Ref. Ares(2025)6963726 - 28/08/2025

ROADVIEW

Robust Automated Driving in Extreme Weather

Project No. 101069576

Deliverable 5.4

SW on Collaborative Perception Solutions – Final report

WP5 – Robust perception system

Authors	Isabelle Morvan (CRF), Hervé Ruellan (CRF), Maxime Cancouet (CRF), Pierre Duthon (CE), Hamed Ouattara (CE), Heikki Hyyti (FGI), Eren Erdal Aksoy (HH)
Lead participant	CRF
Delivery date	28 August 2025
Dissemination level	Public
Type	Document, report

Version 1.0

Funded by the European Union

Innovate UK

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them. Project grant No. 101069576.
UK participants in this project are co-funded by Innovate UK under contract No.10045139.
Swiss participants in this project are co-funded by the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract No. 22.00123.

Revision history

Author(s)	Description	Date
Isabelle Morvan (CRF)	Creation of document from template Document organisation	18/02/2025
Eren Erdal Aksoy (HH)	Study of congestion-aware collective perception added	07/05/2025
Pierre Duthon, Hamed Ouattara (CE)	Weather detection section added.	22/07/2025
Hervé Ruellan (CRF)	Roadside sensor fusion section added	31/07/2025
Maxime Cancouet (CRF)	Collective Perception – Quality Information section added	01/08/2025
Isabelle Morvan (CRF)	Reorganisation of document, added the V2X measurements	05/08/2025
Heikki Hyyti (FGI)	FGI vehicle detection pipeline description	13/08/2025
Isabelle Morvan (CRF)	Move FGI description with the global fusion section	13/08/2025
Maytheewat Aramrattana (VTI)	Revision 1	20/08/2025
Stefan Parzefall (AVL)	Revision 2	20/08/2025
Isabelle Morvan, Maxime Cancouet (CRF)	Integrate the reviewers' feedback	22/08/2025
Pierre Duthon (CE)	Integrate the reviewers' feedback	26/08/2025
Johannes Ripperger (accelCH)	Final checks and formatting	28/08/2025

Contents

List of Figures **4**
List of Tables **5**
Partner short names **6**
Abbreviations **6**
Executive summary **8**
Objectives 8
Methodology and implementation 8
Outcomes 8
Next steps 8
1 Introduction **9**
2 Collaborative perception system setup **10**
2.1 System set-up 10
2.2 Recorded data 12
3 Weather detection **15**
3.1 Introduction 15
3.2 Literature review 16
3.2.1 Survey of existing work on the various tasks 16
3.2.2 Work on weather, visibility and brightness classification 17
3.2.3 Work on ground-state detection 17
3.3 Methodology 18
3.3.1 Description of the two models 18
3.3.2 Training, testing and evaluation 19
3.4 Results 23
3.4.1 Testing result 23
3.4.2 Study on model explainability 23
3.4.3 SOTA comparison: study of generalization capacity 24
3.4.4 Ablation study 25
3.4.5 Table 5Demonstration 25
3.5 Conclusion 27
4 V2X road weather service **28**
5 Roadside sensor fusion **31**
5.1 Roadside sensor evaluation 31
5.1.1 Recorded scenes 31
5.1.2 Sensor evaluation under different weather conditions 33
5.2 Roadside sensor fusion 38
5.2.1 Taking advantage of different sensors 38
5.2.2 Sensor fusion adaptation depending on the weather conditions 39
5.2.3 Radar integration in sensor fusion 40
5.3 Global fusion with vehicle detections 41

5.3.1 Detections from ARVO vehicle 42
5.3.2 Global fusion results 42
5.4 Experimentation results 44
6 V2X collective perception service 47
6.1 Quality information 47
6.1.1 Positional accuracy from experiments 47
6.1.2 Sensor Information 48
6.1.3 Perception Regions 48
6.1.4 Perceived Objects 51
6.1.5 Measurements of end-end latency 52
6.2 Congestion-aware study 54
7 Conclusions 56
8 References 57

List of Figures

Figure 1: Collaborative Perception Solutions and Document Structure 9
Figure 2: Roadside system installation in Finnish Lapland 10
Figure 3: Roadside sensors positioning 11
Figure 4: Roadside camera field of view 11
Figure 5: Roadside camera image during common record with VTT vehicle – snow conditions 12
Figure 6: Roadside LiDAR point cloud during common record with VTT vehicle – snow conditions 13
Figure 7: VTT Vehicle's camera view during common record with roadside camera 13
Figure 8: VTT Vehicle's first LiDAR point cloud during common record with roadside camera 14
Figure 9: VTT Vehicle's second LiDAR point cloud during common record with roadside camera 14
Figure 10 : Weather-condition simulation using CycleGAN 16
Figure 11: Example of weather-attribute detection with PGM 16
Figure 12: Architecture of ResNet50-Truncated-MultiTasks (RTM), where each task has its dedicated attention mechanism and classifier, allowing easy modularity 19
Figure 13: Architecture of PatchGAN-MultiTasks (PGM), where the implicit patches are analysed independently via attention mechanisms before classification 19
Figure 14: Examples of images from our training and testing (TT) dataset. Each image is annotated according to 12 criteria: Weather Type, Intensity, Visibility, Sky Condition, Ground Condition, Presence and Intensity of Precipitation, Lighting Conditions and Glare, Road Spray, Presence of Water/Snow on the Windshield. 20
Figure 15: Distribution of labels for the 12 criteria/tasks in the TT dataset 21
Figure 16: Sample images from the DEMO dataset, created during a measurement campaign for the ROADVIEW project in Finland 22
Figure 17: TSNE applied to the embeddings extracted by the attention mechanism dedicated to: a. weather type classification and b. ground condition on 50,000 test images of the TT dataset, using the RTM model. 24
Figure 18: Calculation of the grad cam (b.) on a sample image (a.) from our DEMO dataset 24
Figure 19: Example of weather detection with a scene recorded with wet road and light snow 28
Figure 20: Roadside Road Weather Message Format 28
Figure 21: Onboard Road Weather Message Format 29
Figure 22: V2X road weather service software modules 29
Figure 23: Awareness of weather for decision-making systems 30
Figure 24: Infrastructure-based perception system 31
Figure 25: Roadside sensor detection for snow covered road, light snow with VTT vehicle 32
Figure 26: Roadside sensor detection for dry road, nightlight with FGI vehicle 32
Figure 27: Roadside sensor detection for dry road, cloudy, daylight with FGI vehicle 32
Figure 28: Roadside sensor detection for wet road, light snow, daylight with FGI vehicle 32

Figure 29: Camera view of the travelled paths considered in the evaluation 33
Figure 30: Vehicle trajectory measurements travelling the full length of the road 34
Figure 31: Vehicle trajectory measurements travelling the bottom part of the road on the camera view 34
Figure 32: Vehicle trajectory measurements travelling the top part of the road on the camera view 35
Figure 33: Performances of the roadside camera perception system depending on the weather conditions 35
Figure 34: Vehicle not detected in the bottom to top path – snow on the road 36
Figure 35: Vehicle not detected in the top part of the road – nightlight 36
Figure 36: Performances of the roadside LiDAR perception system depending on the weather conditions 37
Figure 37: roadside LiDAR recording depending on road surface (left: dry, right: wet) 38
Figure 38: Roadside sensor perception difference – snow conditions 39
Figure 39: Multi-sensor object detection (camera, LiDAR, radar) – CARISSMA outdoor recording 40
Figure 40: Multi-sensor object detection (camera, LiDAR, radar) - CEREMA PAVIN platform recording 41
Figure 41: Combined perception of FGI's LiDAR and CRF's camera, with a pedestrian detected by both sensors .42
Figure 42: Combined perception of FGI's LiDAR and CRF's camera, with a pedestrian only detected by CRF' camera 43
Figure 43: Combined perception of FGI's LiDAR and CRF's camera, with a pedestrian only detected by FGI's LiDAR 43
Figure 44: Perception region with low confidence when the road is covered with snow 44
Figure 45: Perception regions with lower confidence at night 45
Figure 46: Different perception results between roadside's camera and vehicle's LiDAR 46
Figure 47:UTM positional deviations (orange arrows) between roadside (in blue) and vehicle-obtained positions (in red) 47
Figure 48: Perception Region with Low Confidence (snow) 49
Figure 49: Perception Regions with Low and High Confidence (night) 50
Figure 50: Use of perception regions to report lane occupancy 50
Figure 51: Example of video analytics with inaccurate object detection due to snow splash 51
Figure 52: VTT's vehicle positions and confidence ellipses near the intersection (left) and after the intersection (right) 52
Figure 53: Set-up to measure the end-to-end latency 53
Figure 54: End-to-end latency measurements 54

List of Tables

Table 1: Collected data and usage 12
Table 2: Architecture parameters and categories 18
Table 3: F1-scores by task and model, on TT dataset. 23
Table 4: F1 scores (↑) on SOTA datasets, comparison of our RTM model with zero-shot and 10-epochs fine-tune against SOTA models and the classic ResNet-50. *In the SOTA paper, the F1 score is not mentioned, only the accuracy score is mentioned. 25
Table 5: Ablation study: the impact when removing our attention mechanism on a zero-shot detection (here scores are estimated on the BDD100K dataset) 25
Table 6: Consistency of RTM model on the DEMO dataset. We count the number of coherent images, i.e. for which the RTM model gives the same classification on both camera viewpoints. 26
Table 7: Validation of the RTM model on 1000 images from viewpoint 2 of the DEMO dataset. These images were annotated by a human. 26
Table 8: Relative importance of camera and LiDAR perception results for sensor fusion 39
Table 9: Mean error (Euclidean distances) analysis using datasets recorded in Finland 48
Table 10: End-to-end latency of collective perception service 54

Partner short names

accelCH	accelopment Schweiz AG
AVL	AVL Software and Functions GmbH
CE	Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement
CRF	Canon Research Centre France
FGI	Maanmittauslaitos – Finnish Geospatial Research Institute
HH	Hogskolan Halmstad
VTI	Statens Vag- och Transportforskningsinstitut
VTT	VTT Technical Research Centre of Finland

Abbreviations

ADAS	Advanced Driver-Assistance Systems
AUC	Area Under the Curve
C-ITS	Cooperative Intelligent Transport Systems
CPM	Collective Perception Message
CUT	Contrastive Unpaired Translation
D	Deliverable
DENM	Decentralised Environmental Notification Message
ETSI	European Telecommunication Standard Institute
GAN	Generative Adversarial Network
GNSS	Global Navigation Satellite System
GPT	Generative Pre-trained Transformer
GRAD-CAM	Gradient-weighted Class Activation Mapping
IMU	Inertial Motion Unit
LiDAR	Light Detection and Ranging
MQTT	Message Queue Telemetry Transport
MS	Milestone
ODD	Operational Design Domain
PGM	PatchGAN-MultiTasks
RGB	Red Green Blue
RSU	Road Side Unit

RTM	ResNet50-Truncated-MultiTasks
RWM	Road Weather Message
SAE	Society of Automotive Engineers
SOTA	State-of-the-art
SW	Software
TSNE	T-distributed Stochastic Neighbour Embedding
UTM	Universal Transverse Mercator
V2X	Vehicle-to-Everything
VRU	Vulnerable Road User
WP	Work Package

Executive summary

Objectives

The main objective of this deliverable is to present the results of Task 5.2 “Collaborative Perception Solutions Integrating Infrastructure-based Perception Systems” developed within Work Package 5 (WP5) dedicated to “Robust perception system”.

Collaborative Perception Solutions aim to improve the vehicle perception by sharing information provided from roadside infrastructure systems or other connected vehicles using V2X (Vehicle-to-Everything) communication protocols. One of the objectives of this deliverable is to report on the ROADVIEW solutions, which include a camera-based weather detection system, a roadside sensor fusion system, an object detection system using vehicle sensors, and V2X communication protocols that enable sharing the weather and sensor data between roadside systems and connected vehicles. Another important objective of this deliverable is to report on the evaluation conducted in Finland where ROADVIEW roadside sensors were installed at an intersection during winter and spring. This setup allowed testing under various weather conditions and with various types of road users, including Vulnerable Road Users (VRUs).

Methodology and implementation

This deliverable complements the ROADVIEW deliverable D5.3 “SW on Collaborative Perception Solutions – First report” [1] submitted in February 2024. That first report reviewed the relevant C-ITS (Cooperative Intelligent Transport System) services to improve the vehicle perception under complex conditions. It also detailed the software architecture of the collaborative perception solution, which integrates infrastructure-based perception system to provide connected vehicles with additional data, such as weather conditions and perceived objects by the roadside sensors.

The infrastructure-based system was developed based on specifications from D5.3. The onboard V2X module was implemented using the Robot Operating System Application Programming Interface defined among partners in D2.3 to ease th7e integration with various vehicle platforms in the scope of WP8 dedicated to integration and demonstrations.

To evaluate performance in Artic weather conditions, ROADVIEW partners CRF, CE, FGI, and VTT conducted real-world data collection in Finnish Lapland over several months. Data included roadside sensor data from CRF camera and FGI LiDAR, as well as vehicle sensor data from VTT and FGI vehicles. These datasets supported the evaluation of CE’s weather detection, CRF’s sensor fusion and FGI’s object detection in Artic conditions. Additionally, simulation-based studies, such as HH’s work on congestion-aware collective perception, were carried out to complete the real-word evaluation.

Outcomes

The present deliverable D5.4 contributes to ROADVIEW Milestone 12 “Perception algorithms and SW ready”. Main outcomes are:

- Camera-based weather detection software, available as open source on the ROADVIEW GitHub [2], providing good results while remaining cost-effective,
- Roadside sensor fusion with V2X communication software, taking into account the perception conditions,
- Onboard V2X communication software ready for integration with the demonstration vehicles, able to transmit and receive V2X messages describing the perceived environment with weather conditions and perceived objects).
- Papers on the congestion-aware collective perception study,
- Roadside and vehicle datasets with various weather conditions.

Next steps

Task 5.2 is now completed and the software on collaborative perception solutions will be demonstrated in various type of environment. The first demonstration will target highway use case and will demonstrate the collective perception from the roadside unit to a connected automated truck. Second, demonstration will target harsh weather conditions in Finland, while last demonstration will target urban environment with the presence of VRUs.

1 Introduction

This deliverable presents the results of Task 5.2 “Collaborative Perception Solutions Integrating Infrastructure-based Perception Systems” developed within WP5 dedicated to “Robust perception system”.

ROADVIEW project focuses on complex environment (including the presence of Vulnerable Road Users VRUs) and harsh weather traffic conditions which have major impact on the safety and operations of connected and automated vehicles. Weather affects not only the vehicle performance but also the roadway infrastructure, thereby affecting the traffic safety (slippery road, lack of visibility, etc). The ROADVIEW Collaborative Perception Solutions target to provide additional robustness to the vehicle perception system by sharing data provided from roadside infrastructure-based systems or from other connected vehicles.

The next figure illustrates the Collaborative Perception Solutions described in this report:

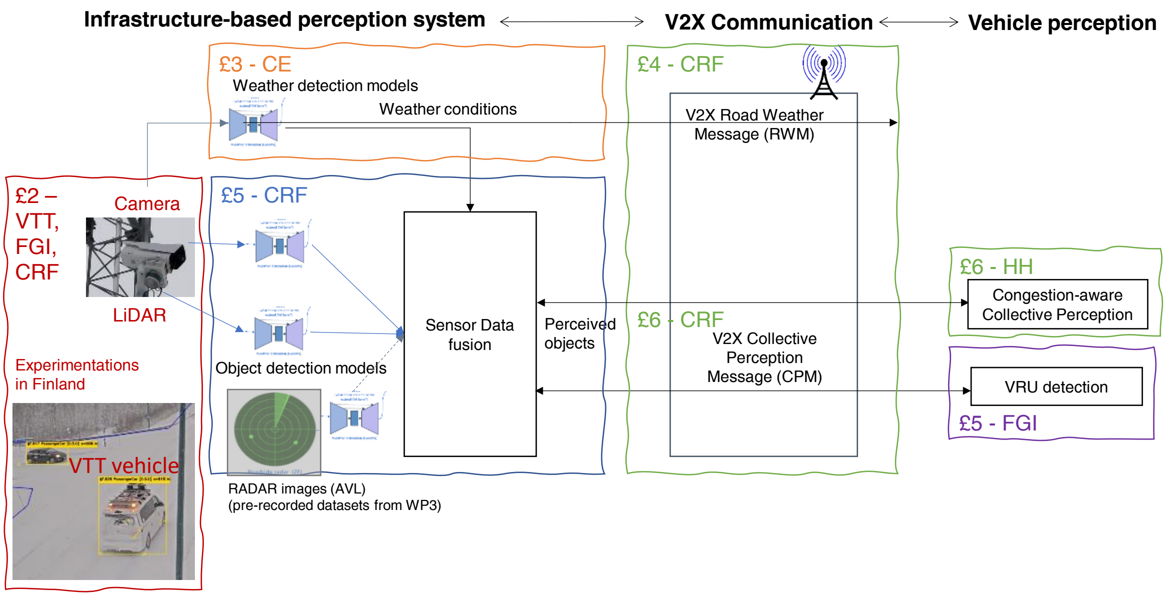


Figure 1: Collaborative Perception Solutions and Document Structure.

First, section 2 of this report describes the setup of an infrastructure-based perception system operating in harsh weather conditions done in Finnish Lapland (VTT, FGI, CRF). This roadside system was used to evaluate the collaborative perception solutions.

Then, this report describes the various Collaborative Perception Solutions studied or developed in ROADVIEW:

- In section 3: weather detection from roadside camera images (CE),
- In section 4: use of V2X communication to share information about road weather (CRF)
- In section 5: object fusion from roadside sensors based on the sensor Operational Design Domain (ODD) (CRF), and global fusion between roadside sensor and vehicle detection in artic weather conditions (FGI)
- In section 6: use of V2X communication to share information about road weather and collective perception (CRF), and a study done in simulation about congestion-aware collective perception (HH).

Last, in the conclusion part, we will discuss about the outcomes and the next steps.

2 Collaborative perception system setup

To test the collaborative perception solutions in harsh weather conditions, an infrastructure-based perception system was installed in Finland.

2.1 System set-up

The perception system comprises two roadside sensors (a camera and a LiDAR) and a roadside unit (RSU). The system was first mounted on a wall of a VTT building in Tampere, where it underwent testing to verify its robustness against winter weather. The roadside camera (Axis Q1656-LE) is equipped with a heating system and a wiper and can operate down to minus 40 °C. The LiDAR (Velodyne VLP-16) can operate only down to minus 10°C based on its datasheet. FGI has prepared a heating system for the LiDAR and for the power supplies installed outdoor on a vertical pole as shown in Figure 2.

The initial testing phase lasted one month, from December 2024 to January 2025, during which the lowest recorded temperature was minus 15°C. After successful validation, the setup was transferred to Finnish Lapland for installation at a road intersection from January 2025 onwards. The lowest recorded temperature was minus 34°C. The roadside sensors and their power supplies installed outdoor were operational at this low temperature.

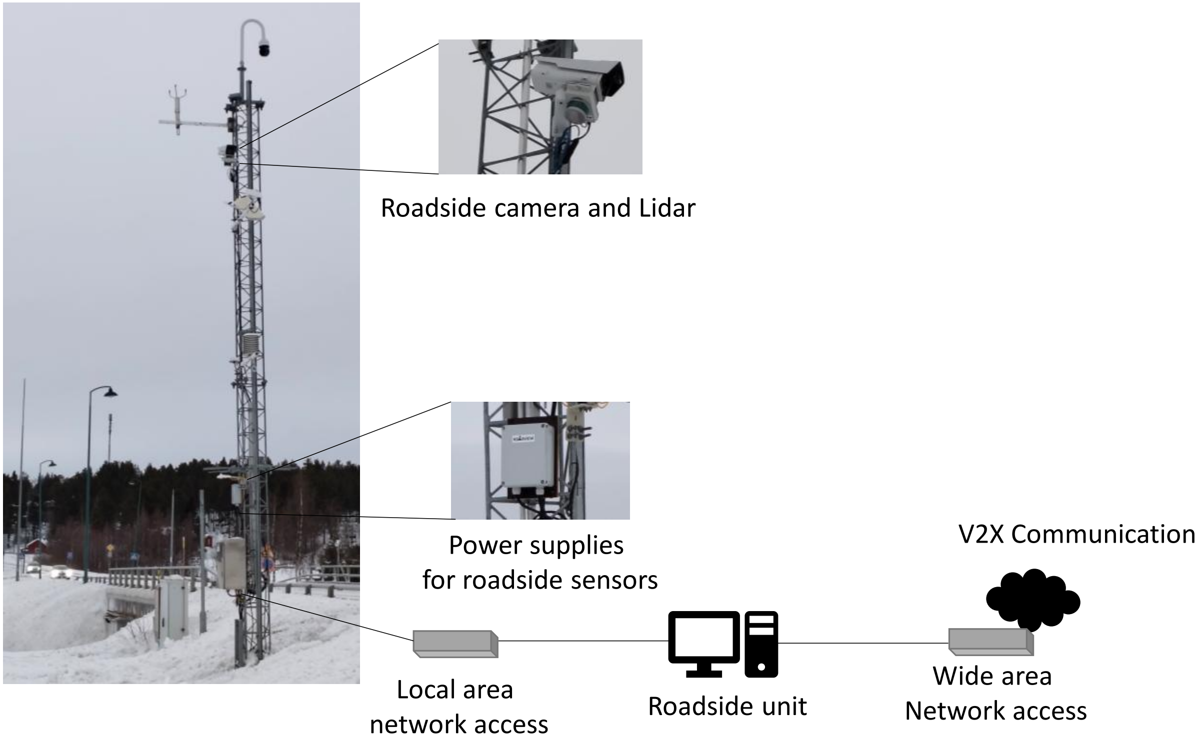


Figure 2: Roadside system installation in Finnish Lapland.

The roadside unit (RSU) was installed indoor and connected to the roadside sensors on the same Local Area Network (LAN) to process the camera and LiDAR data (see sections 3 and 5). The RSU was also connected with a Wide Area Network (WAN) for the V2X communication part (see sections 4 and 6).

The roadside sensors were installed on the Aurora Intelligent Network Road handled by Fintraffic [3] at an uncontrolled road intersection with a cyclist and pedestrian lane near the intersection. This installation enables to analyse real-world scenario with various traffic participants and varying traffic density.

Figure 3 shows the roadside sensor positioning. As explained later in section 5.1.2, the roadside LiDAR range is limited to the intersection area. The roadside camera is able to detect pedestrian or bicycles (VRUs) up to 100 m from the camera and up to 220 m for vehicle detections.

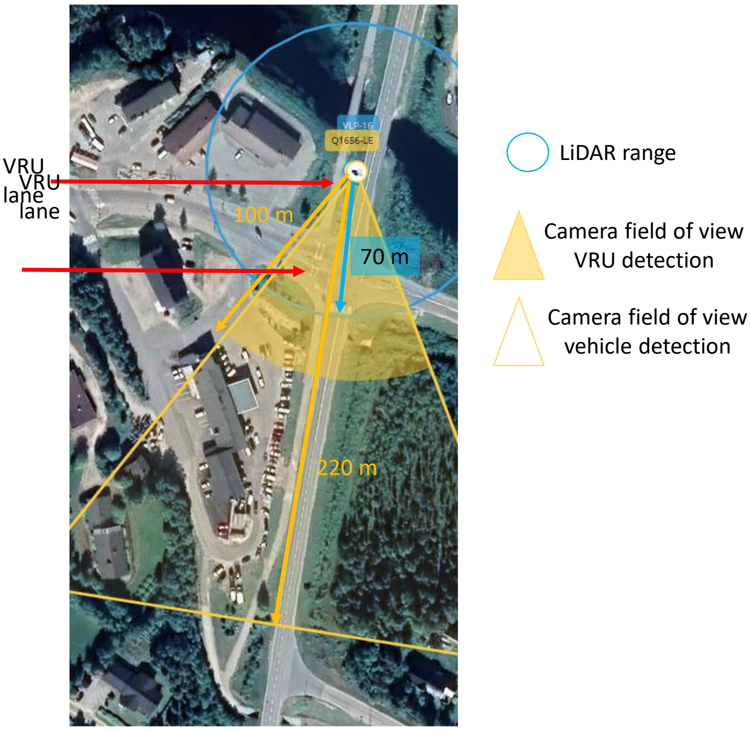


Figure 3: Roadside sensors positioning.

Figure 4 shows the roadside camera view of the monitored intersection.

Figure 4: Roadside camera field of view.

2.2 Recorded data

Several data collection campaigns were organised between partners to evaluate the collaborative perception solutions under various weather conditions as listed in Table 1.

Table 1: Collected data and usage.

Record ID	Period	Recorded data	Weather conditions	Usage
#1	February – May 2025	Roadside camera (CRF, CE)	Winter and spring weather type (snow, rain, cloudy, sunny) – daylight and nightlight	Weather detection (section 3.4.5) and V2X Road weather service (section 4)
#2	January 2025	Roadside camera and LiDAR (CRF) Vehicle camera and LiDAR, GNSS, odometry, IMU data (VTT)	Snow falling with road covered by snow - daylight	Roadside sensor fusion (see section 5) and V2X Collective perception service (see section 6.1)
#3	April 2025	Roadside camera and LiDAR (CRF) Vehicle camera and LiDAR, GNSS (FGI)	Cloudy – nightlight Cloudy – daylight Light snow falling with wet road - daylight	Roadside sensor fusion (see section 5) and V2X Collective perception service (see section 6.1) Vehicle sensor detections (see section 5.3.1)

Figure 5 and Figure 6 show an extract of record #2 for the roadside camera and roadside LiDAR, showing VTT vehicle driving at the monitored intersection with snow fall and the road covered by the snow.

Figure 5: Roadside camera image during common record with VTT vehicle – snow conditions.

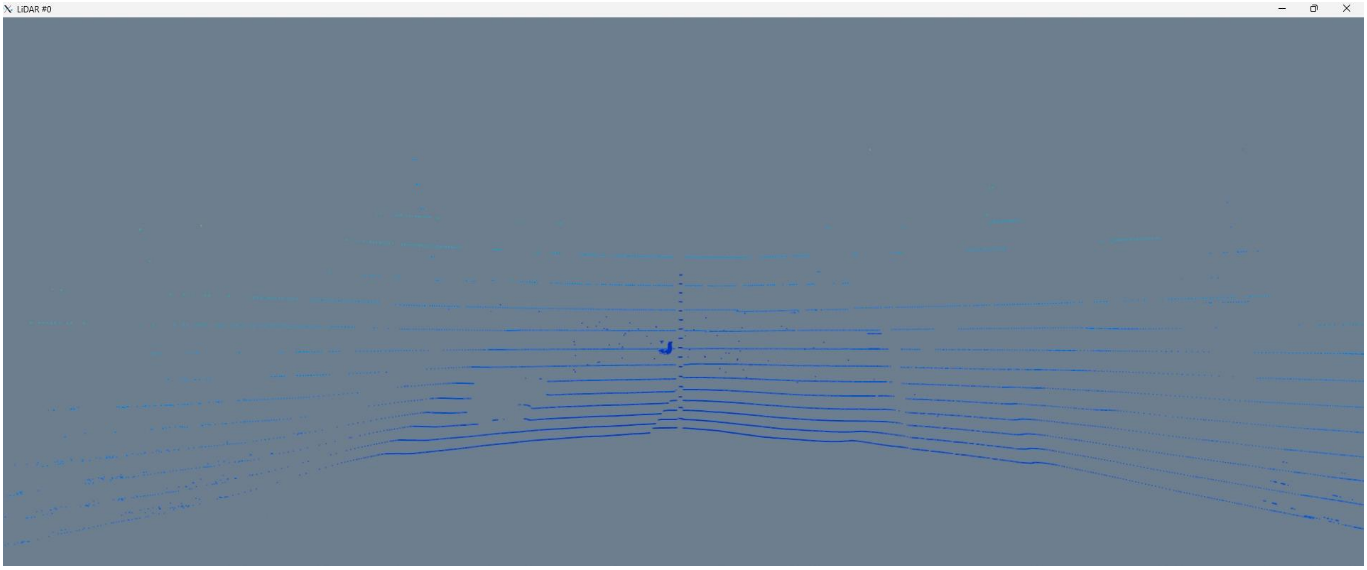


Figure 6: Roadside LiDAR point cloud during common record with VTT vehicle – snow conditions.

We can notice in the VLP-16 point cloud of Figure 6, that the VTT vehicle located at the intersection is not visible whereas the other car visible at the left bottom part of the roadside camera in Figure 5 is visible.

Figure 7, Figure 8, and Figure 9 show the VTT vehicle's sensor view for the camera, and the two LiDARs recorded at the same time than the roadside sensor data.

Figure 7: VTT Vehicle's camera view during common record with roadside camera

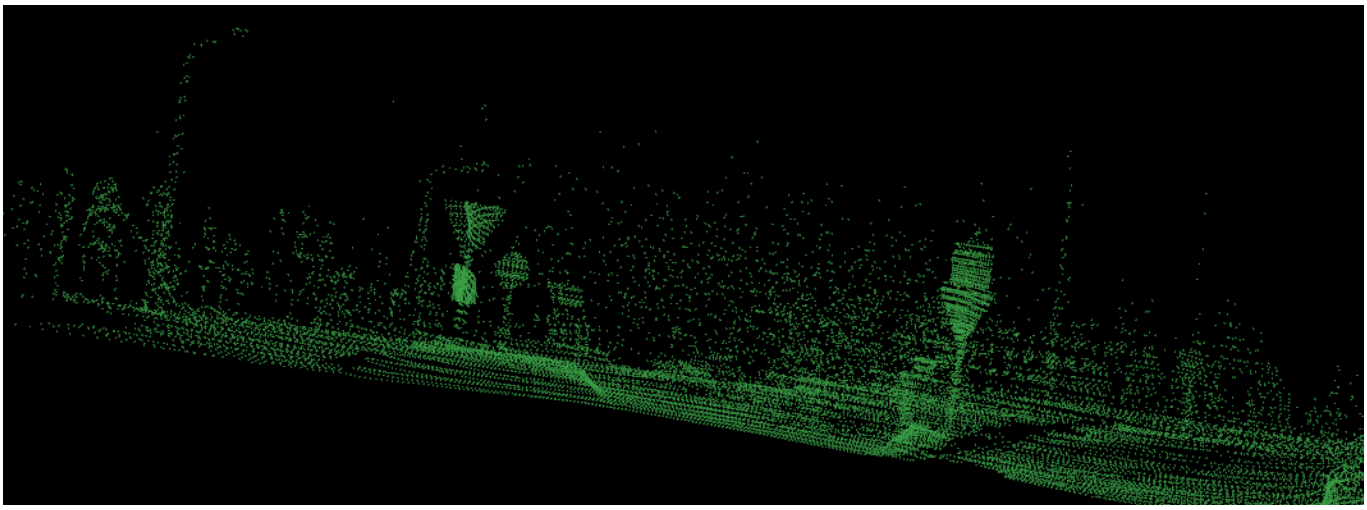


Figure 8: VTT Vehicle's first LiDAR point cloud during common record with roadside camera.

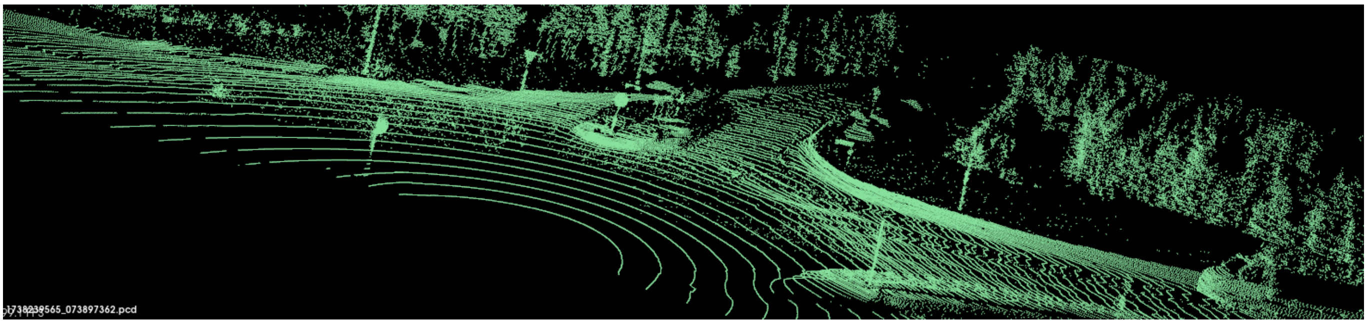


Figure 9: VTT Vehicle's second LiDAR point cloud during common record with roadside camera.

3 Weather detection

3.1 Introduction

As part of the development of collaborative perception systems, we propose to implement a camera-based weather detection system. This system will detect weather conditions and road surface states, enabling the collection of relevant information for autonomous driving. The gathered data will then be transmitted through V2X messages, which is the focus of the following section 4. The present section specifically addresses the challenge of detecting weather conditions using cameras.

The real-time detection of weather conditions (ground state, visibility, precipitation, glare) from RGB images captured both day and night, and their effects, represents a major advance for security, transport, energy, and climate [4], [5], [6], [7], [8], [9], [10]. In the transport domain, this technology improves autonomous driving and ADAS (Advanced Driver-Assistance Systems) by adapting speed, sending warnings, and display signage to indicate actual conditions (rain, snow, fog). Its integration into traffic-management systems helps optimise traffic flow and reduce accidents. Our models, lightweight and fast (with only 3 million parameters), can be deployed on embedded devices (traffic cameras), providing an efficient, low-power solution.

Despite recent progress, visual weather detection remains challenging. As detailed in section 3.2, many approaches are limited to a few classes, ignore intensity and side effects, or require computational resources incompatible with real-time operation. As part of the ROADVIEW project, we want to go one step further. In this report, we propose an original approach based on the following observation: weather mainly alters the visual style of a scene. Style-transfer algorithms such as CycleGAN [11] and CUT [12] show that weather conditions can be simulated realistically (Figure 10) by modifying colour, texture, and brightness while preserving image content. Inspired by these models, we propose two architectures dedicated to detecting weather characteristics: ResNet50-Truncated-MultiTasks (RTM), and PatchGAN-MultiTasks (PGM).

Our models detect not only the weather type but also its side effects (Figure 11), including 12 criteria/tasks:

- **Weather Type**: Sun and Clear, rain, snow, fog, ...,
- **Weather Intensity**: Weak, Average, and High,
- **Visibility**: Low, Average, Good,
- **Sky Condition**: Cloudy, Overcast, Clear Sky, ...,
- **Ground Condition**: Dry, Wet, Snowy, ...,
- **Presence of precipitation**: None, Rain, Snow, …,
- **Intensity of precipitation**: Weak, Average, High,
- **Glare or Reflections**: Present, Absent,
- **Lighting Conditions**: Day, Night, Sunset, …,
- **Road Spray**: Present, Absent,
- **Water On Window**: Present, Absent,
- **Snow On Window**: Present, Absent.

These architectures are flexible: each task is associated with an attention mechanism and a dedicated classifier. For example, the PGM model contains about 3 million parameters and can be extended to new tasks at moderate computational cost. Our models achieve F1 scores above 97 % (Section 3.4) while guaranteeing real-time execution. The source code is available at: https://github.com/roadview-project/Enhanced-Style-Based-Neural-Architectures-for-Real-Time-Weather-Classification-main [2].

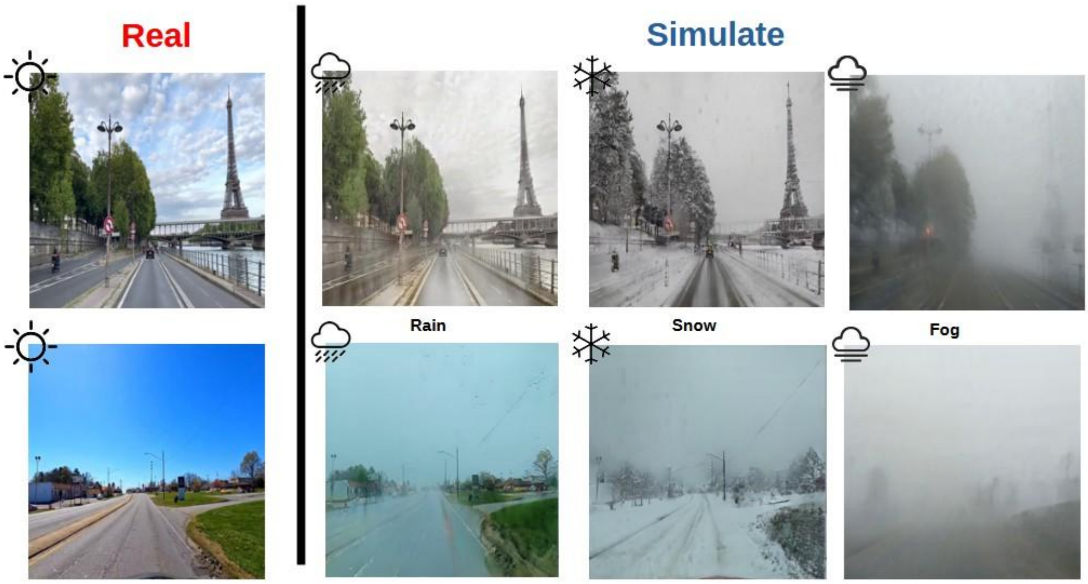


Figure 10 : Weather-condition simulation using CycleGAN.

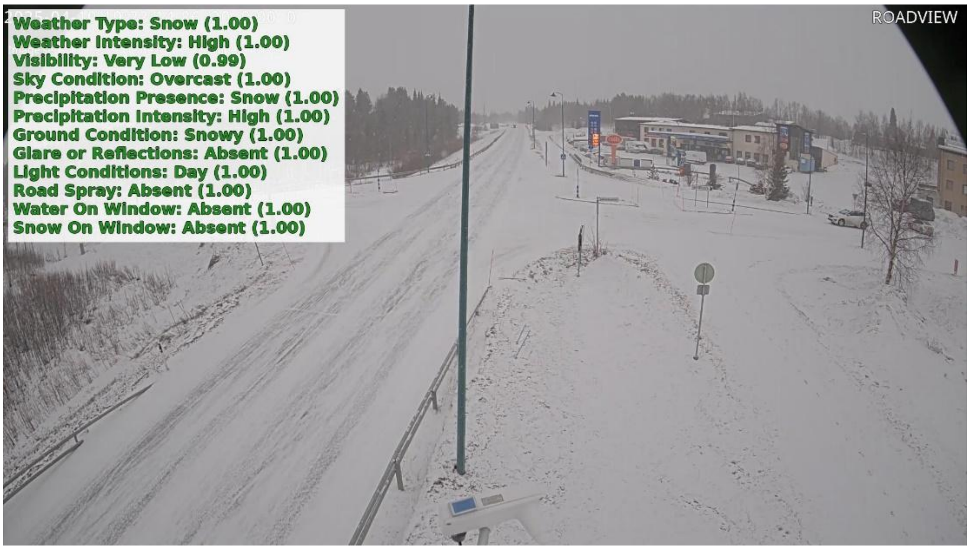


Figure 11: Example of weather-attribute detection with PGM.

3.2 Literature review

3.2.1 Survey of existing work on the various tasks

It is rare to find models specifically trained for the 12 tasks we propose. Our approach is based on a novel hypothesis: the direct influence of weather on the style of objects in an image.

However, several approaches and models address some of the tasks, or a combination of them. The earliest models, using simple CNNs or ResNets, generally only focused on detecting weather conditions in four classes [15], [19], [25]. Among these first methods, some models go a bit further by also handling lighting conditions, for example distinguishing under clear weather between: sunny versus cloudy [16], [18], [20], [26], or by adding other illumination conditions such as sunrise [16], shine [16], glare [21] or night [21]. More recent models, like our method, use attention mechanisms for weather condition classification [17], [22], [23], [24].

Completely independently of weather condition classification, other methods focus on classifying surface state, using smartphone cameras [27], [29], [30], surveillance cameras [28], [31], [33], [34], [13], or even LiDAR data [33], [34]. Once again, our method goes further since it can adapt to images coming from smartphone cameras (pedestrian viewpoint), vehicle cameras, or surveillance cameras, which is our main objective here.

In addition, several large-scale multimodal models, such as GPT-4o [35], [36], [37] are capable of detecting all the attributes studied in this work, with impressive explainability. However, these models require massive amounts of data and considerable compute power, both for training and inference. By contrast, our architectures, although more specialised, offer a resource-efficient solution, optimised in terms of data usage and computation.

The following sections sequentially describe the work related to visibility and brightness qualification, followed by specific studies on ground state detection. These topics are prioritized due to their critical importance for our application.

3.2.2 Work on weather, visibility and brightness classification

Weather-condition classification has been widely investigated classification [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26] with a predominance of approaches based on convolutional neural networks and attention mechanisms. Among these studies, the work by Jingming Xia classification [15] is particularly noteworthy: to reduce computational cost, the author truncated ResNet50 [38] to 15 layers (ResNet15), achieving a significant performance boost. The model's F1 score improved from 85.76 % to 96.03 % for classifying four weather conditions (rain, sunny, snow and fog). Although this improvement is not explained in the paper, it can be partly understood through our approach, which shows that weather primarily affects an image's style and that capturing this style does not necessarily require very deep networks — indeed, overly deep networks can even prove counter-productive. Another study, by Mahesh [39], proposes a model based on ResNet18 to classify three weather types (clear, rain, snow) and three brightness levels (high, medium, low).

With regard to visibility, several studies [40], [41], [42], [43] rely mainly on convolutional networks, some also incorporating physical models, as in [40]. This demonstrates that previous weather detection studies in the literature have remained focused on methods based on convolutional neural networks, without addressing the issue of style variation in their approaches. As shown in the following section, a similar situation exists for the analysis of ground state detection.

3.2.3 Work on ground-state detection

The identification of ground conditions has been the focus of extensive research, particularly for autonomous vehicles condition [27], [28], [29], [30], [31], [32], [33], [4], [34], [13]. Determining whether the road surface is wet, snowy or dry is crucial for adapting vehicle speed and trajectory. Existing approaches exploit various sensors (LiDAR, cameras) and data types (RGB images, segmentation maps, etc.).

Our approach significantly simplifies this problem by directly leveraging the scene's visual appearance. Our models can detect not only the presence of water or snow but also estimate their amount and consequently their impact on driving conditions. This solution combines speed, accuracy and computational efficiency, making it particularly suitable for embedded systems as shown in next sections.

3.3 Methodology

3.3.1 Description of the two models

We hypothesize that weather first modifies the visual appearance (colors, textures, blur, reflections) of a scene. Our pipeline:

- Extraction of stylistic features (global/local Gram or PatchGAN).
- Feature weighting through a dedicated attention mechanism for each task.
- Classification by a lightweight neural network classifier.

For the feature extraction (embedding step), we propose two variants: ResNet50-Truncated-MultiTasks (RTM) and PatchGAN-MultiTasks (PGM). We will show that a single embedding method can be used in preparation for multiple classification tasks. Once this embedding has been carried out, the classification task is simplified and can therefore be carried out with classifiers containing few parameters.

Concerning the classification step, each attribute (type, intensity, visibility, sky, ground, precipitation, brightness, glare, water spray, water/snow on windshield) is handled by a specific attention mechanism + classifier pair. Each head can be activated or deactivated, linearly scaling the parameter count.

We optimize the two models for the classification of ten meteorological characteristics, as explained in the introduction. The patch size and other hyperparameters, are automatically determined via an evolutionary algorithm [44], [45], [46], [47]. This algorithm explores multiple configurations and selects the best-performing ones on the validation set. All hyperparameters and weights are shared on ROADVIEW project github: https://github.com/roadview-project/Enhanced-Style-Based-Neural-Architectures-for-Real-Time-Weather-Classification-main[2].

Table 2: Architecture parameters and categories.

Model	Number of parameters	Category
PatchGAN-MultiTasks (PGM)	3 188 736	Light
ResNet50-Truncated-MultiTasks (RTM)	174 609 524	Average

3.3.1.1 ResNet50-Truncated-MultiTasks (RTM)

The RTM model is inspired by the work of Jingming Xia [15] and the style transfer principles of Gatys [48]. This model leverages the low-level layers of ResNet50, which are known to capture stylistic details. For training, we use pretrained weights from the MoCo-V3 model [49], developed by Facebook's research teams. These weights were trained in an unsupervised manner over 1,000 epochs. The extracted embeddings are then processed by an attention mechanism [50], [51], [52], [53] and a task-specific classifier (Figure 12). The truncated part of ResNet50 has 23,508,032 parameters, each attention mechanism adds 12,582,912 parameters, and each classifier between 6,147 and 14,343 parameters, depending on the number of classes.

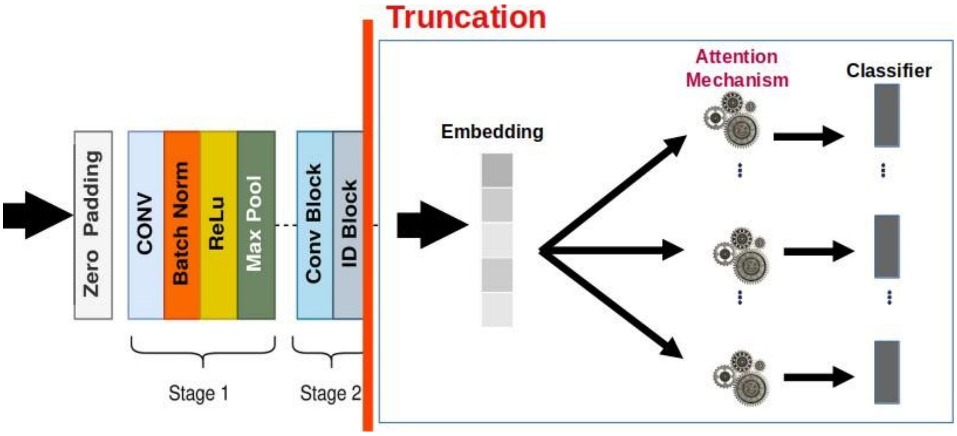


Figure 12: Architecture of ResNet50-Truncated-MultiTasks (RTM), where each task has its dedicated attention mechanism and classifier, allowing easy modularity.

3.3.1.2 PatchGAN-MultiTasks (PGM)

The model PatchGAN-MultiTasks (PGM) follows the classic structure of a PatchGAN. PatchGAN, introduced by Phillip Isola [11], implicitly segments the image into disjoint patches through neurons with distinct receptive fields. Unlike explicit segmentation, these patches are represented by embeddings from independent groups of neurons. The embeddings of the implicit patches are weighted by an attention mechanism [50], [51], [52], [53] for each task, then passed to a classifier (Figure 13).

The part generating the embeddings with disjoint receptive fields comprises 2,756,544 parameters, each attention mechanism adds 10,000 parameters, and each classifier between 41,976 and 64,758 parameters, depending on the number of classes.

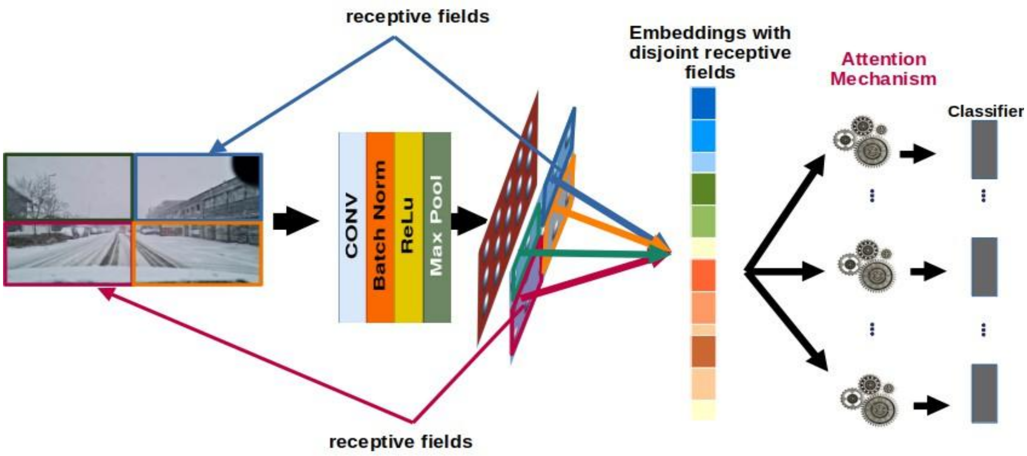


Figure 13: Architecture of PatchGAN-MultiTasks (PGM), where the implicit patches are analysed independently via attention mechanisms before classification.

3.3.2 Training, testing and evaluation

3.3.2.1 Training and testing

As with all methods based on artificial intelligence, we need to start with a training phase. For this, we use an initial database (training and testing dataset – TT dataset) that we have created specifically for our needs. It is necessary to perform training on a dataset that is much larger than the DEMO dataset. In addition, training and testing must be performed on clearly separate datasets. This is why the TT dataset was created. We provide it freely at https://github.com/roadview-project/Enhanced-Style-Based-Neural-Architectures-for-Real-Time-Weather-Classification-main [2]. It contains 503,875 images (Figure 14), human annotated according to our 12 criteria and categorized based on weather conditions. 90% of the images are allocated for training and validation, while 10% are used for testing. The distribution of the data in the dataset is presented on Figure 15. As the figure shows, we have ensured that our classes are fairly balanced across all criteria.

We therefore apply learning to the training base, then calculate the F1 score on the test base. The results of this first phase are presented in section 3.4.

Figure 14: Examples of images from our training and testing (TT) dataset. Each image is annotated according to 12 criteria: Weather Type, Intensity, Visibility, Sky Condition, Ground Condition, Presence and Intensity of Precipitation, Lighting Conditions and Glare, Road Spray, Presence of Water/Snow on the Windshield.

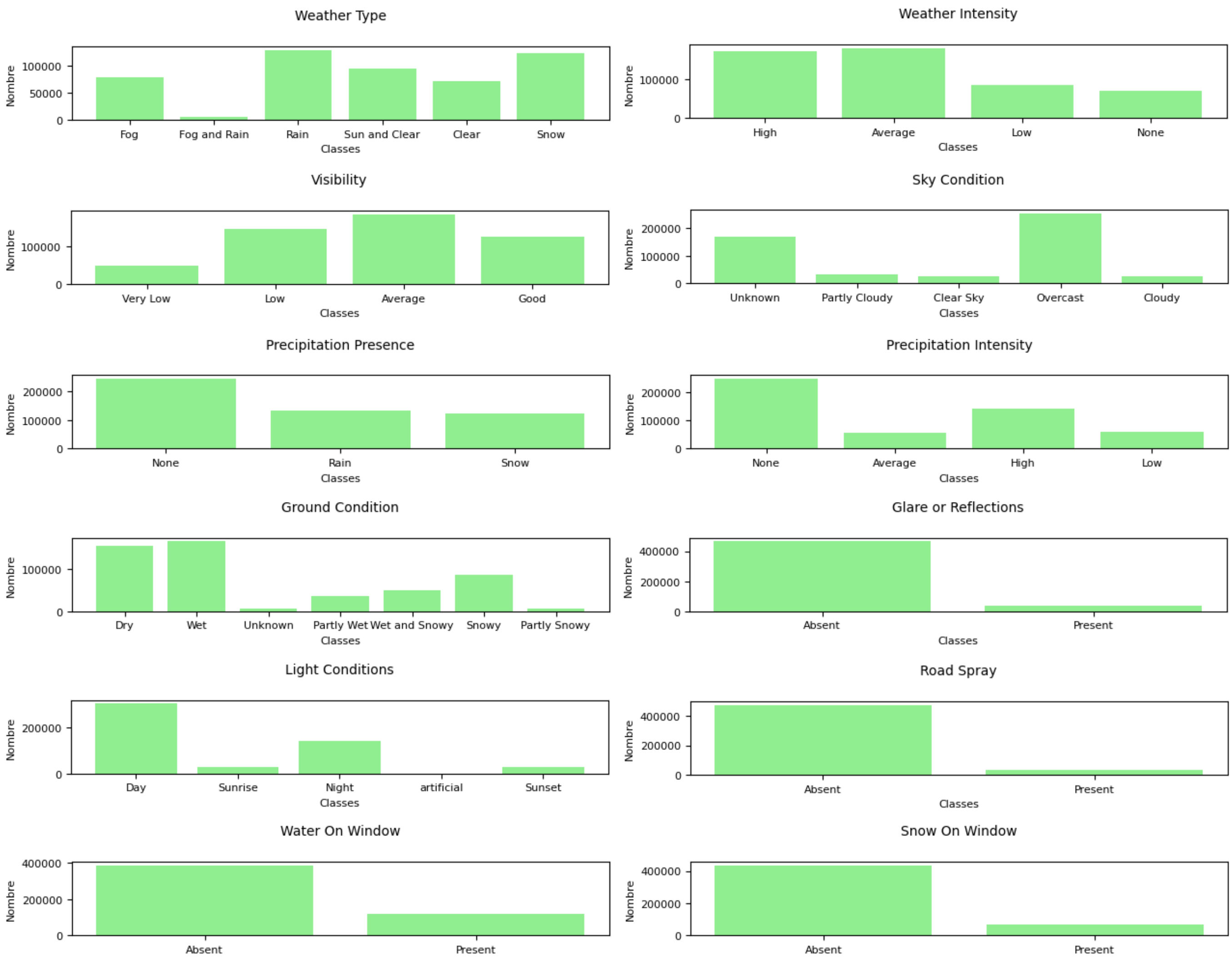


Figure 15: Distribution of labels for the 12 criteria/tasks in the TT dataset.

3.3.2.2 Evaluation: explainability, generalisation capacity, and ablation study

After the testing phase, we examine the explainability of our models in section 3.4.2. We begin with a TSNE analysis of the RTM model on the TT dataset, which allows us to visualize the distribution of intermediate representations (features). If these representations are well clustered by class, it indicates that the first phase of the model—the extraction of stylistic features (embeddings)—is effective. Such a structure ensures that the second phase, classification, is both feasible and simplified.

Following this macroscopic analysis, we deepen the explainability study by focusing on a specific image extracted from the DEMO dataset (see next section 3.3.2.3 for DEMO dataset description). To do so, we use the GRAD-CAM (Gradient-weighted Class Activation Mapping) tool [54] to visualize the regions activated by the attention mechanism.

Section 3.4.3 then evaluates the generalization ability of our two models. We apply them to state-of-the-art (SOTA) datasets from the literature (BDD100K, Image2Weather, WeatherNet-05) without any retraining (zero-shot approach), which allows us to compare their performance to existing work on several of the 12 defined tasks.

Finally, Section 3.4.4 presents an ablation study. This approach involves removing specific components of the model to assess their individual contributions to overall performance. It provides a precise evaluation of the added value of our method compared to the reference approaches on which it is based.

3.3.2.3 Demonstration

After a thorough evaluation of our models, we propose a demonstration to illustrate the capabilities of our solution. This is based on a second dataset created specifically for this study. This second dataset (demonstration dataset – DEMO dataset) was built as part of the ROADVIEW project during experiments conducted in Finland (see section 2). The demonstration covers the entire processing chain: from weather condition detection by the camera, through the V2X communication module, to the adaptation of the autonomous vehicle's behaviour.

In this section, we focus solely on the results obtained from the camera-based weather detection method. For this purpose, the dataset includes two camera viewpoints on the same site. We also have access to weather data measured by sensors installed at the same test location. The complete dataset contains 183,940 images. Figure 16 presents examples of images captured under various conditions and from both viewpoints.

The analysis conducted on this second dataset (DEMO dataset) is structured in two steps: first, we compare the detection results obtained by our model from the two camera angles. In theory, with a perfect model, the two detections should be identical across all images. Then, we use a subset of the dataset manually annotated by a human operator. This subset, which contains 1,000 images, is used to assess the success rate of our algorithms during the demonstration. The results of this demonstration are presented in section 3.4.5.

a. point of view 1 *b. point of view 2*

Figure 16: Sample images from the DEMO dataset, created during a measurement campaign for the ROADVIEW project in Finland.

3.4 Results

3.4.1 Testing result

Table 3 shows the F1 score results obtained by our two algorithms (PGM and RTM) on the TT dataset.

Table 3: F1-scores by task and model, on TT dataset.

Task	PGM model	RTM model
Weather Type	0.9821	0.9916
Weather Intensity	0.9707	0.9828
Visibility	0.9655	0.978
Sky Condition	0.978	0.9836
Precipitation (presence)	0.9867	0.995
Precipitation Intensity	0.9823	0.9916
Ground Condition	0.965	0.9744
Glare	0.9801	0.9817
Lighting Conditions	0.9925	0.9966
Water Spray	0.9727	0.9741
Water on Windshield	0.9859	0.995
Snow on Windshield	0.9944	0.9963
Average F1	0.9797	0.9867

Overall, the models exhibit excellent performance, with F1 scores exceeding 0.97 for most tasks and our two models. These results confirm the relevance of our approach, which considers weather as a key factor in the style and appearance of images.

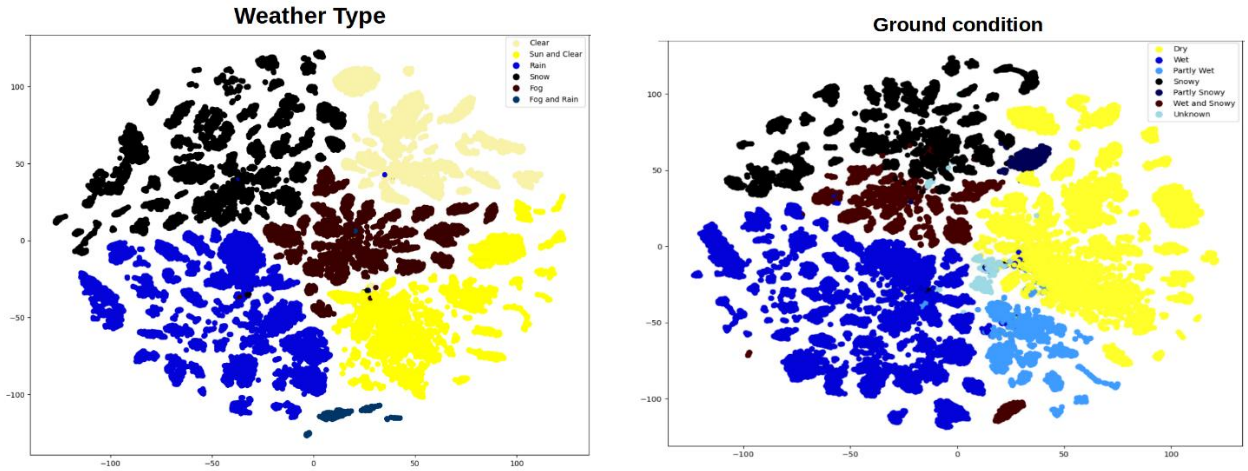
Moreover, the PGM model provides the best trade-off in terms of speed (208 images/second), lightweight architecture (3.3M parameters), and performance (average F1 score of 0.9797). It emerges as the optimal solution, balancing efficiency and low computational cost. The RTM model delivers better performance (average F1 score of 0.9867) and, as we will see, generalizes much more effectively than PatchGAN. However, it has significantly more parameters (12.5M), making it slower (164 images/second) and less suitable for embedded systems.

Beyond the overall scores, it is interesting to consider the explicability and generalization capacity of our two models. This is the subject of the following sections.

3.4.2 Study on model explainability

3.4.2.1 TSNE

In this study we have considered weather as a key factor in the style and appearance of images. This hypothesis is further supported by the analysis of embeddings produced by the models before classification, visualized using TSNE [55], [56]. Examples obtained with the RTM model are illustrated in Figure 17. The analysis reveals a clear separation between different weather conditions (Figure 17.a.), with only a few marginal errors. These results validate our approach, which considers weather as a key factor in the style and appearance of objects in an image. The TSNE shows also that the attention layers of our model —whose embeddings are used for the visualization— separate the different ground conditions quite well (Figure 17.b.). We can clearly distinguish between dry, wet, and snowy surfaces, which demonstrates that our model effectively captures the information related to each characteristic. A single embedding method can therefore be used in preparation for multiple classification tasks. Once this embedding has been carried out, the classification task is simplified and can therefore be carried out with classifiers containing few parameters.

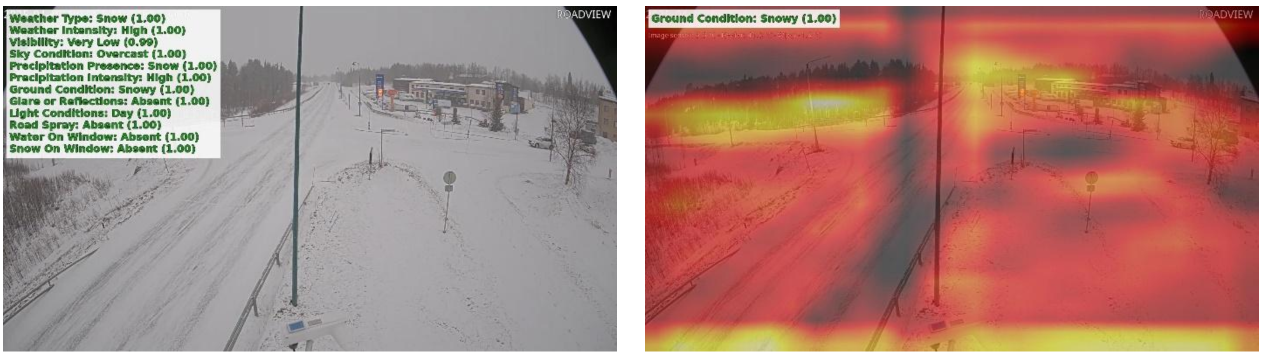


a. weather type *b. ground condition*

Figure 17: TSNE applied to the embeddings extracted by the attention mechanism dedicated to: a. weather type classification and b. ground condition on 50,000 test images of the TT dataset, using the RTM model.

3.4.2.2 GRAD-CAM

We use GRAD-CAM (Gradient-weighted Class Activation Mapping) [54] to visualize the activations of the attention mechanism dedicated to analysing ground conditions. As illustrated in Figure 18.b., the PGM model, focuses on the most relevant areas of the image, namely the ground, to make its prediction (snow ground). These observations further support our hypothesis that each weather condition imprints a distinct style on images. Thus, by leveraging stylistic features through attention mechanisms [50], [51], [52], [53], it is possible to accurately extract specific meteorological information.



a. Original image *b. Grad-CAM*

Figure 18: Calculation of the grad cam (b.) on a sample image (a.) from our DEMO dataset.

3.4.3 SOTA comparison: study of generalization capacity

Table 4 shows that the models scores do not collapse when evaluated on datasets completely outside the training domain. Without any adaptation, we remain within ≤12 F1 points of the in-domain SOTA and outperform a classic ResNet-50 by +18 F1 points. With a short retraining (10 epochs), RTM reaches 0.92 F1 on BDD100K and 0.97 F1 on Image2Weather, narrowing the gap to less than 5 points while keeping the model 5x smaller. Table 4 confirms that the PatchGAN fusion (in the PGM model) is the key to this transferable meteorological style.

Table 4: F1 scores (↑) on SOTA datasets, comparison of our RTM model with zero-shot and 10-epochs fine-tune against SOTA models and the classic ResNet-50. *In the SOTA paper, the F1 score is not mentioned, only the accuracy score is mentioned.

Dataset	SOTA [Reference]	ResNet-50	RTM 0-shot	RTM +10ep
BDD100K	0.973 (F1) [57]	0.560	0.849	0.920
Image2Weather	0.924 (F1) [58]	0.625	0.833	0.972
WeatherNet-05	0.983 (Acc*) [59]	0.652	0.814	0.949

3.4.4 Ablation study

Table 5 shows the ablation study results comparing the standard PatchGAN model (baseline) and the PGM model. As shown on Table 5, adding the attention mechanism to the standard PatchGAN model improves its performance by enabling it to focus on the relevant regions of the image for its predictions (F1 score changes from 0.657 to 0.701, as the same way, Area Under the Curve (AUC) score changes from 0.757 to 0.794). At the same time, the number of parameters increases only slightly (from 3.2M to 3.3M). After demonstrating the relevance of our models on different databases, we present a demonstration of their performance as part of the ROADVIEW project in the next section.

Table 5: Ablation study: the impact when removing our attention mechanism on a zero-shot detection (here scores are estimated on the BDD100K dataset).

Model	Number of parameters	F1 (↑)	AUC (↑)
PatchGAN baseline	3.2M	0.657	0.757
+ Attention → (PGM = ours)	3.3M	0.701	0.794

3.4.5 Demonstration

To evaluate the generalization ability of our models, we use our DEMO dataset (see section 3.3.2.3 for definition). In the DEMO dataset, two cameras are installed at the same site in Finland but oriented from two different viewpoints (see Figure 16). The goal is to verify whether the model consistently predicts the same weather characteristics from both perspectives. The Table 6 reports, out of a total of 13,054 images, the number of cases in which the model produces identical predictions across all weather characteristics.

As shown in Table 6, the model achieves very good results (>86%) on many tasks (Weather Type, Precipitation Presence, Precipitation Presence, Precipitation Intensity, Ground Condition, Light Conditions, Road Spray, Water On Window, Snow On Window). In addition, it achieves acceptable results on all other tasks (>60%) except for Sky Condition. This makes sense, since the camera angles do not capture the sky (see Figure 16). From viewpoint 1, the sky is entirely out of frame, so the model predicts “unknown” most of the time. As shown in the table, the model predicts the same weather type from the two different viewpoints in at least 86% of cases. For more specific characteristics, such as precipitation intensity or ground condition, the results are even better, reaching up to 90% consistency.

Table 6: Consistency of RTM model on the DEMO dataset. We count the number of coherent images, i.e. for which the RTM model gives the same classification on both camera viewpoints.

Task	Number of coherent images	Percentage of coherent images
Weather Type	11 356	86.95
Weather Intensity	8 297	63.53
Visibility	7 889	60.41
Sky Condition	4 717	36.12
Precipitation Presence	12 246	93.77
Precipitation Intensity	11 915	91.23
Ground Condition	11 806	90.40
Glare or Reflections	8 418	64.46
Light Conditions	13 017	99.67
Road Spray	13 048	99.91
Water On Window	12 989	99.46
Snow On Window	13 051	99.93

After checking the consistency of the model, we evaluate the model's performance on 1,000 images captured by the camera at viewpoint 2 and annotated by a human. We observe in Table 7 that its performance remains similar to that obtained when testing on standard benchmarks. The model does not collapse on the other tasks, maintaining an F1 score of around 78% for most of them. The exception is the sky condition, where performance is lower — likely due to the camera capturing only a very limited portion of the sky.

Table 7: Validation of the RTM model on 1000 images from viewpoint 2 of the DEMO dataset. These images were annotated by a human.

Task	Precision (↑)	Recall (↑)	F1 Score (↑)
Weather Type	0.9177	0.7019	0.7839
Weather Intensity	0.9209	0.8149	0.8033
Visibility	0.8960	0.7135	0.7873
Sky Condition	0.2365	0.1124	0.0593
Precipitation Presence	0.9126	0.7619	0.8170
Precipitation Intensity	0.8817	0.8165	0.8478
Ground Condition	0.8578	0.8500	0.8500
Glare or Reflections	0.7590	0.7641	0.7538
Light Conditions	0.8562	0.9248	0.8892
Road Spray	1.0000	0.9990	0.9995
Water On Window	1.0000	0.9802	0.9900
Snow On Window	0.9395	0.9353	0.9050

3.5 Conclusion

Our study shows that considering weather as a visual style is a particularly effective paradigm for real-time, multi-attribute detection of weather conditions from camera images. By combining stylistic feature extraction, attention mechanisms, and lightweight neural architectures, we propose robust models that generalize well across various benchmarks and in real-world conditions.

Key contributions:

- **High accuracy and real-time performance:** The PGM model achieves excellent F1 scores (above 0.97 on average) on most weather-related tasks, while maintaining the speed and efficiency required for embedded systems.
- **Generalization and robustness:** Our models remain stable and performant even when evaluated in zero-shot settings or on entirely new datasets, outperforming standard baselines such as ResNet-50 by a significant margin (+18 F1 points). With minimal retraining, the RTM model achieves state-of-the-art results on public datasets while remaining up to five times more compact than competing approaches.
- **Explainability:** TSNE and GRAD-CAM analyses confirm that our models learn interpretable stylistic embeddings that correspond to human-understandable weather attributes, facilitating transparent decision-making. Our ablation study highlights the importance of each architectural component, particularly the contribution of attention mechanisms to model robustness and accuracy.
- **Consistency in real-world deployment:** When tested from two different camera viewpoints in real conditions, our models provide consistent predictions for the vast majority of attributes, demonstrating their practical reliability and transferability.

Despite these advances, there is still a lot of **perspectives and future work:**

- **Advanced instrumentation:** Building a corpus with instrumented ground truth will allow us to refine intensity estimation and improve model calibration.
- **New attributes and modalities:** Extending our approach to new environmental factors (such as quantifying water or frost on the ground) will enhance operational relevance.
- **Edge deployment:** Ongoing work aims to port these models to embedded GPUs and FPGAs for seamless integration into embedded platforms.

In summary, our results establish style-sensitive, attention-based neural architectures as a new standard for vision-based weather detection, combining accuracy, speed, explainability, and adaptation to real-world needs. In this section, we have described how we measure weather attributes from surveillance camera images. The next section explains how this information is used and merged with other information to produce messages broadcast to surrounding vehicles.

4 V2X road weather service

In the first report D5.3, we have identified relevant C-ITS services for ROADVIEW perception system. One of those services is a Road Weather Service to share weather-related data between road users.

The road weather service generates V2X messages called Road Weather Message (RWM) and contains the weather estimation from the module described in section 3. So far, in Europe, ETSI ITS Technical Committee in charge of specifying ITS protocols has not specified a service for road weather applications similar to SAE J2945 [60] where weather data can be broadcasted every 1 seconds. Weather-related information can be broadcasted using ITS messages in case of harsh weather conditions using DENM (Decentralised Environmental Notification Message) [61]. DENM are event-based whereas the ROADVIEW RWMs are an awareness service with regular broadcast of local weather conditions whatever is the weather situation.

Figure 19 shows an example of detected weather conditions with the roadside camera installed in Finnish Lapland.

Figure 19: Example of weather detection with a scene recorded with wet road and light snow.

The most recent detected weather conditions were transmitted using RWM once per second, so that new vehicles approaching to the RSU can receive the local weather conditions. For our experimentation we select to broadcast the “Weather Type” with “Weather Type Confidence”, “Weather Intensity” and “Ground Condition” using the RWM format shown in Figure 20.

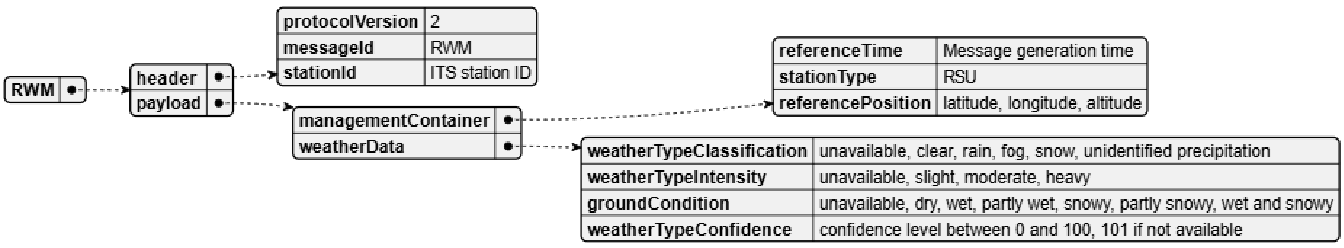


Figure 20: Roadside Road Weather Message Format.

In addition, our Road Weather Service is also able to receive RWMs from vehicles having onboard weather estimator for visibility or slipperiness as described in D5.5 (Environmental and Weather Condition State Estimates) with the RWM format shown in Figure 21.

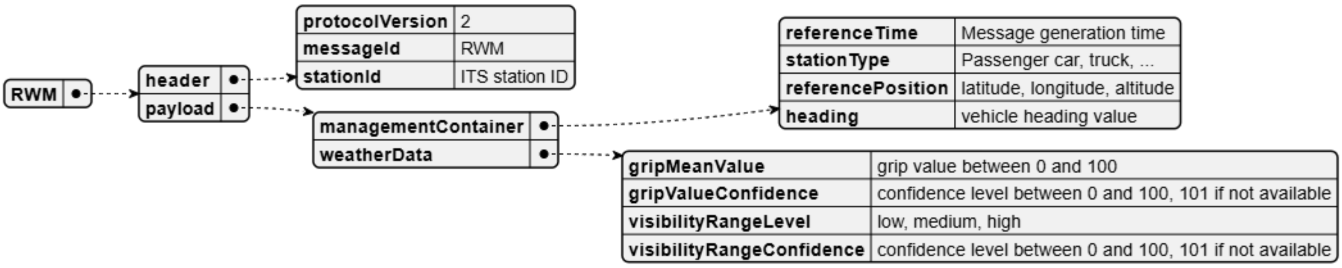


Figure 21: Onboard Road Weather Message Format.

Figure 22 shows the developed software modules to implement the ROADVIEW Road Weather Service with the infrastructure-based system at the roadside and the onboard system to be implemented in the vehicle:

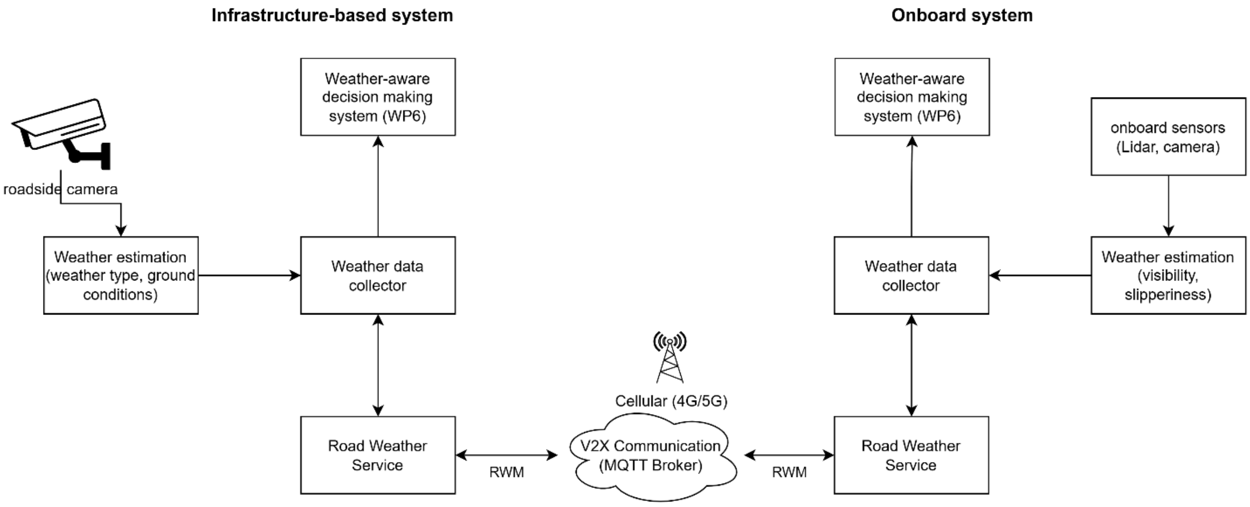


Figure 22: V2X road weather service software modules.

The road weather system is composed of:

- **Weather estimation modules**:
  - Camera-based weather detection (CE), described in section 3,
  - Onboard weather estimations (visibility and slipperiness) developed in task 5.3 of ROADVIEW project,
- **Weather data collector module** collecting the distinct weather estimation both from local sensors or from remote sensors. Other modules such as the weather-aware decision-making system modules developed in WP6, can access to the local weather situation from the weather data collector module. For example, the infrastructure-based system can provide driving advice to surrounding vehicles using V2X messages.
- **Road Weather Service** transmitting and receiving Road Weather Messages through **V2X MQTT (**Message Queue Telemetry Transport) **broker** over cellular network connectivity.

Long range connectivity for road weather service applications was already tested in Finland artic weather conditions as reported in [62] with the installation of road weather sensors along the test track. In the ROADVIEW experiments done in Finnish Lapland (see section 2), the roadside camera images were used both for weather detection with a frame rate of 1 image per minute, and for object detection with a frame rate of 10 image per second (see section 5). Using the same sensor, both types of detections were computed on the same roadside unit system in real-time.

Figure 23 shows an example where a leading vehicle (Weather probe vehicle) can use a RWM to inform the RSU about a portion of the road that is icy. The RSU can then alert the vehicle that intends to drive on the icy road using the V2X messages.

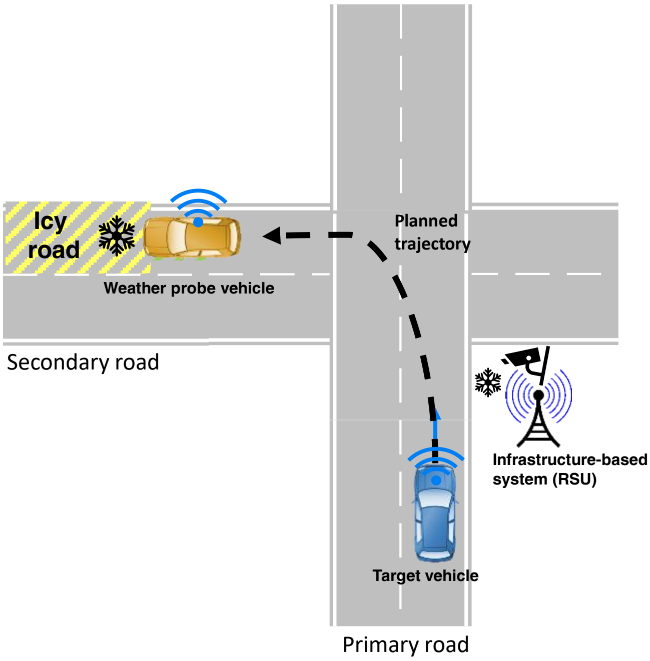


Figure 23: Awareness of weather for decision-making systems.

In the next section we will describe the study done on the roadside sensor fusion system and the limitations caused by difficult perception conditions such as under harsh weather conditions.

5 Roadside sensor fusion

As part of the collaborative perception systems, we propose a roadside sensor fusion system taking into account the roadside sensor Operational Design Domain (ODD).

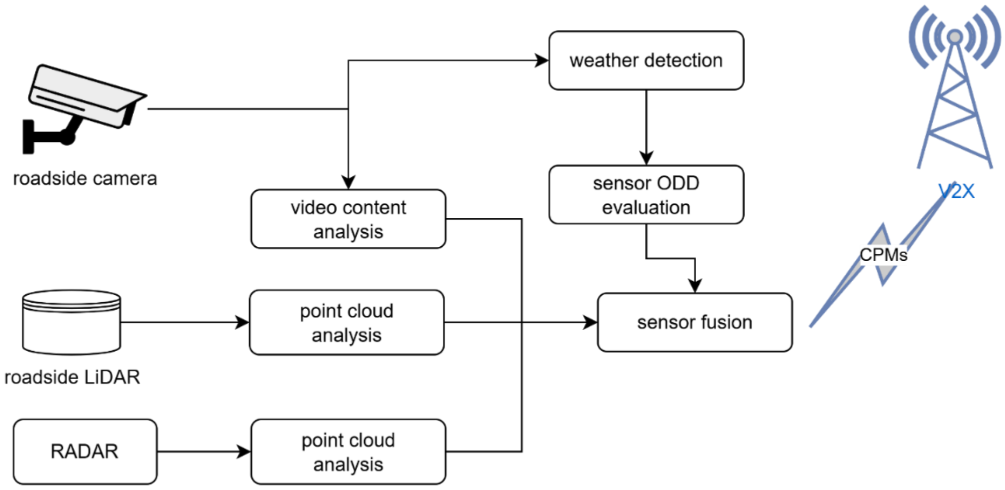


Figure 24: Infrastructure-based perception system.

The infrastructure-based perception system (Figure 24) takes camera images, LiDAR point-clouds and radar point-clouds as input. It processes them to generate perception results in the form of a list of detected objects. The perception results from the different sensors can be combined by the roadside sensor fusion module and transmitted to road users through V2X Collaborative Perception Messages (CPMs). In addition, the roadside infrastructure may receive perception results from vehicles and combine this information with its own perception results to enhance its representation of the traffic state.

The output of roadside sensors may be disrupted by harsh weather conditions. The experimental setup installed in Finland, allowed to record several scenes in different weather conditions, enabling to study the response of the sensors to these weather conditions and to determine how to adapt the processing to best handle these conditions. The camera-based weather detection system described in the previous section can be used to assess in real-time the weather conditions.

5.1 Roadside sensor evaluation

This section describes the evaluation of roadside sensors in different weather conditions. Then it explains how the strengths of the different sensors can be utilized in sensor fusion and how to adapt sensor fusion to the different weather conditions. It further shows the advantages of using a global fusion between the roadside's perception results and a vehicle's perception results. Last, it describes the outcomes of the experimentations realized in Finland.

5.1.1 Recorded scenes

Different sensors have different Operational Design Domains (ODDs). These ODDs may change depending on the weather conditions. For instance, when there is a dense fog, a camera perception distance will be reduced. The experimental setup installed in Finland (see section 2) enabled to record scenes in different weather conditions using two different infrastructure sensors: a camera and a LiDAR. In particular, some scenes were recorded while a vehicle (from VTT or from FGI) was driving in the sensors' field of view in the various possible directions. This enabled creating a dataset with both sensor recordings and accurate measurements of a vehicle position.

This dataset contains a total of 43 scenes with four weather types and ground conditions:

- Snow-covered road, light snow, day; 4 scenes; VTT vehicle;
- Dry road, cloudy, night; 10 scenes; 14 scenes; FGI vehicle;
- Dry road, cloudy, day; 10 scenes; 13 scenes; FGI vehicle;
- Wet road, light snow, day; 12 scenes; FGI vehicle.

Sample of roadside camera and LiDAR detections for the four different conditions are shown in Figure 25, Figure 26, Figure 27 and Figure 28.

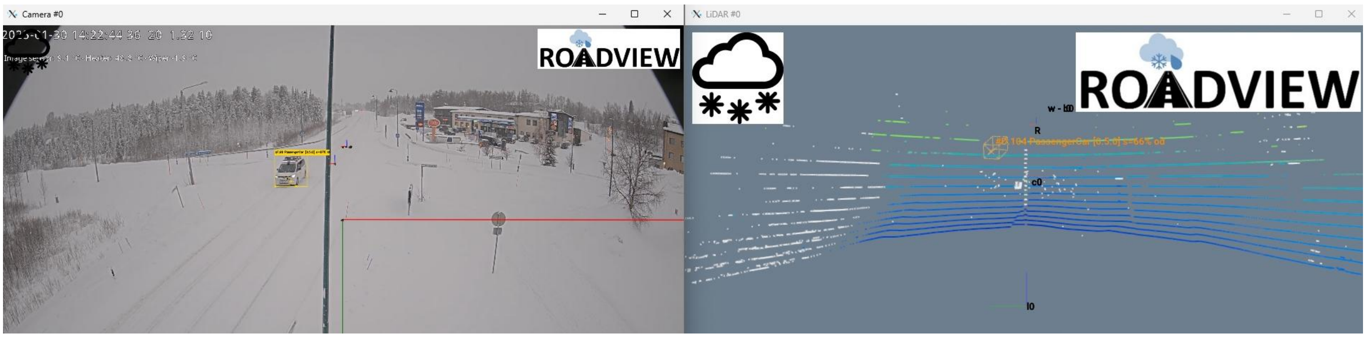


Figure 25: Roadside sensor detection for snow covered road, light snow with VTT vehicle.

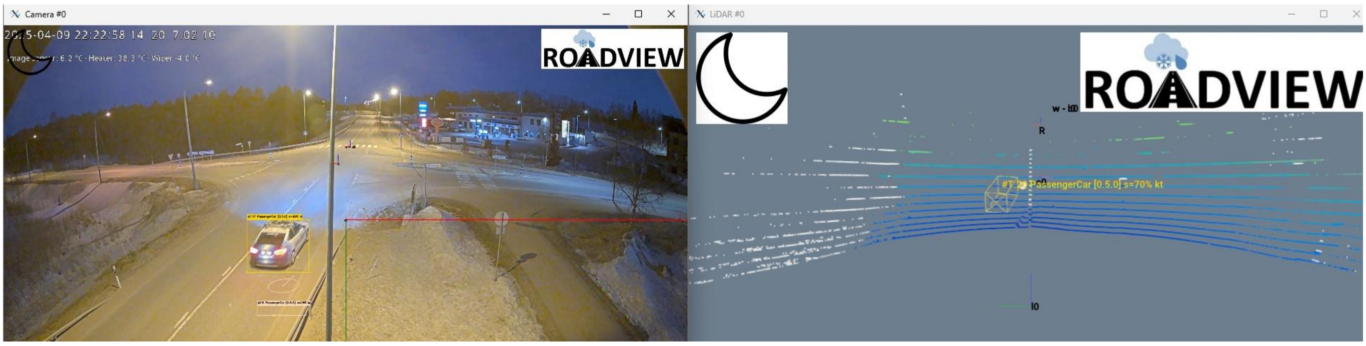


Figure 26: Roadside sensor detection for dry road, nightlight with FGI vehicle.

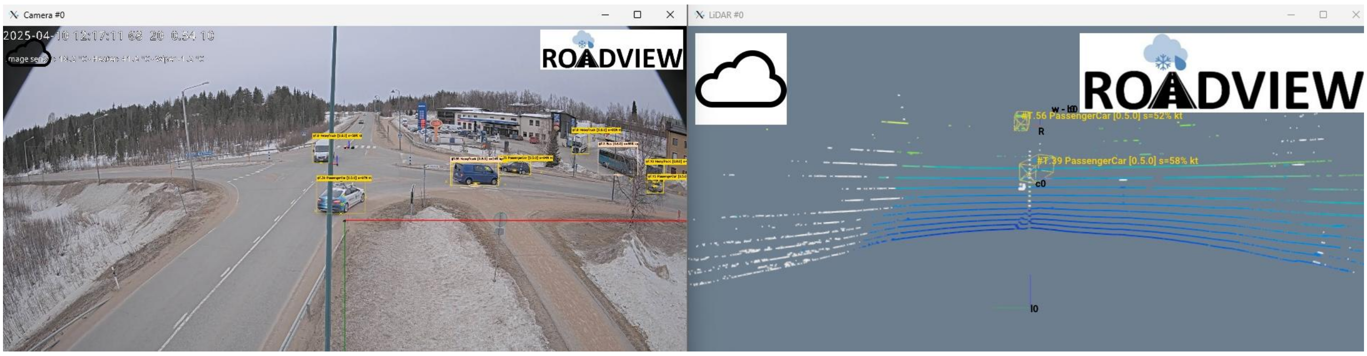


Figure 27: Roadside sensor detection for dry road, cloudy, daylight with FGI vehicle.

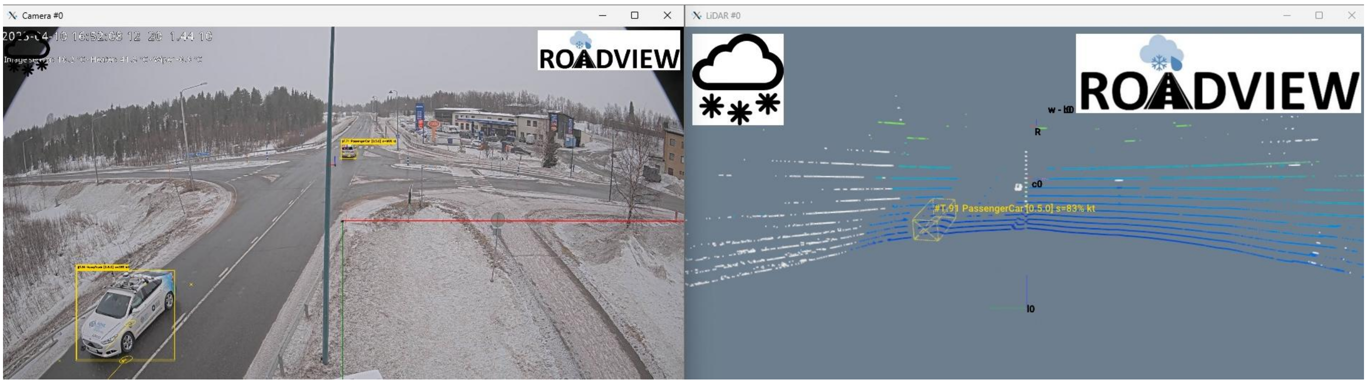


Figure 28: Roadside sensor detection for wet road, light snow, daylight with FGI vehicle.

5.1.2 Sensor evaluation under different weather conditions

The captured scenes were used to study the performances of the infrastructure perception system depending on the weather conditions. For each roadside sensor, the performance was evaluated by measuring the number of detections relative to the number of measured vehicle positions, *i.e.*, the recall for the perception system.

5.1.2.1 Camera evaluation

The evaluation for the roadside camera was restricted to the main road going from the bottom to the top of the image. The crossing road has slopes that are not well handled by the infrastructure perception system. This reduced the number of scenes to 33, discarding scenes where the vehicle stays on the crossing road.

The evaluation results are grouped according to two criteria: the weather conditions and the path travelled by the vehicle. Figure 29 shows the various travelled paths considered for this evaluation on the camera view. The different trajectories were grouped into six different paths, corresponding to the two possible directions and three different parts of the road: travelling the full length of the road (Figure 30), travelling the bottom part of the road (Figure 31), or travelling the top part of the road (Figure 32). Travelling the bottom part of the road corresponds to the vehicle starting from the bottom of the image and then turning either left or right, or inversely, coming from the left or right and turning towards the bottom. Similarly, travelling the top part correspond to a turn to or from the left or right.

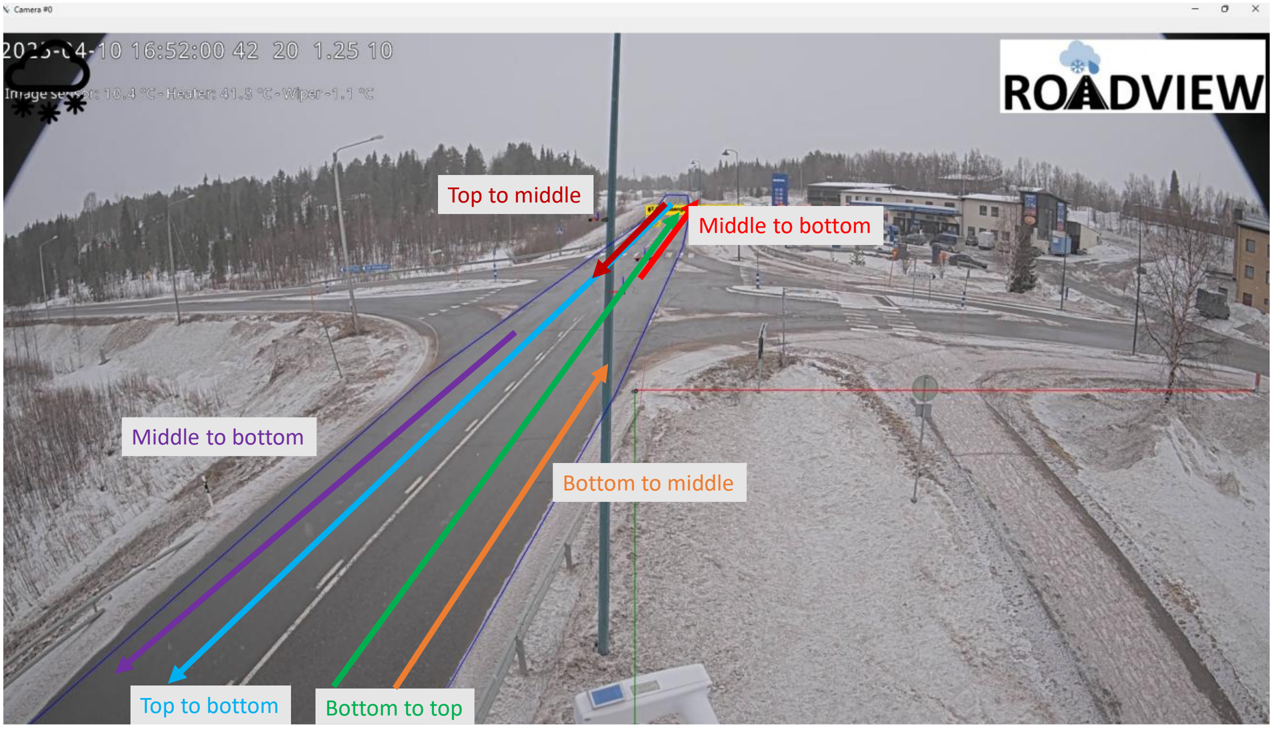
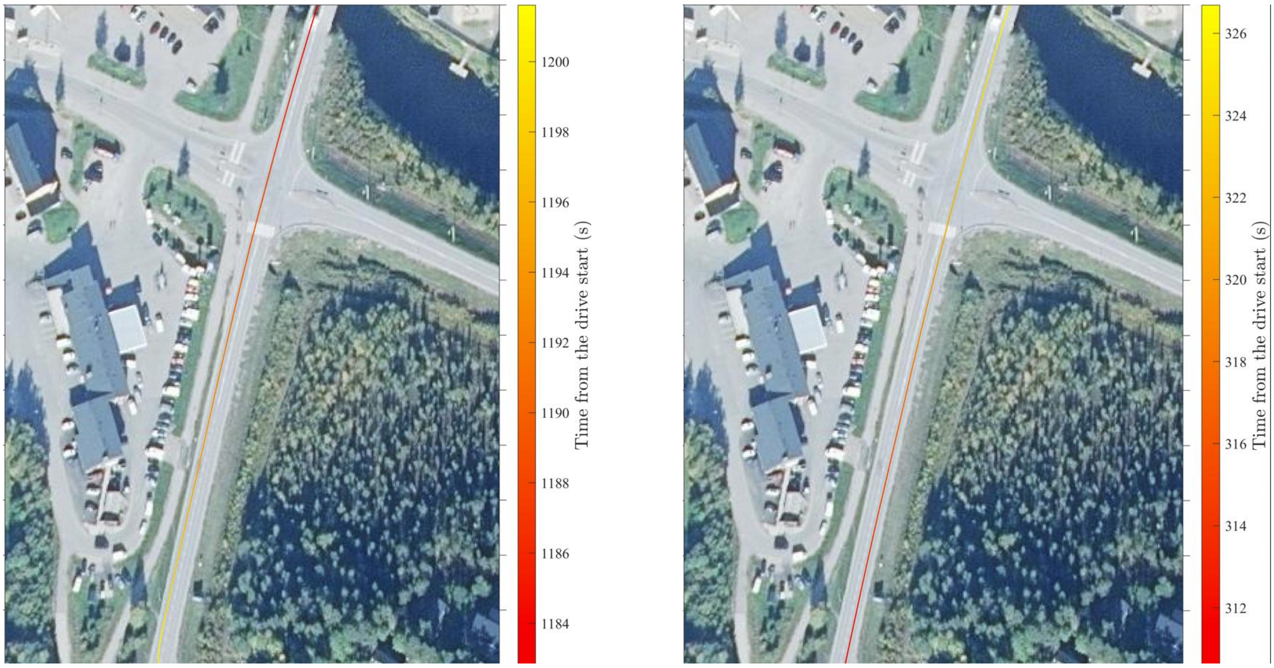


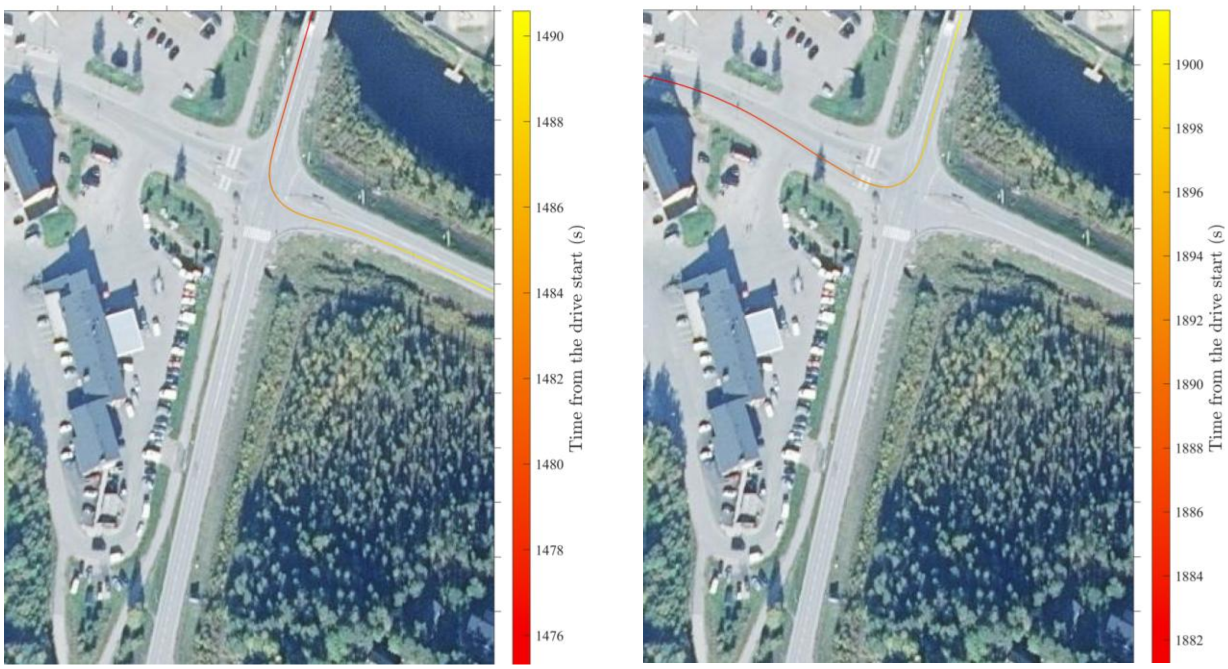
Figure 29: Camera view of the travelled paths considered in the evaluation.

Figure 30, Figure 31 and Figure 32 show some of the vehicle trajectory measurements recorded with VTT and FGI vehicles.



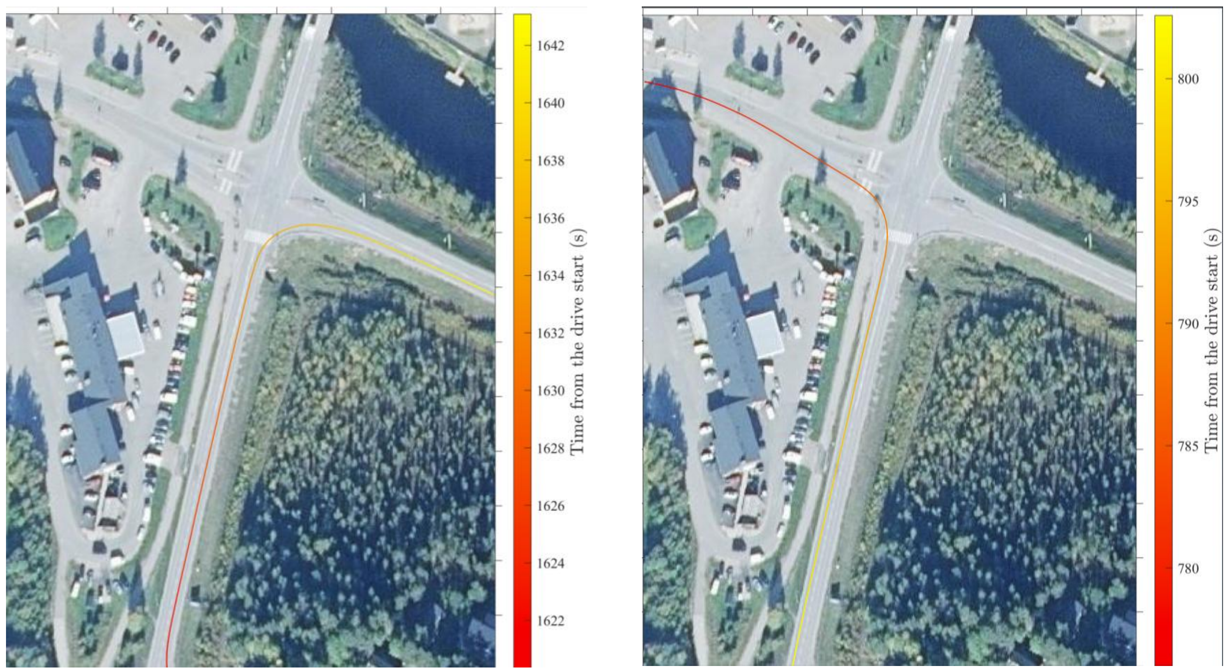
(a) Bottom to top (b) Top to bottom

Figure 30: Vehicle trajectory measurements travelling the full length of the road.



(c) Bottom to middle (d) Middle to bottom

Figure 31: Vehicle trajectory measurements travelling the bottom part of the road on the camera view.



(e) Top to middle (f) Middle to top

Figure 32: Vehicle trajectory measurements travelling the top part of the road on the camera view.

Figure 33 shows the recall value of the roadside camera perception system depending on the weather conditions (shown using different colours) and the trajectory represented on the horizontal axis.

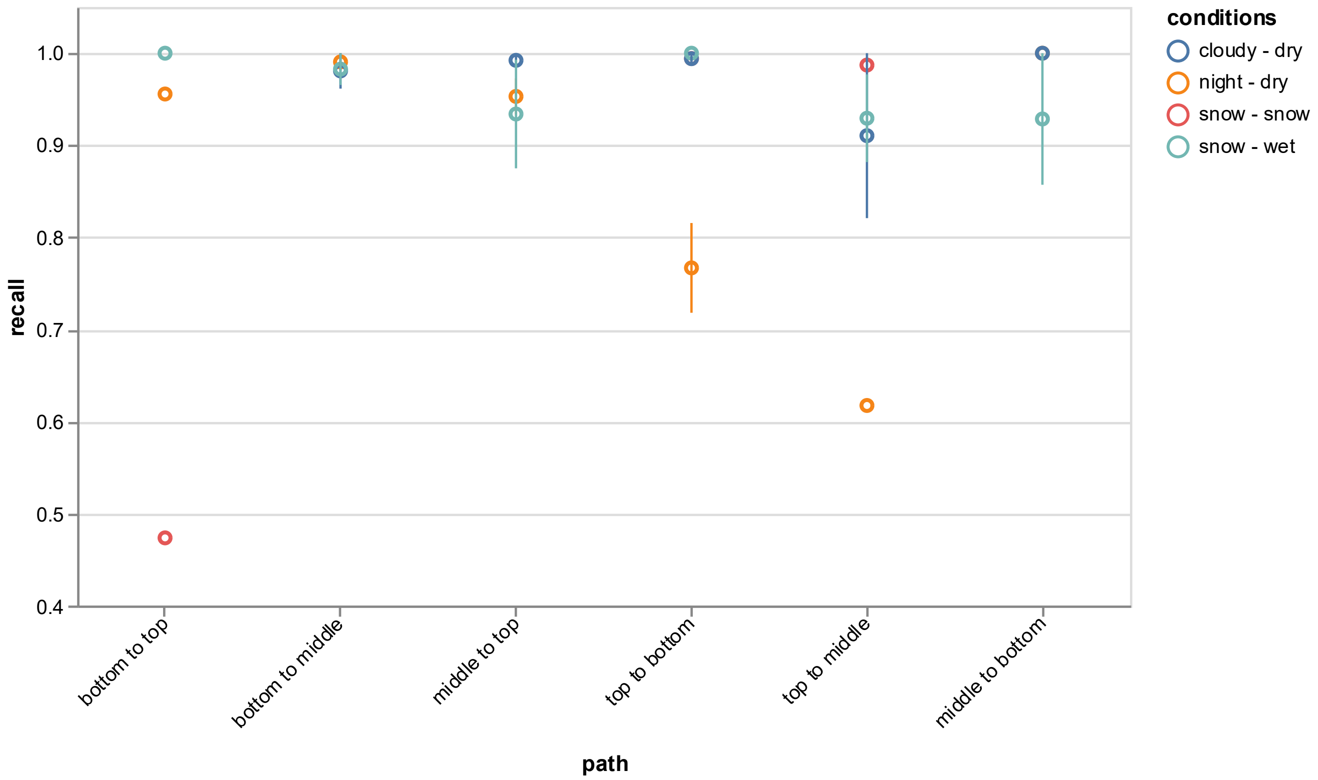


Figure 33: Performances of the roadside camera perception system depending on the weather conditions.

The evaluation results for the roadside camera show that the performance of the perception system is good, usually ranging from 90% to 100% recall. There are two difficult cases for the camera-based perception system. First, when the road is covered with snow, the vehicle is not detected when going away from the camera when far from the sensor (bottom to top case). This seems mostly due to the fact that VTT's vehicle is white and has its rear covered with snow and also partly due to the cloud of snow raised by the vehicles (Figure 34). A comparison with detection results for other vehicles shows that vehicles with darker colours are better detected than those with lighter colours. However, the detections of vehicles going away from the camera (bottom to top or middle to top cases) are not as good as in better weather conditions.

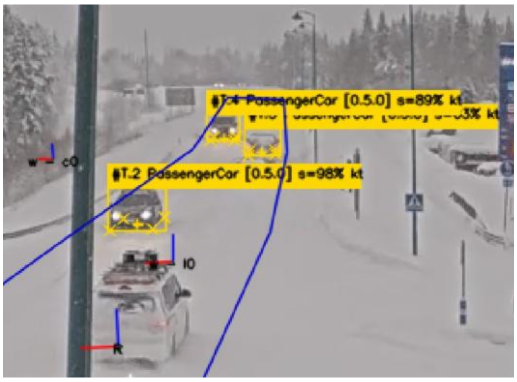


Figure 34: Vehicle not detected in the bottom to top path – snow on the road.

Second, at night, vehicles coming towards the camera are not well detected when far from the camera (top to bottom or top to middle cases). This is due to the vehicle's lights that blind the camera (Figure 35).

Figure 35: Vehicle not detected in the top part of the road – nightlight.

Looking at the scenes recorded with snow-covered or wet road, it seems that the high position of the roadside camera helps reducing the impact of the cloud of snow or of the mist generated by vehicles a lot compared to cameras with lower position such as those mounted on a vehicle.

5.1.2.2 LiDAR evaluation

For the LiDAR, the evaluation is further restricted to a smaller section of the road, corresponding to the volume of the point cloud processed by the detection system. This volume is limited to 70 metres in front of the LiDAR, *i.e.*, stopping just after the intersection. Figure 36 shows the evaluation result for the roadside LiDAR perception system similarly as for the roadside camera showing the recall value per weather conditions, shown using different colours, and the trajectory represented on the horizontal axis.

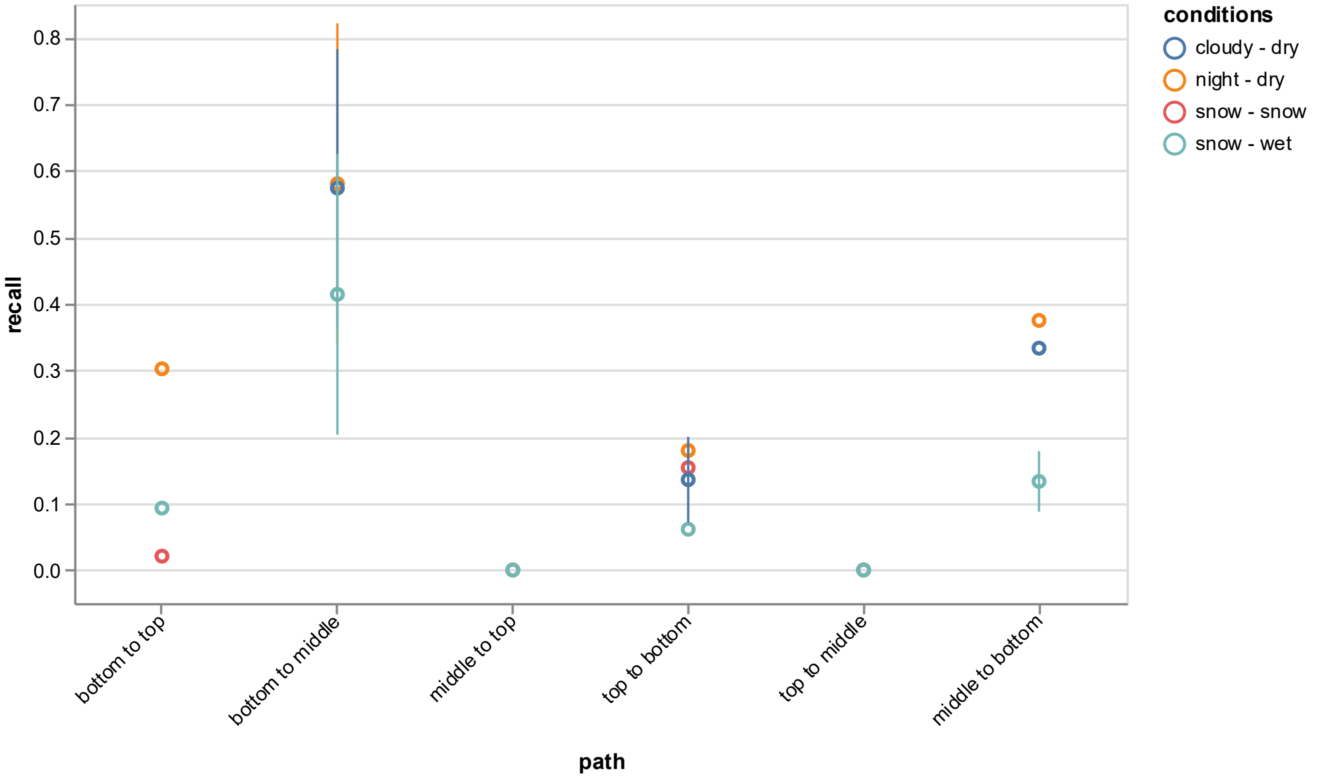


Figure 36: Performances of the roadside LiDAR perception system depending on the weather conditions.

The evaluation results for the LiDAR show different performances compared to the camera results. First, the recall is far lower for the LiDAR than for the camera. In addition, even if the LiDAR's evaluation region doesn't extend as far as the camera's evaluation region, the LiDAR's perception system is not able to detect any vehicle not within a close distance from the sensor's position. In particular, vehicles located at the intersection of the two roads or further away are not detected. This is mainly due to the fact that the LiDAR used is a 16-channel LiDAR, capturing only 16 rows of points vertically. Even if the LiDAR is able to capture points quite far away, it provides no or few details for objects that are not close to its position.

In addition, the perception performance for the LiDAR is better when the vehicle is going away from the sensor than when it is going towards the sensor. A first explanation is that when the vehicle is close to the sensor, it is initially well detected by the roadside perception system. When the vehicle moves further away, the vehicle becomes harder to detect but the tracking algorithm of the roadside perception system can help continuing the perception. This is not the case when the vehicle is going towards the sensor. In addition, when the vehicle drives only at the bottom part of the image, it is on a priority road when coming from the bottom and can turn rather quickly. Inversely, when going towards the bottom, it comes from a non-priority road and has to go more slowly. This means that when coming from the bottom, the vehicle spends less time at the crossroad where is it not well detected than when going towards the bottom.

The LiDAR perception performance is also impacted by the weather conditions. When the road is wet, the performances of the LiDAR is lower than in the other conditions. Indeed, when the road is wet, the road surface is poorly detected by the LiDAR, probably due to a higher reflectivity, making it harder for the perception system to identify vehicles. The difference between the point clouds captured when the road is dry or when it is wet is shown in Figure 37.

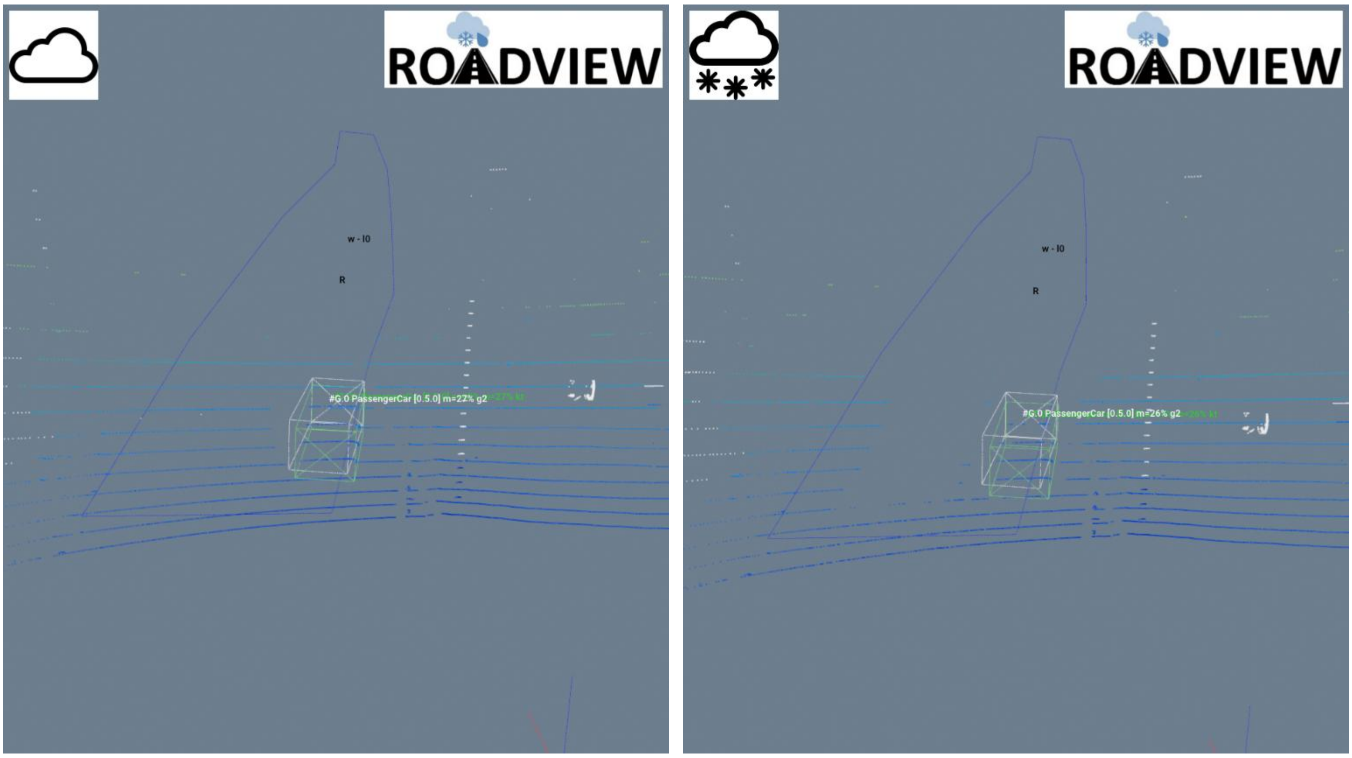


Figure 37: roadside LiDAR recording depending on road surface (left: dry, right: wet).

Lastly, under snowy conditions (a snow-covered road and light active snow), the vehicle is almost not detected when going away from the LiDAR. This is probably because the detection system is not well adapted to detect a minibus from the rear. Indeed, the vehicle in use for snowy conditions was VTT's while under other conditions, it was FGI's sedan.

These results show the variation of perception performances for the camera and the LiDAR under different weather conditions. Unfortunately, during the recordings, there were no harsh weather conditions such as heavy rain, heavy snow or dense fog. However, these results already give some useful indications on how the ODDs of both sensors change under different weather conditions and how their respective perception performances are impacted.

5.2 Roadside sensor fusion

5.2.1 Taking advantage of different sensors

Different types of sensors have different advantages and drawbacks. Therefore, combining several sensors enable to improve the perception performances of the roadside infrastructure. In [63], Wan et al. describe a fusion system where each sensor's output is first processed independently before fusing all the processed outputs into a unified perception output. Their fusion system treats the processed outputs of the different sensors separately. For instance, class labels from the camera are given a higher importance due to the better semantic accuracy of the camera processing system. Similarly, object's shapes from the LiDAR are given a higher importance due to the precision of the 3D measurements of this sensor. Velocity estimates from the radar are given a higher importance because of the direct Doppler-based measurement mechanism of radars.

In their review [64], Yeong et al. propose a comparison of the relative strength of the different sensors. Similarly, in CRF's analytics, the perception results from the camera should be given a higher importance than those from the LiDAR for the objects' type and its confidence. In addition, the perception results from the LiDAR should have a higher importance than those of the camera for the objects' position, speed, orientation and orientation speed. Currently, the analytics used by CRF to process camera images do not provide any information about the objects' size, therefore only the perception results from the LiDAR are used for the objects' size. These results are summarized in the Table 8:

Table 8: Relative importance of camera and LiDAR perception results for sensor fusion.

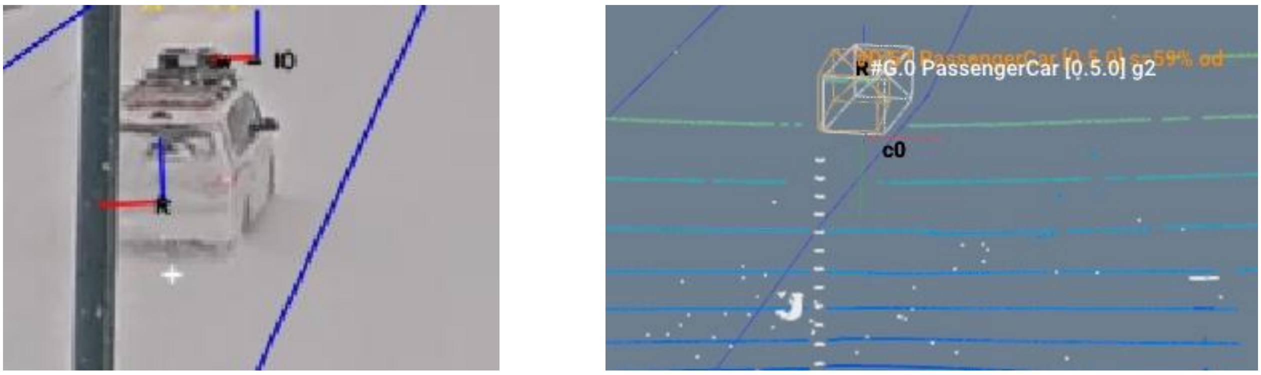
	Camera	LiDAR
Class	+	-
Confidence	+	-
Position	-	+
Size	n.a.	+
Orientation	-	++
Speed	-	+
Orientation Speed	-	++

5.2.2 Sensor fusion adaptation depending on the weather conditions

These relative weights between the sensors' outputs can be adapted depending on the weather. As shown by the results from section 5.1, the performances of the different sensors are changed by different aspects of the weather. There are first the general weather conditions, such a clear, rain, snow, fog, etc. In addition, the road conditions can also impact the sensors' performance. Lastly, even if it is only loosely connected to the weather, the day or night condition also impacts the sensors' performances.

Moreover, the impact of these different weather conditions depends on the location inside the field of view of each sensor, mostly in terms of distance from the sensor and in terms of travelling direction respective to the sensor.

For instance, with a snow-covered road and a light snowfall condition, the importance of the perception results from the camera should be decreased for vehicles going away from the camera at a medium to long range. In Figure 38, we can see a frame where the roadside camera has not detected well the vehicle due to snow splashes at the rear side of the vehicle, while the roadside LiDAR was able to detect the vehicle.



(a) vehicle not detected by the camera-based perception (b) vehicle detected by the LiDAR-based percecption

Figure 38: Roadside sensor perception difference – snow conditions.

Similarly, at night, the importance of the perception results from the camera should be decreased for vehicles going towards the camera at a medium to long range based on the evaluation done in section 5.1.2 .

Inversely, with a wet road, the importance of the perception results from the LiDAR should be decreased at all distances.

The adaptation of the sensor fusion depending on the weather conditions can be realized dynamically if a detailed source of weather information is available. Such a source can be a weather station analysing the weather conditions with different sensors, or the weather analysis tool described in section 3 and developed by CE. The camera-based weather detection system provides detailed information about weather conditions, including the precipitation type and the state of the road surface (such as in Figure 19 that shows the results of this tool for an image extracted from the scene with a wet road and a light snow).

5.2.3 Radar integration in sensor fusion

Both cameras and LiDARs provide data that can be analysed to detect the objects present on the road as well as their characteristics. However, radars provide only sparse data that is hard to process for extracting objects and their characteristics. Nonetheless, radars provide useful information, in particular allowing a direct measurement of an objects' velocity.

The next Figure 39 and Figure 40 show the result of CRF's analytics processing REHEARSE dataset [65] created in ROADVIEW WP3 at CARISSMA and CEREMA PAVIN facilities on different sensors: a LUCID Triton camera (Camera#0), an Innoviz One Falcon LiDAR (LiDAR#1), an Ouster 128 OS1 LiDAR (LiDAR#2) and a ZF ProWave Radar (RADAR#3).

Figure 39: Multi-sensor object detection (camera, LiDAR, radar) – CARISSMA outdoor recording.

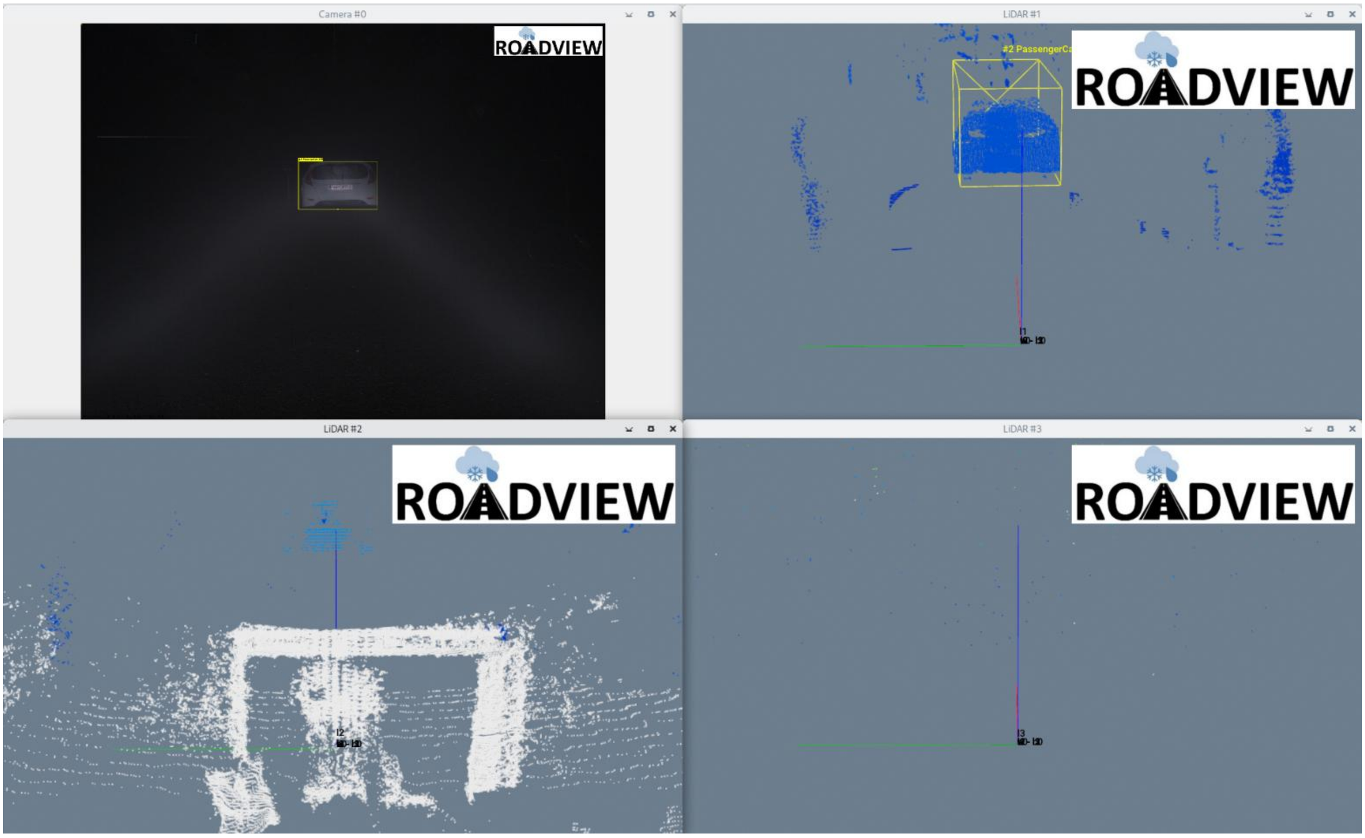


Figure 40: Multi-sensor object detection (camera, LiDAR, radar) - CEREMA PAVIN platform recording.

As shown by these figures, CRF's analytics are able to extract information from the camera and from both LiDARs to detect a passenger car. There is one exception: the car is not detected on the Ouster's LiDAR point cloud (LiDAR#2). This is due to the viewing angle of the car, which is straight on its rear, and which the analytics are not able to process correctly. These figures also show that the radar's point clouds are much sparser than the LiDAR ones and do not provide enough details to extract objects from them.

Having made the same observations, Palladin et al. propose in SAMFusion [66] to combine the radar and LiDAR point clouds with camera images at the feature level. The output of each sensor is first processed by a backbone to extract features from it. Then, these features are combined, using attention and depth-based blending. The combined features are used to extract objects' information. This approach could solve the problem and should be explored in a future work.

5.3 Global fusion with vehicle detections

In their survey, Caillot et al. [67] distinguish three main stages for sharing and combining data in a cooperative Vehicle-to-Everything (V2X) perception pipeline. At the early fusion stage, the raw sensor data is exchanged between cooperating entities. At the mid fusion stage, pre-processed data, *i.e.*, features extracted from the raw data, are exchanged. At the late fusion stage, processed data, *i.e.*, a map representation of the environment, are exchanged.

In a V2X cooperative perception relying on current C-ITS standards, the late fusion strategy is the most relevant. Indeed, Cooperative Awareness Message (CAM) [68] and CPM (Collective Perception Message) [69] messages enable to transmit map-based representations of the road environment respectively as the ego position and as a list of other road users' positions. In particular, combining information from CPM messages can enhance the perception of a road user to areas outside of its sensors' field of view or refine its perception for areas where its sensors have low accuracy.

Recordings realized in Finland show the complementarity of the perception results obtained by different road users. Calibration work realized by FGI enabled to project the point clouds captured by their ARVO vehicle's LiDAR on the images captured by CRF's camera.

5.3.1 Detections from ARVO vehicle

Measurements collected using FGI's autonomous car platform were used to improve the detection of VRU's as well as vehicles in harsh weather conditions. This was achieved with the toolchain developed in ROADVIEW for in-vehicle perception system utilizing precise positioning of the vehicle, time synchronization of sensors, and semantic segmentation of the lidar point cloud. For this purpose, we collected a separate dataset in Finnish Lapland to synchronously save data from the CRF roadside camera and lidar with the sensors of the FGI's autonomous car platform called ARVO.

For the data set, we recorded lidar data from Velodyne Alpha Prime (VLS128) sensor mounted on top of the ARVO vehicle as well as the precise positioning data using Novatel PwrPak7 GNSS+IMU sensor fusion solution. The vehicle pose was improved with a postprocessing solution using the base-station information from the NLS FINPOS positioning service in the Novatel Inertial Explorer software. The point cloud measured with the Velodyne lidar was time synchronized to global time using the Novatel PwrPak7 receiver.

The lidar data was semantically segmented using the Semantic Segmentation AI model developed by HH for 128 channel spinning lidar and documented in ROADVIEW Deliverable D5.2 (Adaptive Sensor Fusion). The time synchronized and semantically segmented lidar data were then projected in the roadside camera coordinates. Then we were able to draw the semantically segmented lidar data on top of the roadside camera image as seen in Figure 41, Figure 42, Figure 43 or Figure 46. The VRU and vehicle labels could then be compared in the road side camera image to verify the quality of VRU and vehicle detection in the roadside sensor system.

5.3.2 Global fusion results

The following figures show the complementarity of the perception results from FGI's vehicle LiDAR and CRF's roadside camera. In these combined representations, FGI's point clouds are coloured according to the semantic labels predicted by HH's semantic segmentation software developed in task 5.1, and the objects detected by CRF's camera analytics are represented with their bounding boxes. In Figure 41, a pedestrian, in pink on the right of the image, is detected by both sensors. FGI's vehicle is next to this pedestrian. FGI ARVO vehicle is identified by a red cross on top of it.

Figure 41: Combined perception of FGI's LiDAR and CRF's camera, with a pedestrian detected by both sensors.

However, in Figure 42, FGI's vehicle is far from the crossroad and the pedestrian is outside of the field of view of its LiDAR. In this case, only the roadside camera is able to detect the pedestrian. In a different scenario where both the pedestrian and the vehicles are going towards the crossroad, informing in advance the vehicle that there is a pedestrian that may cross the road is useful for reducing the risks of accident and ensuring the safety of the pedestrian.

Figure 42: Combined perception of FGI's LiDAR and CRF's camera, with a pedestrian only detected by CRF' camera.

Figure 43 shows the inverse scenario where the pedestrian is hidden from the camera by a tree but is detected by the vehicle. In this case, the vehicle is coming from the right and is still outside the field of view of the camera. This figure shows that perception results from vehicles are a useful complement to the perception results of a roadside system: they enable to extend the perception results to areas outside the field of view of the roadside sensors and to occluded areas.

Figure 43: Combined perception of FGI's LiDAR and CRF's camera, with a pedestrian only detected by FGI's LiDAR.

5.4 Experimentation results

The experiments realized in Finland were useful on several point of views. First, they allowed recording road scenes in different weather conditions jointly from a roadside infrastructure and from a vehicle, as described in section 5.1. These recordings were used to analyse the sensors' performances depending on the weather conditions. Using these results, the relative importance of the perception results of the different sensors can be adapted depending on the weather conditions. Taking as input the information provided by CE's weather detection system (see section 3), CRF's analytics can be adapted in real-time to the current weather conditions.

For instance, the following figures show the adaptation of the roadside camera analytics by defining areas corresponding to lower performances due to the weather and where the confidence in the perception results is lower. These areas can be reported in CPM messages as perception regions with a low confidence value.

Figure 44 shows the adaptation that can be realized when the road is covered with snow and snow is falling. A perception region with a low confidence value, shown in light blue, is included inside the camera's field of view, shown in dark blue. This perception region covers the lane going away from the camera, where the detection quality is decreased as shown by the results of section 5.1. Note that in this example, VTT's vehicle is not detected, while another vehicle further away is detected, albeit imperfectly as a pedestrian is also detected at the same location. This perception region can be used to indicate in a CPM that the perception results of the roadside infrastructure have a low confidence inside this area.

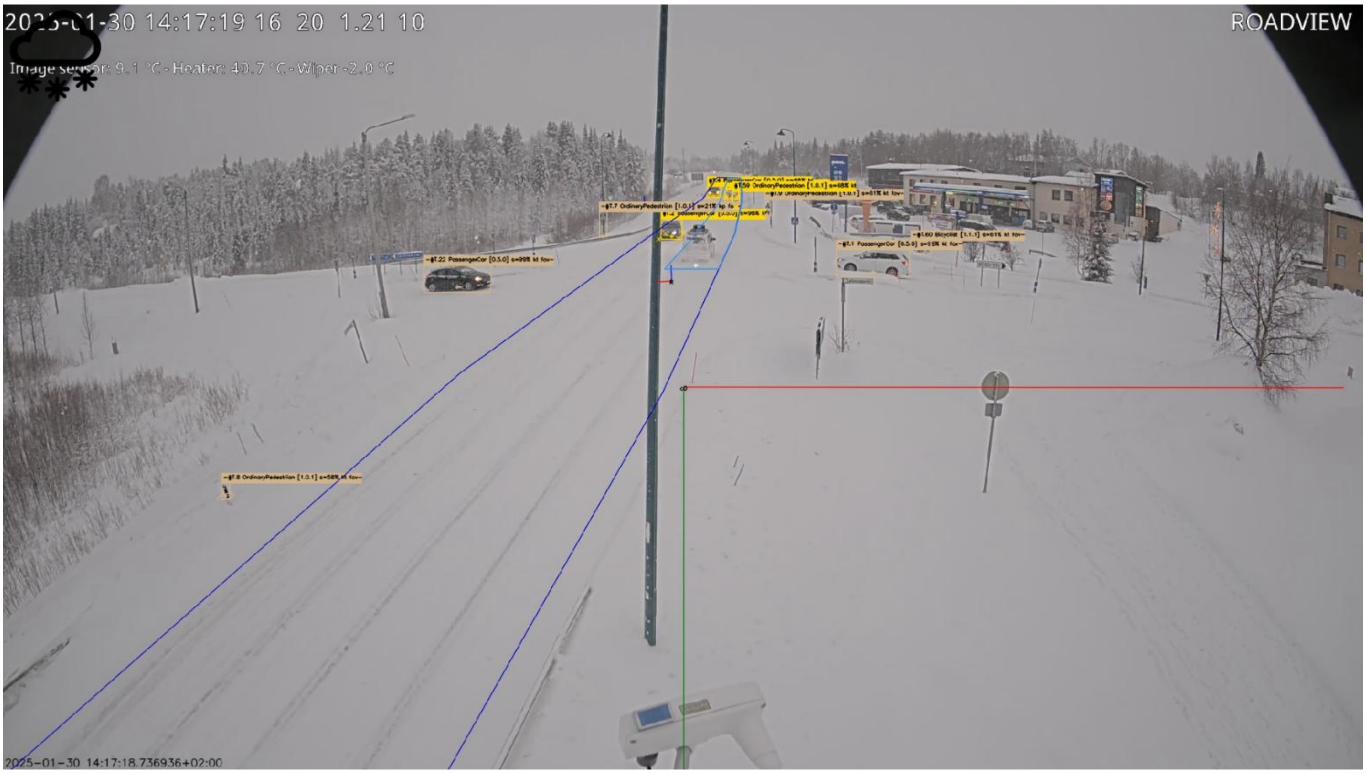


Figure 44: Perception region with low confidence when the road is covered with snow.

Another example is shown in Figure 45, corresponding to night conditions. Here, two perception regions are defined for the lane coming towards the camera. The furthest one has a low confidence value associated to it. Indeed, VTT's vehicle, which is inside this region, is not detected. The analytics are sometimes able to detect the vehicle, in particular when it is in the part of the region closest to the camera. However, globally, the perception results from the analytics inside this region are not good. The closest perception region has a good confidence value associated to it, but still lower than the default confidence value inside the field of view of the camera. Most of the time, vehicles inside this region are correctly detected. However, from time to time, a vehicle may not be detected. This is the case in particular when a vehicle has strong headlights that blind the camera. Using two different perception regions in this case enable to transmit finer information about the perception results confidence inside a CPM message.

Figure 45: Perception regions with lower confidence at night.

In the Finnish installation, unfortunately, the perception range of the roadside LiDAR is not large enough to provide information on areas where the camera's perception may be decreased by the weather conditions. Therefore, it is not possible to show relevant results for the adaptive fusion of the sensors' perception results. This could be improved either by using a LiDAR with more channels (for example with 64 or 128 channels), or by installing the LiDAR closer to the intersection.

For a global fusion, the experimentations showed that the perception results from the roadside infrastructure and from vehicles were complementary, allowing a more complete perception of the traffic when combining them. Figure 46 shows an example where VTT's vehicle (incoming at the bottom) perceives the cyclist as a moving person, while the infrastructure perceives it as both a person and a bicycle. Fusing this information would allow to have a more precise perception of the road users.

Figure 46: Different perception results between roadside's camera and vehicle's LiDAR.

In the next section we will see how the perceived environment obtained by the various sensors can be shared using V2X communication and how the quality of the perception can be included in the V2X messages in case of harsh weather conditions.

6 V2X collective perception service

In this section, we describe the second relevant C-ITS service identified in D5.3. The Collective Perception Message (CPM) enables the interoperable sharing of basic information about the perceived environment by an ITS station (roadside unit or vehicle onboard unit). ETSI ITS Technical Committee has published a first version of this service TS 103 324 [69]. This service is used to share the roadside sensor data as described in section 5 and to receive data from vehicle sensor that can be used to perform global fusion (described in section 5.3).

First in section 6.1, we describe the CRF study done on the quality of the information reported in CPMs. The quality of the reported information contained in V2XCPM) was computed based on the weather conditions, taking into account the ODDs of the sensors and their predictions. The service was evaluated using real-world scenes under different weather conditions recorded in Finnish Lapland.

Second in section 6.2, HH has worked on congestion-aware collective perception to optimize the performance of the V2X collective perception for vehicle-to-vehicle communication. This work was evaluated using simulation.

6.1 Quality information

Basic information shared by the collective perception service includes the sensory capabilities (section 6.1.2), road-related perception regions ((section 6.1.3) and perceived objects (section 6.1.4) as already reported in D5.3.

In this section we will focus on how to take into account the sensor ODD and to report the quality information in V2X CPM. The study was done based on the experiments done in Finland (see section 6.1.1)

6.1.1 Positional accuracy from experiments

Several datasets were recorded in Finland under different weather conditions. They include vehicle sensors data as well as roadside sensors data (LiDAR and camera). In particular, VTT's recordings in snowy conditions and FGI's recordings in daylight and night were of interest for a positional accuracy analysis.

Indeed, it was possible to compare roadside sensor system obtained positions with vehicle-obtained positions (either the GNSS measurements from VTT vehicle, or the GNSS+IMU from FGI vehicle described in section 5.3.1). The analysis focused on two distinct vehicle's paths, namely when the vehicle travels from bottom to top (from the camera point of view) and when it travels from top to bottom of the image. Figure 47 shows an example of UTM position deviation between the vehicle-obtained positions that are used as a ground truth and the positions obtained by roadside camera-based perception system.

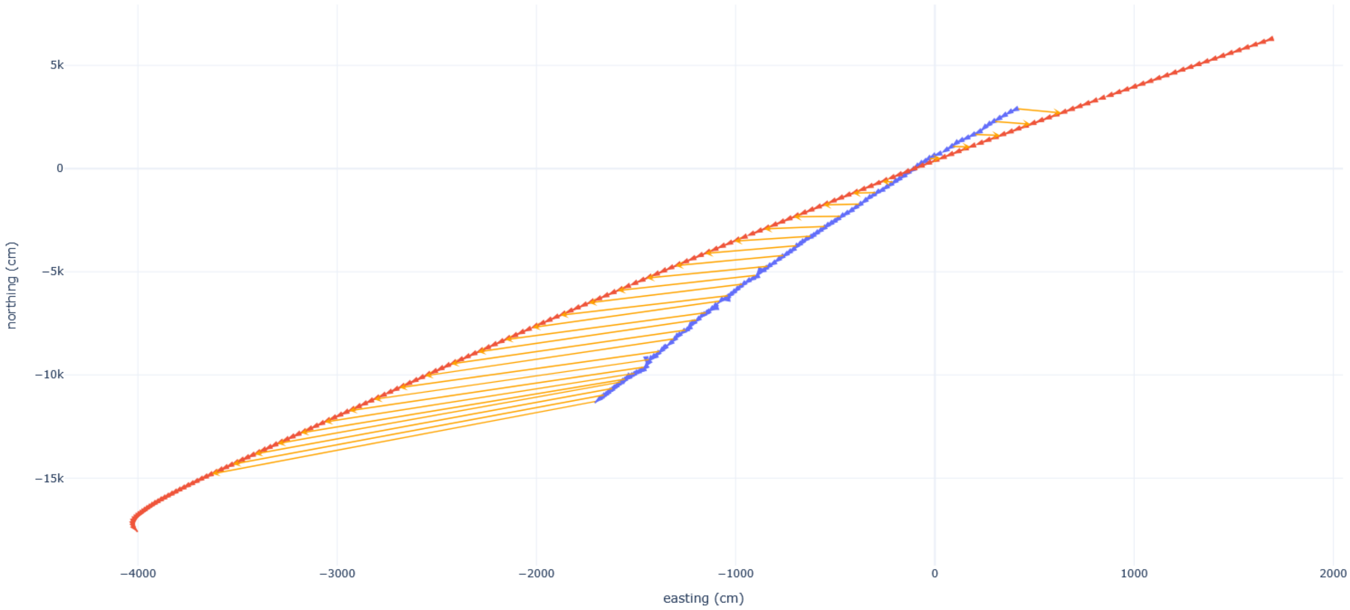


Figure 47: UTM positional deviations (orange arrows) between roadside (in blue) and vehicle-obtained positions (in red).

In each dataset, the vehicle recorded several trajectories, allowing for an accuracy analysis in several directions. In details, the Euclidean distance was calculated between each roadside-obtained position and vehicle-obtained position sharing a same timestamp (or interpolated timestamp). Then, a short-range area was defined as being a circle centred at the middle of the monitored intersection, with a radius of 50m. Finally, the mean of Euclidean distances was calculated for positions inside the short-range area (<50m) and outside of the short-range area (>50m). The results are summed up in Table 9.

Table 9: Mean error (Euclidean distances) analysis using datasets recorded in Finland

Region of interest	Path of the vehicle	Weather and light Conditions			
		Cloudy, dry road, day (FGI)	*Snow falling, wet road, day (FGI)*	*Cloudy, dry road, Night (FGI)*	*Snow falling, snow on the road, day (VTT)*
Short-range area (<50m)	Bottom to top	2.47 m	2.67 m	2.61 m	6.82 m
	Top to bottom	3.32 m	2.75 m	2.84 m	1.62 m
Far (>50m)	Bottom to top	17.01 m	19.65 m	15.60 m	36.62 m
	Top to bottom	16.09 m	16.17 m	12.12 m	12.03 m

In the short-range area (<50m), the mean error stays below 3.5m. One exception being for “bottom to top” in snowy conditions (VTT dataset), where the accuracy is reduced, as the camera-based object detections has trouble to correctly detect the vehicle and therefore estimate its positions, because of the snow raised around VTT’s vehicle, as described in section 5.4. In the other datasets (FGI datasets), accuracy in the short-range area is slightly better in the “bottom to top” direction compared to the “top to bottom” direction. This is due to the tracking algorithm which incrementally enhances consecutive positions, such process working well when initial speed and positions of the object are precise. This is the case in the “bottom to top” trajectory, where the vehicle begins its path close to the roadside sensors (therefore with a better initial position estimation).

In the far field (>50m), errors increase a lot with the object distance to the roadside camera. Indeed, while far objects are correctly classified and detected, the camera-based object detection module struggles to estimate positions. This is partly due to the large camera detection range (>200m), while in these far fields the density of pixels in the camera image is not significant enough. In addition, the difference in accuracy in far fields and short-range area can be explained by camera calibration imprecisions, which grow with distance to the camera. In particular, as for short-range area, the accuracy is greatly decreased for the “bottom to top” path in snowy conditions.

These results highlight the differences of accuracies, showing the need to describe the varying confidences in regions or objects reported in CPMs.

In the next sections, we will describe how to report the quality of information contained in a CPM, first in the sensor information, second in the perception region and last in the perceived object containers.

6.1.2 Sensor Information

The sensor information container in CPM contains the static perception capabilities of the sensors that reflects the nominal setup of the sensor system:

- Sensor type (camera, lidar, radar or fusion),
- Field of view of the sensor (sensor perception region shape)
- Perception confidence level. For example, HOTA (Higher Order Tracking Accuracy) metric [70] can be used to compute the confidence level in the sensor perception region.

In case of varying perception conditions, a perception region may be created as explained in the next section.

6.1.3 Perception Regions

The perception region container is used to transmit information about dynamic changes in perception due to environmental changes. Typically, the weather conditions can affect the sensor perception.

In the case of ROADVIEW Project, the perception region can then be used to reflect the changes in the sensor perception due to harsh weather conditions. For example, in case of low visibility due to fog or heavy rain, camera sensor type can have a reduced perception region shape, and the perception confidence level is decreased compared to clear weather.

The confidence value of 101, meaning that the value is not available, can be used to indicate that a perception region has no changed confidence with respect to the sensor region confidence (or that same value than the one in the sensor region confidence is set). Preassigned confidences can be defined for typical weather and light conditions as recommended by C-ITS Message Profiles such as the one defined by C-ROADS Platform [71].

During the experimentation in Finland, we have characterised the object perception and tracking system based on the weather and road conditions as already described in section 5.4.

Figure 48 shows a representation on a map of the received V2X CPM containing a region with a low confidence due to the snow on the road for the lane where the vehicle moves away from the roadside sensors. Figure 49 shows two regions to indicate different perception confidence levels during the night for vehicles driving towards the camera.

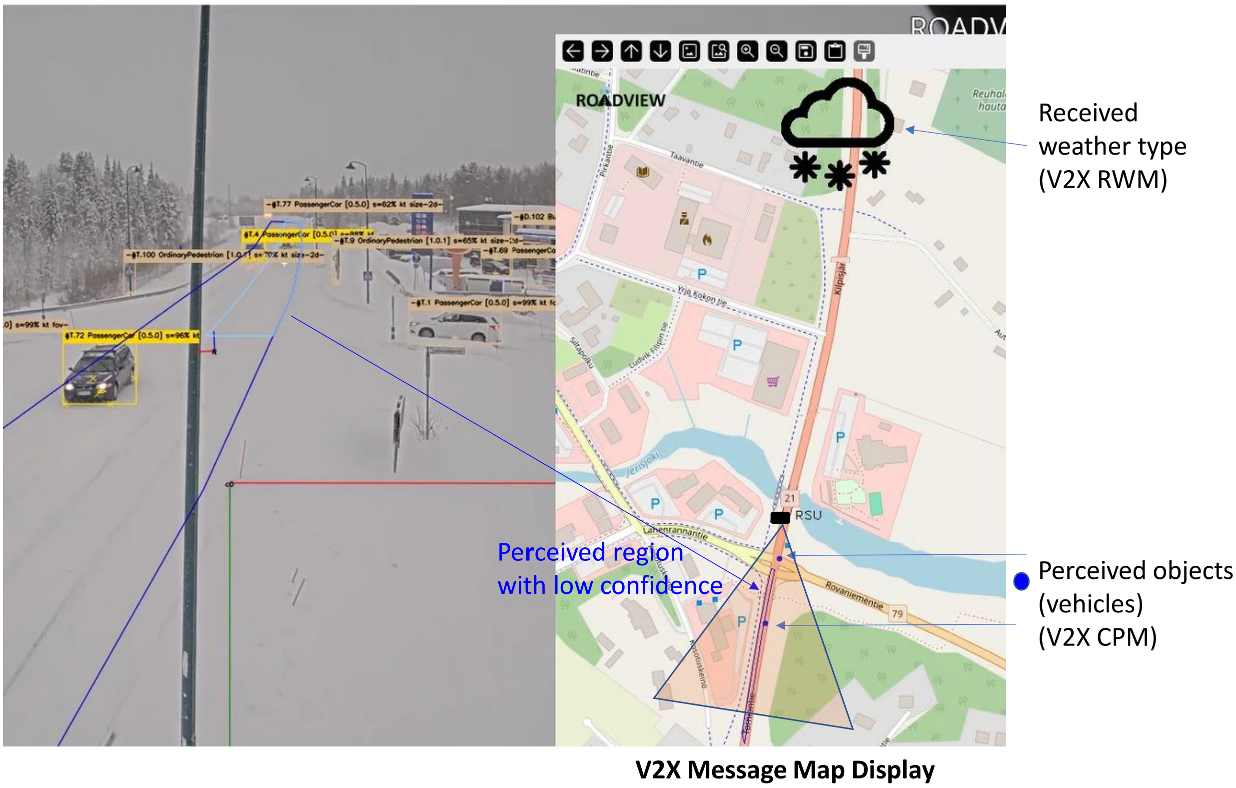


Figure 48: Perception Region with Low Confidence (snow).

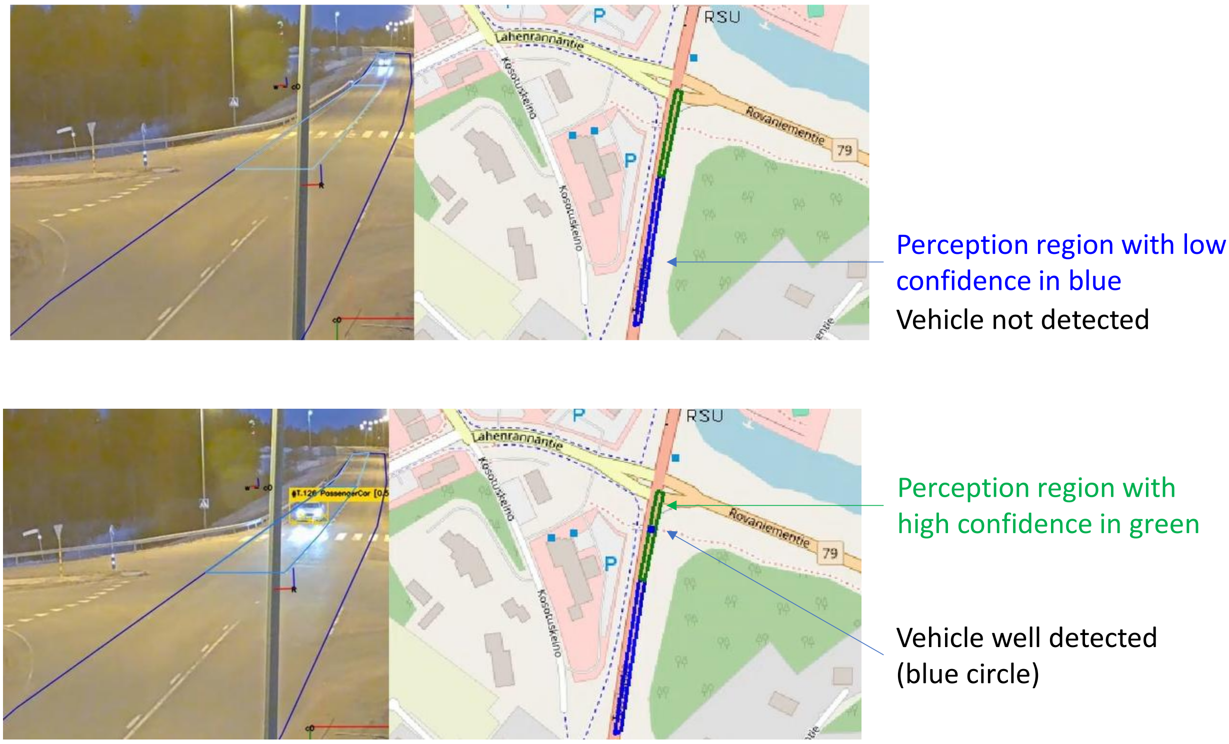


Figure 49: Perception Regions with Low and High Confidence (night).

Another use of perception regions is to report about actual free spaces (areas with no detected objects) depending on the use case or occlusion areas (object shadowing). During the preparation of the demonstration targeting a highway use case, a roadside system composed of a camera and LiDAR was installed at a toll gate. The LiDAR was used to detect the toll gate lane occupancy (free or occupied) and the camera to provide additional object information (such as the presence of VRU). Figure 50 shows the camera object-detection on the left and on the right the representation on a bird's-eye view of the received V2X message (clear weather conditions, toll gate L1 and L2 occupied, toll gate L3 and L4 free, and the position of the detected vehicles).

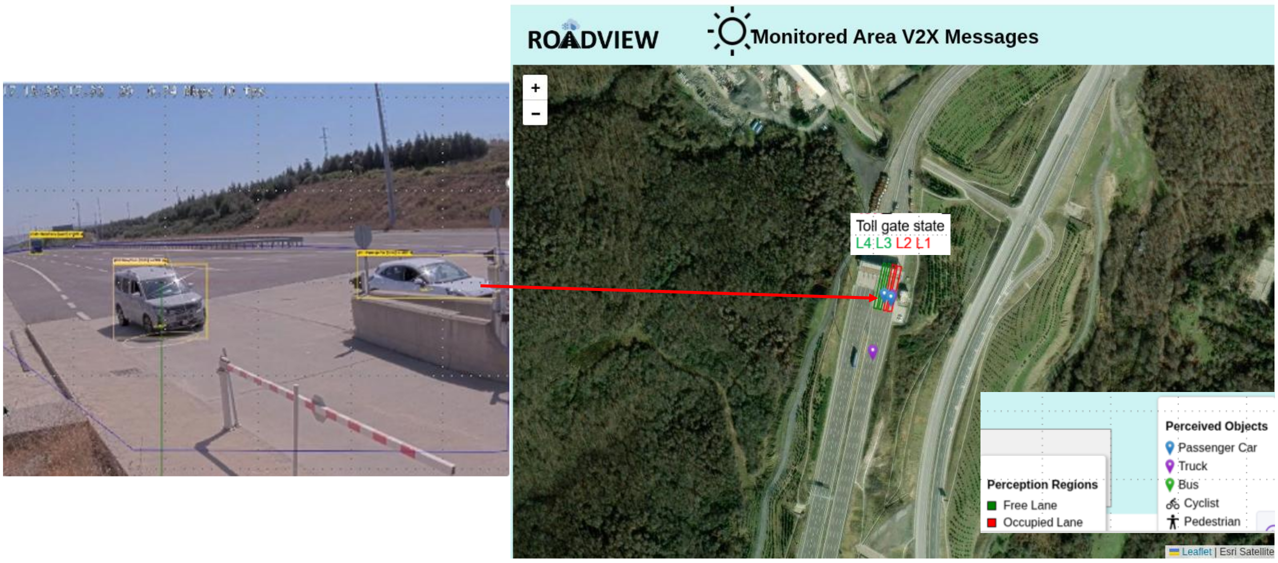


Figure 50: Use of perception regions to report lane occupancy.

6.1.4 Perceived Objects

The CPM also contains the list of perceived objects with their quality information. For each object, quality information can be provided using data elements listed here:

- **Object Perception Quality** representing the confidence that the object actually exists. The value is computed based on the object detection analytics and object tracking algorithm of the sensor fusion system.
- **Object Class Confidence** representing the confidence on the object class**.** Main classes allow to distinguish between vehicles and VRUs. Then subclass such as passenger car, truck, pedestrian, bicycle, etc can be added.
- **Confidence values** on the kinematic state and attitude parameters (e.g., position coordinates (px, py, pz), velocity values (vx, vy, vz))
    - o The confidence value is given at a 95% confidence level, which means there is a 95% probability that the true value of the kinematic parameters lies within an interval centred around the estimated value, with a half-width equal to the confidence value. For example, an estimated coordinate px equal to 5cm with a 95% confidence value equal to 2cm means the true px value has 95% being in the interval [3cm, 7cm].
    - o When considering a parameter following a normal distribution, the confidence values with a 95% confidence levels are approximately equal to 2 times the standard deviation of the parameter.
- **Correlation matrix** representing the relationship and the dependency between pairs of kinematic state and attitude parameters (e.g., a correlation matrix can include corr(px, py), corr(px, vx), etc.).
    - o Such matrix is standardized and therefore each correlation value is contained between -1 and 1.

From the confidence values (therefore from the standard deviations) and the correlation matrix, the covariance matrix can be constructed, which scales the correlation matrix to the actual deviation values of the kinematic state and attitude parameters. This matrix contains all the information necessary to represent the error associated to kinematic and attitude state parameters.

In the context of ROADVIEW, the covariances are generated depending on the weather context. Indeed, harsh weather does not only prevent detections as explained in section 5.1.2, but also impacts the quality of the detection, and therefore the kinematic parameters estimation. For example, in VTT's dataset in Finland (snowy conditions), the video analytics had difficulties detecting the rear of VTT's white car, especially as it raised snow when passing. As a result, the object detection analytics has sometimes detected the roof of the vehicle as a vehicle, with a high detection score (see the orange bounding box with a detection score of 80% in Figure 51). The wrongly estimated bounding box then led the roadside perception system to incorrectly estimate the vehicle true geometric centre, in particular in px's direction.

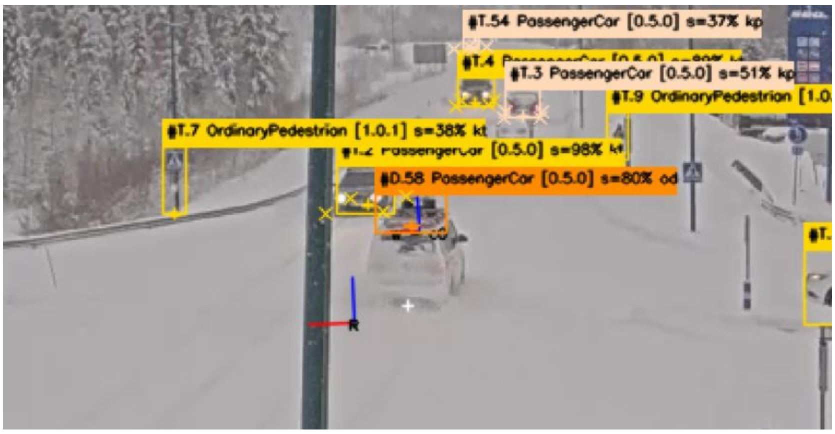


Figure 51: Example of video analytics with inaccurate object detection due to snow splash.

Therefore, these weather particularities are taken into account in the covariance representation. To visualize these changes for two-dimensional position coordinates, a method is to represent the confidence ellipses with a 95% confidence level. In following Figure 52, the covariances of positions coordinates (px shown on the easting horizontal axis, py shown on the northing vertical axis) are represented using confidence ellipses for the VTT's vehicle raising snow example. The left figure represents the vehicle's positions and confidence associated while still near the intersection. The right figure again represents vehicle's positions and confidence but while leaving the intersection.

In the latter case, the distance imprecisions are summed to the snow induced imprecisions, causing a larger ellipse, therefore a larger covariance, in particular in px's direction.

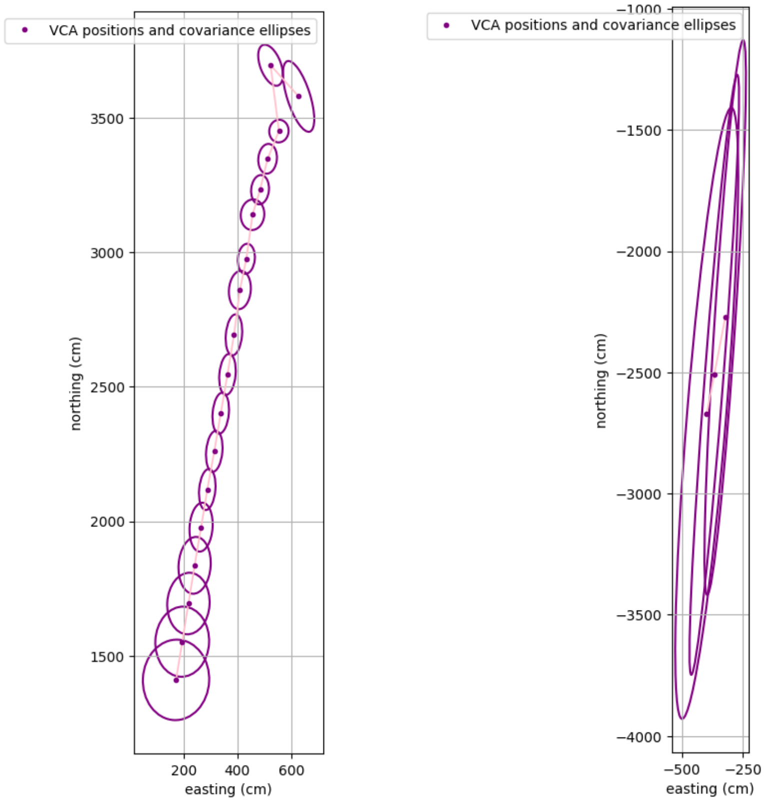


Figure 52: VTT's vehicle positions and confidence ellipses near the intersection (left) and after the intersection (right).

As described in the CPM standard, the resulting covariance matrices can then be decomposed into correlation matrices and confidence values for a meaningful description of uncertainties in CPMs.

In particular, those values can be used as input of the fusion algorithm used to update the objects in the local dynamic map. In CPM, the age of the information for each reported region and each reported object can be obtained by computing the message reference time and the measurement delta time information associated to each region and object. In the next section we will report about the end-to-end latency measurements done to evaluate the age of information of data from the roadside system.

6.1.5 Measurements of end-end latency

Using the roadside system installed in Finland we have measured the end-to-end latency of the infrastructure-based perception system from the object perception time up to the reception in the onboard V2X module. We have analysed the received CPM at the onboard side. The system architecture used during the test is shown in Figure 53, where the roadside system was connected to MQTT broker using wired Internet access, and the receiving onboard V2X was connected to the MQTT broker using 4G network connectivity:

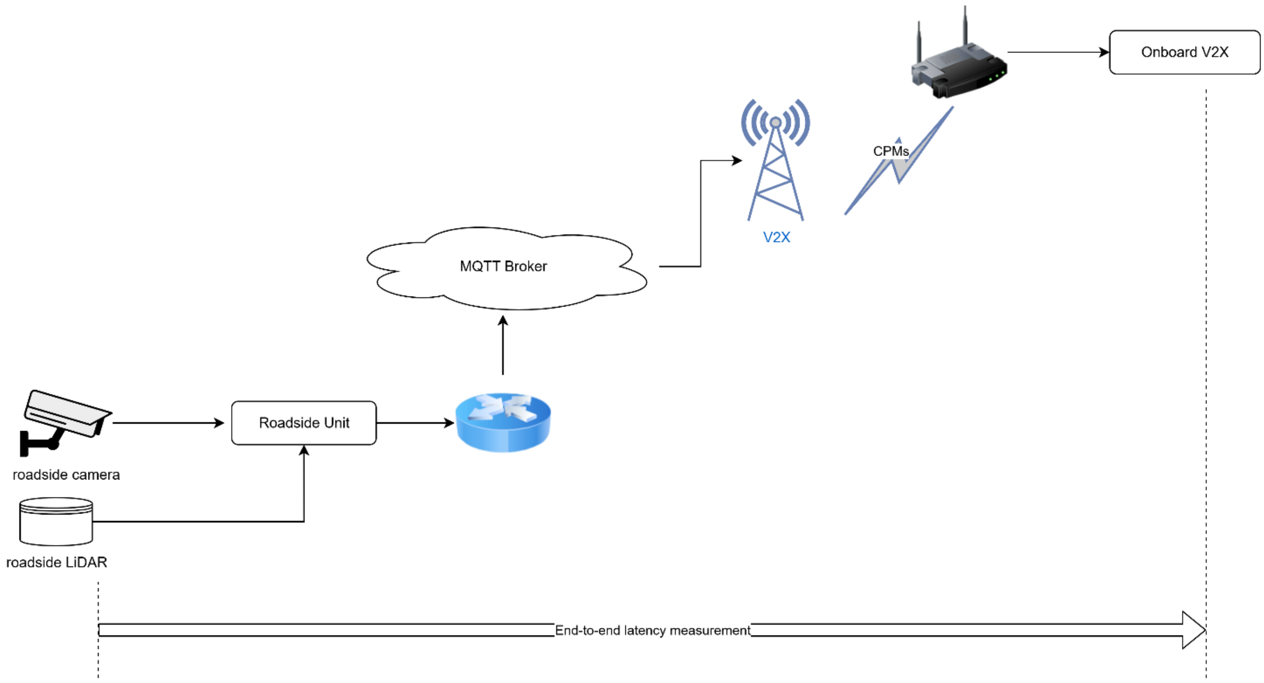


Figure 53: Set-up to measure the end-to-end latency.

Based on the current CPM specifications [69], the CPM generation period can vary between 100 milliseconds and 1 second depending on the situation. The rules to include an object in next CPM generation event are:

- When it is the first time that the object has been detected,
- When changes of the absolute position, the absolute speed or the heading of the object is more than a set of predefined thresholds since its last inclusion in a CPM.
- One second has passed since the object was included in a CPM. For VRUs, the last time inclusion threshold is set to 500 milliseconds.

For our latency measurement tests, we have transmitted CPM with the full object list in the region of interest once every 100 milliseconds. During the tests the quantity of perceived objects included in CPM was varying between 0 and 23 with the selected region of interest to monitor the intersection (we have excluded the parking area as the objects are less relevant for the driving vehicles). In Figure 54, we plot the measurement for end-to-end latency and the number of reported objects during the test.

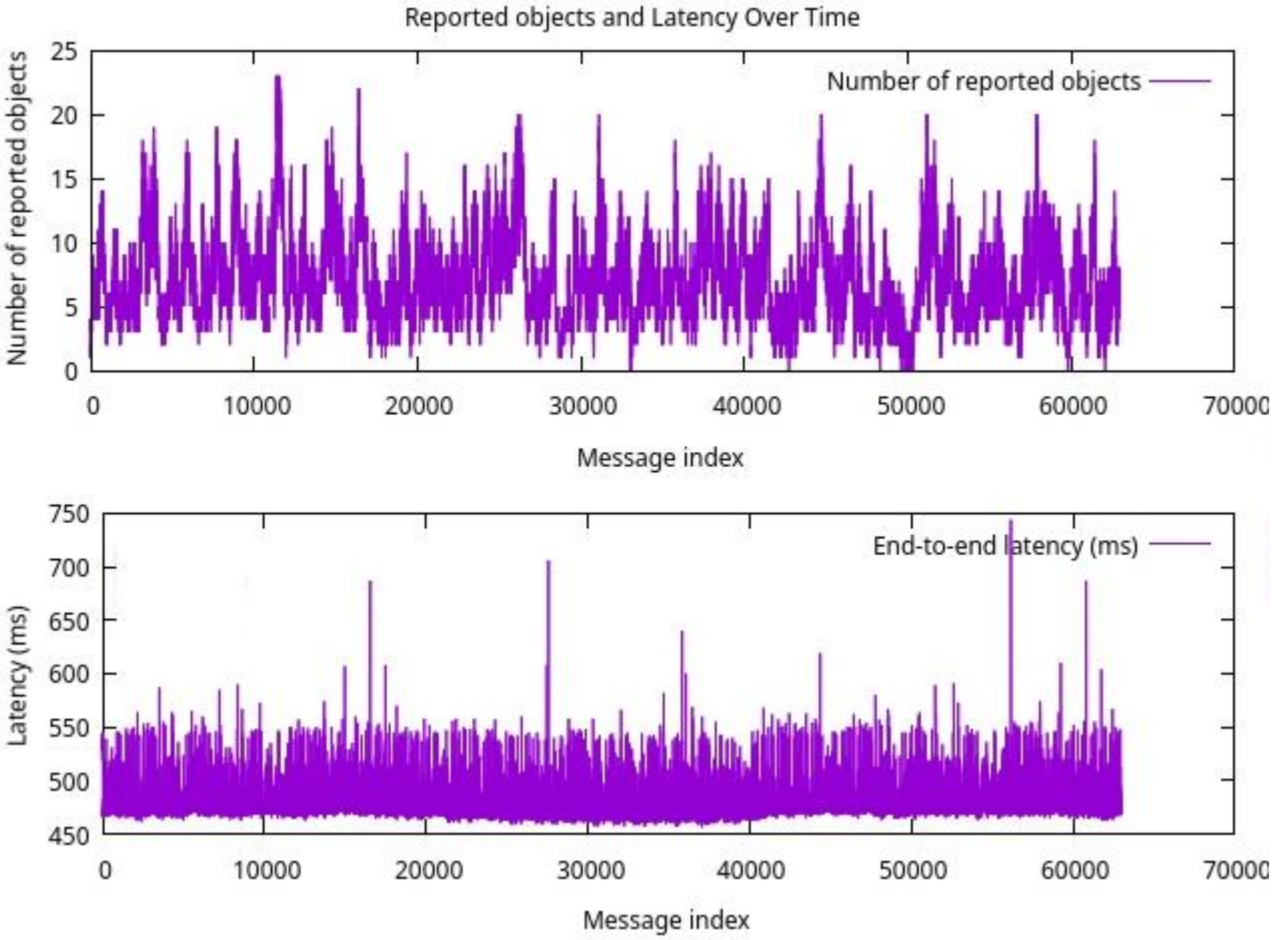


Figure 54: End-to-end latency measurements.

Table 10 shows the end-to-end latency value summary, the maximum value was below 750 ms that was defined as the target value for the ROADVIEW use cases.

Table 10: End-to-end latency of collective perception service.

Minimum value	Median value	Maximum value
456 ms	475 ms	744 ms

6.2 Congestion-aware study

In the previous section, we have explained how to specify the quality of the reported information in a CPM. In case of multiple ITS stations at the same location, the available radio channel resources shall be used carefully to avoid network congestion while keeping sufficient level of information for the traffic safety applications. V2X CPM is not yet deployed on the market, at the early stage of deployment not many ITS stations will generate CPMs. Congestion-aware study is an important topic for a successful deployment of CPM in the field. The object inclusion rules and the value of information of CPM are being studied in [72] or in [73].

HH studied a simulation-based performance optimization method of V2X collective perception by adapting object filtering based on the information redundancy and the available radio channel resources [74]. In this work, HH combined the redundancy mitigation and congestion control-aware filtering and evaluated the performance of the resulting object filtering techniques by realizing realistic, large-scale simulations. In the follow-up work [75], HH investigated the impact of persistence on the spatial and temporal order of radio resource reuse in vehicular highway scenarios, leveraging 5G NR-V2X sidelink communications. In spite of more ordered resource reuse, which is beneficial for performance metrics such as packet delivery ratio and throughput, it turns out that semi-persistent scheduling may provide worse performance with respect to the dynamic scheduling in terms of the age of information.

Last aspect of HH work [76] focused on utilizing paradigms such as Federated Learning (FL) to address the dynamic and decentralized nature of V2V (Vehicle-to-Vehicle) communications. The novel algorithm studied by HH is based on Practical Byzantine Fault Tolerance and provides dynamic vehicle node management, malicious node detection and connectivity interruption recovery features. The simulation shows the ability of this method to reduce the number of V2X redundant messages by optimizing the message broadcasting strategy.

7 Conclusions

This deliverable contains the results of Task 5.2 related to the collaborative perception solutions integrating infrastructure-based perception system. It contributes to Milestone (MS) 12 (due in Month 36) "Perception algorithms and SW ready".

Main outcomes are:

- **A camera-based weather detection system** providing high-accuracy and real-time performance – the software and the datasets are available as open source on the ROADVIEW GitHub [2],
- **A sensor fusion system** taking into account the sensor ODD in case of harsh weather conditions,
    - with a cost-effective solution to obtain the weather conditions in real-time by using the camera images both for weather detection and for object perception.
    - global fusion of object detected by the roadside sensors and by vehicle object detection, including VRU detection in artic weather conditions.
- **A V2X communication system** enabling to share the weather conditions using Road Weather Message (RWM) and the sensor data (sensor capabilities, perception region and perceived objects) using Collective Perception Message (CPM).
    - The study on how to report the information quality using perception regions and covariance matrix; tested with real-world scenes in Finnish Lapland under various weather conditions,
    - Collective perception service congestion-aware study done in simulation.

Thanks to the experiments done in Finnish Lapland, the partners could evaluate each solution with real-world data scenes under various weather conditions and with various type of traffic conditions including the presence of VRUs.

Next steps will be to demonstrate the collaborative perception solutions in the ROADVIEW use cases targeting highway, rural and urban environment as described in ROADVIEW D2.2 [77].

8 References

[1] I. Morvan and P. Duthon, 'ROADVIEW Deliverable 5.3 - SW on Collaborative Perception Solutions - First report'. 2024. [Online]. Available: https://roadview-project.eu/wp-content/uploads/sites/59/2024/05/ROADVIEW_Deliverable-5.3_v02.pdf

[2] 'roadview-project/Enhanced-Style-Based-Neural-Architectures-for-Real-Time-Weather-Classification-main'. Accessed: Aug. 27, 2025. [Online]. Available: https://github.com/roadview-project/Enhanced-Style-Based-Neural-Architectures-for-Real-Time-Weather-Classification-main

[3] 'Fintraffic Aurora Road'. [Online]. Available: https://www.fintraffic.fi/fi/tie/alytie-aurora

[4] T. DiLorenzo and X. Yu, 'Use of ice detection sensors for improving winter road safety', *Accid. Anal. Prev.*, vol. 191, p. 107197, 2023, doi: https://doi.org/10.1016/j.aap.2023.107197.

[5] S. Zang, M. Ding, D. Smith, P. Tyler, T. Rakotoarivelo, and M. A. Kaafar, 'The Impact of Adverse Weather Conditions on Autonomous Vehicles: How Rain, Snow, Fog, and Hail Affect the Performance of a Self-Driving Car', *IEEE Veh. Technol. Mag.*, vol. 14, no. 2, pp. 103–111, 2019, doi: 10.1109/MVT.2019.2892497.

[6] A. Engstle, 'Adverse Weather Conditions in the Validation of ADAS/AD Sensors', in *AmE 2023 – Automotive meets Electronics; 14. GMM Symposium*, 2023, pp. 103–107.

[7] F. Sezgin, D. Vriesman, D. Steinhauser, R. Lugner, and T. Brandmeier, 'Safe Autonomous Driving in Adverse Weather: Sensor Evaluation and Performance Monitoring'. 2023. [Online]. Available: https://arxiv.org/abs/2305.01336

[8] Y. Zhang, A. Carballo, H. Yang, and K. Takeda, 'Perception and sensing for autonomous vehicles under adverse weather conditions: A survey', *ISPRS J. Photogramm. Remote Sens.*, vol. 196, pp. 146–177, 2023, doi: https://doi.org/10.1016/j.isprsjprs.2022.12.021.

[9] M. Z. Shahzad, M. A. Hanif, and M. Shafique, 'Robust ADAS: Enhancing Robustness of Machine Learning-based Advanced Driver Assistance Systems for Adverse Weather'. 2024. [Online]. Available: https://arxiv.org/abs/2407.02581

[10] R. Al-Hindawi, T. I. Alhadidi T, and M. Adas, 'Evaluation and Optimization of Adaptive Cruise Control in Autonomous Vehicles using the CARLA Simulator: A Study on Performance under Wet and Dry Weather Conditions', in *2024 IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET)*, 2024, pp. 1–6. doi: 10.1109/IC_ASET61847.2024.10596222.

[11] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, 'Image-to-Image Translation with Conditional Adversarial Networks'. 2018. [Online]. Available: https://arxiv.org/abs/1611.07004

[12] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, 'Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks'. 2020. [Online]. Available: https://arxiv.org/abs/1703.10593

[13] C. Panhuber, B. Liu, O. Scheickl, R. Wies, and C. Isert, 'Recognition of Road Surface Condition Through an On-Vehicle Camera Using Multiple Classifiers', in *Proceedings of SAE-China Congress 2015: Selected Papers*, Singapore: Springer Singapore, 2016, pp. 267–279.

[14] M. Introvigne, A. Ramazzina, S. Walz, D. Scheuble, and M. Bijelic, 'Real-Time Environment Condition Classification for Autonomous Vehicles'. 2024. [Online]. Available: https://arxiv.org/abs/2405.19305

[15] J. Xia, D. Xuan, L. Tan, and L. Xing, 'ResNet15: Weather Recognition on Traffic Road with Deep Convolutional Neural Network', *Adv. Meteorol.*, vol. 2020, pp. 1–11, May 2020, doi: 10.1155/2020/6972826.

[16] M. Naufal and S. Kusuma, 'Weather image classification using convolutional neural network with transfer learning', in *AIP Conference Proceedings*, Apr. 2022, p. 050004. doi: 10.1063/5.0080195.

[17] O. Papadimitriou, A. Kanavos, P. Mylonas, and M. Maragoudakis, 'Advancing Weather Image Classification Using Deep Convolutional Neural Networks', Sept. 2023, pp. 1–6. doi: 10.1109/SMAP59435.2023.10255190.

[18] M. Patel S. Das and N. Krishnaraj, 'Weather image classification using convolution neural network', *Ann. Romanian Soc. Cell Biol.*, pp. 4156–4166, 2021.

[19] S. Mittal and O. P. Sangwan, 'Classifying Weather Images using Deep Neural Networks for Large Scale Datasets', *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 1, 2023, doi: 10.14569/IJACSA.2023.0140136.

[20] V. Kukreja, R. Sharma, and R. Yadav, 'Multi-Weather Classification using Deep Learning: A CNN-SVM Amalgamated Approach', July 2023, pp. 1–5. doi: 10.1109/WCONF58270.2023.10235097.

[21] M. R. Ibrahim, J. Haworth, and T. Cheng, 'WeatherNet: Recognising Weather and Visual Conditions from Street-Level Images Using Deep Residual Learning', *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 12, 2019, doi: 10.3390/ijgi8120549.

[22] J. Li and X. Luo, 'A Study of Weather-Image Classification Combining VIT and a Dual Enhanced-Attention Module', *Electronics*, vol. 12, no. 5, 2023, doi: 10.3390/electronics12051213.

[23] R. U. Rani, J. Kakarla, and B. Sundar, 'Weather image classification using EfficientNet and Dual Attention Block', in *2023 2nd International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*, 2023, pp. 1–4. doi: 10.1109/ICSTSN57873.2023.10151564.

[24] W. Pikun and W. Ling, 'Image-Based Self-attentive Multi-label Weather Classification Network', in *Proceedings of International Conference on Image, Vision and Intelligent Systems 2022 (ICIVIS 2022)*, P. You, H. Li, and Z. Chen, Eds, Singapore: Springer Nature Singapore, 2023, pp. 497–504.

[25] K. Dahmane, 'Analyse d'images par méthode de Deep Learning appliquée au contexte routier en conditions météorologiques dégradées', Theses, Université Clermont Auvergne [2017-2020], 2020. [Online]. Available: https://theses.hal.science/tel-03022934

[26] J. C. V. Guerra, Z. Khanam, S. Ehsan, R. Stolkin, and K. McDonald-Maier, 'Weather Classification: A new multi-class dataset, data augmentation approach and comprehensive evaluations of Convolutional Neural Networks'. 2018. [Online]. Available: https://arxiv.org/abs/1808.00588

[27] A. Basavaraju, J. Du, F. Zhou, and J. Ji, 'A Machine Learning Approach to Road Surface Anomaly Assessment Using Smartphone Sensors', *IEEE Sens. J.*, vol. 20, no. 5, pp. 2635–2647, 2020, doi: 10.1109/JSEN.2019.2952857.

[28] A. Piroli, V. Dallabetta, J. Kopp, M. Walessa, D. Meissner, and K. Dietmayer, 'SemanticSpray++: A Multimodal Dataset for Autonomous Driving in Wet Surface Conditions'. 2024. [Online]. Available: https://arxiv.org/abs/2406.09945

[29] S. Sattar, S. Li, and M. Chapman, 'Road Surface Monitoring Using Smartphone Sensors: A Review', *Sensors*, vol. 18, no. 11, 2018, doi: 10.3390/s18113845.

[30] S. Jagatheesaperumal, S. Bibri, S. Ganesan, and P. Jeyaraman, 'Artificial Intelligence for Road Quality Assessment in Smart Cities: A Machine Learning Approach to Acoustic Data Analysis', vol. 3, Sept. 2023, doi: 10.1007/s43762-023-00104-y.

[31] D. Gragnaniello, A. Greco, C. Sansone, and B. Vento, 'Real-time joint recognition of weather and ground surface conditions by a multi-task deep network', *Eng. Appl. Artif. Intell.*, vol. 139, p. 109543, 2025, doi: https://doi.org/10.1016/j.engappai.2024.109543.

[32] Y. Chen *et al.*, 'Inexpensive Multimodal Sensor Fusion System for Autonomous Data Acquisition of Road Surface Conditions', *IEEE Sens. J.*, vol. PP, pp. 1–1, Aug. 2016, doi: 10.1109/JSEN.2016.2602871.

[33] Y. A. ; M. T. ; N. T. ; A. Markham, 'Deep learning-based robust positioning for all-weather autonomous driving', *Nat. Mach. Intell.*, vol. 4, no. 9, pp. 749–760, Sept. 2022, doi: 10.1038/s42256-022-00520-5.

[34] R. Changalvala and H. Malik, 'LiDAR Data Integrity Verification for Autonomous Vehicle', *IEEE Access*, vol. 7, pp. 138018–138031, 2019, doi: 10.1109/ACCESS.2019.2943207.

[35] Y. Bang *et al.*, 'A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity'. 2023. [Online]. Available: https://arxiv.org/abs/2302.04023

[36] S. S. Gill and R. Kaur, 'ChatGPT: Vision and challenges', *Internet Things Cyber-Phys. Syst.*, vol. 3, pp. 262–271, 2023, doi: 10.1016/j.iotcps.2023.05.004.

[37] X. Liu, J. Lin, and J. Zhang, 'Beyond Text: Unveiling Multimodal Proficiency of Large Language Models with MultiAPI Benchmark'. 2023. [Online]. Available: https://arxiv.org/abs/2311.13053

[38] J. Dai, Y. Li, K. He, and J. Sun, 'R-FCN: Object Detection via Region-based Fully Convolutional Networks'. 2023. [Online]. Available: https://arxiv.org/abs/1605.06409

[39] D. Mahesh, V. Ravi Kumar, and S. Yogamani, 'Weather and Light Level Classification for Autonomous Driving: Dataset, Baseline and Active Learning', Sept. 2021, pp. 2816–2821. doi: 10.1109/ITSC48978.2021.9564689.

[40] J. You, S. Jia, X. Pei, and D. Yao, 'DMRVisNet: Deep Multi-head Regression Network for Pixel-wise Visibility Estimation Under Foggy Weather'. 2021. [Online]. Available: https://arxiv.org/abs/2112.04278

[41] N. A. F. Neto *et al.*, 'Low-latency perception in off-road dynamical low visibility environments', *Expert Syst. Appl.*, vol. 201, p. 117010, 2022, doi: https://doi.org/10.1016/j.eswa.2022.117010.

[42] Q. A. Al-Haija, M. Gharaibeh, and A. Odeh, 'Detection in Adverse Weather Conditions for Autonomous Vehicles via Deep Learning', *AI*, vol. 3, no. 2, pp. 303–317, 2022, doi: 10.3390/ai3020019.

[43] H. Chaabani, N. Werghi, F. Kamoun, B. Taha, F. Outay, and A.-U.-H. Yasar, 'Estimating meteorological visibility range under foggy weather conditions: A deep learning approach', *Procedia Comput. Sci.*, vol. 141, pp. 478–483, 2018, doi: https://doi.org/10.1016/j.procs.2018.10.139.

[44] J. N. K. Nina de Lacy; Michael J. Ramshaw, 'Integrated Evolutionary Learning: An Artificial Intelligence Approach to Joint Learning of Features and Hyperparameters for Optimized, Explainable Machine Learning', *Front. Artif. Intell.*, vol. 5, p. 832530, 2022, doi: 10.3389/frai.2022.832530.

[45] Y. Huan, F. Wu, M. Basios, L. Kanthan, L. Li, and B. Xu, 'IEO: Intelligent Evolutionary Optimisation for Hyperparameter Tuning'. 2020. [Online]. Available: https://arxiv.org/abs/2009.06390

[46] S. Biswas, A. D. Cobb, A. Sistrunk, N. Ramakrishnan, and B. Jalaian, 'Better call Surrogates: A hybrid Evolutionary Algorithm for Hyperparameter optimization'. 2020. [Online]. Available: https://arxiv.org/abs/2012.06453
[47] L. Zahedi, F. G. Mohammadi, and M. H. Amini, 'HyP-ABC: A Novel Automated Hyper-Parameter Tuning Algorithm Using Evolutionary Optimization'. 2021. [Online]. Available: https://arxiv.org/abs/2109.05319
[48] L. A. Gatys, A. S. Ecker, and M. Bethge, 'Image Style Transfer Using Convolutional Neural Networks', in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2414–2423. doi: 10.1109/CVPR.2016.265.
[49] X. Chen, S. Xie, and K. He, 'An Empirical Study of Training Self-Supervised Vision Transformers'. 2021. [Online]. Available: https://arxiv.org/abs/2104.02057
[50] K. Xu *et al.*, 'Show, Attend and Tell: Neural Image Caption Generation with Visual Attention'. 2016. [Online]. Available: https://arxiv.org/abs/1502.03044
[51] A. Vaswani *et al.*, 'Attention Is All You Need'. 2023. [Online]. Available: https://arxiv.org/abs/1706.03762
[52] J. Hu, L. Shen, and G. Sun, 'Squeeze-and-Excitation Networks', in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141. doi: 10.1109/CVPR.2018.00745.
[53] X. Wang, R. Girshick, A. Gupta, and K. He, 'Non-local Neural Networks', in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7794–7803. doi: 10.1109/CVPR.2018.00813.
[54] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, 'Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization', *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, Oct. 2019, doi: 10.1007/s11263-019-01228-7.
[55] L. van der Maaten and G. Hinton, 'Viualizing data using t-SNE', *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.
[56] T. T. Cai and R. Ma, 'Theoretical Foundations of t-SNE for Visualizing High-Dimensional Clustered Data'. 2022. [Online]. Available: https://arxiv.org/abs/2105.07536
[57] P. Manivannan, P. Sathyaprakash, V. Jayakumar, J. Chandrasekaran, B. S. Ananthanarayanan, and M. S. Sayeed, 'Weather Classification for Autonomous Vehicles under Adverse Conditions Using Multi-Level Knowledge Distillation', *Comput. Mater. Contin.*, vol. 81, no. 3, pp. 4327–4347, 2024, doi: https://doi.org/10.32604/cmc.2024.055628.
[58] S. Chen, T. Shu, H. Zhao, and Y. Y. Tang, 'MASK-CNN-Transformer for real-time multi-label weather recognition', *Knowl.-Based Syst.*, vol. 278, p. 110881, 2023, doi: https://doi.org/10.1016/j.knosys.2023.110881.
[59] M. Mutasodirin and F. Falakh, 'Efficient Weather Classification Using DenseNet and EfficientNet', *J. Inform. J. Pengemb. IT*, vol. 9, pp. 173–179, Aug. 2024, doi: 10.30591/jpit.v9i2.7539.
[60] 'Requirements for Road Weather Applications'. 2022. [Online]. Available: https://www.sae.org/standards/content/j2945/3_202201/
[61] *Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Decentralized Environmental Notification Service; Release 2*, TS 103 831, 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/103800_103899/103831/02.02.01_60/ts_103831v020201p.pdf
[62] M. N. Tahir, P. Leviäkangas, and M. Katz, 'Connected Vehicles: V2V and V2I Road Weather and Traffic Communication Using Cellular Technologies', *Sensors*, p. https://doi.org/10.3390/s22031142, 2022.
[63] L. Wan *et al.*, 'VALISENS: A Validated Innovative Multi-Sensor System for Cooperative Automated Driving'. arXiv, May 2025. doi: 10.48550/arXiv.2505.06980.
[64] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, 'Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review', *Sensors*, vol. 21, no. 6, p. 2140, Mar. 2021, doi: 10.3390/s21062140.
[65] Y. Poledna, M. F. Drechsler, V. Donzella, P. H. Chan, P. Duthon, and W. Huber, 'REHEARSE: adveRse wEatHEr datAset for sensoRy noiSe modEls', in *2024 IEEE Intelligent Vehicles Symposium (IV)*, Jeju Island, Korea, Republic of: IEEE, June 2024, pp. 2451–2457. doi: 10.1109/IV55156.2024.10588491.
[66] E. Palladin, R. Dietze, P. Narayanan, M. Bijelic, and F. Heide, 'SAMFusion: Sensor-Adaptive Multimodal Fusion for 3D Object Detection in Adverse Weather', in *Computer Vision – ECCV 2024*, Cham: Springer Nature Switzerland, 2025.
[67] A. Caillot, S. Ouerghi, P. Vasseur, R. Boutteau, and Y. Dupuis, 'Survey on Cooperative Perception in an Automotive Context', *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 14204–14223, Sept. 2022, doi: 10.1109/TITS.2022.3153815.
[68] *Intelligent Transport Systems (ITS); Facilities Layer; Cooperative Awareness Service; Release 2*, TS 103 900, 2025. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/103900_103999/103900/02.02.01_60/ts_103900v020201p.pdf
[69] *Intelligent Transport System (ITS); Vehicular Communications; Basic Set of Applications; Collective Perception Service; Release 2*, TS 103 324, June 2023. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/103300_103399/103324/02.01.01_60/ts_103324v020101p.pdf

[70] J. Luiten *et al.*, ‘HOTA: A Higher Order Metric for Evaluating Multi-object Tracking’, *Int. J. Comput. Vis.*, vol. 129, no. 2, pp. 548–578, Feb. 2021, doi: 10.1007/s11263-020-01375-2.

[71] ‘C-ITS Message Profiles v2.2.0’. 2025.

[72] A. Figueiredo, P. Rito, M. Luís, and S. Sargento, ‘Enhancing Vehicular Network Efficiency: The Impact of Object Data Inclusion in the Collective Perception Service’, *IEEE Open J. Intell. Transp. Syst.*, vol. 5, pp. 454–468, 2024, doi: 10.1109/OJITS.2024.3437206.

[73] V. A. Wolff, E. Xhoxhi, F. A. Schiegg, and J.-E. Neumann, ‘Uncertainty and Prioritization: Empirical Evaluation of a VoI-Based CPM Generation Pipeline Using Real-World Data’, in *2025 IEEE Vehicular Networking Conference (VNC)*, Porto, Portugal: IEEE, June 2025, pp. 1–8. doi: 10.1109/VNC64509.2025.11054252.

[74] Q. Delooz, A. Festag, A. Vinel, and S. C. Lobo, ‘Simulation-based Performance Optimization of V2X Collective Perception by Adaptive Object Filtering’, in *2023 IEEE Intelligent Vehicles Symposium (IV)*, Anchorage, AK, USA: IEEE, June 2023, pp. 1–8. doi: 10.1109/IV55152.2023.10186788.

[75] A. Rolich, M. Yildiz, I. Turcanu, A. Vinel, and A. Baiocchi, ‘On the Trade-off Between AoI Performance and Resource Reuse Efficiency in 5G NR V2X Sidelink’, in *2025 IEEE Vehicular Networking Conference (VNC)*, Porto, Portugal: IEEE, June 2025, pp. 1–8. doi: 10.1109/VNC64509.2025.11054234.

[76] Z. Xia, J. Zhao, and A. Vinel, ‘Dynamic-PBFT: Enhanced Consensus in Decentralized Federated Averaging for V2V Networks’.

[77] ‘ROADVIEW Deliverable 2.2, Use Cases and Scenarios’. [Online]. Available: https://roadview-project.eu/wp-content/uploads/sites/59/2024/05/ROADVIEW_Deliverable-2.2_v03.pdf