line. Measurements were taken when the ego vehicle was behind the reference vehicle by about 15-20 m. This test was conducted both in clear conditions as a reference and when it was snowing heavily. The point clouds generated from the clear and heavy snow conditions were then compared.

Similarly, in their study, Kim et al. developed a test to measure LiDAR performance in rainy conditions [34]. Their setup consisted of an outdoor test track. The targets were 60 cm x 60 cm in size and for each pair of targets, one target was placed on the left and one on the right. 5 pairs of targets were placed on the test track at 20 m, 40 m, 60 m, 80 m and 100 m away from the starting position of the ego vehicle that was mounted with the LiDAR under test. In each test the ego vehicle would drive to the end of the test track, while the LiDAR collected data at speeds of 20, 40, 60 and 80 km/h. For all speeds, measurements were taken in sunny conditions as a reference and subsequently in rainy conditions at rainfall rates of 10, 20, 30, 40 and 50 mm/h. Moreover, for each of these conditions, the effect of the target material was tested by using targets made from wood, plastic, steel, and aluminium. The main metrics derived from the raw LiDAR data during each test was the number of points in a scan as well as the intensity.

- ## LiDAR Sensor Benchmark Tests

Tests to benchmark different LiDARs have also been proposed in the literature. In their study, Jeffries et al. propose a methodology for benchmarking LiDAR sensors to evaluate range error accuracy and precision for eight automotive grade LiDARs [35]. They note that this test is a part of a series of tests aimed at developing methods, metrics, and targets necessary to develop standardised tests for LiDAR sensors. The test was setup such that multiple targets were placed at various distances away from the position of the LiDAR under test. Both calibrated and other objects were present in the test range. As part of this benchmarking effort the Authors developed a 15 cm × 80 cm flat aluminium target with a Lambertian coating that is 10% reflective from 800 nm to 1600 nm. The size corresponds, approximately, to the cross-section area of a small child when viewed from the side. The LiDARs used in this test were all monostatic, which rely on a common source and receiver location. Hence the Authors

setup the targets such that do not overlap with one another. The arrangement of targets started with aligning the 200 m target to the LiDAR under evaluation, in line with its longitudinal axis. Starting from the reference point (the LiDAR's location), the initial two targets were positioned at 5m intervals, alternating between the right and left sides of a ±30° Field of View (FoV) relative to the longitudinal axis. Additional targets were then situated similarly, spaced at 5m intervals, extending up to 50 m on the left side of the range. The subsequent series of targets were positioned at 10 m intervals, originating from the right side of the range and progressing outward to 100 m. Beyond 100 m, the gaps were increased to 20 m, extending to the final target situated at 200 m. After setting up all targets, a reference LiDAR as well as each LiDAR under test was tested and raw point cloud data each LiDAR was retrieved. Scoring the range of each LiDAR was performed by finding the minimum distance from each LiDAR point reported by the LiDAR under test to the closest coplanar point from the reference point cloud considering only range and cross-range dimensions, without considering differences in elevation. The specific metrics considered to evaluate each LiDAR's range includes:

- **Range accuracy** provided by each LiDAR under test with respect to the reference. It is defined as the difference between reference and test LiDAR range (in cm).
- **Range precision** provided by each LiDAR under test with respect to the reference. It is defined as the standard deviation of test LiDAR range measurements (in cm).
- **Root Mean Square (RMS) plane fit** is also provided and refers to the total fit RMS error between each reported LiDAR under test's target point and the nearest reference point on the target excluding elevation.

3.3. LiDAR Sensor Testing Standards

Standards for testing automotive LiDARs are not mature and such standards remain in the early stages of development/definition. One standard that is currently at the starting stage of development is the P2936 - Standard for Test Methods of Automotive LiDAR Performance standard, from the IEEE Standards Association sponsored committee called IEEE Vehicular Technology Society/Automated Vehicles Standards Committee (VT/AVSC) [36]. This standards aims to define test methods for LiDAR performance which are focused measuring metrics, such as range accuracy/precision/resolution, max/min range, detection probability, angular ac-curacy/precision/resolution, and reflectivity. The test methods within the scope of this standard include: scenarios, use cases, tools, environment, and statistical methods. The Authors of this standard state that its purpose is to, "standardise a suite of objective and subjective test methods for measuring automotive LiDAR performance metrics and specify tools and test methods to facilitate standardised communication and comparison among original equipment manufacturers (OEM), Tier 1 system integrators and component vendors regarding automotive LiDAR performance". Since this standard is still in development, no documentation regarding the details of any tests have been published.

Another standard which was proposed and is still in progress is the ISO/PWI 13228 Test Method for Automotive LiDAR [37]. However, since this standard is still within the proposal stage, no information regarding the content of the standard or any proposed content has been published.

4. Radar Testing Methods

In recent years, the automotive industry has undergone a tech-driven revolution, blending technology and transportation. This shift has ushered in a wave of advanced systems aiming to boost vehicle safety, efficiency, and the overall driving experience. Among these innovations, Radar technology has emerged as a key component to be used in assisted and automated driving applications [38]. Originally developed for military and aerospace, Radar has found a compelling purpose in the automotive domain. Unlike its sensor counterparts, Radar remains immune to adverse weather and offers real-time, long-range object detection of vehicles, pedestrians, or obstacles [39]. From adaptive cruise control (ACC) to automatic emergency braking (AEB), Radar-based systems arm vehicles with the information needed to make quick decisions, mitigate risks and optimise driving conditions.

In this section we will be discussing state-of-the-art Radar testing methods, specifically Radar testing standards. First the use-cases of these Radar testing standards will be covered. Next, the system characteristics and setup for the tests described in each Radar testing standard will be described. Finally, the method of measurement for testing Radars as described in all the Radar testing standards will be described.

4.1. Radar Testing Standards - Use Case

- ### EN 303 396 - V1.1.1

The scope of this standard was to describe possible measurement techniques and procedures for the conformance measurements applicable to automotive and surveillance Radar equipment [40]. Its contents cover testing of automotive Advanced Driver Assistance Systems applications, like Adaptive Cruise Control, Blind Spot Detection, Parking Aid, Autonomous Braking and Pre-collision Systems. This also extends to surveillance Radars for ground based vehicles, like trains, trams and aircraft while taxiing, fixed infrastructure Radars, road crossing obstacle detection Radars and helicopter obstacle detection Radars. This standard is one that was intended to be used as a reference for existing and future ETSI standards covering automotive and surveillance Radar equipment.

- ### EN 302 858 V2.1.1

This standard was prepared under standardisation request of the Commission Implementing Decision C(2015) 5376 final of 4.8.2015 on a standardisation request to the European Committee for Electrotechnical Standardisation and to the European Telecommunications Standards Institute as regards radio equipment in support of Directive 2014/53/EU of the European Parliament and of the Council [41]. The standard is intended to contain the technical characteristics and test methods for narrow-band Radar equipment.

- EN 301 091-1 V2.1.1

The standard specifies technical characteristics and methods of measurements for Radar equipment for ground based vehicle applications in the frequency range from 76 GHz to 77 GHz [42]. It covers integrated transceivers and separate transmit or receive modules. It presents specific requirements for short range devices intended for ground base vehicles in applications such as: Adaptive Cruise Control, Collision Warning, Pre-collision systems, Object Detection and Stop and Go and other applications. The standard mentions that it does not necessarily include all characteristics provided by a user, nor does it contain the method for optimal performance.

- ITU-R M.2057-1

This standard specifies the system characteristics of automotive Radars operating under the frequency band 76-81 GHz [43]. These technical and operational characteristics were intended to be used in compatibility studies between automotive Radars for systems such as Adaptive Cruise Control and Collision Avoidance, against higher resolution sensor applications like Blind Spot Detection.

4.2. Radar Testing Standards - System Characteristics

- ETSI EN 303 396 V1.1.1

Equipment under test (EUTs) of this standard are expected to come with guidance from the manufacturer in the way it is presented and configured [40]. All operating bandwidths are declared by the equipment manufacturer. Where there are EUTs with multiple operating bandwidths, an adequate number of operating bandwidths should be chosen for testing such that the lower and upper limits of the operating frequencies are tested.

Modulation during testing specifies that the Radar should be used in a way that is representative of intended regular use. The manufacturer should employ the mode of operation that results in the highest transmitter activity consistent with the greatest power transmission available during use. This is to ensure that Radar transmissions occur regularly on time and that sequences of transmissions are repeated accurately. Each scheme must be tested for multi-modulation transmitters.

For scanning antennas, EUTs will be divided into one of three types, fixed beam, constant pattern, or variable pattern. Measurements are expected to be made on the device boresight unless otherwise explicitly stated for fixed beam EUTs. For variable pattern EUTs, measurements should be made over a sufficient volume and the direction with the greatest peak power, mean power and duty cycle should be selected in measuring radiated energy.

Normal test conditions include a temperature range from +15°C to +35°C and a relative humidity between 20 to 75 %. The normal test voltage for EUTs should be the nominal mains voltage and the normal voltage of the device shall be the voltage for which the equipment was

designed. The frequency of the test power source is intended to be between 49 to 51 Hz. For lead-acid batteries used on vehicles, the test voltage shall be 1.1 times the nominal voltage of the battery.

Extreme test conditions include multiple temperature categories. These categories range from -40°C up to +70°C. Extreme tests voltages shall be within 10% of the nominal mains voltage. The reference bandwidth for the receiver should be equal to the reference bandwidth shown in table 1 of the standard.

- ETSI EN 301 091-1 V2.1.1

This standard overlaps heavily with ETSI EN 303 396 - V1.1.1 [40]. It covers separate transmit and receive elements, as well as integrated transceivers [41]. The environmental conditions of this standard include the environmental profile of operation of the manufacturer. The operating bandwidth frequency range is determined by the range in which the equipment is transmitting from f_L to the highest frequency f_H from the power envelope. In the case of this standard, the upper frequency limit is 77 GHz and the lower limit is 76 GHz. Additional operating frequency criteria, such as resolution bandwidth, averaging time, video bandwidth, detector/display mode start/stop frequency are also mentioned in the standard.

The standard notes that relevant properties, such as range, speed, azimuth angle and presence, must be declared by the manufacture as well as the RCS (Radar Cross Section). Any EUT (Equipment Under Test) that is considered scannable, or steerable, has an antenna pattern that is electronically or mechanically adjustable [40].

The transmitter must conform to a number of requirements that are set out in ETSI EN 303 396 and include the environmental conditions mentioned above [40]. A description of the mean power is given as well as the method in which it should be recorded. Mean power limits are provided for two options: pulsed Radar and non-pulsed Radar. For pulsed Radar the mean power limit is 23.5 dBm (Decibel meters) to 50 dBm for non-pulsed Radars. For constant pattern scanning antennas in which scanning is inhibited, the mean power is calculated using a result from table 3 of ETSI EN 301 091-1 [41]. Similar descriptive and conformance criteria are present for the peak power of EUTs and unwanted emissions in the out-of-band and spurious domains.

The receiver has its own conformance requirements that are influenced by EN 302 858 - V2.1.1 [42]. The power limits for the receiver antenna are categorised into the narrowband and wideband spectrums and provide a frequency range from 30 MHz to 300 GHz.

The testing criteria for spurious emissions at the receiver antenna is described with Resolution Bandwidth, Video Bandwidth, Detector Mode, Span, Sweep Times, and Amplitude. There is also a test for in band, out-of-band and remote band signals testing, the ability of the receiver to operate as intended when received unwanted signals are present. Identically to the transmitter/transceiver, performance criteria is present for EUTs (in Table 8 of standard EN 301 091-1) that outlines the unwanted signal frequencies with which the receiver is still expected to perform [41].

- ## ETSI EN 302 858 V2.1.1

This standard focuses on short range devices for transport and traffic telematics in the operating frequency of 24.05 to 24.25 GHz. The defines the environmental profile for operation as declared by the manufacturer with normal and extreme test conditions - exactly like the definition in standard ETSI EN 303 396 [40].

General power limits, spectrum access conditions and signal conditions that are stratified between 24.05 to 24.25 GHz and are detailed in this standard for narrow-band Radar devices. Table 3 of this standard shows this for - Low Activity Mode, WLAM, devices [42]. All operating bandwidths of the EUT are expected to be declared by the manufacturer and in the case of EUTs with multiple operating bandwidths, a number of operating bandwidths will be selected such that the upper and lower limits of the operating frequencies are tested. Transmissions from a transmitter are expected to occur regularly in time from EUT modulation and that sequences of transmissions are repeated accurately.

The wanted performance criterion for an EUT are the properties of a given target at a given distance. Properties such as, Presence, Range, Relative Speed, Azimuth angle, all declared by the manufacturer, as well as the RCS. For fixed and scanning antennas, these are the characteristics stated in ETSI EN 303 396, and ETSI EN 302 858 [40][42]. The transmitter conformance includes a description from [ETSI EN 303 396] and must operate within the defined frequency limits in tables 2-3 of ETSI EN 302 858 V2.1.1 [42]. The interpretation of any results mentioned will also follow the requirements stated in ETSI EN 303 396 [40].

For unwanted emissions in the out-of-band domain, descriptions from ETSI EN 303 396 apply within the frequency limits stated above [40]. This is also the case for spurious emissions and is further summarised in table 6 of ETSI EN 302 858 V2.1.1 [42].

Spurious emissions from the receiver will be tested for any mode other than transmission. The power limits for narrow-band and WLAM bands are provided in tables 7 and 8 of ETSI EN 302 858 V2.1.1 [42]. These limits range from -57 dBm for between 30 MHz and 1 GHz for the narrow-band case to -37 dBm for between the 1 GHz to 50 GHz for the WLAM case. The conformance test suites follow defined clauses in ETSI EN 303 396 and follow the same interpretation of the results as the transmitter conformance criteria [40].

The same details are relevant for the out-of-band and remote-band signal receiving. Table 9 of ETSI EN 302 858 V2.1.1proposes unwanted signal limits. All conformance and applicability criteria originate from ETSI EN 303 396 [40][42].

There is an additional requirement for spectrum access that is only applicable to WLAM Radars. The description for such devices is outlined in standard ETSI EN 303 396 [40]. For devices with specific signal categories, characterised by their dwell and repetition time, there is also system characteristics for dwell time and reception time. Frequency Modulation criterion is imposed from table 12 of standard ETSI EN 302 858 V2.1.1, with other relevant descriptions and conformance criteria defined in clauses from ETSI EN 303 396 [40].

Additionally, antenna requirements for WLAM devices that consist of a vertical plane transmitter with emissions as a function of the elevation angle normalised to the maximum

emission boresight. The limits for unwanted vertical radiated emissions between 23.6 and 24 GHz bands are -71 dBm and must achieve at least 20 dBm above 30° for antenna attenuation.

- ITU-R M.2057-1

The system characteristics for this standard includes two categories for automotive Radar systems operating in the frequency band 76 and 81 GHz for intelligent transport systems.

Category one includes Radar systems for driver supporting functions, like adaptive cruise control and collision avoidance. They are expected to be able to detect objects for up to 250 metres. A maximum continuous bandwidth of 1 GHz is also required. The characteristics are defined in table 1 of this standard as Radar A [43].

Category two sensors are ones used in high resolution applications such as blind spot detection, lane change assistance, traffic crossing alerts, and close proximity object detection for ranges up to 100 m. For high resolution applications, a wider bandwidth of 4 GHz is required. These Radars typically are used for active and passive safety. Their system characteristics are also in table 1 of this standard and are referred to as Radars B, C, D and E [43].

This standard notes that all Radar devices should be able to disseminate objects with a separation capability of 10 cm and be able to detect objects within 15m, considered close proximity [43]. It is expected that devices may be combined into a suite for use on an ego vehicle, where based on sensor information, data processing will occur that triggers the appropriate Radar.

4.3. Radar Testing Standards - Methods of Measurement

All Radar standards for automotive applications surveyed in this report, when methods of measurement are mentioned, follow the testing methodologies outlined in standard ETSI EN 303 396 V1.1.1, hence we only report the methods of measurement for each test described in ETSI EN 303 396 V1.1.1 [40].

- ETSI EN 303 396 V1.1.1

Initial steps to measurements should include observations with peak and mean measuring modes over the full bandwidth to establish where the maximum values are. This is the foundation for establishing subsequent measurement within a narrow frequency band.

Descriptions and diagrams for Radar testing is given in annex A of ETSI EN 303 396 [40]. Test sites should be verified with valid certification where further information can be found in standard ETSI TR 102 273 [44]. Any mounting brackets used should be made from a low conductivity and low relative dielectric constant. Examples of expanded polystyrene and balsawood are suggested in ETSI EN 303 396 [40].

The range length for all type of tests should be sufficient for testing in the far field. The approximate far field distances are given in table 3 of the ETSI EN 303 396 V1.1.1 standard [40].

There are two test methods (B.1 and B.2) outlines across all Radar standards investigated in this report. B.1 is a test method by using Rx links. The calibration for this is given in annex C of the ETSI EN 303 396 V1.1.1 standard [40]. The summary of test B.1 is as follows: first the EUT is placed within an anechoic chamber for spherical evaluation. The EUT is then placed in an orientation that is closest to intended operation on an automotive vehicle. The test antenna is initially oriented vertically, for vertical plane polarisation corresponding to the transmitter frequency. The output antenna is connected to the spectrum analyser to render a measurable signal. The EUT is first tested without modulation and the analyser tuned to the transmitter frequency under test. The test antenna is then adjusted through lowering and raising to find the maximum signal level by the analyser. Alternatively, the EUT can be tilted through a sufficient azimuthal angle. After which the EUT is rotated through the horizontal plane for 360° until the maximum signal level is detected by the analyser. Alternatively, the EUT can be tilted similar to the previous case. This process is repeated with the test antenna orientated for horizontal polarisation. The maximum signal from the spectrum analyser is noted and converted into radiated power.

The second method, B.2, is a radiated test set-up calibrated by using substitution. A test site, outlined in annex A of the ETSI EN 303 396 V1.1.1 standard, is configured with equipment placed at a specified height on a support and in a position closest to intended regular use [40]. The steps follow identically with the process mentioned in B.1. Divergence ensues when the substitution antenna is connected to a calibrated signal generator. The sensitivity of the spectrum analyser is adjusted, if necessary. The substituted test antenna is then adjusted in height until the maximum signal is received. The antenna can also be tilted through a suitable range. The input signal to the substitution antenna is also adjusted to produce a level detected by the spectrum analyser that is equal to the level detected when the transmitter radiated power was measured. Any corrected changes must be accounted for when adjusting the spectrum analyser attenuation. The measurement is then repeated in the horizontal plane and the measure of radiated power of the radio device is the larger of the two values recorded at the substitution antenna input.

5. Sensor Fusion Testing

Sensor Fusion is increasingly becoming an important technology for improving the performance of autonomous systems. Fusion technologies are applied in various areas, such as advanced driver assistance systems (ADAS) [45], simultaneous localisation and mapping (SLAM) [46], object recognition and tracking [47], [48], human-machine interaction [49], robot systems [50], human health monitoring [51], remote sensing [52] and fault detection [53], [54]. Moreover, as they systems utilise and fuse together information obtained by different sensors, they have the advantage of avoiding the perceptual limitations and uncertainties that would normally be present in a single sensor [55]. Given the advantages of such systems and their growing number of applications where sensor fusion is used, testing these systems for benchmarking and verifying performance is crucial.

In their paper Elgharbawy et al. highlight three key problems associated with the development and implementation of ADAS in the automotive industry [56]. First, ADAS face challenges in coping with the dynamic traffic situations encountered in daily life. This includes difficulties in predicting and responding to the often unpredictable motion of vulnerable road users like pedestrians and cyclists. Second, the collaborative multi-sensor data fusion, which is essential for environmental perception in ADAS, brings about software complexity due to the interconnection of multiple sensors. Third, conducting exhaustive testing of the fusion module under real driving conditions is practically impossible due to cost and time-to-market constraints. These challenges collectively contribute to the need for alternative methods to systematically test and improve the robustness, reliability, and safety of ADAS software.

In response to these identified problems, the authors propose a solution in the form of a general and modular framework for benchmarking multi-sensor fusion algorithms via driving simulation. The main contribution of this framework is the development of a real-time Hardware-in-the-Loop (HiL) co-simulation approach with model-based reactive testing. This approach aims to address the challenges highlighted by detecting and interpreting the dynamic driving environment, with a specific focus on factors like the presence and movement of pedestrians. Through iterative development of a sophisticated sensor model within this framework, the testing system can gradually introduce sensor failures to achieve increasing degrees of realism. The authors note that by reducing the gap between simulated and real-life environments, their approach offers an alternative method for systematically testing complex automotive software systems, thereby contributing to the improvement of ADAS in a highly safety-critical context.

The framework begins with a virtual 3D environment containing the EGO vehicle, dynamic road objects, and EGO road contacts. The sensor models receive data from this environment and generate object lists. These lists are then processed in a model-based test automation block, which selects test cases and introduces faults like latency, detection failures, and false object labelling. The faulty object lists from each sensor model are fused using a sensor fusion algorithm to produce a final object list containing tracked objects around the EGO vehicle. This information is then sent to a central ADAS ECU, which implements an ADAS function using the tracked object list, ultimately altering the EGO vehicle's dynamics. This testing framework

assesses the sensor models and fusion algorithm's robustness based on the EGO vehicle's response to various test cases, such as a pedestrian suddenly entering the road and being detected by both camera and radar sensors. In this example the ADAS ECU evaluates results, deciding whether to take action (i.e. avoid the pedestrian). If a fault, such as a time delay/low latency occurs between the radar and camera sensor detections and the ECU still reacts, it shows that the system can tolerate such sensor faults. However, if the delay causes a lack of reaction, this this would indicate that the delay between sensor detections was too large, hence improvements are needed at the sensor suite level to reduce this delay to ensure safe responses.

6. Perception Sensor Standards for Object Detection Applications

This section lists the current standards relating to perception sensor testing specifically in the context of object detection applications.

6.1. BS EN ISO 16001:2017 Earth-moving machinery. Object detection systems and visibility aids. Performance requirements and tests

ISO 16001:2017 specifies general requirements and describes methods for evaluating and testing the performance of object detection systems (ODSs) and visibility aids (VAs) used on earth-moving machines [57]. It covers the following aspects:

- detection or visibility or both of objects including people in the detection zone;
- visual, audible, or both warnings to the operator and if appropriate to the persons in the detection zone;
- operational reliability of the system;
- compatibility and environmental specifications of the system.

The standard specifies testing requirements for object detection application using RADAR, Ultrasonic, Camera Systems (CCTV) and electromagnetic (EM) signal transceiver systems. For RADAR, the test in the standard describes a method for determining the detection zone for radar systems used to detect obstacles near earth-moving machinery. The procedure is designed to identify the geometry of the three-dimensional zone within which a person can be reliably detected by an earth moving machine. For ultrasonic detection systems, the standard described tests that are designed to measure the performance of ultrasonic detection systems intended for use on earth moving machinery. The test description mentions performance requirements for ultrasonic systems, namely:

- overall performance criteria of the system;
- the performance criteria and limits of detection zone;
- the criteria for location and fixing of the components;
- the operational system's reliability;
- deactivation of the system;
- the detection time;
- physical environment conditions (vibration, shock, temperature, humidity).

The tests described for ultrasonic systems are done so specifically for earth-moving machinery with operational reverse speeds of up to 10 km/h. For ultrasonic transponder systems, the standard describes a test method for verifying the detection zone necessary to prevent accidental contact between the machine and the worker and confirming that the worker who enters the detection zone is detected without fail.

For camera systems (CCTV), the standard describes tests are designed to measure the performance of CCTV systems intended for use on earth-moving machinery. Aspects of the performance to be determined are:

- overall quality of the image presented in terms of TV lines resolved or equivalent, based on the scale on the test body reading that can be resolved
- limit of operational light levels to maintain minimum proposed resolution
- vertical and horizontal fields of vision for the system
- detection distance (range)
- masking due to direct exposure to high intensity light
- time taken for the system to respond fully to rapid changes in light levels

The Standard also describes performance tests for visual object detection systems to detect a person or test body approaching an earth-moving vehicle under operation for visual object detection systems (i.e. camera based systems). The tests described by the standard include: Testing object detection system performance when a test body approaching from the front side of the earth-mover, rear side of the earth-mover, left-hand side of the earth-mover and the right hand side of the earth-mover. The standard also describes how the results of these four tests can be merged to measure the detection zone of a camera based system.

Finally, for camera based systems, the standard describes tests for object detection tasks where the system needs to distinguish between obstacles and people. These tests described assess the following aspects of a camera system:

- Obstacle detection performance;
- The human detection and recognition (differentiation human/other object) performance and reliability;
- Measure the limits of the detection zone;
- Measure the detection time;
- Check the operational system's reliability.

6.2. ISO 17386:2023 Intelligent transport systems. Manoeuvring aids for low-speed operation (MALSO) . Performance requirements and test procedures.

ISO 17386:2023 addresses light-duty vehicles, such as passenger cars, pick-up trucks, light vans, and sport utility vehicles (motorcycles excluded) equipped with Manoeuvring Aids for Low Speed Operation (MALSO) systems [58]. It specifies the minimum functionality requirements which the driver can generally expect of the device, i.e. detection of and information on the presence of relevant obstacles within a defined (short) detection range. It defines minimum requirements for failure indication as well as performance test procedures; it includes rules for the general information strategy but does not restrict the kind of information or display system.

MALSO systems considered in this standard use object-detection devices (sensors) for ranging in order to provide the driver with information based on the distance to obstacles. The current test objects are defined based on systems using ultrasonic sensors, which reflect the most commonly used technology at the time of publication of the standard. For other sensing technologies which will potentially emerge in the future, these test objects shall be checked and changed if required. Visibility-enhancement systems like video-camera aids without distance ranging and warning are not covered by this standard. Reversing aids and obstacle-detection devices on heavy commercial vehicles are not addressed by this document.

The ISO 17386:2023 standard also specifies object detection system performance criteria for ultrasonic and radar based systems. It covers performance criteria for the dynamic performance of object detection systems, which include: requirements for the relative velocity of detected objects, allowable start up delays for detection, and detection latency. It also covers performance requirements for sensor range coverage - horizontal, vertical, rear, front and corner coverage as well as minimum coverage ratios. It also includes information about sensor system failure tests.

In addition to performance requirements tests for object detection, this standard includes operational tests for object detection using ultrasonic and radar sensors. The tests cover, requirements for standard test objects to be used when testing a system's obstacle detection capability for ultrasonic and radar systems. The two operational tests described are tests for Coverage of horizontal areas of relevance and Coverage of vertical areas of relevance.

7. Euro NCAP Protocols

Euro NCAP, short for the European New Car Assessment Program, is an independent organization dedicated to evaluating the safety of new cars sold in Europe. This initiative was established to provide consumers with reliable information about the safety performance of various vehicles on the market. Euro NCAP conducts rigorous testing protocols designed to assess a car's safety in various scenarios. These protocols encompass a range of evaluations, including crash tests that simulate real-life accident scenarios, pedestrian protection tests, assessments of safety assist systems, and evaluations of child occupant protection features. The results of these tests are then used to generate safety ratings, allowing consumers to make informed decisions about the safety of the vehicles they are considering purchasing.

7.1. NCAP AEB VRU Systems Test Protocol

The Euro NCAP AEB VRU (Autonomous Emergency Braking Vulnerable Road User) Testing Protocol is a comprehensive set of guidelines and procedures designed to assess the effectiveness of AEB systems in vehicles when it comes to avoiding collisions with vulnerable road users, such as pedestrians, bicyclists, and motorcyclists [59]. This protocol outlines various test scenarios and conditions to ensure a thorough evaluation of these safety systems. This standard is relevant for perception system testing as the AEB systems that would be tested according to the NCAP protocol would rely on perception sensors to inform the decisions that they make.

The testing protocol includes detailed definitions of terms used throughout the document to ensure clarity and consistency. It then proceeds to describe the different test scenarios that the AEB system will undergo, including car-to-pedestrian, car-to-bicyclist, and car-to-motorcyclist scenarios. Each scenario has specific procedures and conditions to simulate real-world situations accurately.

To conduct these tests, the protocol outlines the necessary measuring equipment, including devices for longitudinal and lateral path error measurement, as well as profiles for determining impact speeds. This ensures precision and reliability in the evaluation process. Additionally, the protocol specifies the conditions under which the tests should be conducted, such as the test track layout, weather conditions, and surrounding environment. For instance, the protocol details the preparation of Vulnerable Road User Targets (VUTs) used in the tests, specifying their dimensions and setup. It also includes procedures for testing under low ambient lighting conditions, ensuring that the AEB systems' performance is evaluated in various lighting situations, crucial for real-world scenarios where visibility may be reduced.

Throughout the protocol, there are annexes providing additional details on obstruction dimensions, testing at low lighting conditions, brake application procedures, and more. These detailed guidelines ensure that the AEB VRU Testing Protocol provides a thorough and standardised method for assessing the safety performance of AEB systems in vehicles concerning vulnerable road users. The goal of this NCAP protocol is to provide consumers with reliable

information about a vehicle's ability to prevent collisions and protect those most at risk on the road.

7.2. NCAP VRU Test Protocol

The Euro NCAP VRU (Vulnerable Road User) Testing Protocol is a detailed set of guidelines and procedures developed to assess the safety of vehicles concerning pedestrian and cyclist collisions [60]. This protocol covers various aspects of vehicle preparation and assessment, focusing on the effectiveness of deployable systems in detecting and protecting vulnerable road users. One key area of evaluation is the detection of pedestrians, ensuring that vehicles equipped with these systems can accurately identify individuals on the road. The protocol also specifies the timing of system deployment, assessing how quickly and effectively the vehicle responds to the presence of a pedestrian or cyclist.

Furthermore, the protocol delves into the protection offered by these systems at different impact speeds, considering scenarios both below and above the deployment threshold. It also evaluates the bonnet deflection due to body loading, which is crucial for understanding how well a vehicle's structure absorbs impact forces in a collision with a vulnerable road user. Additionally, detailed vehicle marking guidelines are provided, outlining specific reference lines and points for impact testing with headforms and legforms.

This protocol is immensely valuable for testing perception sensors in vehicles because it provides a standardised and comprehensive framework for evaluating their effectiveness in detecting and reacting to vulnerable road users. By following these guidelines, manufacturers can ensure that their vehicles meet safety standards regarding pedestrian and cyclist protection. The protocol's focus on specific impact points, timing of deployment, and vehicle preparation allows for consistent and reliable testing across different vehicles and sensor systems.

7.3. NCAP AEB Car-to-Car Systems Test Protocol

The Euro NCAP AEB Car-to-Car Systems Testing Protocol is a detailed set of guidelines and procedures aimed at evaluating the performance of Autonomous Emergency Braking (AEB) systems in vehicles during car-to-car scenarios [61]. This protocol outlines specific test scenarios and conditions to assess how effectively AEB systems can detect and react to potential collisions with other vehicles on the road. It provides a standardised framework for manufacturers to test the functionality and reliability of their AEB systems, ultimately aiming to enhance vehicle safety and reduce the risk of accidents.

The protocol covers various aspects of AEB system testing, including definitions of terms used throughout the document and detailed descriptions of test scenarios. It specifies the reference system used for measurements, such as longitudinal and lateral path errors, as well as the profiles for determining impact speeds during car-to-car situations. Measuring equipment requirements are outlined to ensure accurate and consistent testing procedures, with guidelines for data filtering to maintain reliability.

Furthermore, the Euro NCAP AEB Car-to-Car Systems Testing Protocol considers various test conditions such as the required test track layout, weather conditions, and surrounding environment. It details the preparation of the Vehicle Under Test (VUT) for testing, ensuring that all parameters are set correctly for accurate evaluation. This protocol is especially valuable for perception sensors testing because it provides a structured and standardised approach to assess the capabilities of AEB systems in detecting and responding to potential collisions with other vehicles. By following these guidelines, manufacturers can identify areas for improvement in their AEB systems, leading to safer vehicles on the road. In summary, the protocol's focus on test scenarios, conditions, and measurement equipment helps ensure that AEB systems perform reliably in real-world car-to-car scenarios.

7.4. NCAP Lane Support Systems Test Protocol

The Euro NCAP Test Protocol for Lane Support Systems is a comprehensive set of guidelines and procedures developed to evaluate the performance of Lane Keeping Assistance (LKA) and Lane Departure Warning (LDW) systems in vehicles [61]. This protocol outlines the specific testing cases and conditions required to assess how effectively these systems assist drivers in maintaining their lane and warn them of unintended lane departures. The goal is to provide a standardised framework for manufacturers to test and improve the functionality and reliability of their lane support systems, ultimately enhancing vehicle safety.

The protocol begins with definitions of terms used throughout the document, ensuring clarity and consistency in understanding the testing procedures. It then describes the reference system used for measurements, focusing on parameters such as lateral path error. Measuring equipment requirements are detailed, including specifications for accurate measurements and data filtering processes to maintain reliability in the results. Furthermore, the Euro NCAP Lane Support Systems Testing Protocol considers various test conditions critical for evaluating the performance of these systems. This includes specifications for the test track layout, which is essential for simulating real-world driving scenarios. Additionally, guidelines for Vehicle Under Test (VUT) preparation are provided, ensuring that the vehicle's systems are properly calibrated and set up for testing.

This protocol is particularly valuable for perception sensors testing because it provides a structured and standardised approach to assess the capabilities of Lane Keeping Assistance and Lane Departure Warning systems. By following these guidelines, manufacturers can identify areas of improvement in their systems, such as accuracy in detecting lane markings and providing timely alerts to drivers. Ultimately, the Euro NCAP Lane Support Systems Testing Protocol aims to enhance road safety by encouraging the development of reliable and effective lane support technologies in vehicles. The focus on test scenarios, conditions, and measurement equipment ensures that these systems perform consistently and reliably.

8. Sim4CAMSens Workshop Summary

Participants and Advisory Board

Participants							
WMG	NPL	rFpro	CSAC	Claytex	Syselek	Oxford RF	AESIN
Advisory Board							
Jaguar Land Rover (JLR)							

8.1. Summary of the Workshop

The Sim4CAMSens Noise Factors workshop was held at WMG, University of Warwick on 5th December 2023.

The workshop started with a presentation by Professor Valentina Donzella, from WMG on Noise Factor Analysis for Sim4CAMSens. The presentation covered the following topics:

1) An overview of the expected outcomes of the workshop, as per points below:
    a. Raising awareness of structured **tools** and **techniques** that can support a thorough **noise factor analysis** of automotive perception sensors.
    b. Prioritise noise factors for each main sensor technology (RADAR, LiDAR, Camera)
    c. Understand if the current Sim4CAMSens test plan can support the testing of the identified key noise factors.
2) Motivation for Noise Factor Analysis – The goal is to left shift sensor testing from the real world to simulation. Left shifting requires adequate model for the system, environment, and sensors. Therefore, a good understanding of the noise factors that affect sensor data and of **how** they affect the data is required to support high fidelity sensor models. To assess the fidelity of the developed sensor models, there is the need of: data quality metrics to evaluate noisy and clean sensor data; noisy and clean collected real sensors data to be used as a comparison for the models. There is no standardised method to evaluate data quality degradation for all perception sensors (Camera, LiDAR, and Radar).
3) Research from WMG covering the effect of sensor noise on the performance of various perception tasks (object detection, segmentation, etc) [6], [62] [63], [64], [65] – the research was presented to highlight the importance of noise factor analysis in understanding its effect on perception performance.
4) Validating models of sensors in simulation which generate simulated sensor data – Data quality metrics were highlighted as a tool to be used to compare the sensor data generated in simulation to real sensor data to validate the sensor models, as well as noise models created in simulation.
5) Research from WMG covering data quality metrics that could be used, as well as noise factors to consider, were presented. This included some research examples on noise factor analysis from WMG mainly on camera sensors, but also on LiDARs and Radars.
6) An introduction to the planned Sim4CAMSens winter collection was provided (see Figure 1 below).

Subsequently, participants engaged in a comprehensive review of WMG's tables (see Table 1, Table 2, and Table 3 below) describing noise factors for each type of sensor. Following this review, participants were divided into two groups to assess, supplement, and score noise factors based on their impact on sensor data. The exercise aimed to prioritise the most influential noise factors for each sensor in terms of their impact on sensor data degradation. The final phase involved a discussion among participants to align Sim4CAMSens with identified crucial noise factors and testing plans.

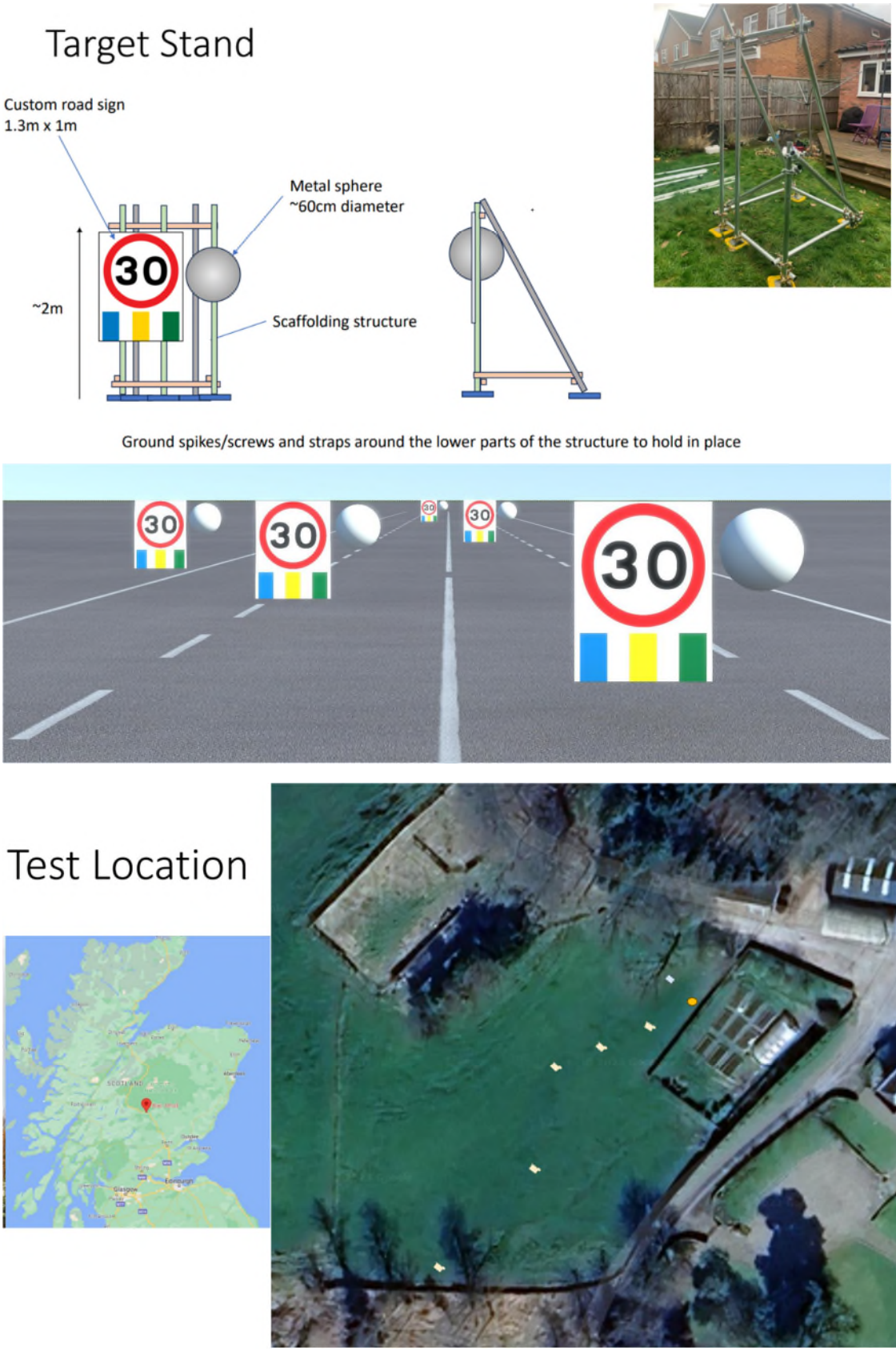


Figure 1: Winter Testing Targets, Location and Layout.

Table 1: Camera Noise Factors [6].

Factor Type	Noise Factor	FR	I_{RGB}	$P_{(X,Y)}$	DF	Description
Piece to Piece	Alignment		✓	✓		Misalignment between sensor and lens during assembly
	Fabrication Quality	✓	✓			CMOS fabrication variability (photodiode, circuitry, Analogue to Digital Converters)
	Lens Shape Purity		✓	✓		Lens fabrication variability, resulting in non-ideal absorption and refractions.
	Dark Current Variability		✓			There are mechanisms of compensating the current generated by the photodiodes in presence of no light, but usually there is a variability of this current from pixel to pixel.
	ISP	✓	✓	✓		ISP alters the data gathered by the image sensor; functions implemented can include: denoising, demosaicing, colour correction, white balancing, sharpening edges, etc.
Change over Time	Ageing or Electronics	✓	✓			Degradation of the performance of the electronic components, resulting in effects such as increased/decreased resistance, leaking currents, etc.
	Degradation of Lens		✓	✓		Lens wear out and ageing resulting in attenuation and refractions.
	Vibration of Mounting		✓	✓		Long term effect of vehicle vibrations resulting in loosening of mounting.
	Pollutant Ingress		✓	✓	✓	Ingress of particulates such as dust, water, condensation.
	Pixel Degradation		✓			Exposure to electromagnetic waves resulting in degradation of silicon doping and reduction of pixel performance.
	Board Aging	✓	✓		✓	Printed circuit board degradation over time such as whisker/dendritic, connector pin contact degradation.
Usage	Misplacement of Sensor		✓	✓		Change in the positioning of the sensor due to terrain and vehicle, causing a variation of the sensor coordinate system (axes and/or angle) with respect to original calibration.
	Vehicle Impact		✓	✓		Impacts on the camera unit or vehicle which results in misalignment of the sensor/lens.
	Chemicals and Contaminants		✓	✓		Cleaning materials and chemicals may react with the lens surface and cause irreversible damage.
	Obstructions		✓	✓		During driving, materials/particles (e.g. water, stains, etc.) can obscure/refract the incoming light
	Lens Scratch		✓	✓		Scratches can reflect and attenuate the incoming light differently to intended.
	Vehicle Dynamic Settings		✓	✓		Adjusting height of vehicle through weight, tyres, pitch, loaded weight etc. thereby changing the sensor coordinate system with respect to original calibration.
Environment	Sensor Saturation or depletion		✓			Scenes with extreme brightness or very low luminosity (i.e. sunrise, sunset, exiting tunnel) can cause saturation or depletion of areas of pixels and therefore an inaccurate rendering of the scene.
	Extreme Temperatures	✓	✓		✓	Sensor operating in conditions outside of manufacturer recommended temperature.
	Adverse Weather		✓			Conditions such as rain, snow, fog, sleet, frost, mist, etc.
	Optic Obstructions		✓	✓		Obstructions, such as mud, stains, frost, water spray, flies, etc., which are partially or fully on the lens or windshield which can block or refract the light
	Low Illumination		✓			Low light resulting in required high pixel gain, creating a larger difference in output intensity between adjacent pixels due to increased noise

	Sun		✓			The sun can cause local saturation, lens flare, IR detection into the colour channels.
System Interactions	Malicious Attacks	✓	✓	✓	✓	Artificial alteration of the image (externally or internally) i.e. cyberattacks, external light source attacks, etc.
	Windshield Distortion			✓		Curvature of windshield which changes the angle at which light enters the sensor.
	Power Supply	✓	✓		✓	Unstable or varying supplied power causing variations in generated signals.
	EMI	✓	✓		✓	Electromagnetic Interference (EMI) from start-up/shut-down of electronics, motors, etc. inducing current within the sensor wires and connections.
	Saturation of Buffer		✓	✓	✓	Sensor internal buffer saturation, causing problems on transmitted data flow, e.g. inability to process incoming data, incorrect stored data, etc.
	LED Flicker		✓			Pulsing LEDs in the environment resulting in fluctuations in the generated images
	Localised Light Source		✓			Headlight, flashlight, high beams, laser beams, etc.

Table 2: LiDAR Noise Factors [62].

Factor Type	Noise Factor	I	ToF	ΨΘ	X,Y,Z	Description
Piece to Piece	Laser Diode	✓	✓			Light emission is affected by the variability of fabrication parameters
	Mounting			✓		Can affect the emission direction
Change over Time	Emitter	✓	✓			Fluctuation/degradation of emitter power, bias, wavelength shift
	Mechanics			✓		Wear in mechanical parts resulting in offsets and misplacement
	Receiver	✓	✓			Degradation could result in a responsivity wavelength shift and could result in lower or higher intensity recorded for a specific wavelength
	Circuits	✓	✓			Electronic circuit components degradation/aging over time Usage
Usage	Multiple Returns	✓			✓	From multiple objects in beam path, ground, beam divergence
	Motion			✓		Vehicle vibration, speed, acceleration, ground holes, etc.
	Clock Speed		✓			The clock is used as reference for the ToF (instability, errors)
	Lens Damage	✓			✓	Dispersion effects reducing intensity and refraction may result in a return from a location that is not expected from the beam path
Environment	Weather	✓	✓			LiDAR is affected by weather conditions, such as rain, snow, fog, etc.
	Obstruction	✓			✓	Lens can be obstructed by objects, rain, mud, etc. Water drops can result in lensing effect, reduce intensity, etc. Mud can occlude the laser beam.
	Ambient Conditions	✓	✓			These conditions can affect light propagation. Temperature affects optical, electronic, mechanical components. Luminosity affects detector performance.
System Interactions	Malicious Attacks	✓			✓	External systems can disrupt the emissions and/or reception, e.g. by absorbing and reemitting at altered times or other methods
	Other LiDARs (Interference, False Detection)				✓	Other LiDAR units can cause interference, false detection, etc.
	EMI	✓	✓		✓	Internal and external electrical components interactions

Table 3: RADAR Noise Factors [63].

Factor Type	Noise Factor	Location (Max Range)	RCS	Velocity	Target Power (NF)	Description
Piece to piece	Alignment of Components	(✓)	✓		✓(✓)	Misalignments due to assembly and manufacturing tolerances, affecting RF emission. Silicon and System-on-a-Chip fabrication (containing antenna, modulator, etc.). Alignment/shape of radome affecting emission angle/power. Quantisation error, component efficiencies.
	Fabrication Quality	✓(✓)	✓	✓	✓(✓)	
	Radome	✓(✓)	✓		✓	
	Signal Processing	✓(✓)	✓	✓	✓	
Change over time	Ageing of Electronics	✓(✓)	✓	✓	✓(✓)	Aging of critical components can reduce effectiveness of the processing/emission/reception. Vibrations over time causing internal/external alignment issues. Contaminants and degradation will affect dielectric properties.
	Misalignment with Radome	✓(✓)	✓		✓	
	Degradation of Radome		✓		✓	
Usage	Misplacement of the Sensor	✓(✓)			✓	Sensor located in different location compared to original calibration. Heavy impact resulting in internal/external alignment issues Ingress of humidity/particles affecting attenuation/transmission, chemical reaction/damage to radome. Material near field, water, leaves, snow, on radome or bodywork. These settings affect the number of vibrations and may shift relative position of the radar. Corruption can occur in the memory, dynamic or static, which can cause profound effects on the firmware and software, as well as the signal processing. Each component will naturally have variability and with usage there might be unexpected behaviours.
	Vehicle Impact	✓(✓)	✓		✓(✓)	
	Chemicals/Contaminants/Component-Damage	✓(✓)	✓		✓(✓)	
	Obstructions	✓(✓)	✓		✓(✓)	
	Vehicle Dynamics Settings	✓(✓)		✓	✓(✓)	
	Firmware/Memory/Software Corruption	✓(✓)	✓	✓	✓	
	Component Noise	✓	✓		✓(✓)	
Environment	Extreme Temperature	✓(✓)	✓	✓	✓(✓)	Electrical/mechanical components extreme variations. Weather particles can cause attenuation, back-scattering, absorption. Clutter might create reflections and ghost objects. Reflection of RF waves due to side lobes, causing increased noise
	Adverse Weather	✓(✓)	✓	✓	✓(✓)	
	Object Clutter	✓	✓	✓	✓(✓)	
	Multipath Reflections	✓(✓)	✓	✓	✓(✓)	

	Road Surface Conditions (e.g. Wet Surface)	✓(✓)	✓	✓	✓(✓)	floor and potential for detection of false points [60], [61]. Road surface can be covered by water, ice, sand etc. which can reflect and attenuate the RF wave.
System Interaction	Vehicle Bodywork (Bumper)	✓(✓)	✓		✓	Bodywork might reflect/absorb RF, altering beam profile. Instability and fluctuations of power supplied. Internal EMI, inducing noise into analogue Circuit. Automotive RADARs operating in the same spectral region.
	Power Supply		✓(✓)	✓	✓	
	EMI	✓	✓	✓	✓(✓)	
	Cross Device Interference	✓(✓)	✓	✓	✓(✓)	

8.2. Outcomes of the Workshop

The tables below summarise the final key noise factors identified as an outcome of this workshop.

1) The first table (Table 4) lists the key/most important noise factors for all perception sensors (Cameras, LiDARs and RADARs) as identified during the workshop. Colour coding is used to identify noise factors which will not be tested, those that will be tested, those which are partially tested as a part of the Sim4CAMSens winter campaign.
2) The second table (Table 5) present the same list but includes other measurement campaigns which are part of Sim4CAMSens.
3) Finally, the third (Table 6) shows which noise factors can and cannot be simulated.

- All Sensors (Summary of Most Important Noise Factors)

Table 4: All Sensors (Summary of Most Important Noise Factors).

Noise Factor	Camera	LiDAR	RADAR
Weather (Snow)			
Weather (Rain)			
Weather (Fog)			
Particulates (Dust, Particles, etc)			
Weather (Temperature)			
Road Surface Conditions (Wet, Snowy)			
Obstructions (on the sensor – near field)			
Obstructions (within FOV)			
Interference			
Multipath			
Target Material Property			
Vehicle Dynamics			
Potholes			
Lighting Conditions			

Legend:

Red = Not Testing

Orange = Partially Tested

Green = Testing

Black = Noise factor not applicable to sensor

- Noise factors that could be tested by the Consortium (Most Important Noise Factors)

Table 5: Noise factors that could be tested by the Consortium (Most Important Noise Factors).

Noise Factor	Camera	LiDAR	RADAR
Weather (Snow)			
Weather (Rain)			
Weather (Fog)			
Particulates (Dust, Particles, etc.)			
Weather (Temperature)			
Road Surface Conditions (Wet, Snowy)			
Obstructions (on the sensor – near field)			
Obstructions (within FOV)			
Interference			
Multipath			
Target Material Property			
Vehicle Dynamics			
Potholes			
Lighting Conditions			

Legend:

Red = Not Testing

Orange = Partially Tested

Green = Testing

Blue = What we can plan to capture

Black = Noise factor not applicable to sensor

- What noise factors can be simulated? (Most Important Noise Factors)

Table 6: What noise factors can be simulated? (Most Important Noise Factors).

Noise Factor	Camera	LiDAR	RADAR
Weather (Snow)			
Weather (Rain)			
Weather (Fog)			
Particulates (Dust, Particles, etc)			
Weather (Temperature)			
Road Surface Conditions (Wet, Snowy)			
Obstructions (on the sensor – near field)			
Obstructions (within FOV)			
Interference			
Multipath			
Target Material Property			
Vehicle Dynamics			
Potholes			
Lighting Conditions			

Legend:

Red = Cannot simulate

Orange = Might be able to simulate

Green = Can simulate

Black = Noise factor not applicable to sensor

9. Winter Testing Campaign

9.1. Site Setup

All equipment was transported to Scotland on the 4th of January 2024, and test site setup started on the 5th of January, with 3 days spent on setting up the targets and the enclosure, and 4 days spent on setting up the sensors and data logging.

Figure 2: Setting up the testing site.

Figure 3: Sensor enclosure in the field.

Figure 4: Custom designed sensor target.

9.2. Weather Stations

The following weather stations/sensors were used during the Winter Testing Campaign:

- Davis Weather Station for temperature, pressure, humidity, and other “basics”
- OTT Parsivel2 disdrometer
- OTT VS2K present weather sensor

Detailed weather data is only available from 1st February 2024. In January, the following events were missed: 7 days of rain, 2 days of rain/sleet, and 4 days of snow. Table 1 presents various types of weather events as detected by the weather sensor, with the number of their occurrences and their total duration (each occurrence covers a time period of 5 minutes). As can be seen in the table, the highest number of occurrences had no precipitation (7635), followed by light rain (913), light drizzle with rain (315), and light snow (156). There were 37 occurrences of moderate/heavy soft hail, and other intensities of drizzle, rain, or snow had even less occurrences than that.

Table 7: Different types of weather events and their occurrences.

Classification	SYNOP Code	Occurrences	Time Period
No precipitation	00	7635	26d 12h 15m
Light drizzle (<0.1mm/h)	51	35	0d 2h 55m
Moderate drizzle (<0.5mm/h)	52	14	0d 1h 10m
Heavy drizzle	53	0	0
Light drizzle with rain (<2.5mm/h)	57	315	1d 2h 15m
Moderate drizzle with rain (<10mm/h)	58	5	0d 0h 25m
Heavy drizzle with rain	58	0	0
Light rain (<2.5mm/h)	61	913	3d 4h 5m
Moderate rain (<10mm/h)	62	29	0d 2h 25m
Heavy rain	64	0	0
Light rain, drizzle with snow	67	10	0d 0h 50m
Moderate rain, drizzle with snow	68	9	0d 0h 45m
Heavy rain, drizzle with snow	68	0	0
Light snow	71	156	0d 13h 0m
Moderate snow	72	32	0d 2h 40m
Heavy snow	73	3	0d 0h 15m
Light soft hail (<1.0mm/h)	87	0	0
Mod/Heavy soft hail	88	37	0d 3h 5m

9.3. Data Logging

In total, 15 sensors were used:

- Oxford RF radar + 4 other radars.
- 5 LiDARs (see Figure 6)
- 4 cameras (see Figure 5)
- 1 thermal imaging camera.

Most sensors recorded data every 5 minutes:

- 30 seconds data for radar and LiDAR.
- 25 frames for cameras.

NPL sensors logged every 30 s using the same logging tools used at Cardington.

Some sensors have not been 100% reliable, namely:

- One of the LiDARs needed frequent reboots.
- One of the radars got stuck frequently.
- One imaging camera has stopped working.

Figure 5: Image capture with the automotive camera.

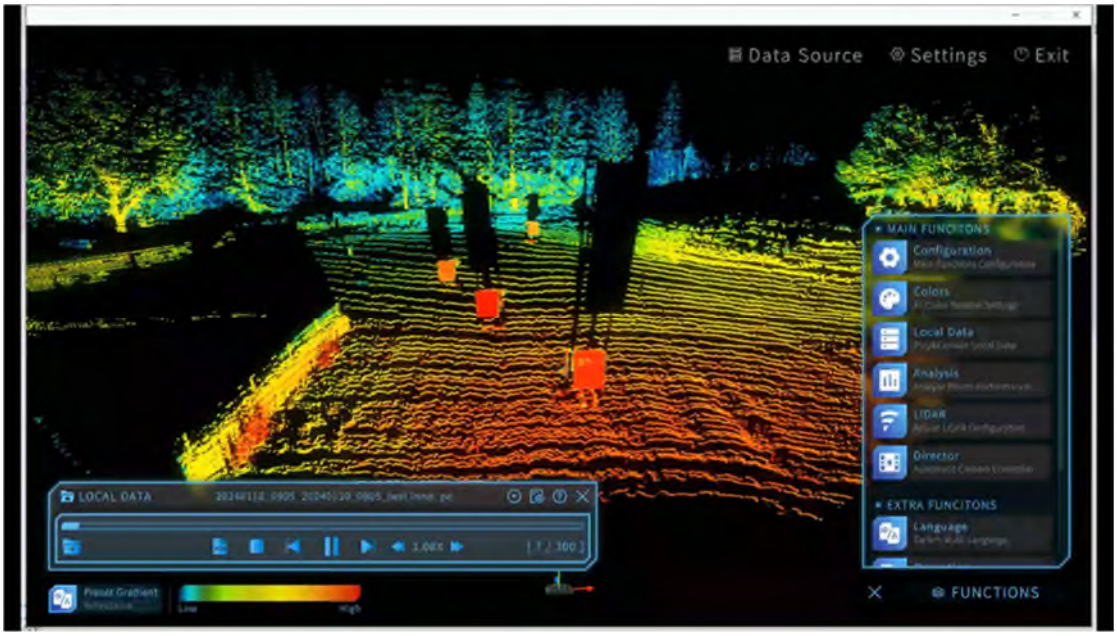


Figure 6: Point cloud visualisation for one of the LiDARs.

9.4. Robustness Analysis by Quantifying Sensor Failures

To analyse each sensor's ability to consistently output and store data, the output files – which were set to be generated each 5 minutes – were analysed for their file size. Files with zero or near-zero size were deemed to be empty or corrupted. The percentage of such files in the output directory was taken as an indication of sensor failures, which is an important characterisation of sensor robustness, or lack thereof.

▪ Automotive LiDAR A

LiDAR A point cloud data was saved every 5 minutes, with the file size of each output file plotted against the timeline in Figure 7. The number and percentage of files with file sizes in different ranges are given in Table 8.

These results show that the LiDAR A was failing a lot when outputting point cloud data, as almost 60% of all the files had their file size close to zero. When data outputting worked, there were two distinct file sizes of the point cloud: ~825 MB and ~360 MB, which also appears suspicious since there were no major changes in the scene.

These sensor failures suggest that this particular unit used in the campaign likely had a fault.

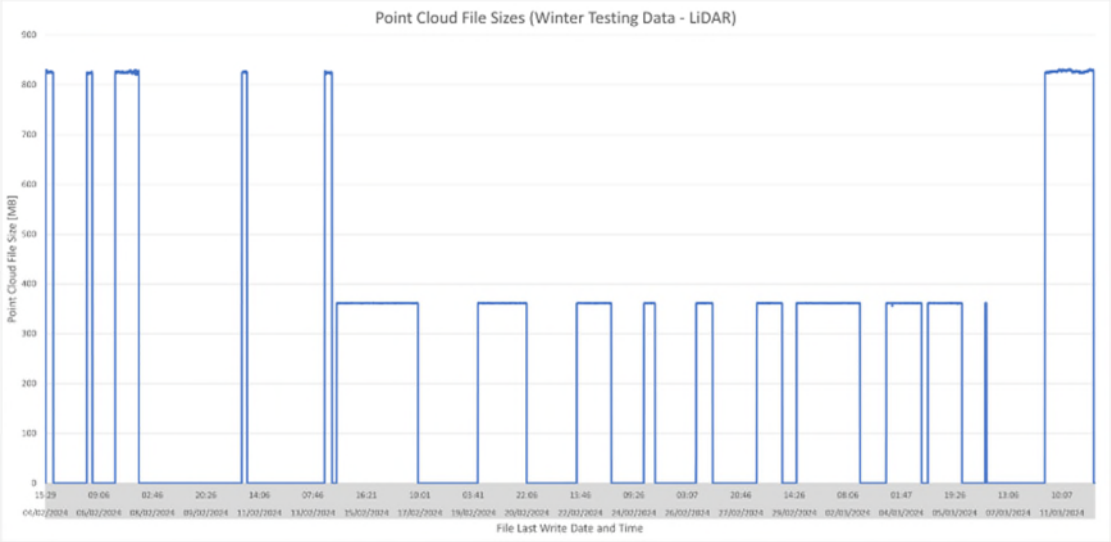


Figure 7: Point cloud file sizes over time.

Table 8: Point cloud failures (near-zero file-size point clouds).

Point Cloud File Size	Number of Files	Percentage
>800 MB	917	9.35%
0 MB < & < 800 MB	3292	33.56%
~0 MB	5600	57.09%
SUM	9809	100.0%

Automotive Visible Camera

RAW images of the automotive (visible light) camera were recorded every 5 minutes, with their file size plotted over time in Figure 9. The number and percentage of images with their file size falling into different ranges are presented in Table 10.

The camera failed very rarely - in only 0.04% cases - when the file size of the image was found to be close to zero. The average file size was 5.11 MB with a standard deviation of 0.10 MB (most images had the file size of exactly 5356800 Bytes = 5.11 MB).

Overall, this sensor was robust in terms of its ability to consistently output data files.

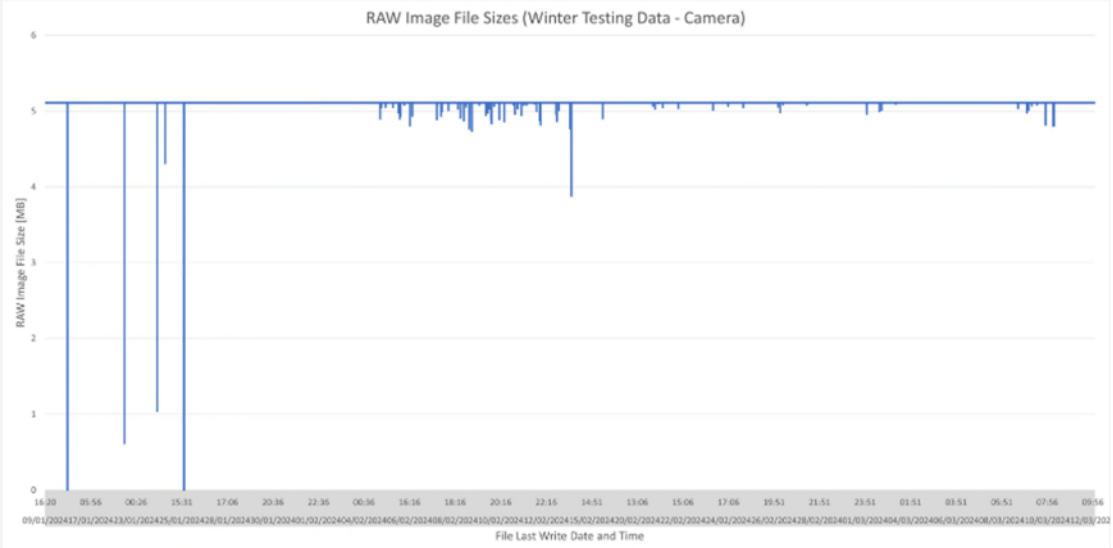


Figure 8: Image file sizes over time.

Table 9: Camera failures (near-zero file-size images).

Image File Size	Number of Files	Percentage
> 80 MB	17212	99.78%
0 MB < & < 80 MB	29	0.17%
~0 MB	9	0.05%
SUM	17250	100.00%

9.5. Qualitative Analysis of Inclement Weather and Other Noise Factors

- ### Automotive Visible Camera

Figure 10 shows the scene captured by the automotive camera during clement weather, identified by the weather sensor codes for no precipitation (SYNOP Code 00), as well as high visibility or high meteorological optical range (MOR 20000 m). This image was captured at 13:20, and it was taken as the baseline image. However, the image shows some presence of cloud cover, which has reduced the overall brightness and contrast in the scene.

Figure 9: Clement weather (no precipitation and good visibility).

Figure 11 was taken at 10am during no precipitation (SYNOP Code 00) but with somewhat reduced visibility (MOR 1192 m), indicating presence of mist or fog in some areas. This can be particularly noticed in the left part of the image over the sky and in the background where parts of the hill are washed out by a white layer of fog. Otherwise, the scene seems to be well illuminated by the sun with overall high contrast, especially in the area covered by the grass.

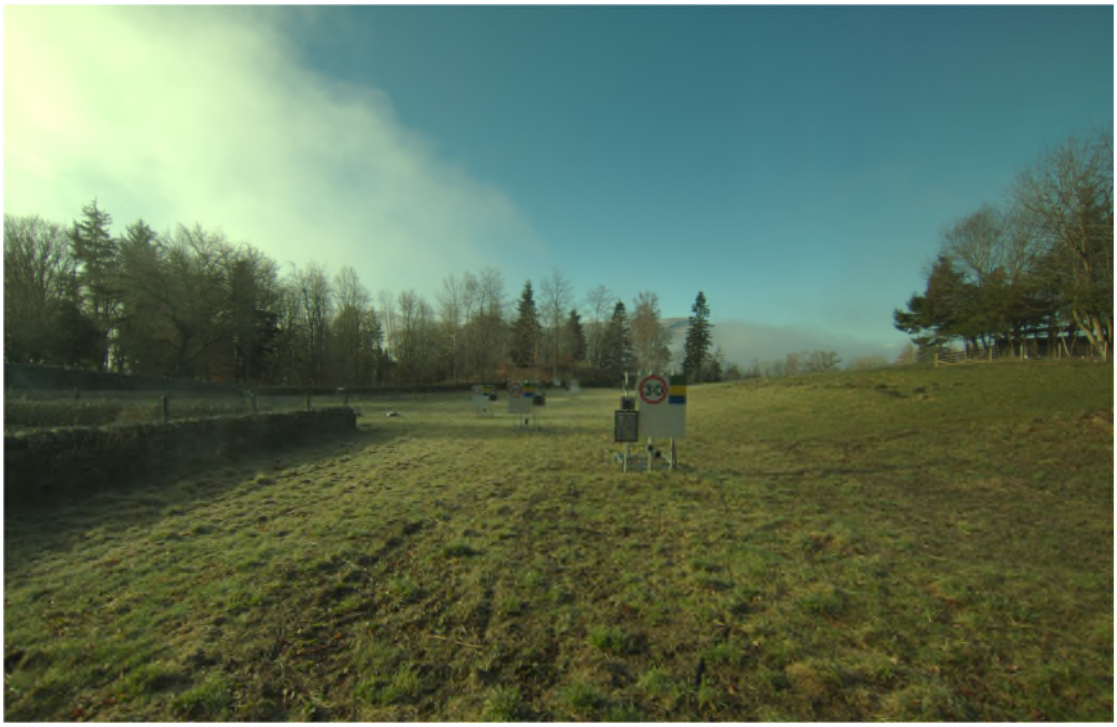


Figure 10: Mist or fog reduces the visibility of the hills in the background (1 March, 10:00).

Figure 12 was taken somewhat earlier on the same day when, due to the position of the sun, the trees on the left were casting a shadow that darkened a large part of the image.

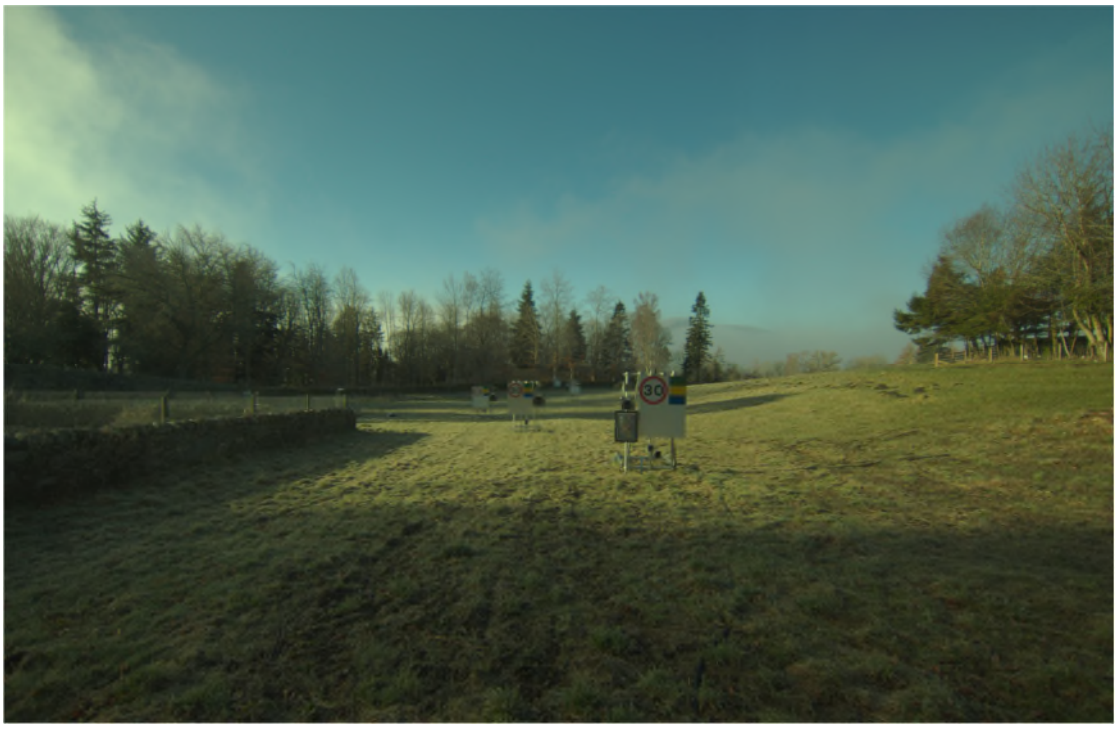


Figure 11: With the sun on the left, the trees are casting a shadow into the scene (1 March, 9:25).

As the sun moved forward into the scene, lens flare was observed (see Figure 13). Lens flare has the potential to obscure objects/targets in the scene in the areas where it appears. Figure 14 shows lens flare obscuring the zero on the speed limit sign.

Figure 12: The sun is causing lens flare (1 March, 11:45).

Figure 13: Lens flare is obscuring a part of the speed limit sign (2 March, 16:00).

During the campaign there were some days with snow, however, most of the snowfall occurred during night, so it was not possible to see the actual snowfall in any of the images. Figure 15 shows snow cover visible on the ground after a bout of light to medium snow. Snowflakes can latch onto the glass of the sensor enclosure, demonstrating their potential to create direct occlusions on sensor surfaces (see Figure 16).

Figure 14: Snow cover on the ground after a bout of light to medium snow (9 February, 10:05).

Figure 15: Snowflakes visible on the surface of the sensor enclosure (9 February, 11:45).

Figure 17 depicts the effects of light rain with droplets appearing on the sensor enclosure. These images were taken consecutively with a 5-minute distance between each of them. The images show that rain droplets can potentially obscure targets in the scene, but the problem is exacerbated even further when the sun peaks through the clouds, shining directly through the droplets which makes them appear whiter and opaquer.

Figure 16: Light rain causing droplets on the sensor enclosure (29 Feb, 13:20-13:55).

- ## Automotive LiDAR B

Due to LiDAR A producing erroneous outputs, the following evaluation was performed on data collected with a long-range automotive LiDAR B.

Figure 22 shows the visualisation of the point cloud recorded during no precipitation, resulting in a clear view of various targets in the scene. Figure 23 shows the visualisation of the point cloud recorded during a bout of heavy snow, as identified by the disdrometer's SYNOP Code 73. It can be seen that a lot of snowflakes are registered in the near range (10-20 m) as points of relatively low reflectivity.

This demonstrates that heavy snowfall has the potential to impact the quality of point cloud data, especially in the close range which is important for the detection of pedestrians, cyclists and other objects that may be interacting with vehicles in such a close proximity.

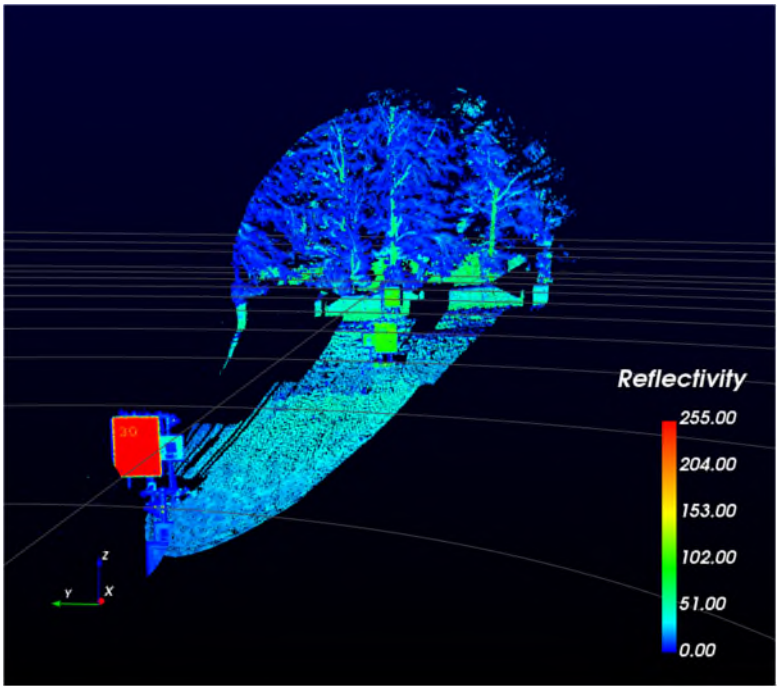


Figure 17: Visualisation of the point cloud during no snow (8 February, 20:15).

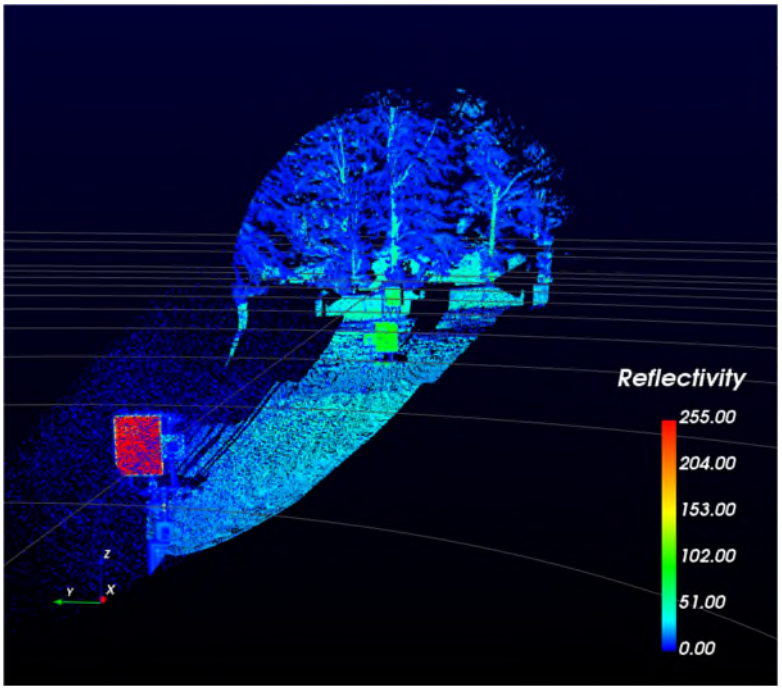


Figure 18: Visualisation of the point cloud during heavy snow (8 February, 20:40).

10. Quantitative Analysis of Winter Testing Campaign Data

In the Winter Testing Campaign data, we identified six datasets for image quality assessment, for the automotive visible light camera. Each datasets contains a baseline image captured during no precipitation, as well as several images captured immediately after the baseline (with 5 min time separation) during various adverse weather conditions. These conditions include: Hail, Light Rain, No Snow[1], Light Snow, Moderate Snow, and Heavy Snow. The relationship between each baseline and weather conditions captured in the period immediately following each baseline are shown in Table 1.

Table 10 Adverse Weathers in Each Baseline (N/A represents no paired image, otherwise index of the paired image is provided.)

Adverse Weather / Baseline	Hail	Light Rain	No Snow	Light Snow	Moderate Snow	Heavy Snow
Baseline A	**A1**	N/A	N/A	**A2**	N/A	N/A
Baseline B	N/A	N/A	N/A	**B1-B5**	N/A	N/A
Baseline C	N/A	N/A	N/A	**C1-C4**	N/A	N/A
Baseline D	N/A	**D1-D7**	N/A	N/A	N/A	N/A
Baseline E	N/A	N/A	**E1**	**E3**	**E2, E4, E6-8**	**E5**
Baseline F	**F2, F6-7**	N/A	**F1, F3**	**F4-5, F8**	N/A	N/A

These datasets obtained with the visible light camera are presented in Appendix I. Note that the images have been lightened up for presentation purposes in this report, while the analysis has been undertaken using the original images as taken by the cameras.

[1] The 'No Snow' condition means that there was no active precipitation in the moment of capture. However, there may have been precipitation 5 minutes before and/or after. This means that the 'No Snow' photos are not necessarily representative of baseline conditions as there might be snowflake occlusion present from a previous bout of snow.

10.1. Image Quality Assessment (IQA)

Image Quality Assessment (IQA) is the process of evaluating and measuring the perceptual quality of images. Two main approaches are Subjective IQA and Objective IQA. The Subjective IQA is a human evaluation of images using subjective scores provided by people, based on their preferences or perceived quality. The Objective IQA, on the other hand, is computer based, using various algorithms to compute quantitative metrics.

The Objective IQA Metrics can be further divided into the following types:

- Full Reference (FR) – Requires the original/pristine version of the image that serves as a reference for computing the metric for the target/distorted image.
- No Reference (NR) – Does not require a reference image to compute the metric for the target/distorted image.
- Reduced Reference (RR) – While the reference image is not required, some characteristic information about the original image must be provided.

Table 11 Full Reference (FR) IQA Metrics

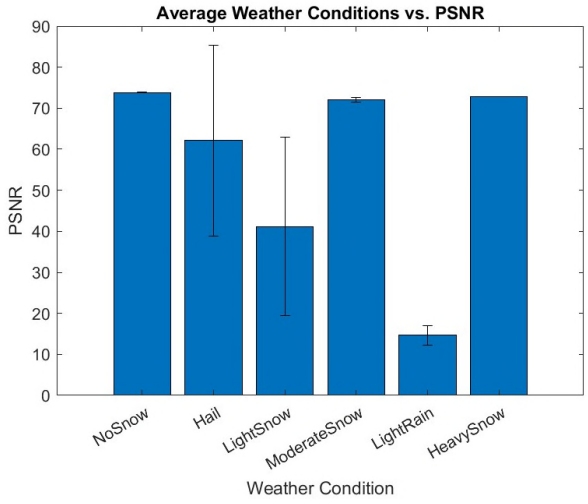
Metric Acronym	Year Introduced	Metric Name & Reference
MSE	-	Mean Squared Error
PSNR	-	Peak Signal-to-Noise Ratio
SSIM	2004	Structural Similarity
MS-SSIM	2004	Multi-Scale Structural Similarity
IW-SSIM	2011	Information Content Weighted Structural Similarity Index
CW-SSIM	2009	Complex Wavelet Structural Similarity Index
VIF	2004	Visual Information Fidelity
FSIM	2011	Feature Similarity Index Measure
SR-SIM	2012	Spectral Residual Based Similarity
GMSD	2013	Gradient Magnitude Similarity Deviation
MS-GMSD	2017	Multi-Scale Gradient Magnitude Similarity Deviation
VSI	2014	Visual Saliency-induced Index
DSS	2015	DCT Subband Similarity Index
HaarPSI	2016	Haar Perceptual Similarity Index
MDSI	2016	Mean Deviation Similarity Index

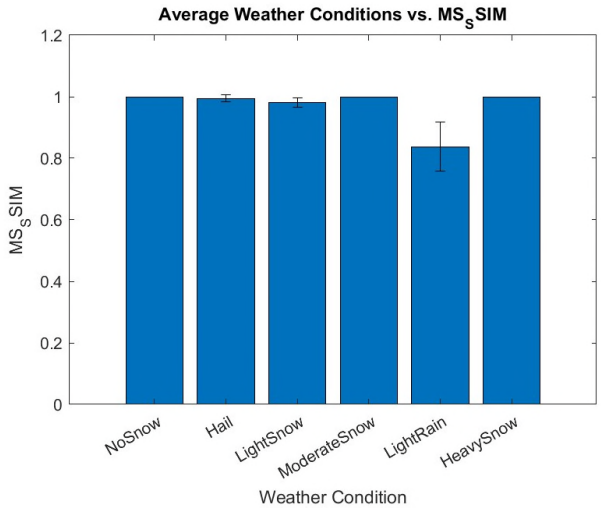
Table 12 No Reference (NR) IQA Metrics

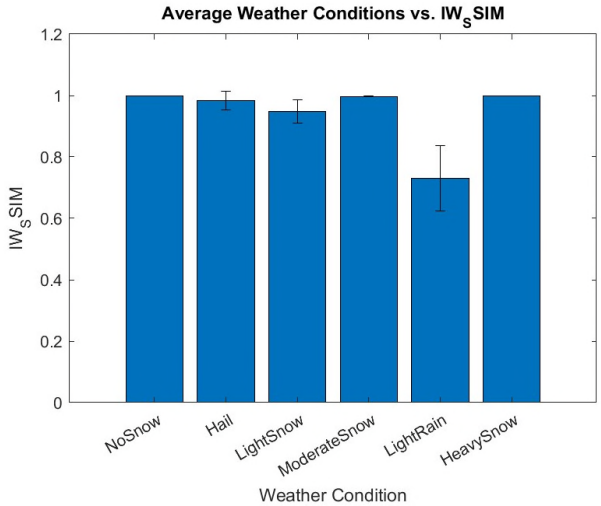
Metric Acronym	Year Introduced	Metric Name & Reference
BRISQUE	2012	Blind/Referenceless Image Spatial Quality Evaluator
NIQE	2013	Natural Image Quality Evaluator
PIQE	2015	Perception-based Image Quality Evaluator

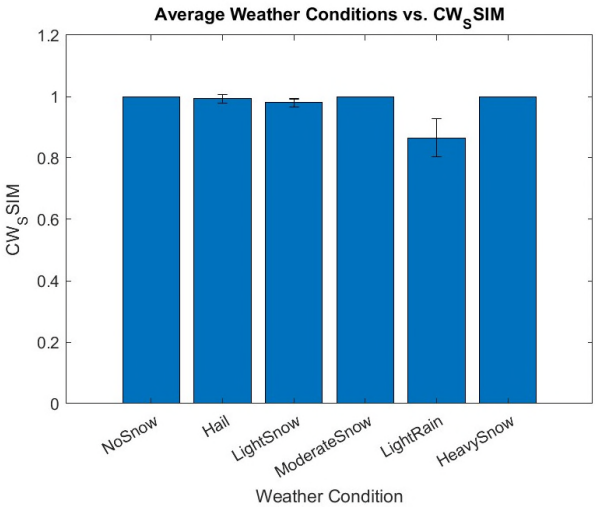
10.2. Automotive Visible Camera Analysis

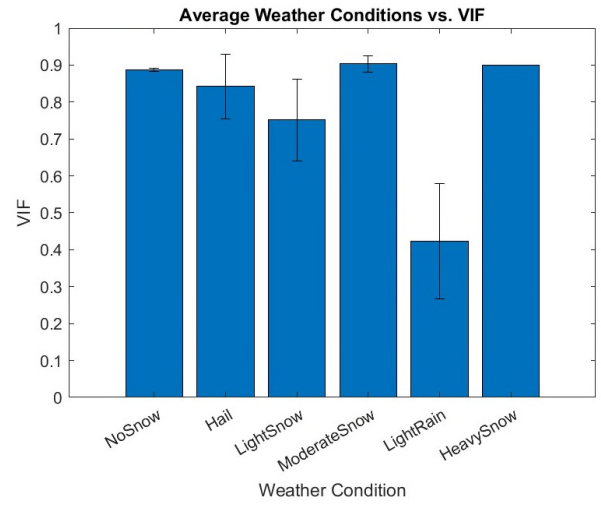
- Average Visuals (FR & NR IQA Metrics)

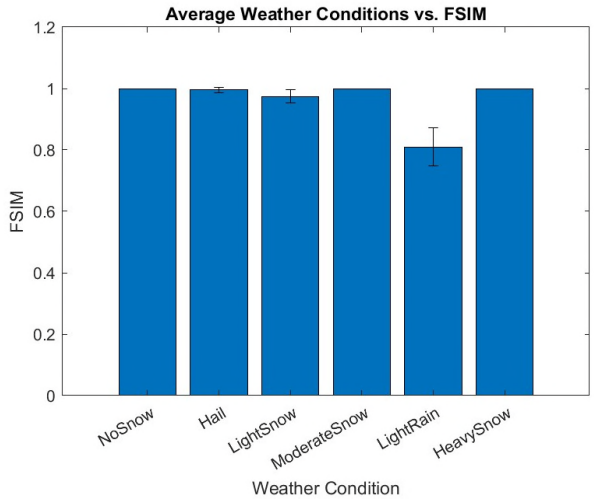


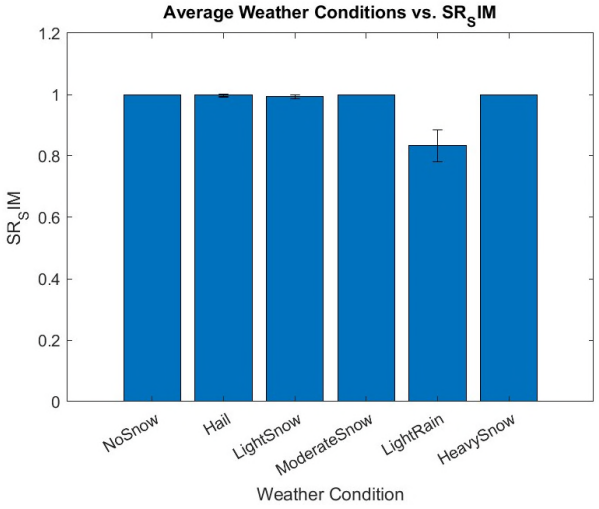


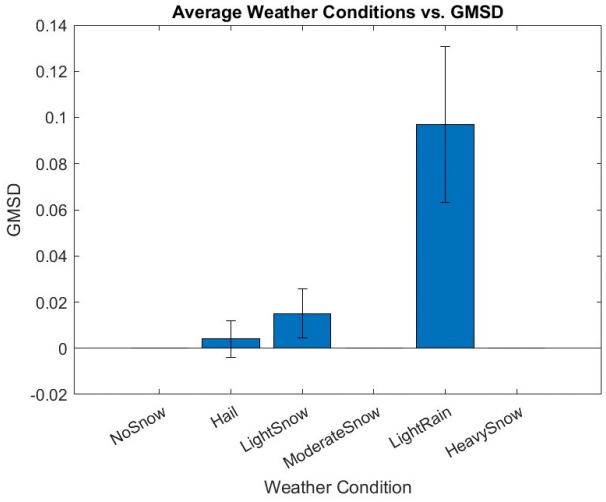


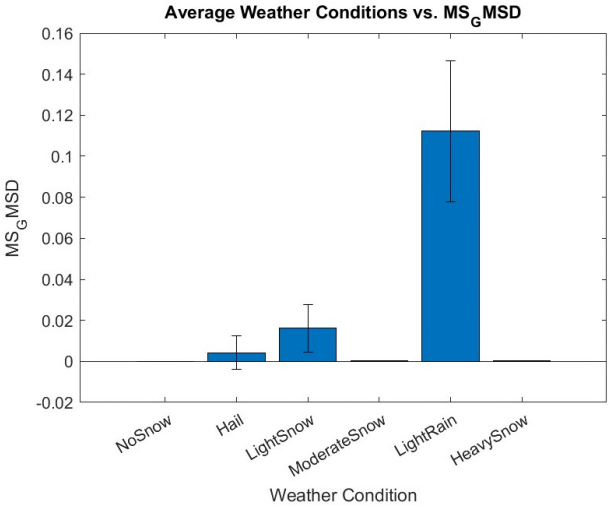


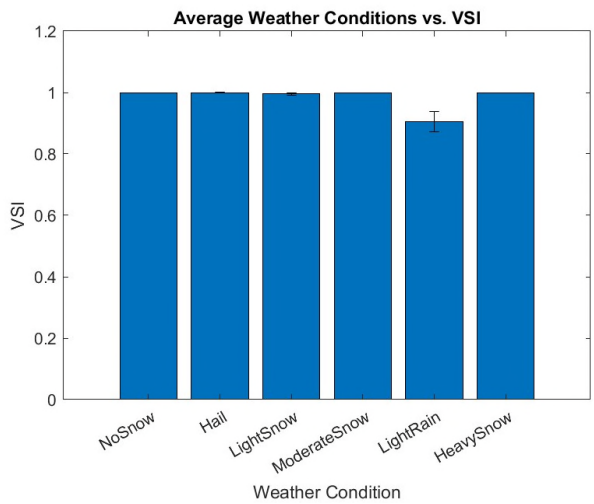

Average Weather Conditions vs. VIF
VIF
Weather Condition
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow

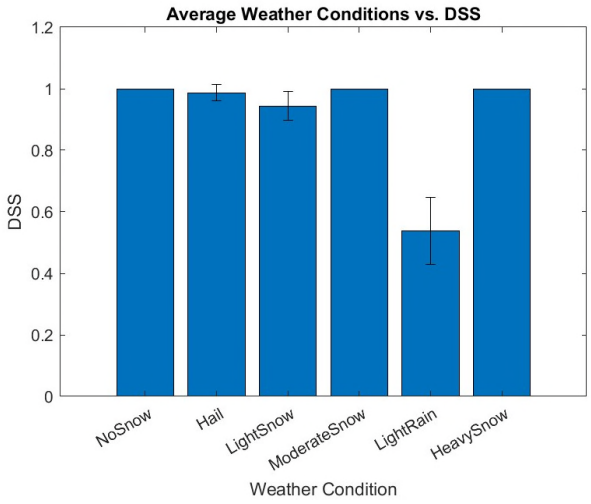

Average Weather Conditions vs. FSIM
FSIM
Weather Condition
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow

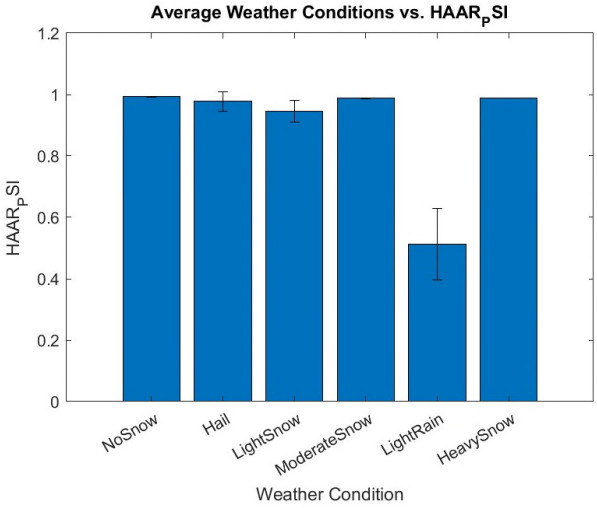

Average Weather Conditions vs. SR_SIM
SR_SIM
Weather Condition
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow

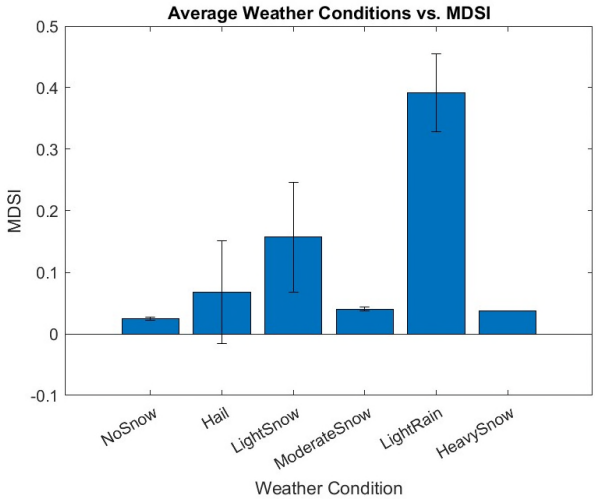

Average Weather Conditions vs. GMSD
GMSD
Weather Condition
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow

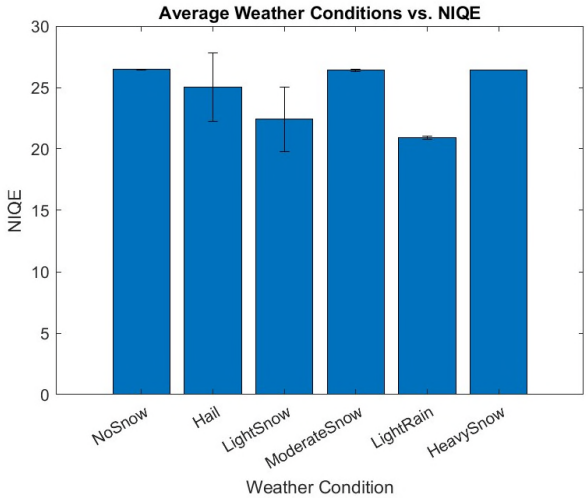

Average Weather Conditions vs. MS_GMSD
MS_GMSD
Weather Condition
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow

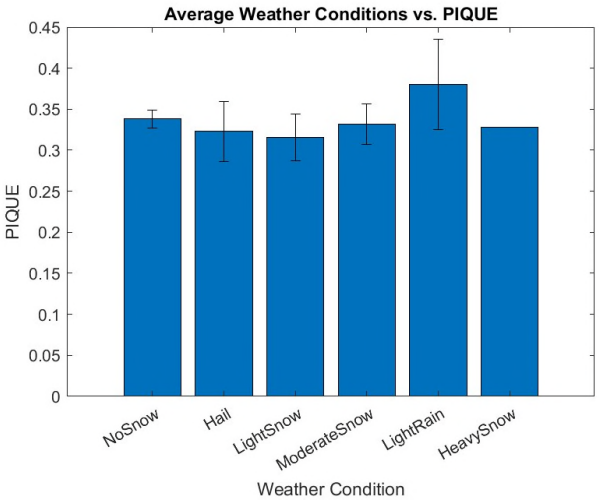

Average Weather Conditions vs. VSI
VSI
Weather Condition
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow


Average Weather Conditions vs. DSS
DSS
1.2
1
0.8
0.6
0.4
0.2
0
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition

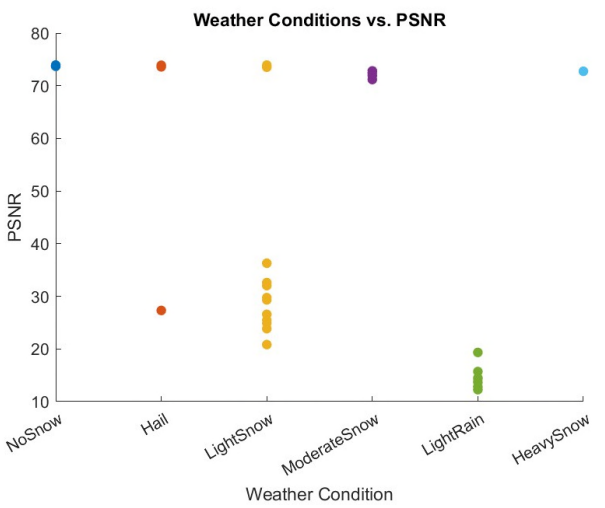

Average Weather Conditions vs. HAARpSI
HAARpSI
1.2
1
0.8
0.6
0.4
0.2
0
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition

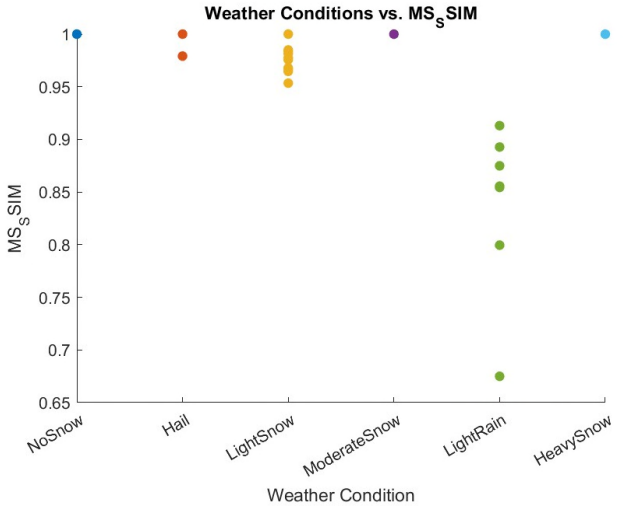

Average Weather Conditions vs. MDSI
MDSI
0.5
0.4
0.3
0.2
0.1
0
-0.1
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition

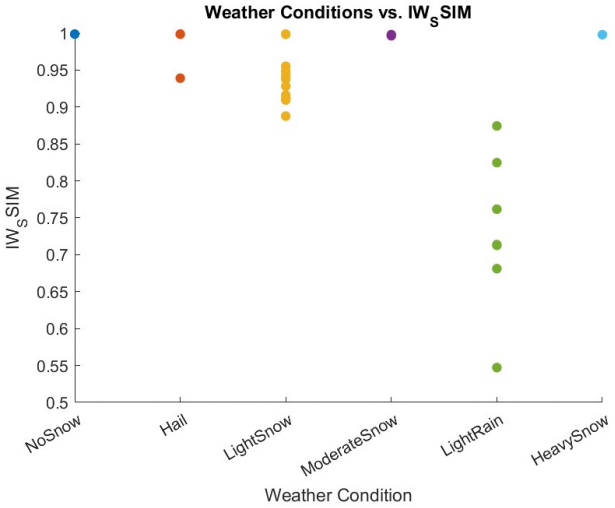

Average Weather Conditions vs. NIQE
NIQE
30
25
20
15
10
5
0
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition

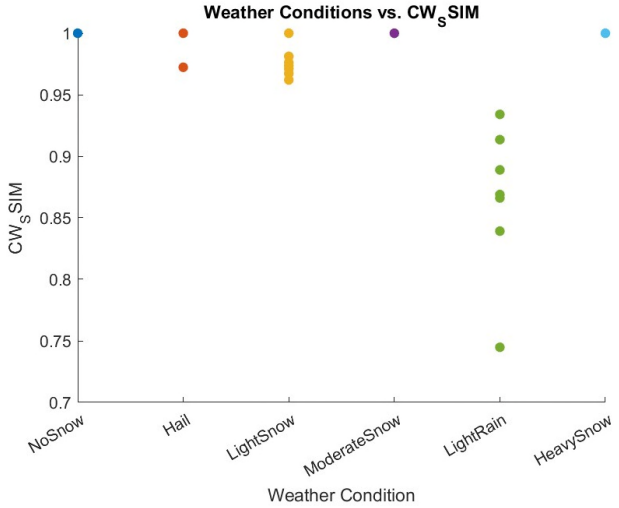

Average Weather Conditions vs. PIQUE
PIQUE
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition

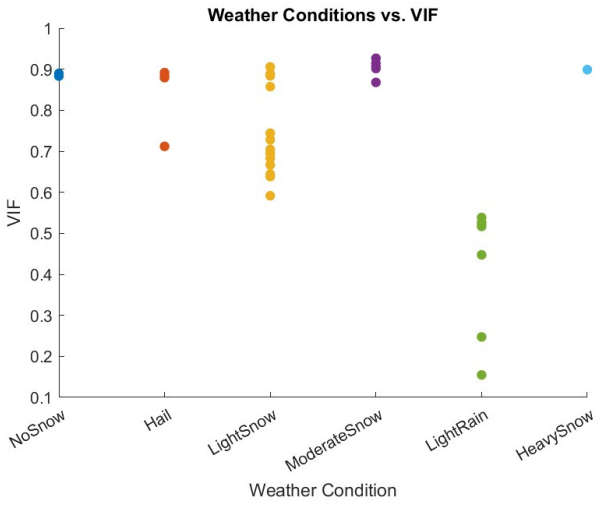
- Cluster Visuals (FR & NR IQA Metrics)

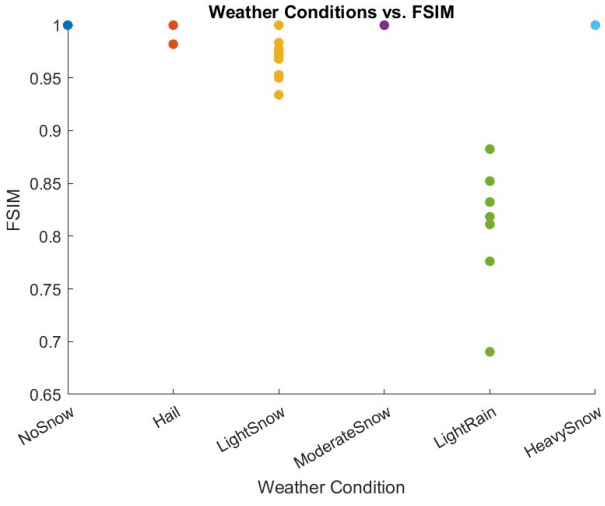


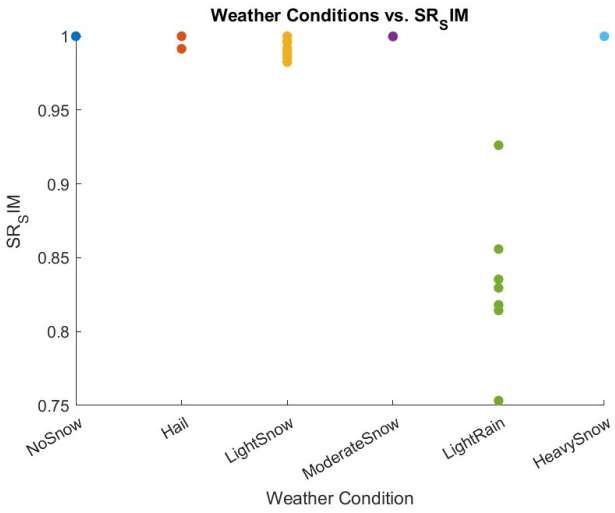


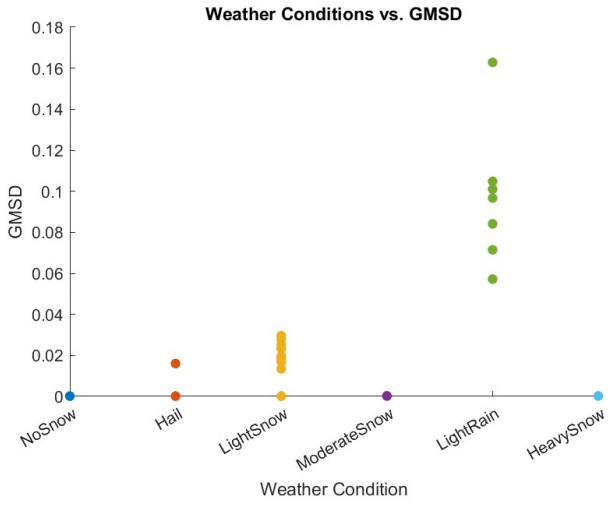


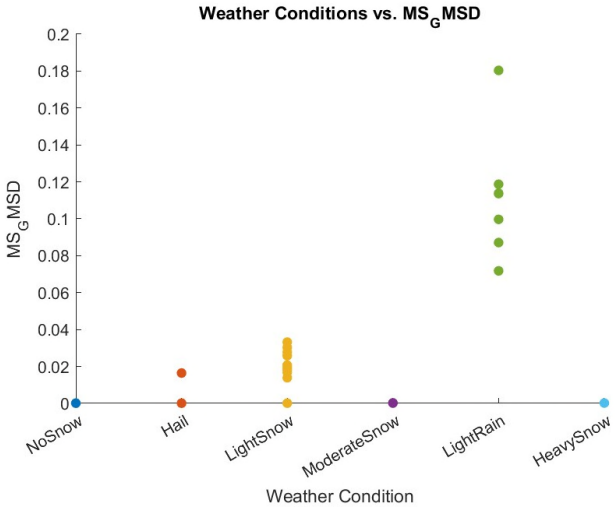


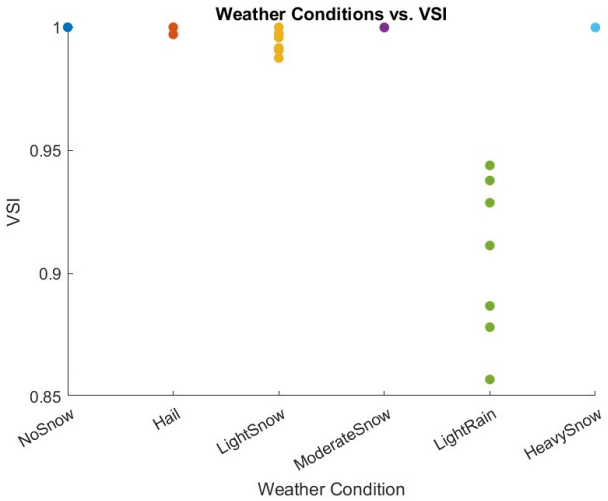


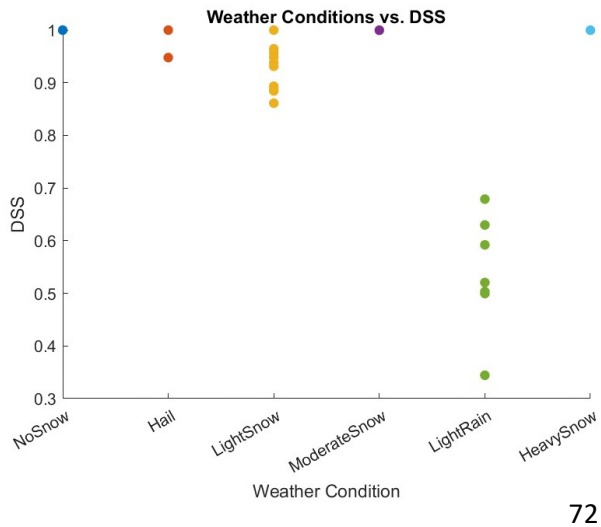


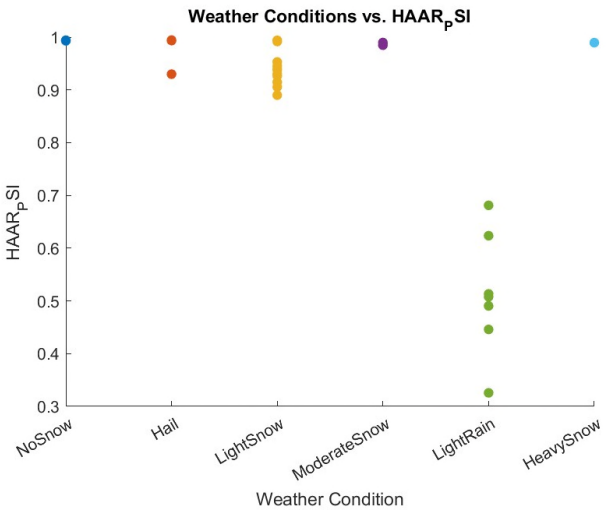

Weather Conditions vs. SR_SIM
SR_SIM
1
0.95
0.9
0.85
0.8
0.75
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition

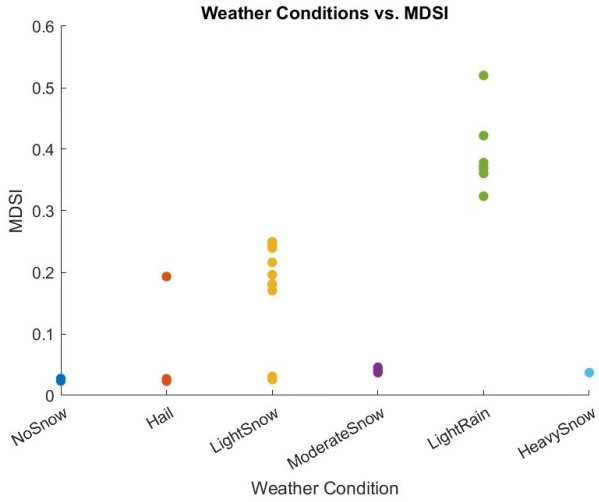

Weather Conditions vs. GMSD
GMSD
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition

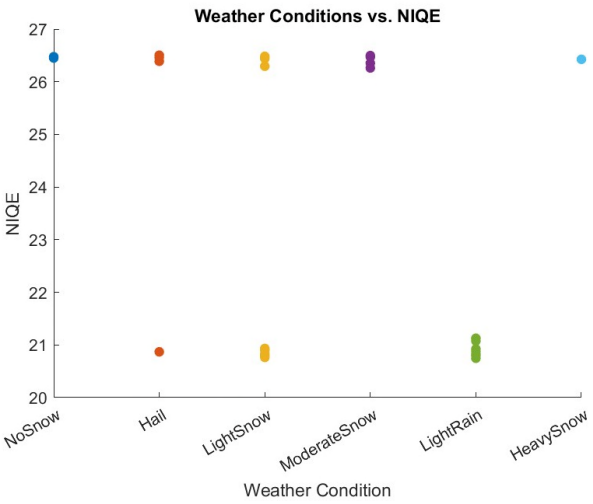

Weather Conditions vs. MS_GMSD
MS_GMSD
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition

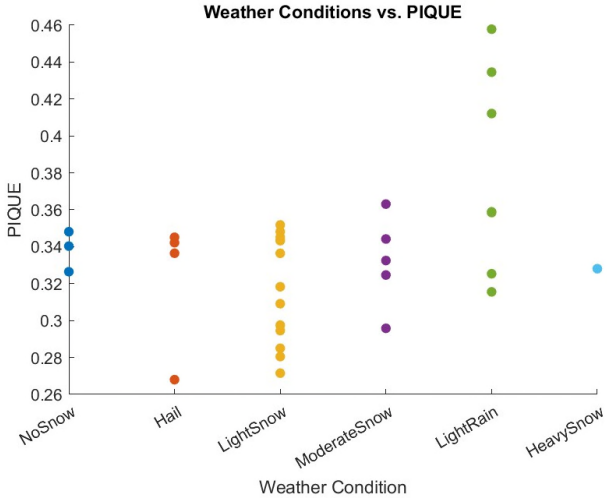

Weather Conditions vs. VSI
VSI
1
0.95
0.9
0.85
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition


Weather Conditions vs. DSS
DSS
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition


Weather Conditions vs. HAAR_PSI
HAAR_PSI
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
NoSnow
Hail
LightSnow
ModerateSnow
LightRain
HeavySnow
Weather Condition







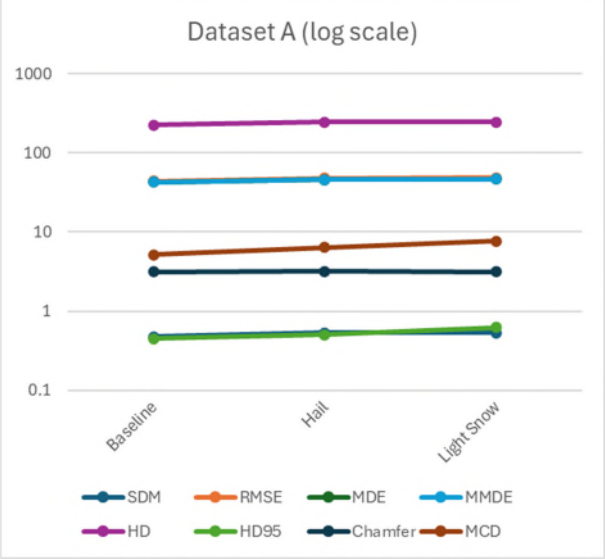
The results of IQA of the visible camera datasets show that various types of precipitation can have an impact of image quality, as detected by various full reference and no reference metrics. The biggest difference in image quality was observed for 'Light Rain'. However, the dataset containing the occurrences of light rain was also impacted by the sun, changing levels of cloud cover, and water droplets on the sensor enclosure. It was therefore not possible to disentangle the effect of light rain from these other noise factors.

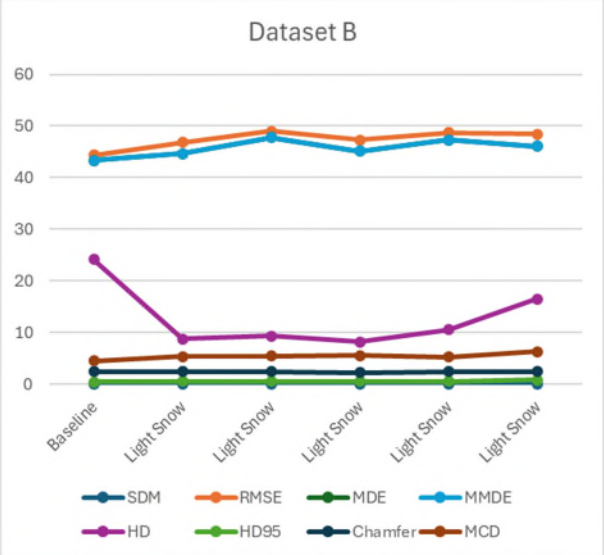
Images taken during 'Moderate Snow' and 'Heavy Snow' did not have major decline in image quality as they were largely black due to being taken during the night. The winter testing dataset did not have instances of moderate or heavy snow during daylight.

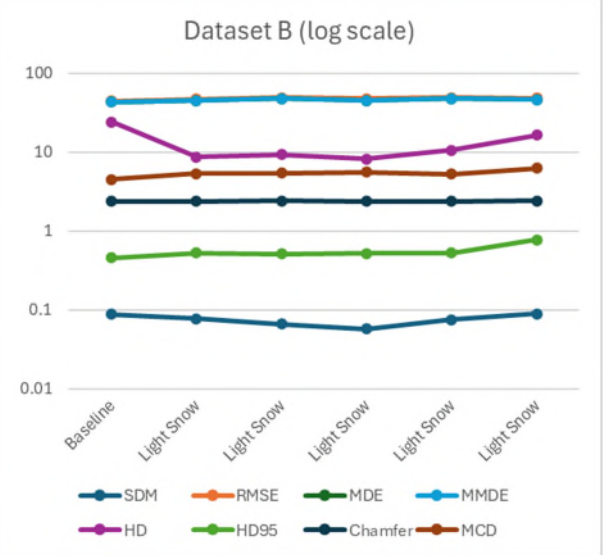
These results highlight the challenges of evaluating the quality of images that were taken in an outside environment due to the contemporaneous effects of multiple noise factors present in that environment. In particular, time of day, position of the sun, and cloud cover, seem to have major impact on visible-light-camera images, which can overshadow the effects of various types of precipitation (rain, snow, hail).

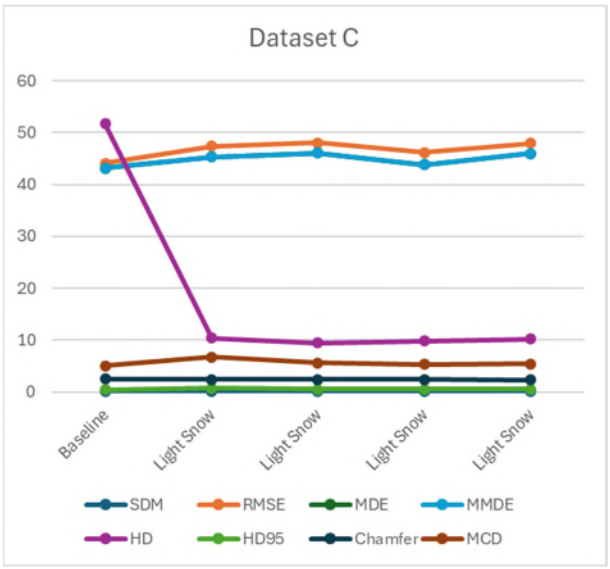
10.3. LiDAR Point Cloud Analysis

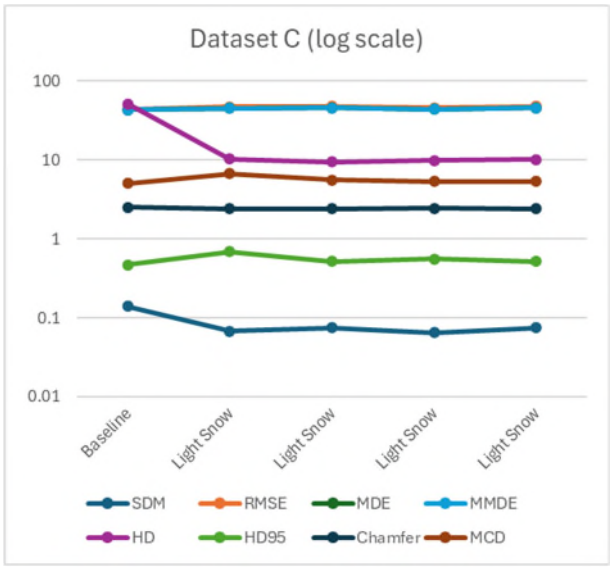
- Point Cloud Quality Assessment (PCQA)

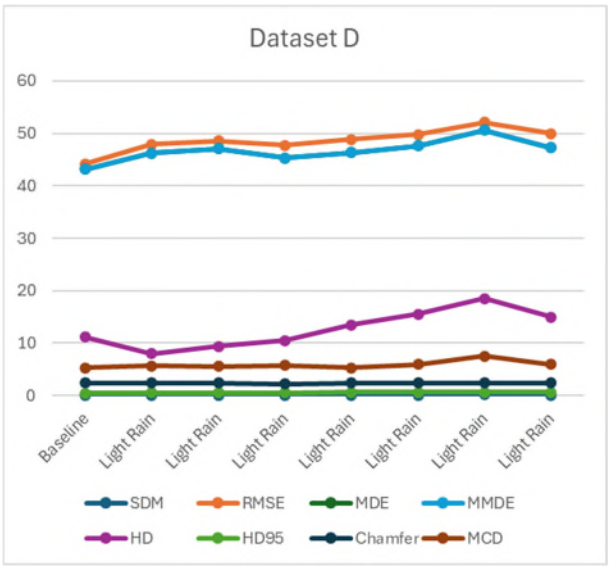


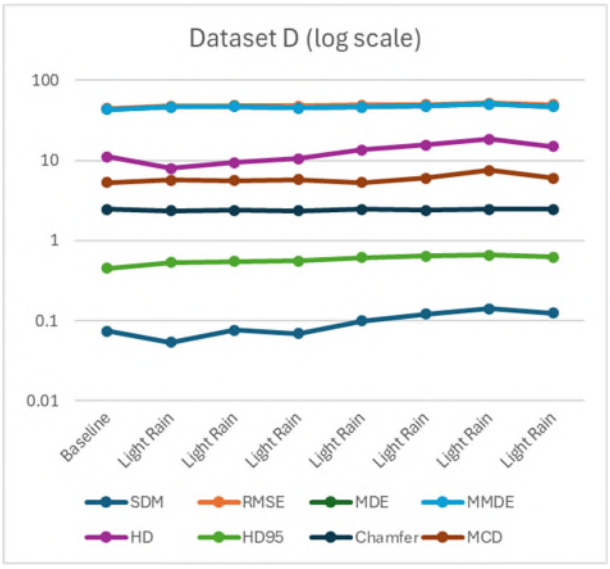


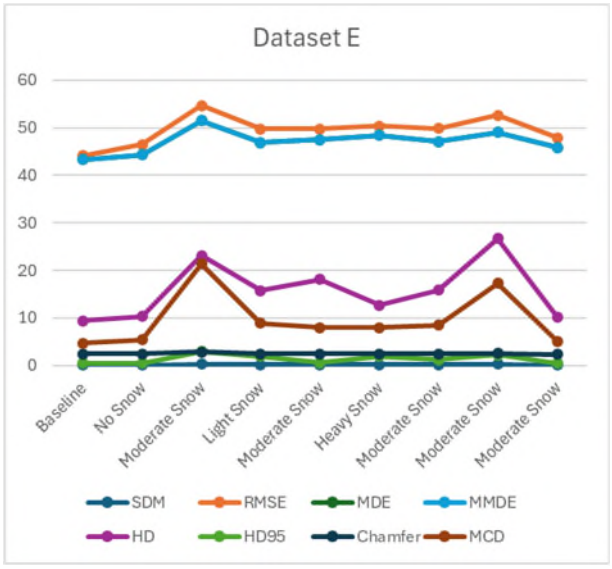


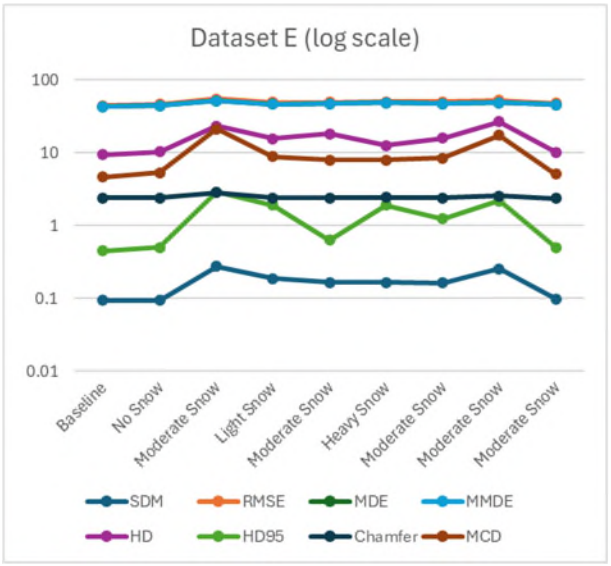

Dataset C
60
50
40
30
20
10
0
Baseline
Light Snow
Light Snow
Light Snow
Light Snow
SDM
RMSE
MDE
MMDE
HD
HD95
Chamfer
MCD

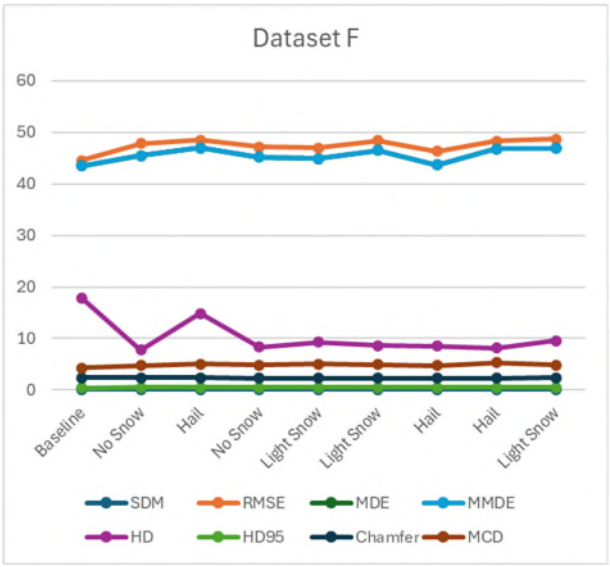

Dataset C (log scale)
100
10
1
0.1
0.01
Baseline
Light Snow
Light Snow
Light Snow
Light Snow
SDM
RMSE
MDE
MMDE
HD
HD95
Chamfer
MCD

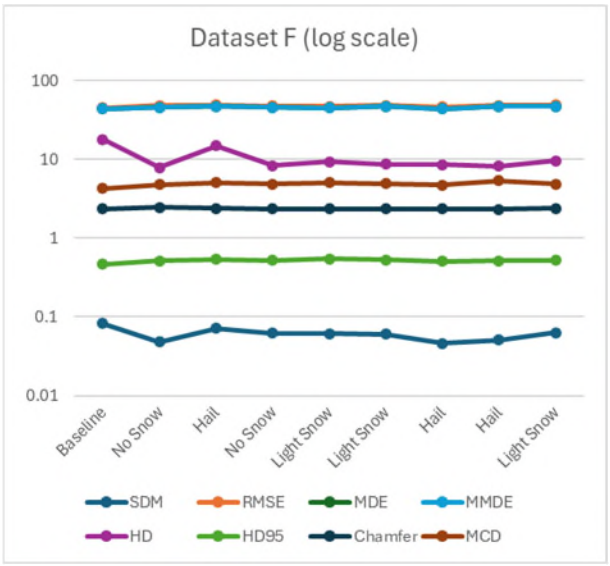

Dataset D
60
50
40
30
20
10
0
Baseline
Light Rain
Light Rain
Light Rain
Light Rain
Light Rain
Light Rain
Light Rain
SDM
RMSE
MDE
MMDE
HD
HD95
Chamfer
MCD


Dataset D (log scale)
100
10
1
0.1
0.01
Baseline
Light Rain
Light Rain
Light Rain
Light Rain
Light Rain
Light Rain
Light Rain
SDM
RMSE
MDE
MMDE
HD
HD95
Chamfer
MCD









The Full Reference PCQA metrics were applied to six point cloud datasets, each containing between three and nine point clouds. These datasets correspond to the image datasets obtained with the visible light camera at the same time. Point clouds collected during various types of precipitation were compared to their respective baselines when there was no precipitation.

The general finding is that during various types of precipitation, various metrics tend to have higher values than for the baseline, although there are some exceptions. This suggests that precipitation can be detected in the point cloud, which leads to larger discrepancies from the baseline point cloud that served as the reference. Other possible explanation for discrepancies stems from the fact that point clouds are different because the scene is not entirely static, with likely movement in the grass, trees, branches, and leaves.

11. Insights from the 2nd Winter Testing Campaign

The 2nd Winter Testing Campaign has collected more outdoor data in another location, using a slightly different sensor stack and targets. This has resulted in additional datasets which have led to novel insights that are presented in this section.

11.1. Image Quality Assessment for Visible Camera

To investigate the impact of inclement weather on image quality, a sequence of 20-60 consecutive images has been identified for each of the following conditions, based on the synoptic codes from the present weather sensor (disdrometer), the visibility data, and the image time stamps:

- Baseline day – No precipitation and maximum visibility; during daylight hours;
- Rain day – Rain intensity ranging from light rain to heavy rain; during daylight hours;
- Snow day – Snow intensity from light snow to heavy snow; during daylight hours;
- Low visibility – Part of the dataset when visibility was among the lowest in the dataset (but still above 400 m); during daylight hours.
- Baseline night – No precipitation and maximum visibility; during nighttime;
- Hail evening – Hail identified by the present weather sensor; during evening hours.

The image quality assessment (IQ) was completed using the no-reference metrics presented in Table 13, namely the BRISQUE, NIQE, and PIQE. Figure 24 presents the box plot of the BRISQUE metric for identified weather conditions. The lowest score (the highest image quality) was computed for the rain sequence, while the highest score (the lowest image quality) was computed for the baseline night sequence. Given that the applied BRISQUE model was trained on an image database with common image distortions such as compression artifacts, blurring, and noise, it is not surprising that it did not perform very well on different types of distortions caused by environmental factors.

The box plot for the NIQE metric is presented in Figure 25. NIQE model was trained on a database of pristine images of natural scenes, however, it can be applied to images with arbitrary distortions. With this metrics, the lowest score (the highest image quality) was obtained for baseline day sequence and low visibility day sequence, while the highest score (the lowest image quality) was obtained for baseline nighttime sequence. There is some decline in image quality for rain and snow, compared to day baseline.

Finally, the results for the PIQE score are presented in Figure 26. This model is untrained and can be used on arbitrary types of distortions. The results with this metric are the most logical, with the lowest score (the highest image quality) obtained for baseline day sequence, and the highest score (the lowest image quality) obtained for baseline nighttime sequence.

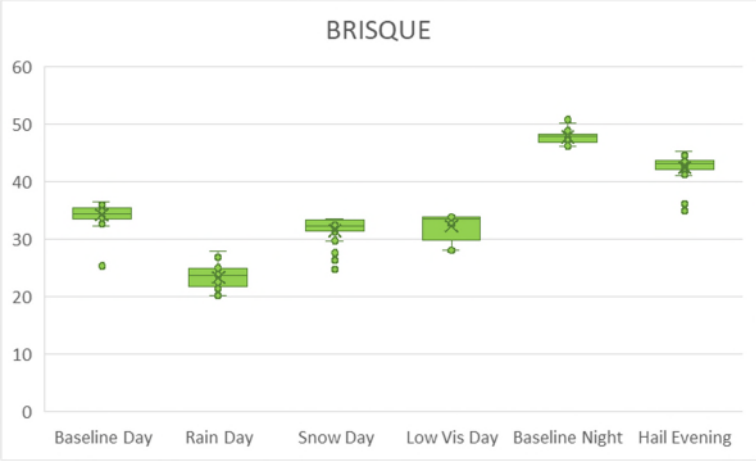


Figure 19: Image Quality Assessment using the BRISQUE metric.

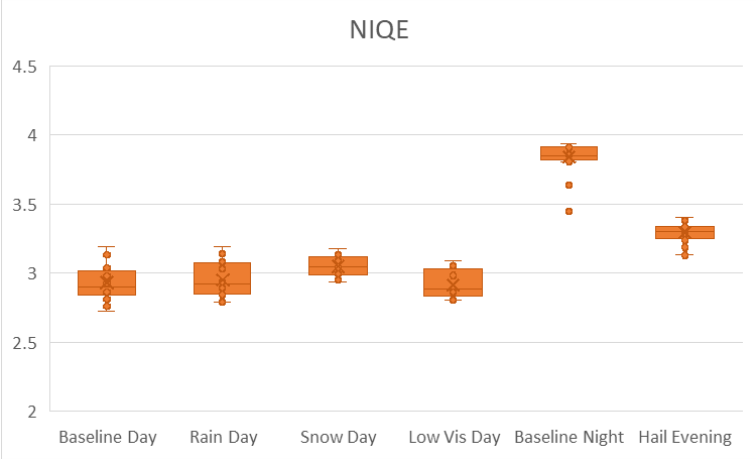


Figure 20: Image Quality Assessment using the NIQE metric.

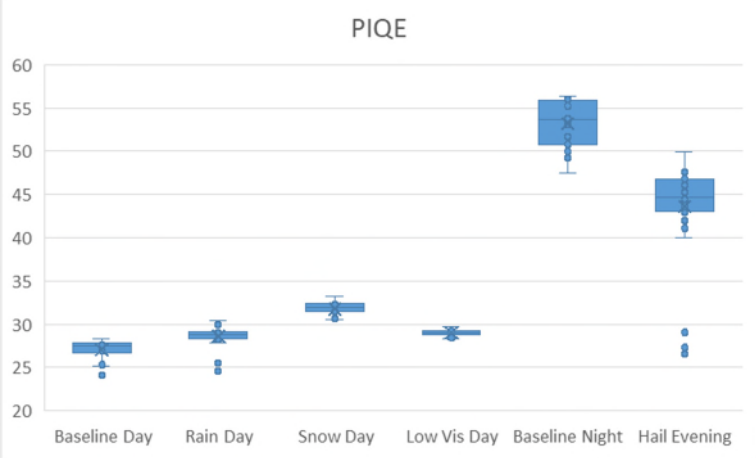


Figure 21: Image Quality Assessment using the PIQE metric.

11.2. Traffic Sign Recognition from Visible Camera

To study the potential impact of adverse weather on object detection, first we identified the baseline image and images with more prominent levels of precipitation, as identified by the weather sensors.

Identifying the baseline image is always a challenge as there is no such a thing as ideal environmental conditions (for example, the position of the sun and the level of cloud cover are always impacting the image), so the baseline was selected during daylight when there was no precipitation (based on the disdrometer) and the visibility was maximum (based on the two sensors that provide visibility) – see Figure 27.

For inclement weather examples, Figure 28 was taken in daylight during moderate (to heavy) snow. In the 1st Winter Testing Campaign, the only instance of falling snow was during the night, so there were no useful images taken with the visible light camera. In the 2nd Winter Testing Campaign, there were many examples of snow precipitation captured during daylight. When zooming into the figure with moderate (to heavy) snow, the snowflakes are visible in front of the target with the speed limit (see Figure 32). The zoom also reveals shot noise in RGB colours, which is present in all images.

Figure 29 was taken in daylight during heavy rain, and the image also shows the presence of water droplet occlusion. This occurred despite the cover over the sensors, probably due to wind. Figure 30 was recorded in the moment of minimum visibility in the entire dataset (99 metres) and during moderate snow. Since some of the targets were less than 99 metres away from the sensors, they are still visible. Finally, Figure 31 was capture during hail at nighttime.

Figure 22: Baseline (during daylight, no precipitation, and maximum visibility).

Figure 23: Moderate to Heavy Snow

Figure 24: Heavy Rain

Figure 25: Lowest Visibility (99 m) & Moderate Snow

Figure 26: Hail at Night

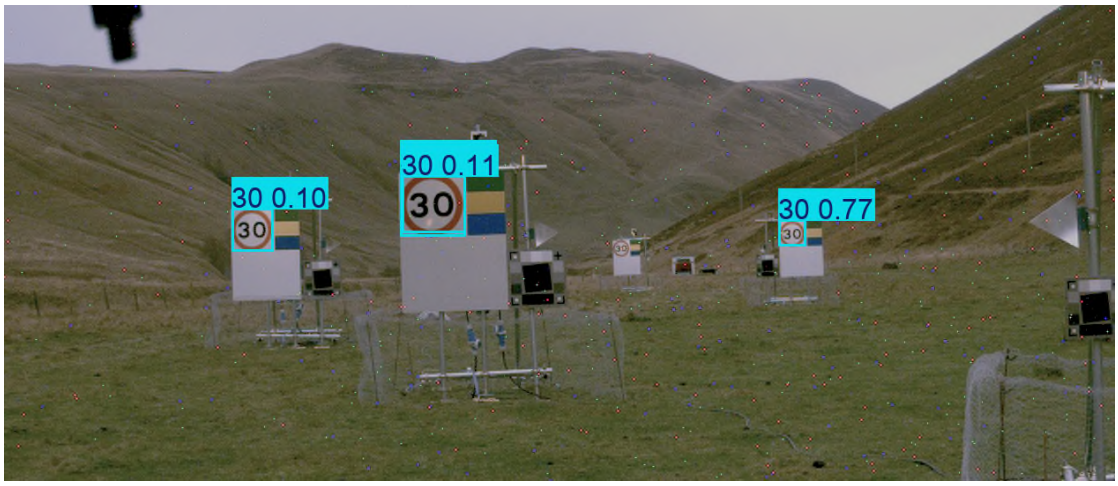
Figure 27: Moderate to Heavy Snow with visible snowflakes in front of the target (zoomed in).

For the traffic sign recognition, we employed a pre-trained open-source model, based on the YOLOv8 architecture and trained on a Turkish dataset of 30,000+ labelled images of traffic signs[2]. The result of object detection including confidence are presented in Figure 33.

In the baseline image, three out of four speed limit signs were recognised. In moderate to heavy snow, only two of the four traffic signs are recognised, while in heavy rain with droplet occlusion, only one traffic sign was detected. In the low visibility with moderate snow case, three traffic signs were detected. Although the visibility was the lowest in the entire dataset (99 m), this was still sufficient to detect the traffic sings closer to the camera. Furthermore, in this example, the confidence for the two closer signs were higher than in the baseline case. It is not obvious why that might be the case; perhaps due to traffic signs being placed on a larger white board which blends better with the background in the low visibility case. Finally, in the hail instance captured in the night, red and green pixels were falsely recognised as traffic lights, demonstrating a negative impact of shot noise on object detection.

Although retraining the model with our own dataset and pre-processing the images to remove the shot noise would likely improve the object detection performance, the results presented here demonstrate that adverse weather and other noise factors can indeed have a profound effect on perception using automotive camera.

[2] https://huggingface.co/nezahatkorkmaz/traffic-sign-detection

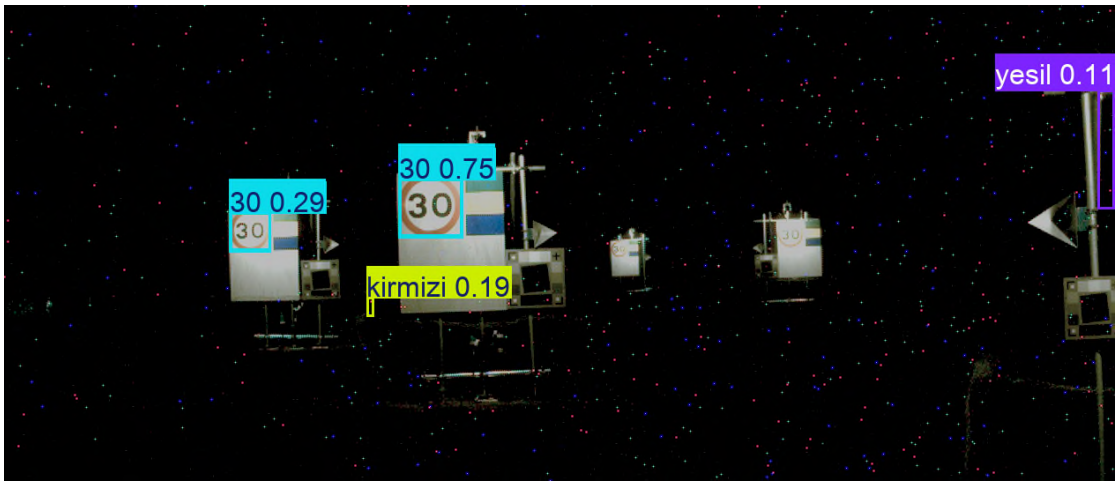


(a)

(b)

(c)

(d)



(e)

Figure 28: The result of object detection on selected images: (a) Baseline; (b) Moderate to Heavy Snow; (c) Heavy Rain; (d) Low visibility (99 m) and Moderate Snow; (e) Hail at night.

11.3. LiDAR Blooming Effect

LiDARs can be susceptible to the 'blooming effect', which can occur when highly reflective objects saturate the sensor, making the returned signal appear distorted and larger than the actual object. The blooming effect was noticed during the 2nd Winter Testing Campaign in the point cloud of one of the deployed LiDARs as an 'aura' of ghost points around a retroreflective target. On a specific occasion, it was noticed that the blooming effect was particularly strong, manifesting as a wall of ghost points extending several metres around the retroreflective target (see Figure 34).

This blooming effect appeared to be amplified in high humidity likely due to increased scattering of light on water molecules (of water vapour), as well as on any water particles that might be suspended in air. The weather station registered the relative humidity of 94% when the point cloud shown in Figure 34 was recorded. In addition to scattering, other possible causes of LiDAR beam broadening and spatial errors include local atmospheric turbulences that can induce fluctuations in the refractive index of air. All of these are examples of how multiple noise factors can work synergistically to boost each other's detrimental effects on LiDAR data quality.

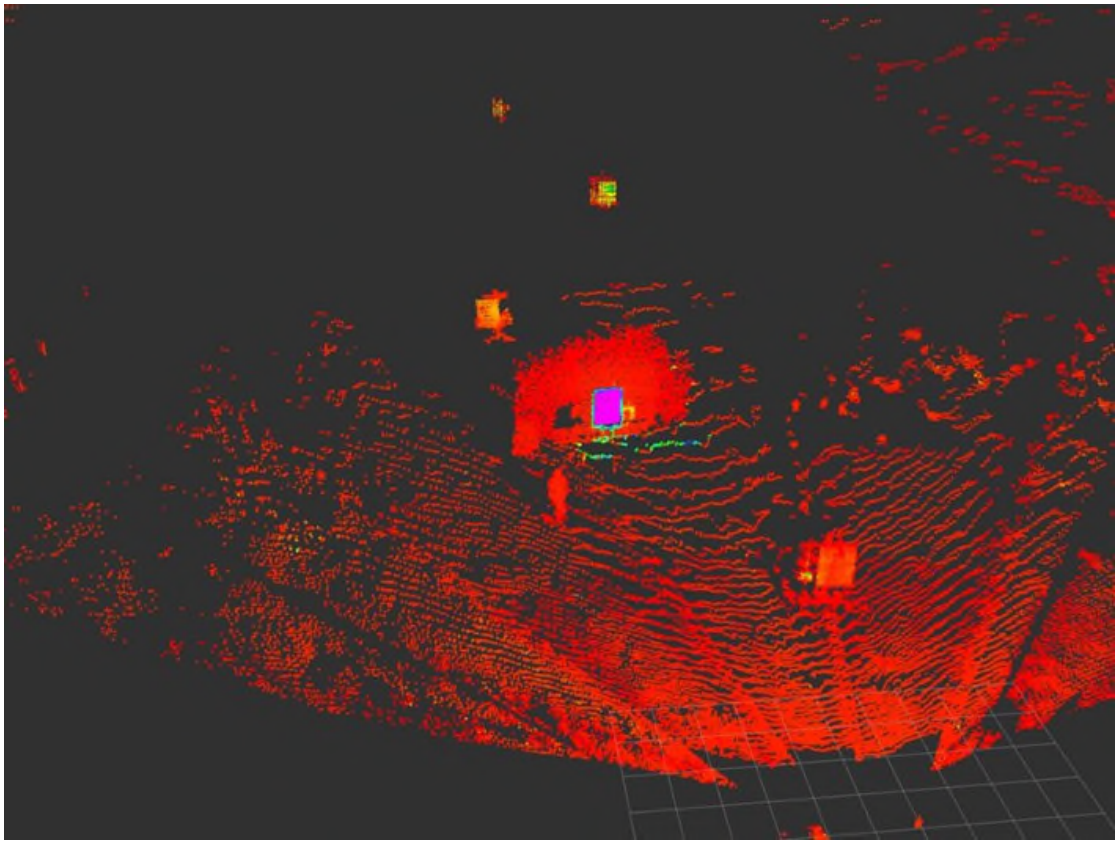


Figure 29: LiDAR point cloud shows the 'blooming effect' as a (non-existent) red wall around the retroreflective target whose actual size is shown in purple.

12. Sensor Noise Factor Testing in the Lab

12.1. Development of Novel Methodologies

As part of the Sim4CAMSens project, novel methodologies for sensor noise factor testing in the laboratory have been developed. Within the newly established WMG Perception Sensors Testing Lab (see Figure 35), experiments have been conducted to test the robustness of LiDAR sensors with respect to presence of water droplets and mud occlusion on the sensor surface.

Figure 30: WMG Perception Sensors Testing Lab.

In real-world scenarios, water droplets are likely to occur on the sensor surface any time there is rain or drizzle outside, which means there is the potential for compounding effects of two noise factors – the actual precipitation and the water droplets on the sensor surface. Water droplets are domed (or half-spherical) in shape and they may exhibit lensing effects and refract LiDAR beams, which in turn has the potential to distort the point cloud.

LiDAR sensors exposed to the environment, especially those mounted in the bumper area, could also be subject to mud occlusion. Mud occlusion could take the form of a hazy layer of fine-particle dust on the LiDAR sensor's surface, muddy water droplets, or fully opaque splotches that can block parts of the sensor surface. To what extent different levels of mud occlusion could impact LiDAR performance is an open research question, with effects likely varying between different LiDAR makes and models. For example, the result could be missing points in the point cloud, warnings being generated by the device, or LiDARs not working at all due to occlusion being detected.

12.1. Equipment and Experimental Setup

The set of LiDARs used for the experiments including their main characteristics is presented in Table 14. All three LiDARs are solid-state LiDARs, which is the technology of LiDARs most likely to be used on vehicles to support automated and assisted driving functions. The two main wavelengths used are 906 nm and 1550 nm, while horizontal and vertical fields of view (FOV) are similar for the three LiDARs. There are some differences in other characteristics, such as their range, range precision, points rate, blind spot, and mass. The ingress protection rating of these devices allowed us to conduct experiments with water spray and mud occlusion.

Table 13: LiDAR sensors used in the experiments and their main characteristics.

Sensor Model	Technology	Wave-length [nm]	Range [m]	Range Precision [cm]	Points Rate [pts/s]	FOV	Blind Spot [m]	Mass [kg]	Ingres Protection Rating
LiDAR A	Non-repetitive scanning	905	150	± 3 cm	452,000	120° - Horizontal 25° - Vertical	< 0.5	1120	IP67
LiDAR B	Semi-solid state	1550	250	± 5 cm	900,000	120° (40°) - Horizontal (ROI) 25° (4.8°) - Vertical (ROI)	< 1.5	1700	IP67, IP69K
LiDAR C	MEMS solid-state	905	180	± 5 cm	787,500 - Single 1,575,000 - Dual	120° - Horizontal 25° - Vertical ± 0.2°	< 0.5	690	IP67, IP6K9K

The following materials and equipment were used to generate noise factors (water droplets and mud occlusion) on the sensors (see Figure 36):

- Atomiser spraying bottle – BRAND® atomiser PE-HD, 400 ml
- Clay powder – Sigma-Aldrich bentonite 285234
- Mixing spatulas – Levgos® smartSpatula® disposable spatulas 210 mm and 310 mm
- Petri dishes – BRAND® Petri dish, glass, 100 mm diameter, 20 mm height
- Precision scale – Kern EHA 500-2 Precision Balance Weighing Scale

Appropriate personal protective equipment (PPE) was used, according to the requirements from the chemical safety data sheet and the activity risk assessment, consisting of certified:

- Eye safety glasses
- Nitrile gloves
- Face mask (dust protector)

Figure 31: Materials and equipment for creating noise factors.

LiDAR targets used for the experiments (see Figure 37) consisted of:

- Calibrated reflectance boards – SphereOptics Zenith-Lite, 90%, 0.3 m and 0.5 m
- Traffic cones – Laserscanning reflective PVC traffic cones with a metal disc, 0.54 m tall.

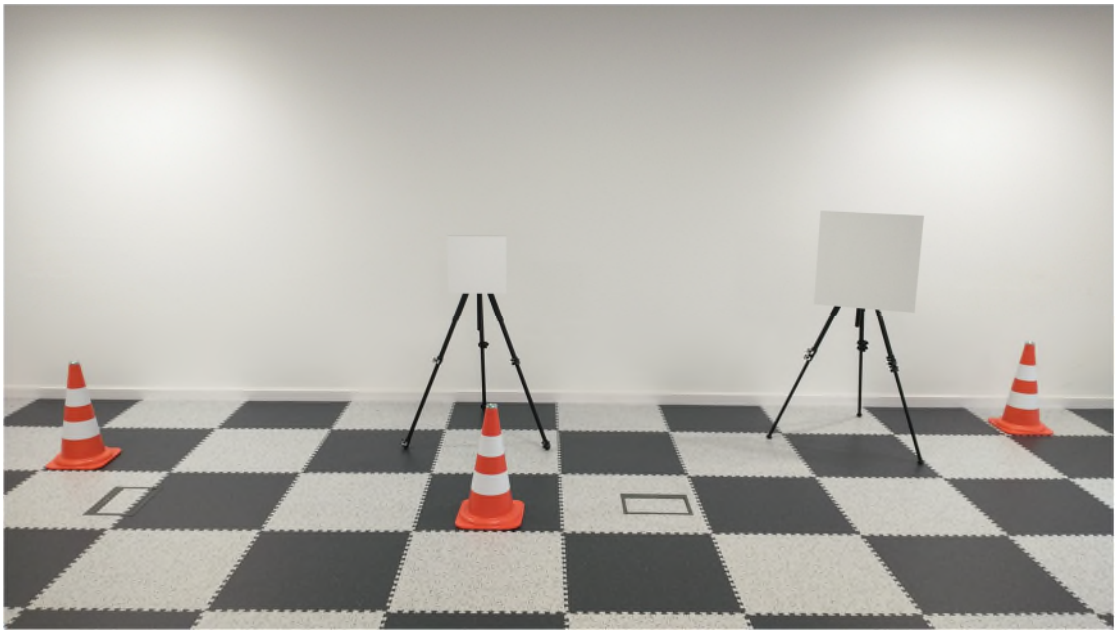


Figure 32: Experimental setup of LiDAR targets - calibrated reflectance boards and traffic cones.

12.1. Generating Water Droplets on LiDAR Sensor

Water droplets were created using a spraying bottle with an adjustable nozzle which allows for two settings: a fine spray and a solid stream. For these experiments, a fine spray setting was used. Several sprays were needed to cover the entire LiDAR sensor (see Figure 38) and there was a variation in the number and size of water droplets created in different experimental runs.

To quantify the amount of water droplets on a LiDAR sensor, an image of the LiDAR sensor was taken using a DSLR camera set on a tripod in front of the LiDAR (see Figure 39), and various image processing methods were used to calculate the percentage of water droplet coverage.

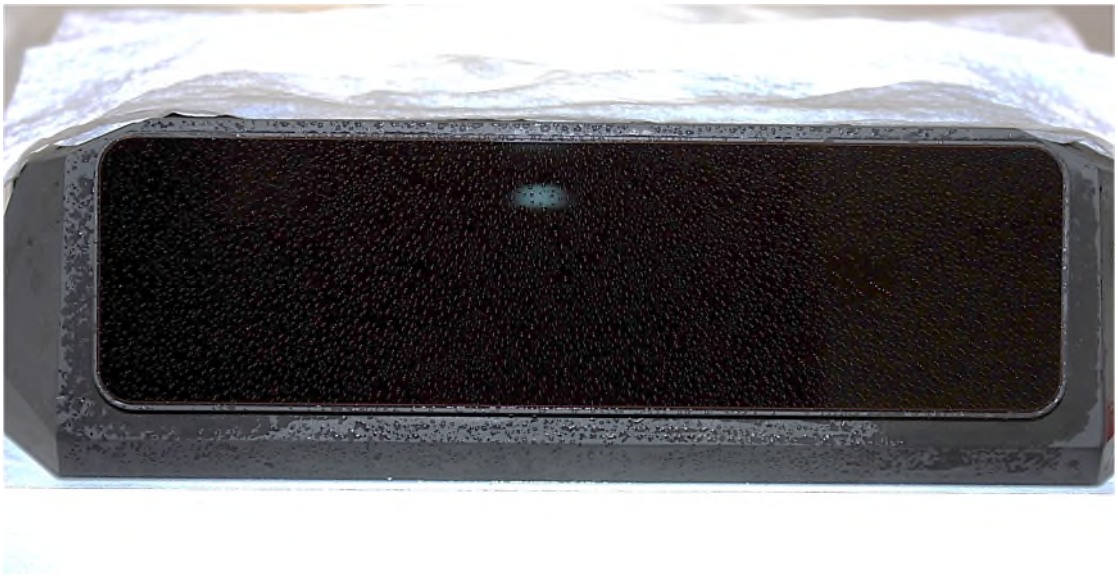


Figure 33: An example of clear water droplets on the LiDAR.

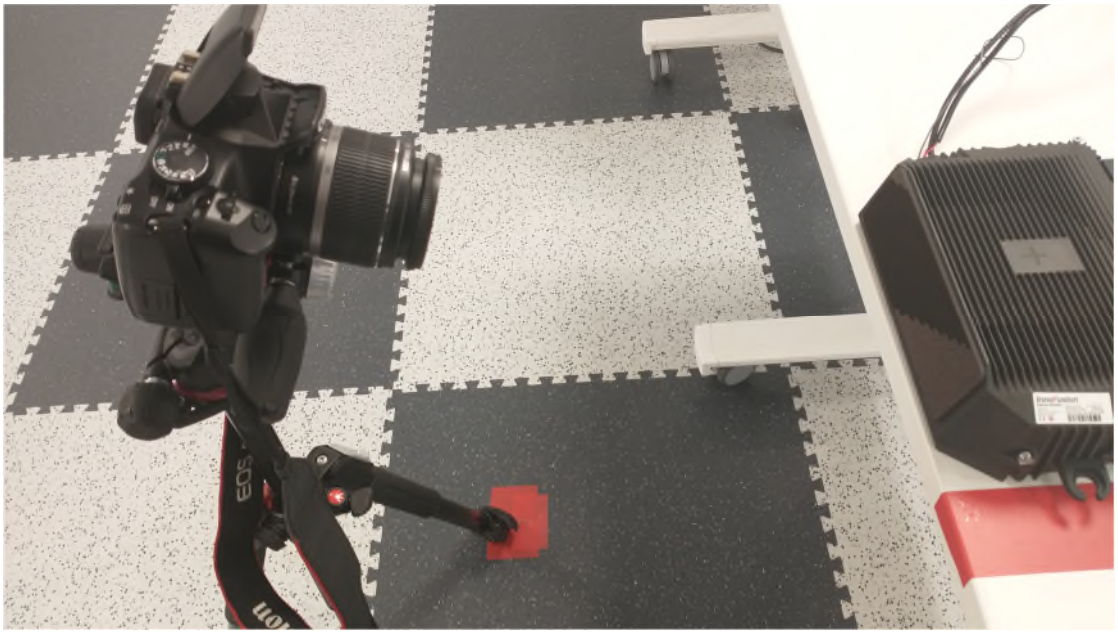


Figure 34: Using a DSLR camera to take the picture of the LiDAR sensor plate.

12.2. Generating Mud Occlusion on LiDAR Sensor

Two generate mud occlusion on LiDAR sensors, two procedures were initially investigated: muddy water droplets and mud splotches.

For muddy water droplets, different amounts of clay powder were mixed with water inside the atomiser spray and well shaken. Precision scale was used to measure the mass of water and the mass of added clay powder, to compute the clay concentration. The atomiser was then used to spray muddy droplets onto the surface of the sensor in the same way as clear water droplets (see Figure 40).

Similar to the quantification of water droplet coverage, an image of the front of the sensors was taken using a DSLR camera on a tripod, and images were processed to compute the mud coverage of the LiDAR sensor. The full analysis of the experiments with clear and muddy water droplets is reported in D2.2 Test Methodology Report.

To create mud splotches, a paste was created by mixing clay powder with water and then applied with a lab spatula onto the surface of the sensor (see Figure 41). Various sizes and locations of mud splotches were created.

Due to the opaqueness of mud splotches, they were found to be able to knock out large areas of the point cloud in the live viewer (see Figure 42). This makes them particularly dangerous in real life situations, where they have the potential to obscure objects in the scene and unpredictably impact automated and assisted driving functions.

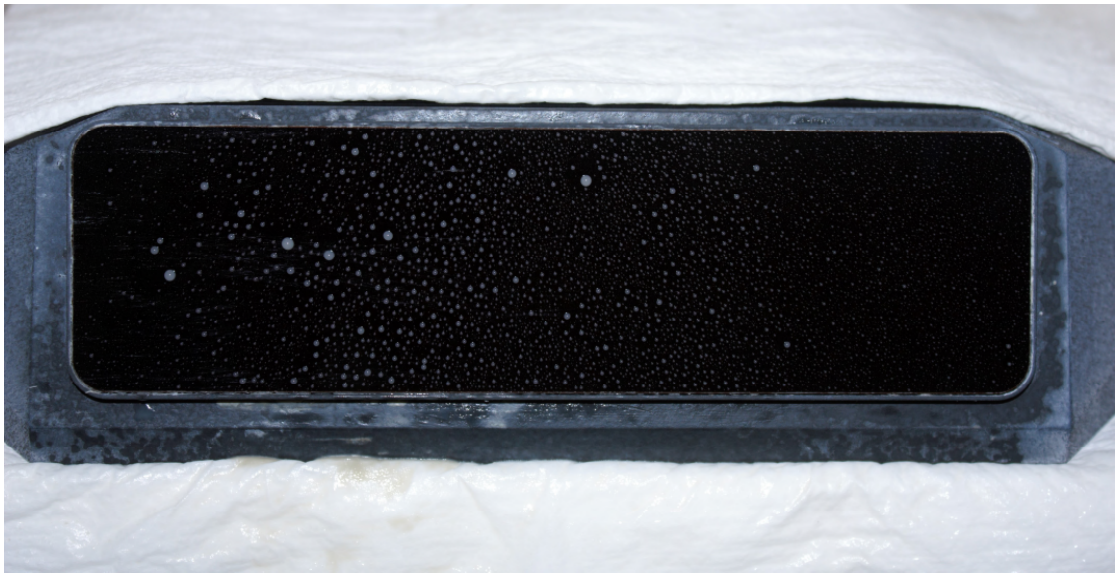


Figure 35: An example of muddy water droplets on the LiDAR.

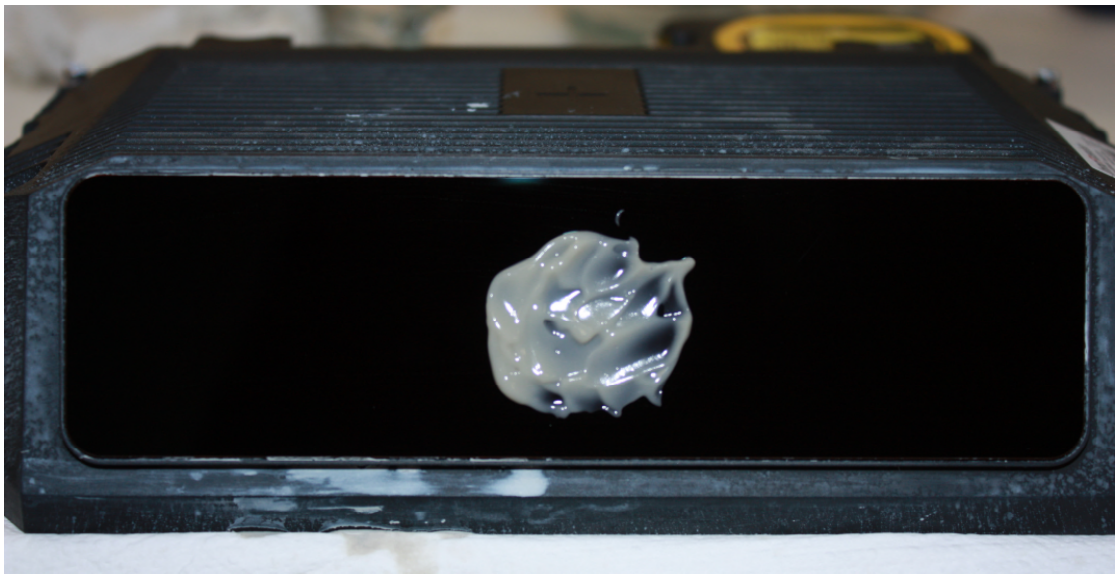


Figure 36: An example of a mud splotch on the LiDAR.

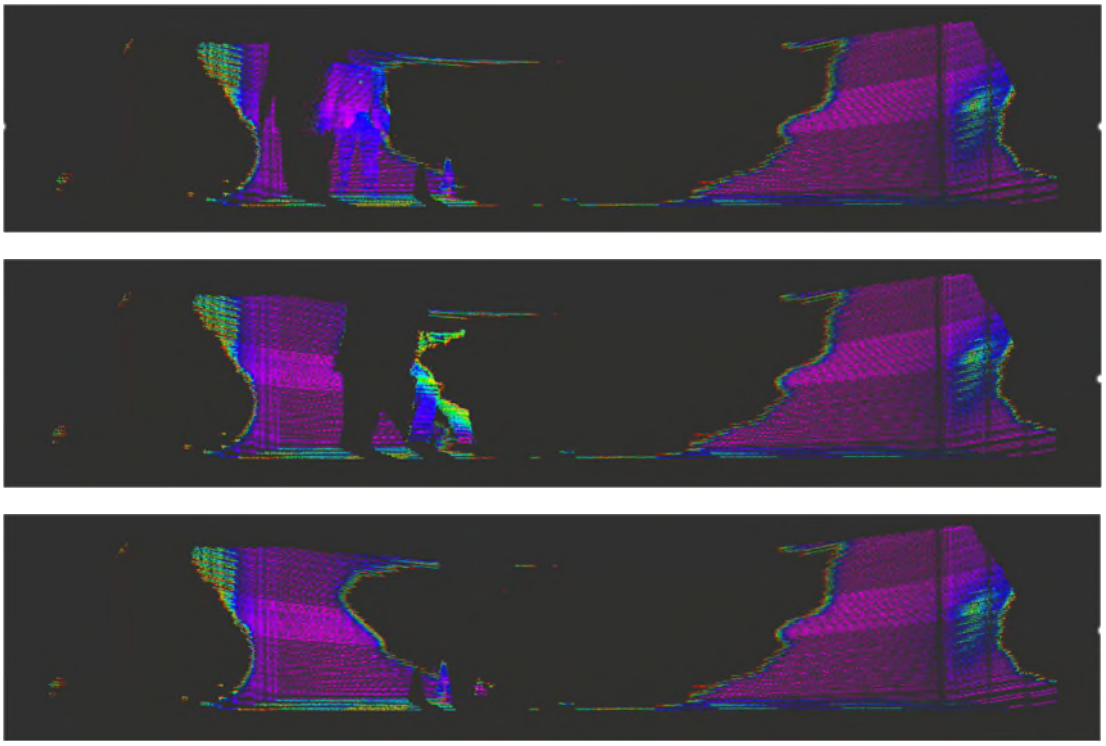


Figure 37: A pedestrian disappears in the area of the point cloud blocked by a mud occlusion.

13. Conclusions

The state-of-the-art review of sensor testing methodologies for automotive applications presented in Sections 2-7 reveals a diverse landscape with both advancements and limitations. The findings indicate a wealth of testing standards and methods across cameras, LiDARs, and radars, each with their own focus and developmental trajectories:

- For cameras there are a plethora of available testing standards, predominantly focused on intrinsic parameters such as resolution, sharpness, and field of view. Notably, the upcoming P2020 standard aims to introduce image quality testing procedures and metrics specifically tailored for automotive cameras, aligning with the needs of machine vision applications.
- LiDAR testing, although lacking standardised tests, exhibits comprehensive performance evaluation methods in academic literature. These evaluations cover a spectrum of parameters including range, field of view, and environmental conditions. However, the absence of standardised tests remains a notable gap, especially considering the variability in LiDAR technologies.
- Conversely, radar testing benefits from established standards designed for automotive applications, particularly in Advanced Driver Assistance Systems (ADAS) functions like adaptive cruise control and blind spot detection. Yet, limitations persist in benchmarking tests and the consideration of various noise factors.
- Sensor fusion testing, crucial for integrating multiple sensor inputs in ADAS systems, remains underdeveloped in terms of standardised methodologies. The described hardware-in-the-loop simulation testing in this report provides a glimpse into potential evaluation frameworks, yet it primarily focuses on ADAS system responses rather than individual sensor performance.
- Furthermore, while object detection applications testing standards offer valuable insights, their applicability to automotive settings remains a concern. These standards, originally designed for other vehicular applications, lack specificity for automotive needs and fail to address the impact of diverse environmental noise on sensor performance.
- Commonalities across sensor testing methodologies include the assessment of both intrinsic and extrinsic sensor parameters such as range, resolution, and field of view. However, differences emerge notably in the availability of testing standards, the consideration of adverse weather conditions (more prevalent in LiDAR testing), and the setup of lab-based tests for cameras versus both lab and outdoor settings for radars and LiDARs.

Despite these advancements, notable limitations persist in current testing standards and methods across sensor types:

- For cameras, the focus on still image cameras neglects the diversity of camera sensors crucial for automotive applications.
- LiDAR testing suffers from a lack of standardization, particularly in benchmarking different LiDAR types and assessing their performance under varied conditions.

- Radar testing standards lack comprehensive benchmarking tests and overlook considerations for radar interference and the impact of environmental noise.
- Sensor fusion testing methodologies, while promising, currently offer limited insights into noise impacts and sensor suite performance nuances.
- Lastly, object detection application testing, while informative, requires adaptations for automotive contexts, especially regarding environmental noise effects.

The Sim4CAMSens consortium conducted the Winter Testing Campaign in an outdoor location in Scotland, using a variety of sensor technologies (camera, LiDAR, radar). This campaign has produced numerous datasets that provide valuable insights into the robustness of various perception sensors, as well as examples of the plethora of noise factors that those sensors may be exposed to in a real-world environment:

- Camera sensors were found to be particularly sensitive to environmental noise factors. For example, the level of lighting (impacted by the time of day, the position of the sun, and the level of cloud coverage), the visibility range (which can be reduced due to presence of mist or fog), and different types of precipitation (rain, snow, hail). Rain droplets and snowflakes can also catch onto the surface of the sensors or the sensor enclosure and present themselves as an occlusion.
- Rain droplets caught on the sensor or sensor enclosure surface usually appear as semi-transparent round circles in images. However, when they are directly illuminated by the sun, they appear as fully opaque white circles. This highlights the potential for multiple noise factors to interact with each other and multiply their negative effects on the sensor performance.
- Cameras may also be impacted by the shadows that can darken parts of the image, as well as by the lens flare. Lens flare has the ability to block objects in the image (e.g., traffic signs), which can then potentially also impact the performance of object detection algorithms.
- LiDAR sensors, unlike cameras, are active sensors which emit their own light beams in the infrared spectrum, and as such they can operate in complete darkness. They can still, however, be impacted by various environmental noise factors. The point clouds recorded in heavy snow demonstrate that heavy snowfall can be detected as points of low reflectivity (~10%), especially in the 10-30 m range.
- For one of the LiDARs, the 'blooming effect' was noticed around a retroreflective target, indicating potential issues with highly reflective traffic signs in high humidity conditions.
- Adverse weather conditions can also have a marked impact on object detection performance, as shown with the example of traffic sign recognition using visible light camera images.

The literature review presented in Sections 2-7, the qualitative and quantitative analysis of datasets from the Winter Testing Campaign presented in Sections 8-11, as well as the novel laboratory methodologies for sensor testing presented in Section 12, underscore the necessity for continued development and refinement of sensor testing standards and methodologies that are specifically tailored for automotive applications. Addressing these gaps will be pivotal

in enhancing the safety, reliability, and performance of sensors integrated into modern vehicle technologies.

References

[1] S. and W. B. and S. N. and H. M. Punke Martin and Menzel, “Automotive Camera (Hardware),” in *Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort*, S. and L. F. and S. C. Winner Hermann and Hakuli, Ed., Cham: Springer International Publishing, 2016, pp. 431–460. doi: 10.1007/978-3-319-12352-3_20.

[2] I. 42 Photography, “ISO 12233:2023 Photography — Electronic still picture imaging — Resolution and spatial frequency responses,” 2023. [Online]. Available: https://www.iso.org/standard/79169.html

[3] I. 42 Photography, “ISO 12232:2019/Amd 1:2020 Photography — Digital still cameras — Determination of exposure index, ISO speed ratings, standard output sensitivity, and recommended exposure index — Amendment 1: Determination of encoding-relative sensitivity (ERS),” 2020. [Online]. Available: https://www.iso.org/standard/79168.html

[4] I. 42 Photography, “ISO 15739:2023 Photography — Electronic still-picture imaging — Noise measurements,” 2023. [Online]. Available: https://www.iso.org/standard/82233.html

[5] I. 42 Photography, “ISO 17850:2015 Photography — Digital cameras — Geometric distortion (GD) measurements,” 2015. [Online]. Available: https://www.iso.org/standard/60819.html

[6] B. Li, P. H. Chan, G. Baris, M. D. Higgins, and V. Donzella, “Analysis of Automotive Camera Sensor Noise Factors and Impact on Object Detection,” *IEEE Sens J*, vol. 22, no. 22, pp. 22210–22219, 2022, doi: 10.1109/JSEN.2022.3211406.

[7] I. 42 Photography, “ISO 17957:2015 Photography — Digital cameras — Shading measurements,” 2015. [Online]. Available: https://www.iso.org/standard/31974.html

[8] I. 42 Photography, “ISO 18844:2017 Photography — Digital cameras — Image flare measurement,” 2017. [Online]. Available: https://www.iso.org/standard/63552.html

[9] I. 42 Photography, “ISO 19084:2015 Photography — Digital cameras — Chromatic displacement measurements,” 2015. [Online]. Available: https://www.iso.org/standard/63894.html

[10] I. 42 Photography, “ISO/TS 19567-1:2016 Photography — Digital cameras — Texture reproduction measurements — Part 1: Frequency characteristics measurements using cyclic pattern,” 2016. [Online]. Available: https://www.iso.org/standard/65356.html

[11] I. 42 Photography, “ISO 16505:2019 Road vehicles — Ergonomic and performance aspects of Camera Monitor Systems — Requirements and test procedures,” 2019. [Online]. Available: https://www.iso.org/standard/72000.html

[12] I. P. A. I. Q. W. Group, “IEEE P2020 Automotive Image Quality Working Group,” 2020. [Online]. Available: https://site.ieee.org/sagroups-2020/

[13] “IEEE P2020 Automotive Imaging,” *IEEE P2020 Automotive Imaging*, pp. 1–32, 2018.

[14] S. Roos, T. Brühl, M. Pfeffer, L. Ewecker, and W. Stork, “A Method for Evaluation and Optimization of Automotive Camera Systems based on Simulated Raw Sensor Data,” in *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2022, pp. 1334–1341. doi: 10.1109/SMC53654.2022.9945613.

[15] C. Team, “Carla.” [Online]. Available: https://carla.org/

[16] S. Zhao, L. Zhang, Y. Shen, and Y. Zhai, “Research on Benchmarking of Smart Camera Based on Hardware-in-Loop (HiL),” in *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2020, pp. 1819–1823. doi: 10.1109/ITNEC48623.2020.9085054.

[17] I. 42 Photography, “ISO 19093:2018 Photography — Digital cameras — Measuring low-light performance,” 2018. [Online]. Available: https://www.iso.org/standard/63942.html

[18] I. 42 Photography, “ISO 14524:2009 Photography — Electronic still-picture cameras — Methods for measuring opto-electronic conversion functions (OECFs),” 2009. [Online]. Available: https://www.iso.org/standard/43527.html

[19] T. C. 79- Alarm and electronic security systems, “IEC 62676-5:2018 Video surveillance systems for use in security applications - Part 5: Data specifications and image quality performance for camera devices,” 2018. [Online]. Available: https://webstore.iec.ch/publication/34391

[20] R. Roriz, J. Cabral, and T. Gomes, “Automotive LiDAR Technology: A Survey,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6282–6297, 2022, doi: 10.1109/TITS.2021.3086804.

[21] S. Shi, X. Wang, and H. Li, “PointRCNN: 3D Object Proposal Generation and Detection From Point Cloud,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 770–779. doi: 10.1109/CVPR.2019.00086.

[22] E. Arnold, O. Y. Al-Jarrah, M. Dianati, S. Fallah, D. Oxtoby, and A. Mouzakitis, “A Survey on 3D Object Detection Methods for Autonomous Driving Applications,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3782–3795, 2019, doi: 10.1109/TITS.2019.2892405.

[23] J. Wu, H. Xu, Y. Tian, R. Pi, and R. Yue, “Vehicle Detection under Adverse Weather from Roadside LiDAR Data,” *Sensors*, vol. 20, no. 12, 2020, doi: 10.3390/s20123433.

[24] H. Wang, B. Wang, B. Liu, X. Meng, and G. Yang, “Pedestrian recognition and tracking using 3D LiDAR for autonomous vehicle,” *Rob Auton Syst*, vol. 88, pp. 71–78, 2017, doi: https://doi.org/10.1016/j.robot.2016.11.014.

[25] X. Peng and J. Shan, “Detection and Tracking of Pedestrians Using Doppler LiDAR,” *Remote Sens (Basel)*, vol. 13, no. 15, 2021, doi: 10.3390/rs13152952.

[26] W. Huang *et al.*, “A Fast Point Cloud Ground Segmentation Approach Based on Coarse-To-Fine Markov Random Field,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7841–7854, 2022, doi: 10.1109/TITS.2021.3073151.

[27] R. Karlsson, D. R. Wong, K. Kawabata, S. Thompson, and N. Sakai, “Probabilistic Rainfall Estimation from Automotive Lidar,” in *2022 IEEE Intelligent Vehicles Symposium (IV)*, 2022, pp. 37–44. doi: 10.1109/IV51971.2022.9827119.

[28] D. Bastos, P. P. Monteiro, A. S. R. Oliveira, and M. V Drummond, “An Overview of LiDAR Requirements and Techniques for Autonomous Driving,” in *2021 Telecoms Conference (ConfTELE)*, 2021, pp. 1–6. doi: 10.1109/ConfTELE50222.2021.9435580.

[29] ANSI, “ANSI Z136.1-2022: Safe Use of Lasers Standard,” 2022. [Online]. Available: https://webstore.ansi.org/standards/lia/ansiz1362022?source=blog&_gl=1*1gygwfp*_gcl_au*MTExNTEyMTc2NC4xNjkyNjI1OTcy

[30] ANSI, “ANSI Z136.6-2015: Safe Use Of Lasers Outdoors Standard,” 2015. [Online]. Available: https://webstore.ansi.org/standards/lia/ansiz1362015

[31] T. Gomes *et al.*, “Evaluation and Testing System for Automotive LiDAR Sensors,” *Applied Sciences*, vol. 12, no. 24, 2022, doi: 10.3390/app122413003.

[32] J. Park, J. Cho, S. Lee, S. Bak, and Y. Kim, “An Automotive LiDAR Performance Test Method in Dynamic Driving Conditions,” *Sensors*, vol. 23, no. 8, 2023, doi: 10.3390/s23083892.

[33] M. Jokela, M. Kutila, and P. Pyykönen, “Testing and Validation of Automotive Point-Cloud Sensors in Adverse Weather Conditions,” *Applied Sciences*, vol. 9, no. 11, 2019, doi: 10.3390/app9112341.

[34] J. Kim, B. Park, C. Roh, and Y. Kim, “Performance of Mobile LiDAR in Real Road Driving Conditions,” *Sensors*, vol. 21, no. 22, 2021, doi: 10.3390/s21227461.

[35] Z. Jeffries, J. P. Bos, P. McManamon, C. Kershner, and A. Kurup, “Toward open benchmark tests for automotive lidars, Year 1: Static range error, accuracy, and precision,” *Optical Engineering*, vol. 62, no. 03, Jan. 2023, doi: 10.1117/1.oe.62.3.031211.

[36] VT/AVSC/ALWG, “P2936 - Standard for Test Methods of Automotive LiDAR Performance,” 2023. [Online]. Available: https://standards.ieee.org/ieee/2936/10419/

[37] I. 22/SC 32/AHG 1, “ISO/PWI 13228 Test Method for Automotive Lidar,” 2023. [Online]. Available: https://genorma.com/en/project/show/iso:proj:84347,

[38] I. Bilik, O. Longman, S. Villeval, and J. Tabrikian, “The Rise of Radar for Autonomous Vehicles: Signal Processing Solutions and Future Research Directions,” *IEEE Signal Process Mag*, vol. 36, no. 5, pp. 20–31, 2019, doi: 10.1109/MSP.2019.2926573.

[39] S. Chaudhary *et al.*, “Coherent detection-based photonic radar for autonomous vehicles under diverse weather conditions,” *PLoS One*, vol. 16, no. 11, p. e0259438, Nov. 2021, doi: 10.1371/journal.pone.0259438.

[40] ETSI, “EN 303 396 - V1.1.1: Short Range Devices; Measurement Techniques for Automotive and Surveillance Radar Equipment,” 2016. [Online]. Available: https://www.etsi.org/deliver/etsi_en/303300_303399/303396/01.01.01_60/en_303396v010101p.pdf

[41] ETSI, “EN 302 858 V2.1.1 Short Range Devices; Transport and Traffic Telematics (TTT); Radar equipment operating in the 24,05 GHz to 24,25 GHz or 24,05 GHz to 24,50 GHz range; Harmonised Standard covering the essential requirements of article 3.2 of the Directive 2014/53/EU,” 2017. [Online]. Available: https://standards.globalspec.com/std/10070385/DS/EN%20302%20858%20V2.1.1

[42] ETSI, “ETSI EN 301 091 - v1.1.1: Short Range Devices; Transport and Traffic Telematics (TTT); Radar equipment operating in the 76 GHz to 77 GHz range; Harmonised Standard covering the essential requirements of article 3.2 of Directive 2014/53/EU; Part 2: Fixed infrastructure radar equipment,” 2017. [Online]. Available: https://www.etsi.org/deliver/etsi_en/301000_301099/30109102/02.01.01_60/en_30109102v020101p.pdf

[43] ITU, “ITU-R M.2057-1: Systems characteristics of automotive radars operating in the frequency band 76-81 GHz for intelligent transport systems applications,” 2018. [Online]. Available: https://www.itu.int/rec/R-REC-M.2057/en

[44] ETSI, “ETSI TR 102 273-4 V1.2.1: Electromagnetic compatibility and Radio spectrum Matters (ERM); Improvement on Radiated Methods of Measurement (using test site) and evaluation of the corresponding measurement uncertainties; Part 4: Open area test site ,” 2001. [Online]. Available: https://www.etsi.org/deliver/etsi_tr/102200_102299/10227304/01.02.01_60/tr_10227304v010201p.pdf

[45] X. Jia, Z. Hu, and H. Guan, “A new multi-sensor platform for adaptive driving assistance system (ADAS),” in *2011 9th World Congress on Intelligent Control and Automation*, 2011, pp. 1224–1230. doi: 10.1109/WCICA.2011.5970711.

[46] S. Saponara, “Sensing and Connection Systems for Assisted and Autonomous Driving and Unmanned Vehicles,” *Sensors*, vol. 18, no. 7, 2018, doi: 10.3390/s18071999.

[47] Carvalho Henrique, Vale Alberto, Marques Rúben, Ventura Rodrigo, Brouwer Yoeri, and Gonçalves Bruno, “Remote inspection with multi-copters, radiological sensors

and SLAM techniques," *EPJ Web Conf.*, vol. 170, p. 7014, 2018, doi: 10.1051/epjconf/201817007014.

[48] M. Nakagawa *et al.*, "SEAMLESS NAVIGATION USING VARIOUS SENSORS: AN OVERVIEW OF THE SEAMLESS NAVIGATION CAMPAIGN," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXIX-B4, pp. 35–38, 2012, doi: 10.5194/isprsarchives-XXXIX-B4-35-2012.

[49] Y. Chen and Q. Zhao, "A Novel Square-Root Cubature Information Weighted Consensus Filter Algorithm for Multi-Target Tracking in Distributed Camera Networks," *Sensors*, vol. 15, no. 5, pp. 10526–10546, 2015, doi: 10.3390/s150510526.

[50] X. Mao, W. Li, C. Lei, J. Jin, F. Duan, and S. Chen, "A Brain–Robot Interaction System by Fusing Human and Machine Intelligence," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 3, pp. 533–542, 2019, doi: 10.1109/TNSRE.2019.2897323.

[51] G. HernáNdez-PeñAloza, A. Belmonte-Hernández, M. Quintana, and F. ÁLvarez, "A Multi-Sensor Fusion Scheme to Increase Life Autonomy of Elderly People With Cognitive Problems," *IEEE Access*, vol. 6, pp. 12775–12789, 2018, doi: 10.1109/ACCESS.2017.2735809.

[52] S. Siachalou, G. Mallinis, and M. Tsakiri-Strati, "A Hidden Markov Models Approach for Crop Classification: Linking Crop Phenology to Time Series of Multi-Sensor Remote Sensing Data," *Remote Sens (Basel)*, vol. 7, no. 4, pp. 3633–3650, 2015, doi: 10.3390/rs70403633.

[53] V. Ankarao, V. Sowmya, and K. P. Soman, "Multi-sensor data fusion using NIHS transform and decomposition algorithms," *Multimed Tools Appl*, vol. 77, no. 23, pp. 30381–30402, 2018, doi: 10.1007/s11042-018-6114-2.

[54] L. Song and R. Yan, "Bearing fault diagnosis based on Cluster-contraction Stage-wise Orthogonal-Matching-Pursuit," *Measurement*, vol. 140, pp. 240–253, 2019, doi: https://doi.org/10.1016/j.measurement.2019.03.061.

[55] Z. Wang, Y. Wu, and Q. Niu, "Multi-Sensor Fusion in Automated Driving: A Survey," *IEEE Access*, vol. 8, pp. 2847–2868, 2020, doi: 10.1109/ACCESS.2019.2962554.

[56] M. Elgharbawy, A. Schwarzhaupt, G. Scheike, M. Frey, and F. Gauterin, "A generic architecture of ADAS sensor fault injection for virtual tests," in *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, 2016, pp. 1–7. doi: 10.1109/AICCSA.2016.7945680.

[57] T. C. I. 127/SC 1, "BS EN ISO 16001:2017 Earth-moving machinery - Object detection systems and visibility aids. Performance requirements and tests," 2017. [Online]. Available: https://www.iso.org/standard/63688.html

[58] T. C. I. 204, "BS ISO 17386:2023 Intelligent transport systems - Manoeuvring aids for low-speed operation (MALSO) Performance requirements and test procedures," 2023. [Online]. Available: https://www.iso.org/standard/82952.html

[59] E. NCAP, “EURO NCAP TEST PROTOCOL – AEB VRU SYSTEMS,” 2024. [Online]. Available: https://www.euroncap.com/media/79884/euro-ncap-aeb-lss-vru-test-protocol-v45.pdf

[60] E. NCAP, “EURO NCAP TEST PROTOCOL – VRU TESTING,” 2024. [Online]. Available: https://www.euroncap.com/media/79878/euro-ncap-vru-testing-protocol-v91.pdf

[61] E. NCAP, “EURO NCAP TEST PROTOCOL – LSS SYSTEMS TESTING,” 2023. [Online]. Available: https://www.euroncap.com/media/79865/euro-ncap-lss-test-protocol-v43.pdf

[62] P. H. Chan, G. Dhadyalla, and V. Donzella, “A Framework to Analyze Noise Factors of Automotive Perception Sensors,” *IEEE Sens Lett*, vol. 4, no. 6, Jun. 2020, doi: 10.1109/LSENS.2020.2996428.

[63] P. H. Chan, S. S. Roudposhti, X. Ye, V. Donzella, and S. Roudposhti, “A noise analysis of 4D RADAR: robust sensing for automotive?,” 2023, doi: 10.36227/techrxiv.24517249.v1.

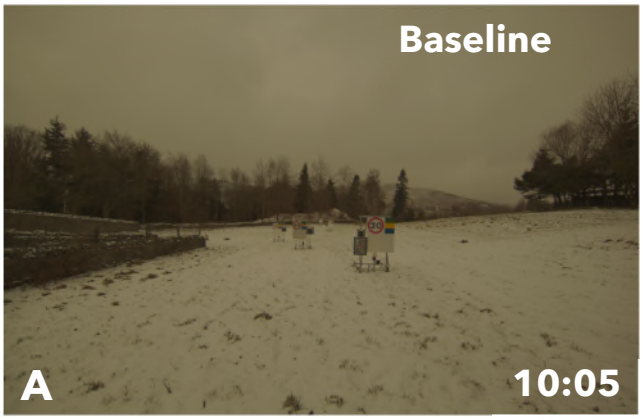
[64] D. Gummadi, P. H. Chan, H. Wang, V. Donzella, and H. Chan, “Correlating traditional image quality metrics and DNN-based object detection: a case study with compressed camera data,” 2023, doi: 10.36227/techrxiv.24566371.v1.

[65] Y. Wang, H. Zhao, K. Debattista, and V. Donzella, “The Effect of Camera Data Degradation Factors on Panoptic Segmentation for Automated Driving.” [Online]. Available: http://wrap.warwick.ac.uk/177340

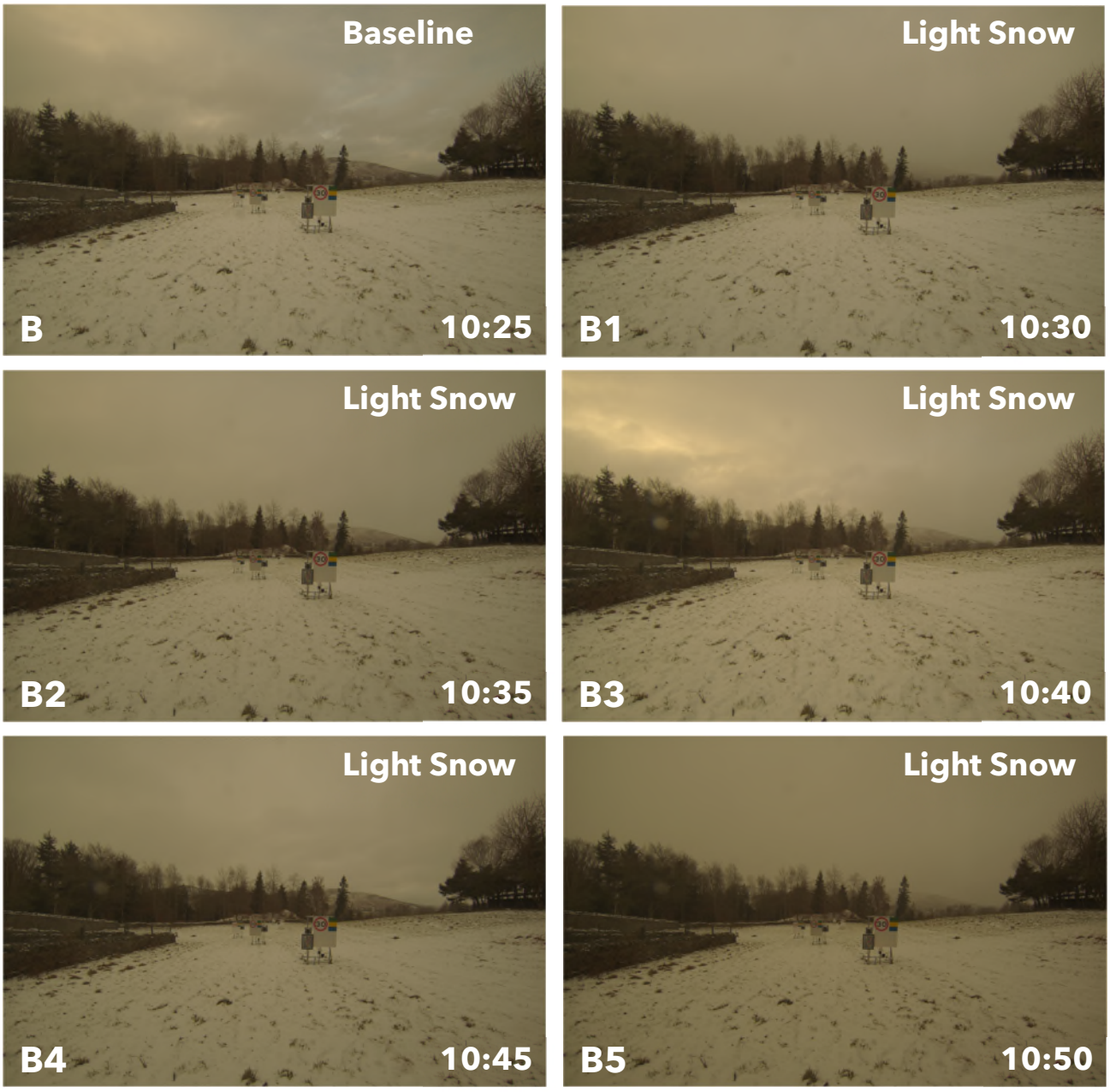
Appendix I – Image Datasets for IQA

13.1. Automotive Visible Camera

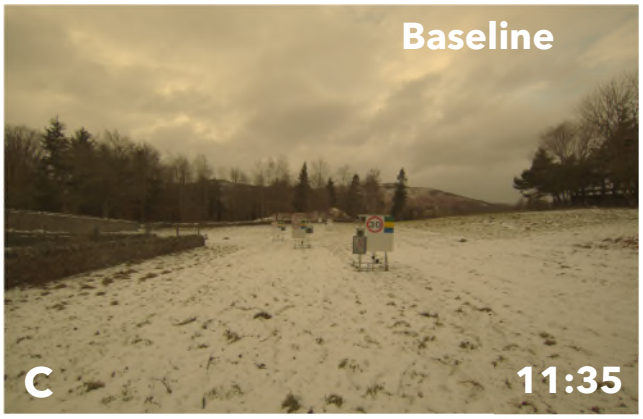
- Visible Camera - Dataset A



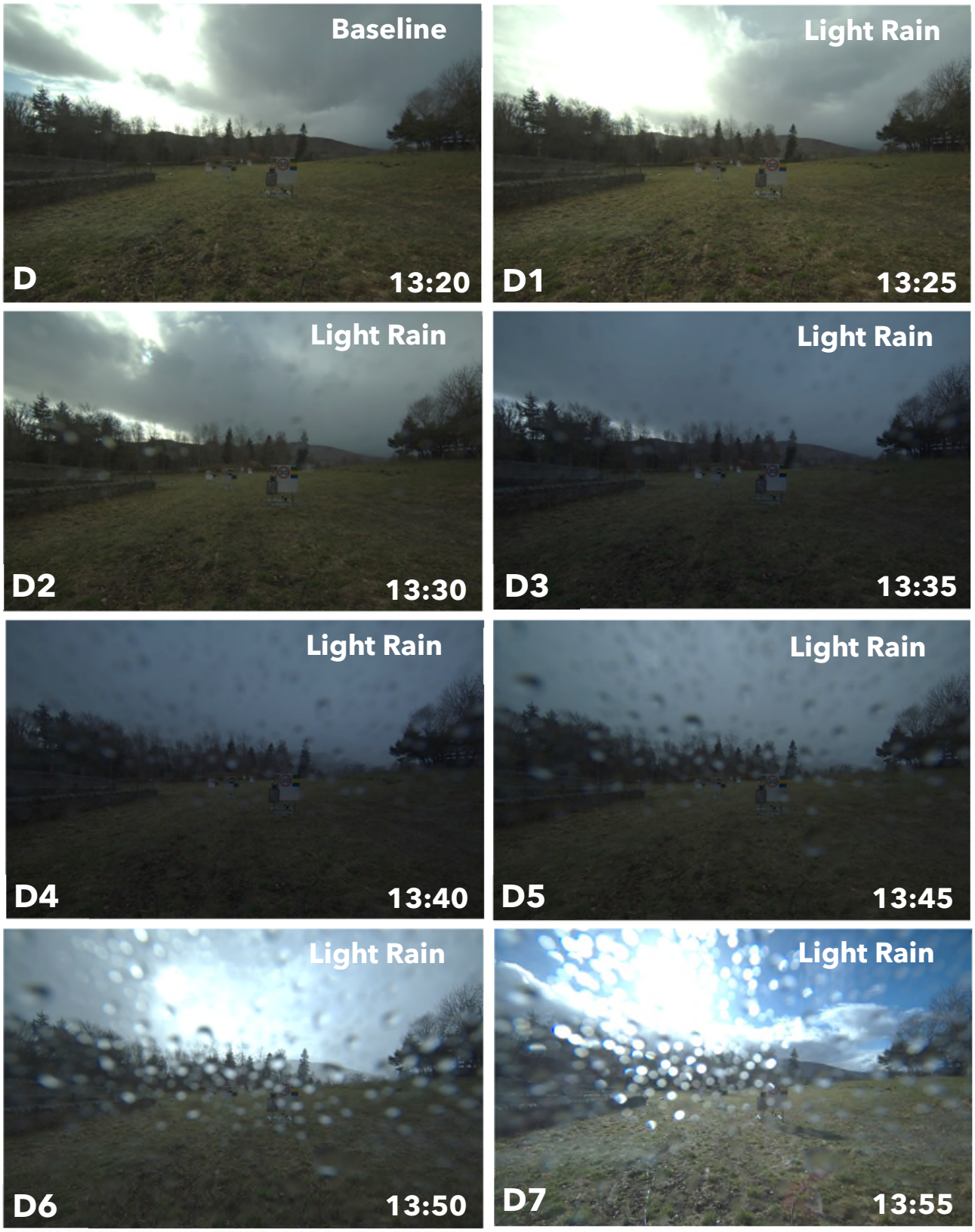
- Visible Camera - Dataset B



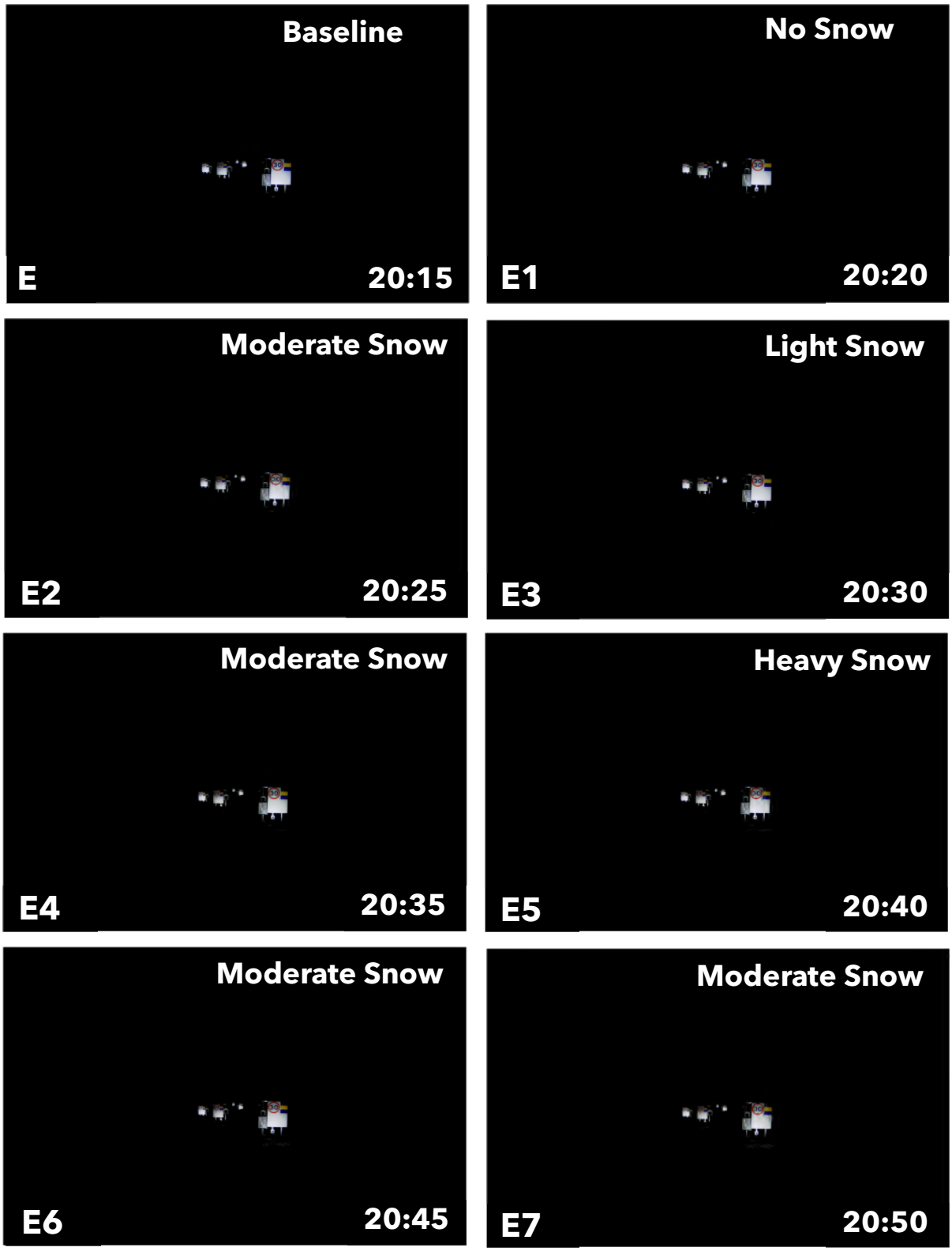
- Visible Camera – Dataset C



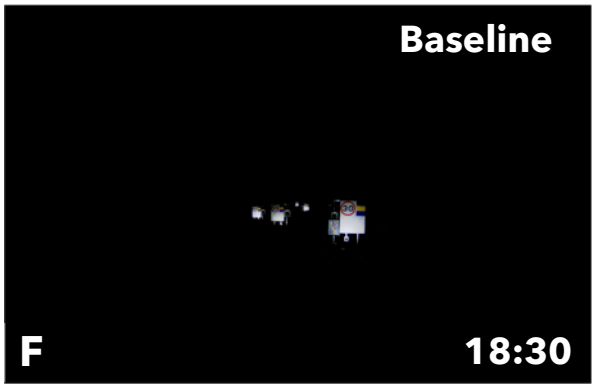
- Visible Camera – Dataset D

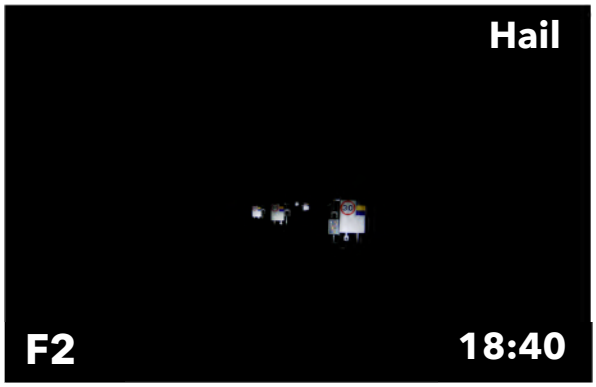


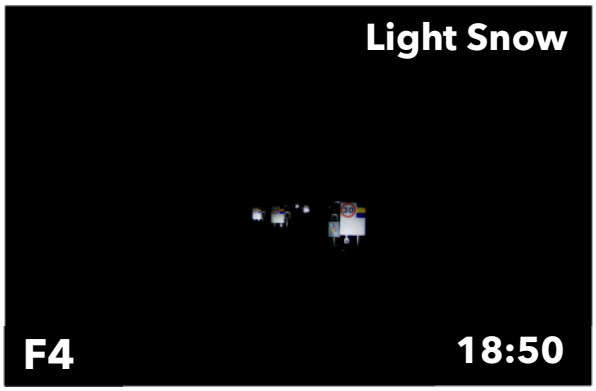
- Visible Camera – Dataset E

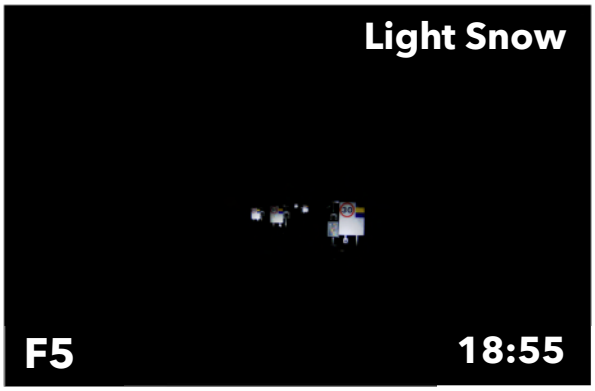


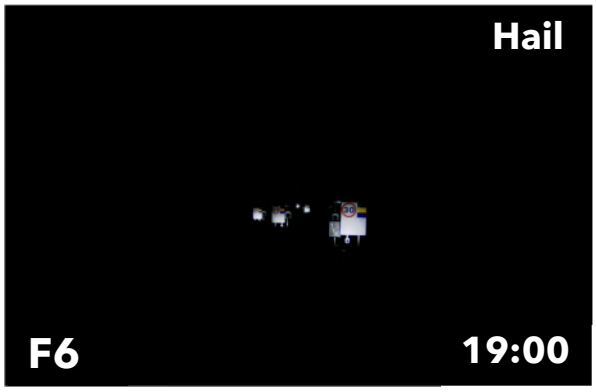
- Visible Camera – Dataset F

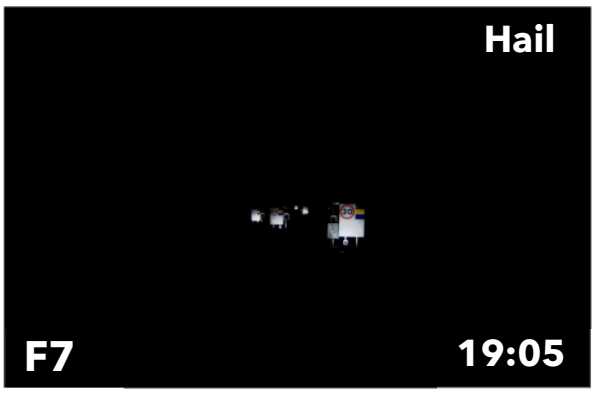


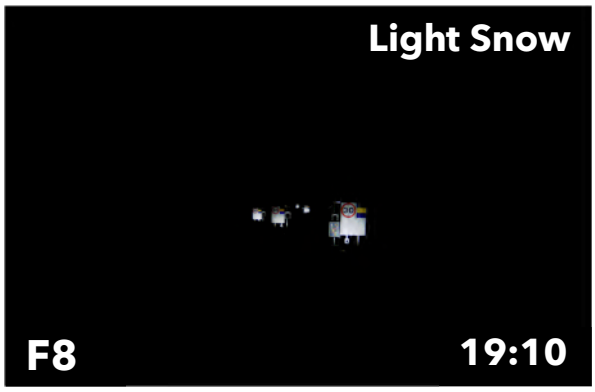



Light Snow
F4
18:50


Light Snow
F5
18:55


Hail
F6
19:00


Hail
F7
19:05


Light Snow
F8
19:10