Designing cooperative interaction of automated vehicles with other road users in mixed traffic environments

interACT D5.1 Basis Sensor Data Fusion and Driver Monitoring

Work package	WP5: Integration, Testing and Demonstration
Task	Task 5.1
Authors	**Drakoulis, Richardos – ICCS** (Main author), Drainakis, Giorgos – ICCS, Portouli, Evangelia - ICCS, **Tango, Fabio – CRF** (Main author)**, Johannes Ruenz – Bosch** (Main author)
Dissemination level	Public (PU)
Status	Final
Due date	31/10/2018
Document date	31/10/2018
Version number	1.0
	This work is part of the interACT project. interACT has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395. Content reflects only the authors' view. The Innovation and Networks Executive Agency (INEA) is not responsible for any use that may be made of the information it contains

This is a draft version of deliverable D5.1
which has not been approved by the EC, yet.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Table of contents

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Index of figures

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

Index of tables

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Glossary of terms

Term	Description
Term	Description
Automated vehicle (AV)	Vehicle that provides automation of longitudinal and lateral vehicle control and can free the driver from the driving task
Cooperation and Communication Planning Unit	interACT central software unit that plans AV behaviour and explicit HMI control in an integrated, timely, and synchronised manner
Electronic horizon	Actual map information & road characteristics translated into data
Interaction	Within interACT interaction is understood as the complex process where multiple traffic participants perceive one another and react towards the continuously changing conditions of the situation resulting from actions of the other TP, to achieve a cooperative solution. These actions and reactions involve various means of communication
Motorised TP	Vehicles (cars, trucks, etc.) travelling on the road
Non-motorised TP	Pedestrians or cyclists
On-board user	Human on-board of the AV who acts as a driver in all cases the AV cannot handle (SAE level 3) or is a passenger for all SAE 4 and 5 applications
Other road user	All possible road users from the perspective of the ego vehicle (the AV) i.e. pedestrians, bicyclists, motorcyclists, vehicles, automated vehicles

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

List of abbreviations and acronyms

Abbreviation	Meaning
AV	Automated vehicle
CCPU	Cooperation and Communication Planning Unit
CF	Confusion Matrix
CR	Correct Rate
DF	Data Fusion
DD/DDC	Driver Distraction / Driver Distraction Classifier
DMS	Driver Monitoring System
FP	False Positive
FN	False Negative
HMI	Human Machine Interface/Interaction
Lidar	Light detection and ranging sensor. Used for the detection of distances and velocities of objects
Radar	Radio detection and ranging sensor
SDF	Sensor Data Fusion
TRP	TRaffic Participant
TP	True Positive
TN	True Negative
UML	Unified Modelling Language
WP	Work Package

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Executive Summary

The interACT project aims to study and model interactions among human traffic participants and develop software and hardware, which will enable an automated vehicle to interact with other traffic participants.

In its Work Package 5, the interACT project has revised the requirements (as derived in WP1 from target scenarios, from system architecture and from legal aspects). The deliverable D5.1 provides a direct association of the revised requirements to the different technical tasks, in order to make easier to derive an appropriate sensors setup, focusing on object detection and traffic participants tracking (both static and dynamic), as well as pedestrians intention features recognition.

In addition, we describe the work done for the on-board user's monitoring system, which will be installed on the CRF prototype vehicle, to provide valuable inputs to the HMI (developed inside WP4). In details, we use the driver monitoring data to adapt the on-board HMI strategy. Different amount of information will be displayed, depending on the driver distraction level. The ultimate goal is that the driver feels safe and well informed about the next actions of the automated vehicles and does not feel the need to intervene in situations that are well handled by the automated vehicle.

The next steps inside WP5 (task T5.1) are the complete integration of the perception platform on the prototype vehicles and the data collection, in order to perform the validation of the sensor data fusion. In addition to that, the on-board driver monitoring system will be installed and data will be available for WP4 (HMI) and possibly WP3 (for modules such as "Interaction Planning" and "Safety Layer"). These activities will guide the development work in the next WPs of the project.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

1. Introduction

The interACT project aims to study and model interactions among human traffic participants and to develop software and hardware, which will enable an AV to interact with other traffic participants.

Automation of the driving task is expected to increase road safety and improve traffic flow, among other possible benefits. Therefore, many efforts focus on the development and market deployment of vehicles that can automatically perform several parts of the driving task. An issue that has not been studied in depth yet, refers to the interaction of AVs with other traffic participants. The interaction between human traffic participants is a significant part of the driving task[1]. An AV needs to interact with other traffic participants, in order to efficiently and safely share the road infrastructure with them.

The following sections provide more details, such as the purpose and the scope of this document, as well as the interdependencies with the other deliverables.

1.1 Purpose and scope

As stated in the Description of Work (DoW), the main outcomes of WP5 are focused on:

- The integration of the interACT outputs from WP2/3/4 (sensors and perception, CCP Unit, HMI element) for testing in driving simulators and demonstrator vehicles.
- The adaptation of existing sensor–based data fusion and control algorithms to the project needs.
- The demonstration of the interACT vehicles.

Therefore, WP5 is about the implementation and integration of the CCPU (from WP3), the perception algorithms (from WP2) and the HMI elements (from WP4) in the prototype vehicles, in order to design and prepare advance functionalities, which are able to interact with the on-board user and other traffic participants. In order to ensure parallel research for the implementation and testing of the various components, the two vehicles used for the interACT demonstration focus on different scenarios and use-cases, whose currently development is based on the results of WP1. The main activity in WP5 includes the technical testing and the validation of the components, modules and systems of each demonstrator vehicle, in order to verify that their functionalities are fulfilled by the requirements defined in WP1. User evaluations of these will be conducted in WP6.

Based on that, the work presented in this document D5.1 focused on the description of the driver monitoring and of the design of the basic sensor data fusion, which serves as input for WP2. Overall, this deliverable is structured in the following way. After this introduction, chapter 2 deals with

[1] This includes the communication of own intent and anticipation of others' intent in order to mutually agree on a common future motion plan.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

the description of sensor data fusion (SDF) system, to be used inside the interACT project. Here, a short overview of the current literature is provided, together with the definition of the requirements for the SDF system. In addition, preliminary evaluation and data analysis are given. Section 3 is dedicated to the other main topic of deliverable D5.1, namely the driver monitoring system (DMS) to be used in the interACT project (the task is still in progress). A short description of the concept with an overview of the related work is provided, together with the data analysis and the main results achieved so far. Finally, the document ends with the conclusion and the reference sections.

1.2 Intended readership

This deliverable reports the scientific and technical activities carried out in Task 5.1 (T5.1), whose name is “Basic sensor fusion adaptation”. Its specific goal is to adapt the basic data fusion algorithms of the demonstrator vehicles, so that they comply with the requirements from the perception algorithms from WP2. This task will therefore involve an adaptation of the existing fusion modules, to extract the data needed by WP2 to recognise the intentions of the other road users. In addition, T5.1 will provide some insight into driver state, by providing driver-monitoring data, with specific focus on driver distraction. Internal sensors and Machine Learning (ML) algorithms will be used to adapt the HMI (e.g. to support different information strategies).

This document focuses on four different readerships. At first, it addresses mainly the interACT WP 5 partners, since it presents the current status of the basic data fusion and driver monitoring and it serves as a basis for its further development, integration and finalisation. Secondly, the content of this document also influences the technical work within the other interACT WPs, therefore it also addresses all other interACT partners who are involved in development or integration tasks. Thirdly, it gives the European Commission Project Officer of the interACT project an overview of the work conducted in the WP. Fourthly, since this is a public document, it is expected to serve as a useful reference to all interested researchers in academia and automotive industry.

1.3 Relationship with other interACT deliverables

This deliverable is part of the activity in Work Package (WP) 5: “Integration, Testing and Demonstration”. It is directly based on the deliverables D1.1 “Definition of interACT use cases and scenarios” (for the selected use cases) and D1.2 “Requirements and system architecture and interfaces for software modules” (for the extraction of the requirements, especially for the perception aspects). Furthermore, both the sensor data-fusion system and the driver monitoring system (including their requirements and related architecture of sub-systems) will directly influence the technical development of WP2 (related to the perception system) and of WP3 (in particular, related to the trajectory planning and safety layer activities), as well as the design of the HMI components in WP4.

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

2. Basic Sensor Data Fusion for the interACT system

The target in the interACT project is that automated vehicles interact with other human road users in their environment like pedestrians and manually driven cars. In this context the sensing and perception is a key to plan and control interaction in traffic. Especially the prediction of the behavior of other road users as well as the detection of interaction features, like hand waving, are necessary. To sense the automated vehicle's environment, different sensor technologies like video, radar, LiDAR and ultrasonic sensors are utilized. A sensor data fusion is used to combine different the measurements of different sensors technologies and benefit from the advantages of each sensor.

Generally, the focus of the interACT project in the sensing and perception layer is to fulfill new requirements on traffic participant's behavior prediction and new required features recognition.

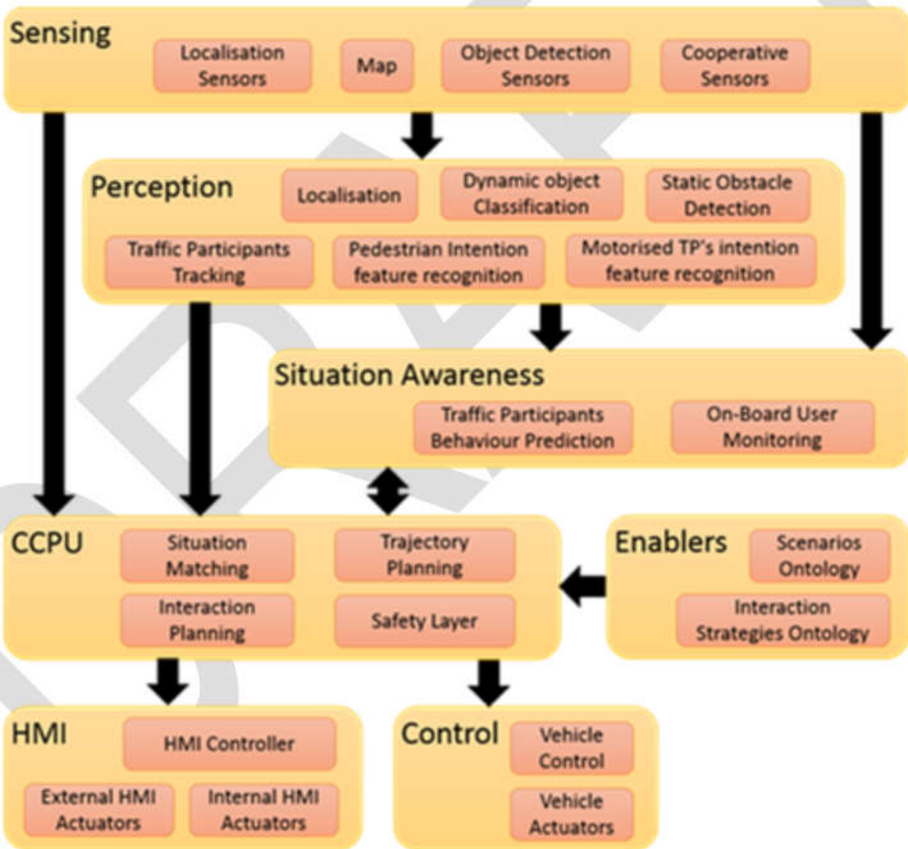

Figure 1: Sub-system and component in the functional blocks diagram of interACT deliverable D3.1

Figure 1 shows the basic signal processing structure in the interACT project. The sensing layer represents the fundament of the environment perception with sensors for ego vehicle localization, a digital map, object detection and cooperative sensors. The overlying perception layer processes the

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

information from sensing layer and provides derived signals such as localization, dynamic object classification and tracking, static obstacle detection and interaction features of pedestrians and motorized traffic participants. With information from sensing and perception layer the situation awareness layer determines the behavior prediction of other traffic participants and the on board user monitoring. Finally, the CCPU layer controls the external and internal HMI and behavior of the automated vehicle through vehicle actors.

The goal of this section is to present a reasonable setup and architecture of the sensing and perception layer that is consistent to the other interACT deliverables, especially D1.1 and D1.2, which will then be integrated in the CRF experimental vehicle. The focus of the experimental vehicle by BMW is to concentrate on the external HMI concept.

2.1 Procedure, methodology and structure

In section 2 of this deliverable, an appropriate sensor setup for the CRF experimental vehicle is derived by considering different requirements from different sources like the interACT deliverables from WP1 and 9. Deliverable D9.1 describes risks due to personal rights. Figure 2 shows the basic procedure in the derivation of the final sensor setup.

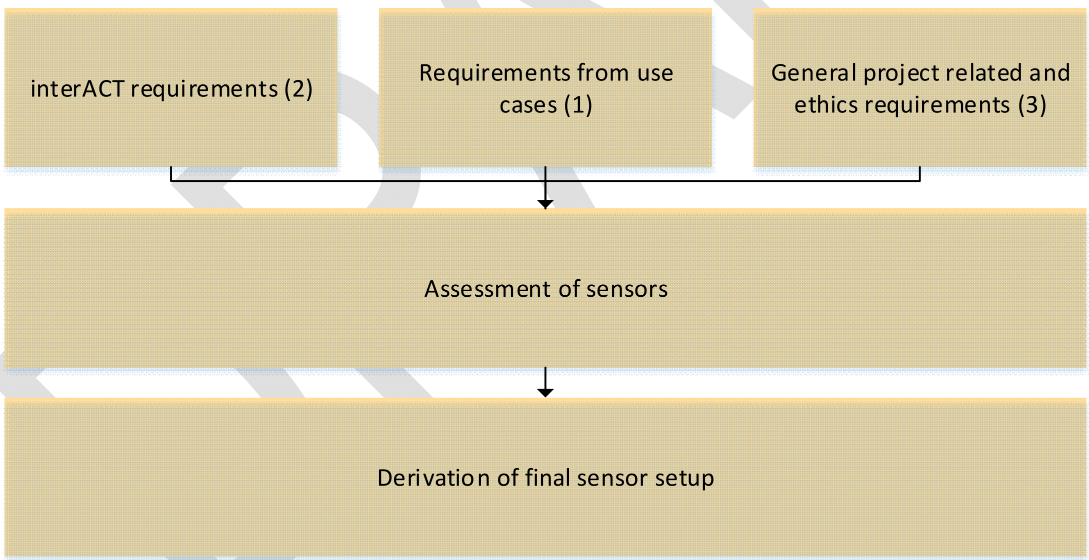

Figure 2: Basic Procedure of sensor setup derivation.

Generally, three main sources of requirements are given. The requirements on the sensing layer are derived by means of the interACT project’s target scenarios described in interACT deliverable 1.1 (1), the interACT partner’s requirements and system architecture described in interACT deliverable

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

D1.2 (2). Furthermore, there are basic requirements (3) due to legal aspects, special context in a public funded project and ethical constraints, which are described in interACT deliverable D9.1.

In the following section, the collected requirements from all interACT partners in deliverable D1.2 are clustered to the different interACT technical layer modules of D3.1. This gives a direct association of the requirement to the different technical tasks and makes it easier to derive an appropriate sensor setup. Requirements, which constraints the environment and addresses all technical layer modules, are collected separately. Afterwards, the interACT target scenarios described in D1.1 are analyzed and further requirements on the sensing layer are derived. Especially, further requirements on the opening angle of sensors are deduced. Further limitations due to legal aspects, special context in a public funded project and ethical constraints, which are described in interACT deliverable D9.1 are approached.

In context of all these requirements, constraints and limitations, an appropriate sensor setup is derived and the system architecture of the CRF interACT demo vehicle is shown. The CRF vehicle is the project internal platform the show the full functionality of the CCPU and all components and will demonstrate the complete technical signal chain from sensing to actor and HMI control. Compared to this, the BMW demo vehicle will be specialized on the close-to-market integration of external HMI components and will be set up in a way that WP 6 partners can do evaluation studies on real roads. This is why the following report focuses on the description of the CRF demo vehicle.

Finally, two preliminary experiments are described. The first experiment shows an evaluation of different GNSS INS systems for a purchase decision process. The second experiment is a feasibility study of detecting hand gestures with an advanced high resolution radar prototype sensor.

2.2 Partners Requirements on Sensing and Perception layer

In the beginning of the interACT project requirements were collected and presented in deliverable D1.2. To make the process of defining a sensor setup easier, the requirements are reclustered according to the interACT system architecture, described in D3.1. This leads to the following categories of the modules from the sensing and perception layers.

The requirements are shown in the following form. The ID, name, description and metric of the requirements are related to deliverable D1.2. Furthermore, the degree of fulfillment by the selected sensor setup described in section 2.6 is anticipated. The degree of fulfillment is discretized in fully fulfilled, partial fulfilled and not fulfilled.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

2.2.1 Sensing Layer: Localisation Sensors and Perception Layer: Localization

WP5_OPE_REQ_v01		**Fully fulfilled**
Speed range	The system shall work in the operative range (for speed), including HMI	(0 ÷ 30) km/h for CRF demo
WP5_PER_REQ_v03		**Fully fulfilled**
Digital map and Localization	The system may include digital map + localization sensor	Yes/Not

2.2.2 Sensing Layer: Map

WP5_PER_REQ_v01		**Partial through digital map**
Slot detection 1	The system should detect 2 parking lines and a third parking line or curb	Yes/Not
WP5_PER_REQ_v03		**Fully fulfilled**
Digital map and Localization	The system may include digital map + localization sensor	Yes/Not
WP2_PER_REQ_v15		**Partial through digital map**
Road perception	The system shall be able to detect and recognize road boundaries, intersections and zebra crossings	Relative position from road boundaries, intersections or zebra crossings

2.2.3 Sensing Layer: Object detecting sensors, Perception Layer: Static obstacle detection and Perception Layer: Traffic participants tracking

WP5_OPE_REQ_v01		**Fully fulfilled**
Speed range	The system shall work in the operative range (for speed), including HMI	(0 ÷ 50) km/h
WP3_OPE_REQ_v25		**Fully fulfilled**

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

TP Detection and Classification	The system shall be able to detect and classify other TPs	example classes/objects: - pedestrian - cyclist - vehicle
WP5_PER_REQ_v02		**Fully fulfilled**
Slot detection 2	The system shall be able to determine if a parking slot is valid to park	Yes/Not Parallel: length + 1m (longitudinal) Cross: length + 1m (lateral)
WP3_PER_REQ_v07		**Fully fulfilled**
Sensor location	Sensors shall be installed on the demo vehicle to cover 360deg of perception	Front, side, rear (requested)
WP3_PER_REQ_v08		**Fully fulfilled**
Surrounding traffic (e.g. leading vehicle required)	The system shall be able to detect surrounding traffic (in front and behind the vehicle)	FOV = 180° obstacle distance >10 cm minimal distance 2 m front and rear of vehicle
WP2_PER_REQ_v09		**Fully fulfilled**
Object Tracking	Tracking of other objects around the vehicle	• provision of position and velocity • no dense group of people • distance of objects > 1,5 m • objects with direct line of sight • tracking of pedestrian, vehicle, bicycle • determination of their centre of gravity
WP2_PER_REQ_v16		**Partial**
Intersection state	The system shall be able to determine the state of an intersection (open or blocked)	Intersection is blocked or not

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

WP3_PER_REQ_v21		**Fully fulfilled**
CCPU 6	The CCPU receives pre-classified objects from perception.	object classes with type, position, direction, velocity, etc.

2.2.4 Sensing Layer: Cooperative Sensors

WP2_OPE_REQ_v23		**Fully fulfilled**
Smartphone integration	The system should be able to receive data from pedestrian's smartphone	
WP5_PER_REQ_v04		**Fully fulfilled**
V2I/V2V	The system may include communication	Yes/Not

2.2.5 Perception Layer: Dynamic object classification

WP3_OPE_REQ_v32		**Not fullfilled**
No unclassified moving obstacles	In the scenarios, no moving objects, which do not fit to the category car, cyclist, single pedestrians, may appear.	
WP3_OPE_REQ_v25		**Fully fulfilled**
TP Detection and Classification	The system shall be able to detect and classify other TPs	example classes/objects:
WP3_PER_REQ_v10		**Fully fulfilled**
Type of Objects (Object Classification)	The system shall classify the (tracked) objects around	object distance > 2m classes: pedestrian, vehicle, bicycle
WP3_PER_REQ_v21		**Fully fulfilled**
CCPU 6	The CCPU receives pre-classified objects from perception.	object classes with type, position, direction, velocity, etc.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

2.2.6 Perception Layer: pedestrian intention feature recognition

WP2_PER_REQ_v12		Fully fulfilled
Recognition of intention and interaction feature "head pose"	Determination of pedestrians' head pose	• pedestrian distance to vehicle > 1,5m • position of the pedestrian must be in front of the vehicle
WP2_PER_REQ_v14		**Fully fulfilled**
Recognition of interaction feature "waving"	For intention determination and interaction planning; Relevance depends on the results of the observational studies.	• pedestrian distance to vehicle > 1,5m • position of the pedestrian must be in front of the vehicle

2.3 Requirements of interACT scenarios and use cases

In the interACT project, in deliverable D1.1, all partners identified relevant scenarios and use cases for the project. The scenarios were collected and rated relating to

- Relevance for safety,
- Relevance for traffic flow,
- need for interaction with human road users,
- Realisation in the vehicles and
- Realisation in driving simulators.

Afterwards, the interACT partners decided to focus on the must-have use cases listed in Table 1. Moreover, in addition to the must have use cases further use cases were identified, which have a research interest within the interACT project. The optional use cases are listed in

Table 2.

Table 1: Must-have use cases in interACT

Must have use cases
React to crossing non-motorised TP at crossings without traffic lights
React to an ambiguous situation at an unsignalised intersection
React to non-motorised TP at a parking space
React to vehicles at a parking space

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Table 2: Optional use cases in interACT

Optional use cases
React to vehicles in turning situations
React to crossing non-motorised TP at signalised crossings

For the detection angle of the object tracking there have been already hard requirements from the partners listed in the section. A tracking system, which covers 360 degree around the ego vehicle, is necessary.

Besides, there are only minor requirements registered in D1.2 on the range of the tracking system and also the detection area for intention feature recognition is not yet defined. In the following, these limits are derived with the interACT project scenarios.

2.3.1 Object detection range

In our use cases it is obvious that the requirements on the detection range are most challenging with the oncoming vehicle in the scenario "React to a vehicle while turning- Other vehicle yields" visualized in the following figure:

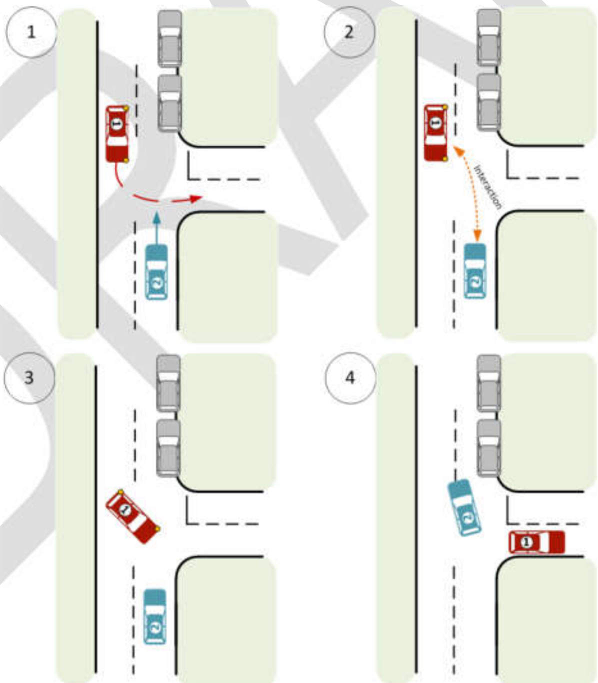

Figure 3: Scenario "React to a vehicle while turning- Other vehicle yields"

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

We can assume that both vehicles are driving with $50\frac{km}{h} = 13.9\frac{m}{s}$, decelerate to the standstill situation with comfortable $3\frac{m}{s^2}$ and have in the standstill situation a distance of $15\,m$. In this simple case the maximal needed detection range $r_{min,Tracking}$ is:

$$r_{min,Tracking} = \frac{\left(13.9\frac{m}{s}\right)^2}{3\frac{m}{s^2}} + 15\,m = 79\,m.$$

2.3.2 Detection area for interaction feature recognition sensors

In the interACT project the term "intention features" means indicators for the future behaviour of other traffic participants. For vehicles, this can be the recognition of braking manoeuvres. For pedestrians this can be for example the pedestrian's head orientation.

With interaction features, explicit gestures and indicators are meant. For vehicles, this can be a turn indicator recognition and for pedestrians for example hand gesture recognition.

With the assumption that the automated vehicle only drives forward, in the interACT project scenarios this leads to the detection area for and interaction feature recognition illustrated in the following figure:

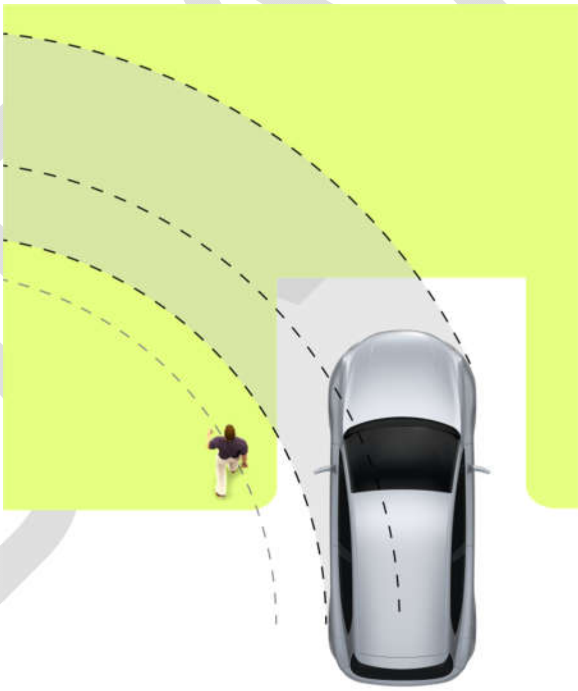

Figure 4: Qualitative illustration of the detection area for interaction feature recognition.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

The dark dashed lines represent the minimum turning circle for the vehicle and the grey dashed line corresponds to a safe distance to the pedestrian. The green area is the required detection range.

With the assumption that there are sensors which are looking in driving direction, the worst case is that the vehicle must interact with pedestrian next to it. Through a standstill manoeuvre it is possible to provoke the illustrated scene. In that case the coloured area needs to be covered by the interaction feature recognition sensors.

The use case examples in interACT deliverable D1.1 “React to multiple non-motorised TP (two from left one from right) at a parking space” in Figure 5 represents the worst case scenario. In this case the derived detection range is required.

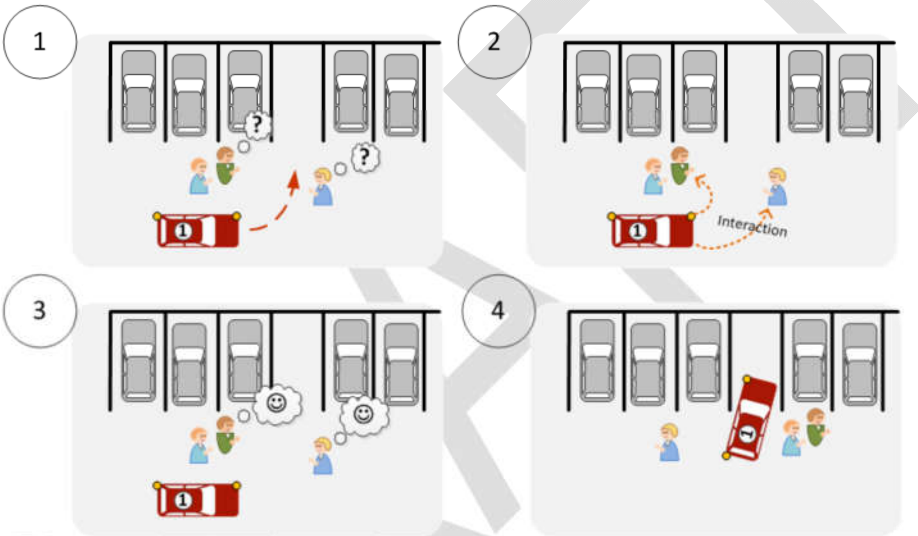

Figure 5: React to multiple non-motorised TP (two from left one from right) at a parking space.

2.4 Further general and project specific requirements

Besides technical requirements, it is also necessary to follow ethical requirements as well as legal rules as described in deliverable D9.1. Especially in conjunction with personal rights, care should be taken when using video images. It is possible to record other traffic participants from a moving vehicle with high resolution video images but the distribution of these video images must be strictly controlled.

For simplification, this leads to the usage of other sensor technology for main tasks like object tracking to allow a simple distribution of data, which empowers the recipient to deduce ground-truth data. Furthermore, in the interACT project a very complex and between the partner hard cross-linked system will be developed. The usage of sensors, which are freely available in the market, permit a simple distribution of measurements without limitation due to product protection of partners.

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

2.5 Perception Fusion requirements to support intention recognition of other vehicles

Knowing the intention and future behaviour of the other TPs is crucial for the AV to understand the encountered situation and react in the most appropriate way. In the interACT architecture two components that deal with motorised TP intentions and behaviour have been defined, namely the *Motorised TPs' intention feature recognition* and the *TPs' behaviour prediction*. The first aims to recognise the intentions of the motorised TPs that are perceived by the AV, while the second to predict their trajectories based on these recognised intentions. The adopted approaches that solve both intention recognition and trajectory prediction problems will be presented in D2.3.

Based on the recent literature, intention of a motorised TP is determined considering mainly its kinematic state (e.g. velocity, heading), its signals (e.g. turn indicator is flashing) the road topology (e.g. traffic signs, existence of adjacent lanes), possible interactions with other TPs (e.g. vehicle in front is slowing down) and the behaviour of its driver (e.g. head movement, driving style). In interACT, information on signals and driver behaviour is not available for the other TPs. The main sources of information for Motorized TPs' intention recognition are the Perception modules and the Map. From these components, the following information will be processed and fused:

Table 3: Sketch about which info is provided by which component.

Information	**Component(s)**
Road topology (static info)	Map (AV's Electronic Horizon: Road topology information, i.e. digital map information including specific urban and parking areas will be available for selected road segments on which WP5 and WP6 tests will be performed following the project's Common Road format).
AV motion state (dynamic info)	Localisation
TPs motion state (dynamic info)	TPs tracking
TP type (dynamic info)	Dynamic object detection and classification

The number of possible intentions of a motorised TP is very large and depends on the driving context. Since in interACT we are focusing mostly on urban and parking areas, a reduced set of manoeuvers will be considered as the possible motorised TP intentions to recognise which will be reported in D2.3.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

To support intention recognition, this type of information needs to be transformed to useful indicators, that we will call them *observations*. For the urban case, the following observations are considered:

Table 4: Table of info and variables used for urban scenario.

Observation	Code name	Information used for calculation
Existence of lane at left	ELL	AV motion state, Road topology
Existence of lane at right	ELR	
Existence of left turn	ETL	
Existence of right turn	ETR	
Existence of object in front	EOF	AV motion state, TPs motion state, Road topology
Time-to-collision with object in front	TTCF	
Relative velocity to object in front	RVOF	
Left lane is blocked (e.g. by another TP)	LBL	
Right lane is blocked	LBR	
Time to left lane crossing	TCL	
Time to right lane crossing	TCR	
Time to intersection	TTI	
Distance to intersection	DTI	
Vehicle shape intersects parking space boundaries (used only for parking scenario)	INPS	
Acceleration	ACC	TPs motion state
Velocity	VEL	
Heading relative to the road direction	HDN	
Yaw rate	YRT	

It is noted that the observations included in the above table have been considered for the "must-have" interACT usage cases that concern interactions in not-signalised intersections. Therefore, observations regarding the existence of traffic signs and lights have not been included. This simplification does not affect the methodology, since the observations associated with the signalised intersections are not depending on the others and thus they can be introduced as additional parameters in the calculation of each of the interactions without affecting the existing terms. Further information will be given in D2.3.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

2.6 Derivation of sensor setup

As experimental vehicle basis a Jeep Renegade, provided by CRF, is used. The vehicle is equipped with an ultrasonic system to detect obstacles, which are close to the vehicle. Furthermore, the vehicle is equipped with radar sensor that provides objects in the front of the vehicle.

The main focus of the interACT project is not the sensing and layer. Because of that ready-made solutions are preferred. Moreover, some information which is not available through ready-made sensors will be provided by a digital map. In the perception layer especially the intention and interaction feature recognition will be a new development in the interACT project. In these challenges also the sense of a sensor data fusion will be evaluated.

For the detection, tracking and classification of static and dynamic objects a freely available laser-scanner system that consists of six laser-scanners and an ECU is used. In the following figure the covered area by the equipped laser-scanners are shown:

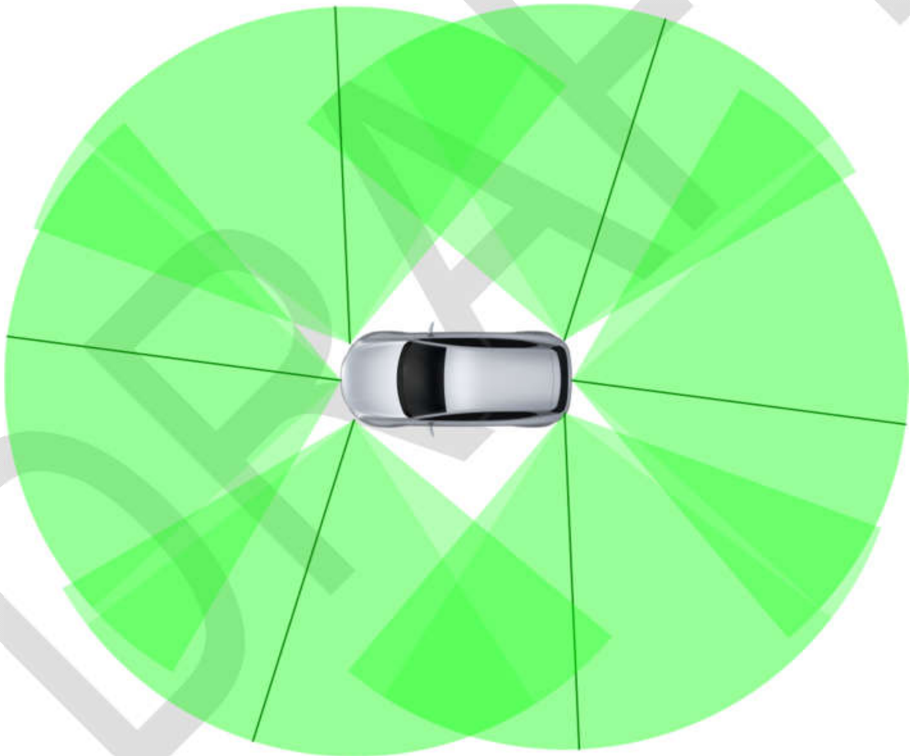

Figure 6: Visualization of the laser-scanner system's viewing field.

The advantages of the laser-scanner system is the high range and the covered area, no conflicts with personal rights of other traffic participants and the ability to share the measurements with partners without any issues because of one partners product protection. Furthermore, laser-scanners like radar sensors are independent from the environment light availability.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

For the detection of intention and interaction features like head orientation of pedestrians, waving of pedestrians and turn indicator of vehicles the potential of different sensor technologies will be evaluated. Therefore, the CRF vehicle will be equipped with two freely available stereo video cameras with a viewing angle of 90°. The cameras will be rotated and not looking in the driving direction of the vehicle. The following figure illustrates the viewing field of the cameras:

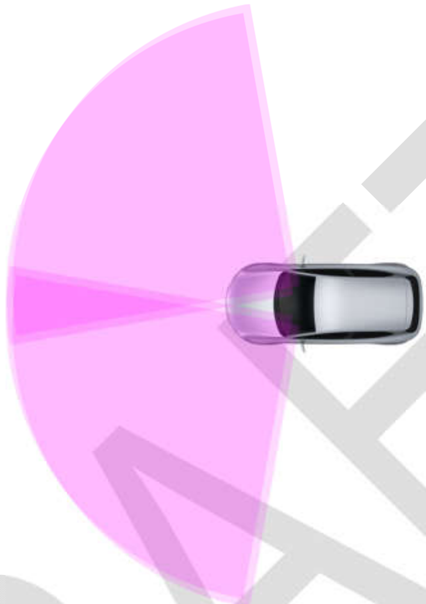

Figure 7: Visualization of the cameras' viewing field.

Besides the cameras, two advanced high-resolution radar sensors will be equipped to the CRF vehicle to detect hand gestures of pedestrians. The following figure shows the field of view of these sensors. Because of limited vertical opening angle, the mounting positions of the sensors are limited.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

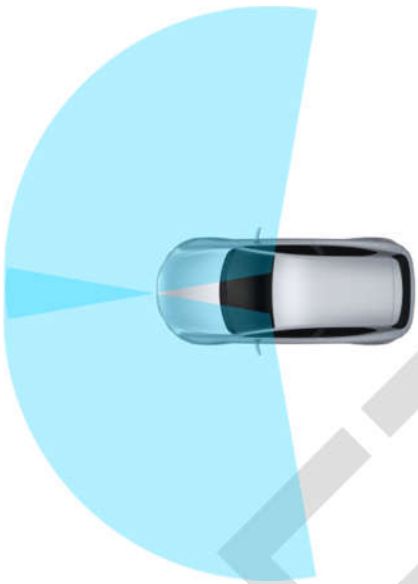

Figure 8: Visualization of the high resolution radar sensors' viewing field.

Moreover, a stereo video camera from BOSCH Group will be mounted on the CRF experimental vehicle that provide lane markings and Object detection, tracking and classification in a limited opening angle. The viewing field is shown in the following figure:

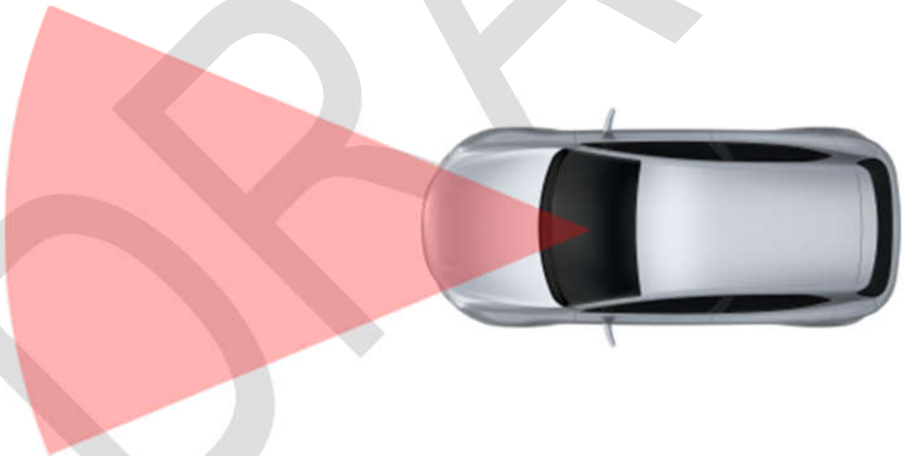

Figure 9: Viewing field of the stereo video camera.

For the localisation of the experimental vehicle, a GNSS INS system is used. Because of the low velocities in parking scenarios, a dual antenna setup that provides an accurate heading is used. With a correction service or a GNSS base station a position accuracy in the range of 2cm is possible. The GNSS

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

INS system is also equipped with an IMU that can provide accelerations and rotation rates of the vehicle.

In the following figure, the combined sensor setup with the opining angle of all sensors is visualized:

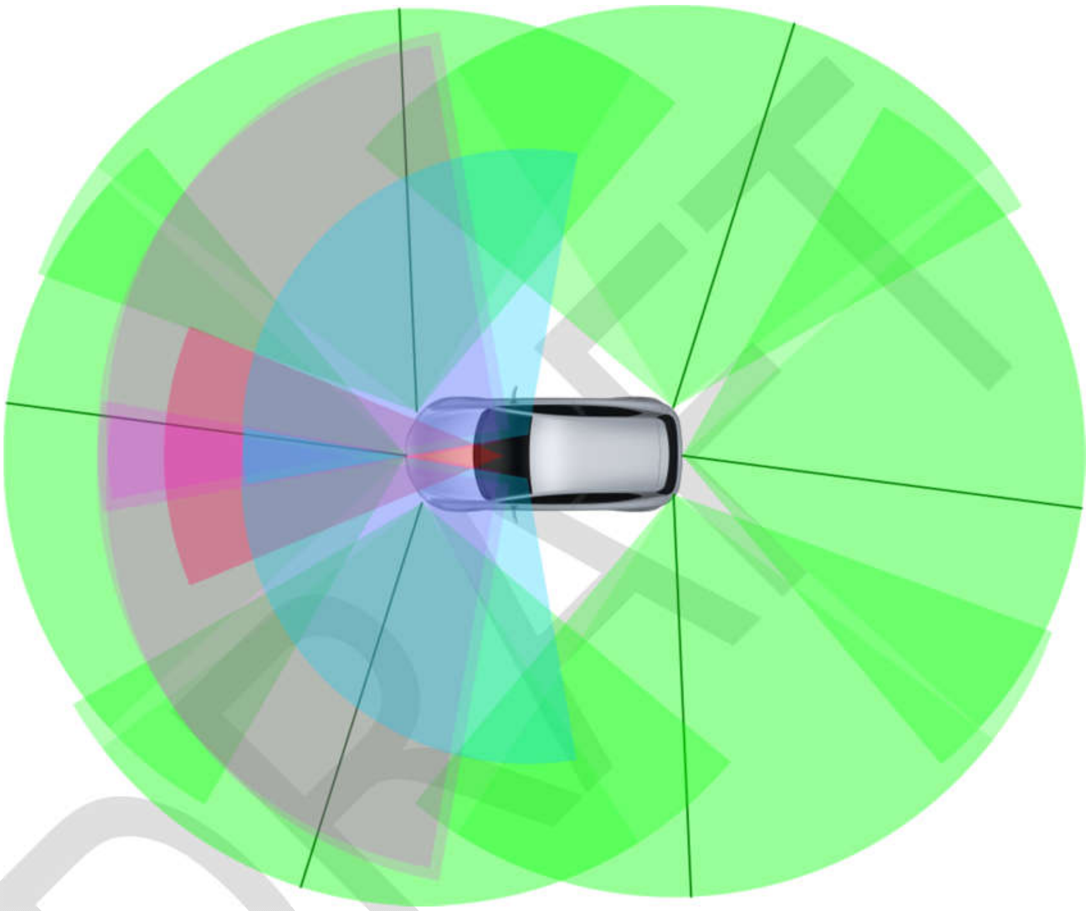

Figure 10: Viewing field of all sensors.

2.7 CRF vehicle hardware system architecture

The figure below shows the connection plan of all components in the CRF interACT vehicle except the driver monitoring system which is presented in section 3. The plan is divided into components located at the trunk (1), components located in the cockpit (2) and components located around the vehicles (3) like sensors.

In the trunk on the left side is the power management system which is connected to the vehicle 12V board network through a protection circuit. Every load like sensors or the measurement PC is

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

connected directly to the power management system. This makes a current protection of each load possible.

As core element a measurement PC receives the measurement from the sensors though a CAN interface and an Ethernet network. Moreover, the GNSS INS system is strongly fixed to the car body to allow a precise localization with a minimum of errors of the mounting position relative to the GNSS antennas and vehicle body. Furthermore, a PTP Grandmasterclock makes time synchronized measurements between all sensors possible. As cooperative sensors only Smartphones are evaluated in the interACT project. Therefore, a WLAN access point is mounted in the trunk to allow a simple connection as the project focus is not on the transmission channel.

In the cockpit a control element for the power management with an emergency switch for the measurement system is mounted. Besides, also a monitor and keyboard is available to control the measurement PC.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

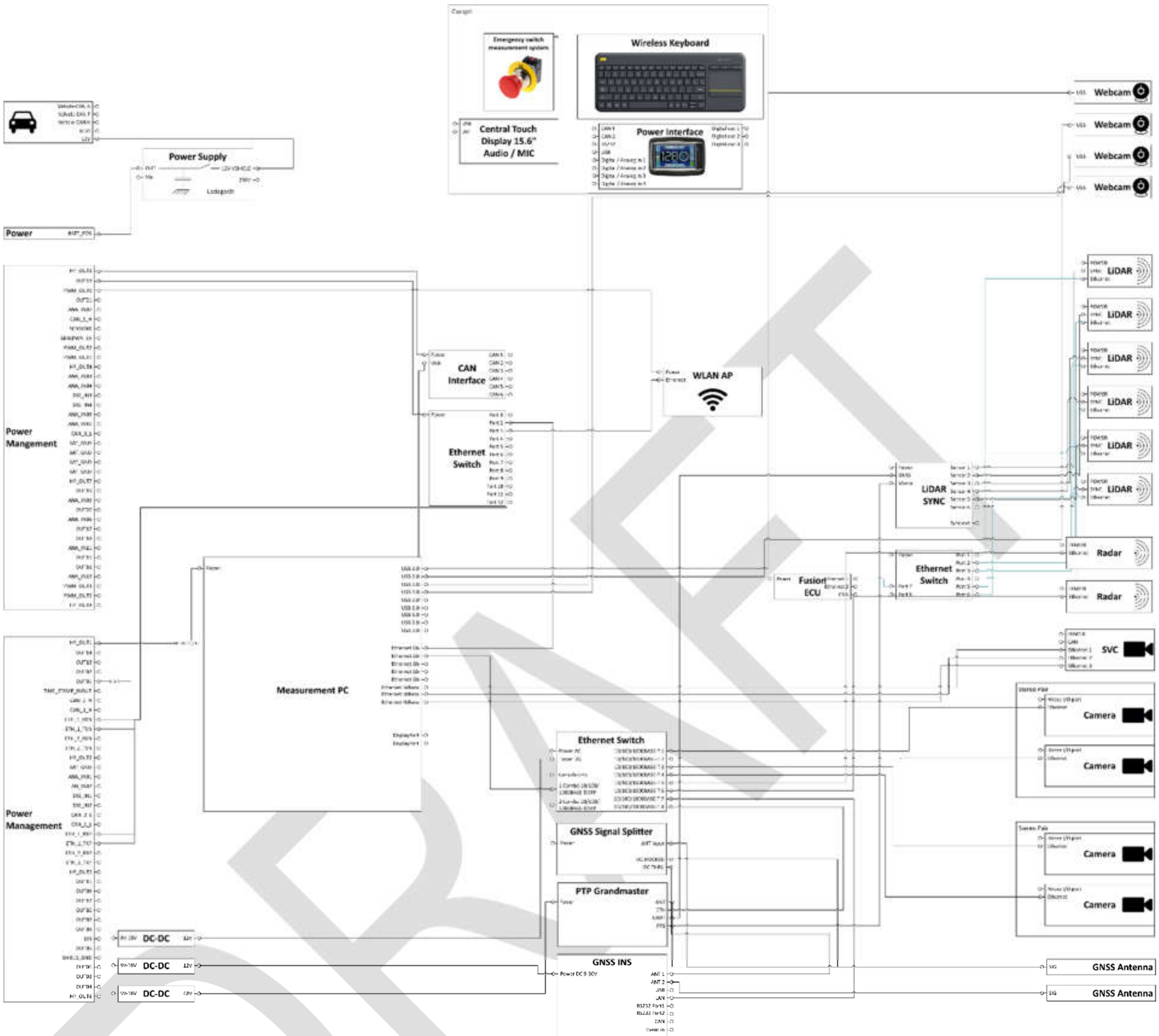

Figure 11: Hardware system architecture for sensing and perception layer of the CRF experimental vehicle

For simplification of the sensor data fusion signal processing, time synchronized sensor measurements will make sense. With an acceptable effort, the most accurate time source is the time received from a GNSS receiver.

The laser-scanner system can be synchronized by a GNSS PPS time pulse signal paired with a NMEA messages. This guarantees that measurements of all sensors are in driving direction at seconds change. The update frequency of the laser-scanner system is set to 25 Hz.

The high-resolution radars and GigE cameras can be synchronized by PTP IEEE 1588 protocol. To synchronize the PTP master to GNSS clock a GNSS PTP Grandmasterclock is used. The cameras and radar are also set to an update frequency of 25 Hz.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

2.8 Experiments for preliminary Evaluation

In the following sections, we provide an overview of the preliminary evaluation, carried out on the components illustrated in the previous chapters.

2.8.1 Comparison of GNSS INS systems

The system architecture of the interACT project completely depends on an accurate localization of the vehicle. Because of that, the temporary equipment with cost expensive GNSS INS systems does not make sense. Instead, it was decided that we try to equip the vehicle with a less expensive GNSS INS system, which allows the testing of the vehicles in good GNSS environments. Therefore a comparison of less expensive GNSS INS systems with a highly accurate system was done.

In sum, three systems were equipped to a vehicle to the same GNSS antennas over signal splitters. Every system receives the correction information from the German SAPOS provider. Afterwards, the initial setup processes for all three systems were performed and a test run was including highway to urban scenarios.

In the following figure, the horizontal positioning accuracy in relation to the reference system is illustrated in a histogram. Especially in good GNSS environments, the position accuracy is sufficient and system A is used in the CRF vehicle.

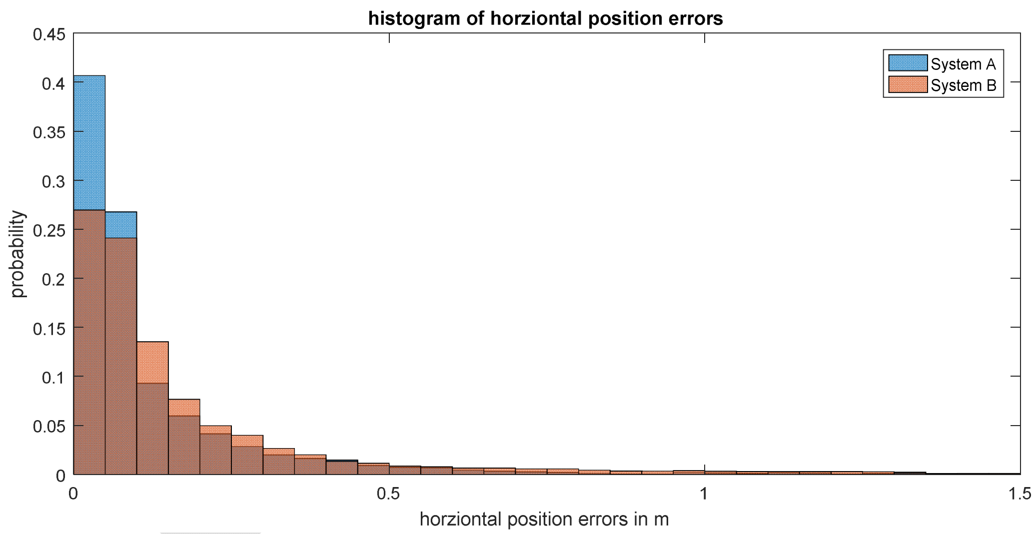

Figure 12: Histogram of horizontal position errors.

.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

2.8.2 Feasibility study: Test sensitivity of high resolution radar

In the interACT project it is planned to support the detection of pedestians' hand gestures with radar sensors. Therefore, it is necessary to detect radar reflections of the hand and arm and a feasibility study was realized.

In the following figure, a snap-shot of this measurement is printed. A test person stands in a distance of approximately 20m to the radar sensor and executes hand waving gestures. Below the video image the radar detection are illustrated in the birds eye view. The main cloud represents detections on the test person and on the right radar detections of the side vegetation. Finally, also the distance to velocity diagram is plotted. The velocity is directly measured through the Doppler frequency of the reflection. The detections of the arm and hand have velocities unequal to zero.

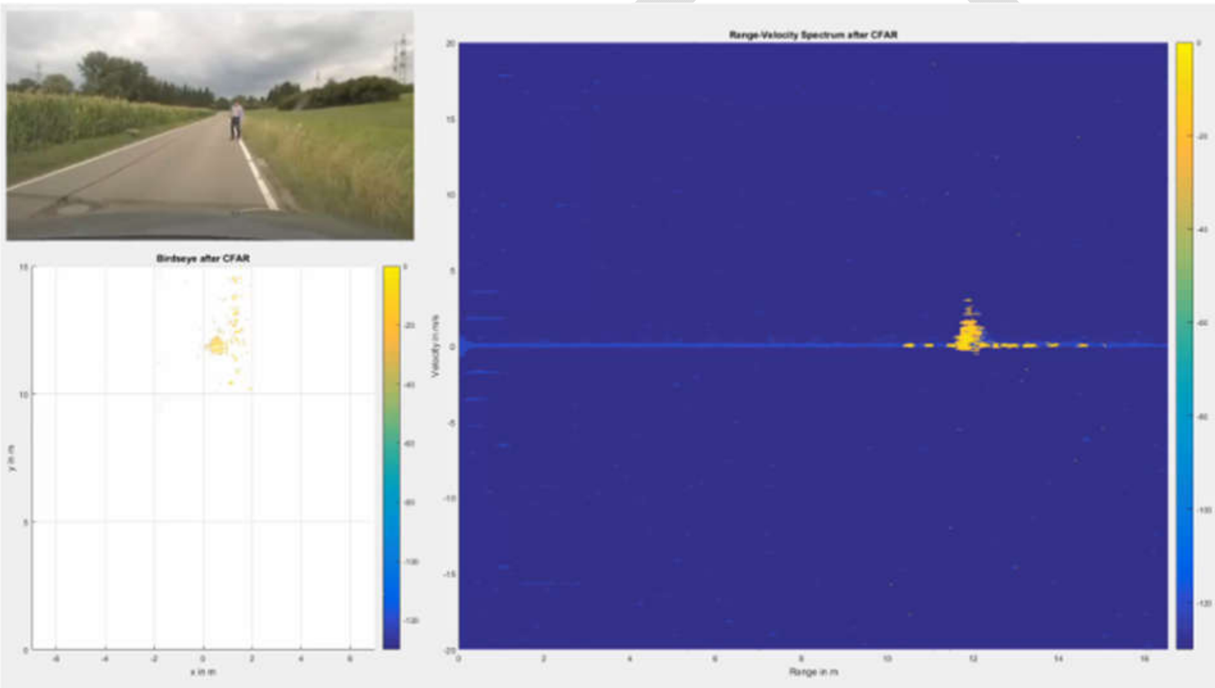

Figure 13: Snap-shot of the feasibility study of the high resolution radar.

This leads to the assumption that It seems possible to fusion radar measurements with video measurements to detect hand gestures of pedestrians, but further testing and development has to be done.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

3. On-board User Monitoring for the interACT system

Driver distraction – and inattention – is an important safety concern. With increasing automation of the driving functions this safety issue, is somehow addressed by the vehicle automation and give new freedom to the driver to do other things while driving. This is one main benefit of vehicle automation.

However it is necessary to still know about the driver state for two reasons. First, in automation level SAE 1-3, the driver might be asked to take over in critical situations, thus it is important to ensure that the driver is able to take over and not completely out of the control loop (e.g. not heavily distracted or sleeping). Second, driver monitoring data can be used to adapt the HMI and the interaction strategy of the AV depending on the driver status. In particular:

- If the driver is distracted less information about the environment might be needed.
- Information could be presented at different locations in the vehicle depending on the task the driver is involved in, e.g. information on the tablet if the driver is using this.
- The driving strategy of the vehicle itself might be adapted depending on the distraction level – e.g. more comfortable driving for distracted drivers, more sporty driving for attentive drivers.
- The time to take back the driver in the control loop could be adapted depending on the distraction level, distracted drivers will need more time and eventually other/enriched information to get back in the loop

In the interACT project, we use the driver monitoring data to adapt the on-board HMI strategy. Different amount of information will be displayed, depending on the driver distraction level. The objectives of the on-board HMI design in the project are to reach the appropriate information level for the driver to feel safe and comfortable and to identify the most relevant information depending on different levels of distraction. The overall goal of the design task is that the driver feels safe and well informed about the next actions of the automated vehicles and does not feel the need to intervene in situations that are well handled by the automated vehicle.

The following figure shows where the on-board user monitoring component is collocated inside the system architecture (as proposed in deliverable D3.1 "Cooperation and Communication Planning Unit Concept"):

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

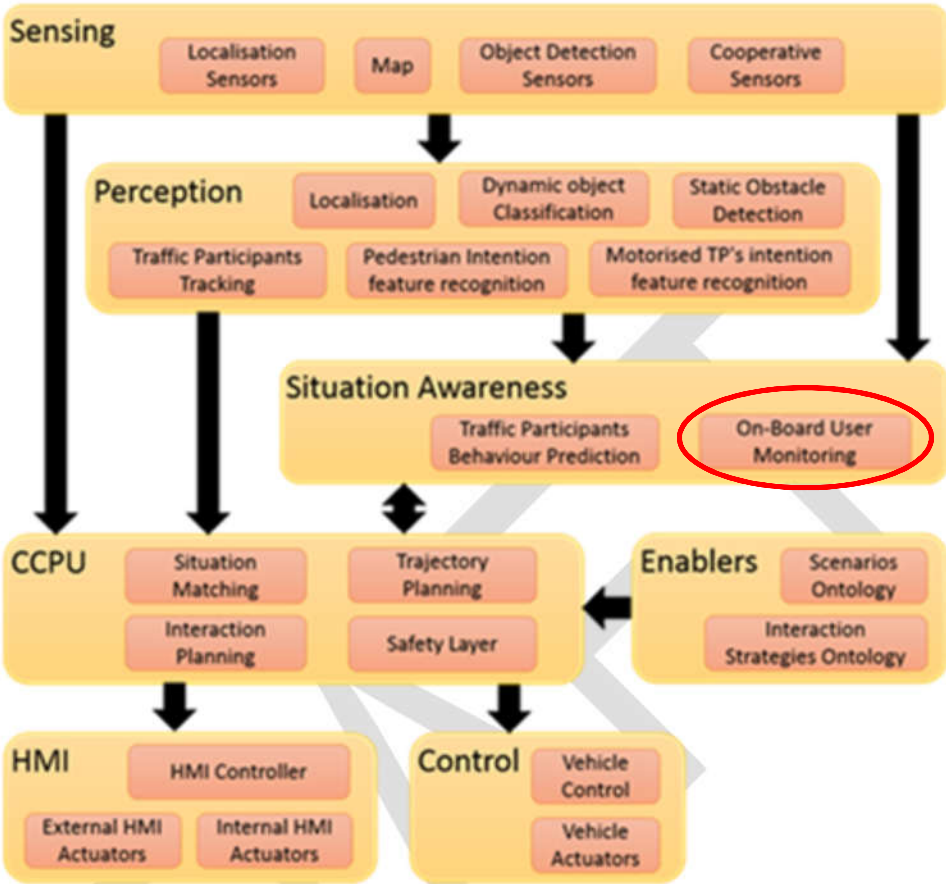

Figure 14: Sub-systems and components in the functional blocks diagram, highlighting the role of driver monitoring system.

This picture of the components can help to understand that the User Monitoring component is part of the situation Awareness module and gives input to the CCPU (in particular to "Interaction Planning" and "Safety Layer" modules).

3.1 Context Description

Advanced Driver Assistance Systems (ADASs) have been progressively introduced into vehicles, in order to make driving task safer, more efficient and more comfortable [1]. In addition, all these systems are paving the road to the autonomous vehicles (AVs) which will be a reality in a couple of decades.

It is well known that, in normal driving mode, the majority of road accidents (> 80%) are due to human error, or anyway human (wrong) behaviour ([2], [3]). Among these, distraction is a key factor for driving safety: it was responsible for 391,000 injuries and 3,477 fatalities in 2015 (see [4] and also [5]-[8]). Many sources can lead to distraction, but the most likely ones usually come from inside the vehicle. Thus, drivers do many things while driving manually, but now with automated driving

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

functions they might be allowed to do this. Anyway, what happens when they have to take back into the control loop of the vehicle (take-over request, e.g. for reaching the system limits) and they are not ready to do that? So, in such a context, the development and the evaluation of a robust driver distraction classifier (DDC) can be of paramount importance for developing more efficient Autonomous Driving Functions (ADFs).

In order to evaluate the system able to monitor the driver's state, we adopted the following definition for distraction: "*anything that delays the recognition of information necessary to safety maintain the lateral and longitudinal control of the vehicle (driver's primary task) due to some event, activity, object or person, within or out-side the vehicle that compels or tends to induce the driver's shifting attention away from the fundamental driving task by compromising the driver's auditory, biomechanical, cognitive or visual faculties or combinations thereof*". Activities not related to primary tasks, that driver performs while driving, are defined as secondary activities [13]. This means that a driver is considered to be distracted when there is an activity that attracts his/her attention away from the task of driving. It is worth to note that more types of distractions occur at the same time: manual and visual distraction, as well as cognitive distraction. In this project, we focus on visual distraction, which is *the diversion of attention toward a competing activity that requires the driver to look at a secondary target inside the vehicle instead of looking at the road*.

3.2 Related Works

Although much research interests have been attracted in recent years, today there are still no accurate evaluation technologies for DD, unless in combination with the DMS devices. In the next subsections, we provide an overview of the existing technologies and techniques to detect and classify driver's distraction.

3.2.1 Sensing Technologies

Many different attributes were proposed for driver's distraction detection, which can be classified based on the sensor modality and the detection methods ([16]-[19]). In existing systems, three types of sensing modalities can be identified:

- psychological;
- vehicle dynamics/behaviour;
- visual.

The psychological signals are corporal parameters that can be altered when the driver is being distracted, including electroencephalogram (EEG) and electrocardiogram (ECG). In this context, most used parameters are heart rate, blood pressure, electro-dermal activity, or brain activity. In general, these types of measures can be very accurate and give fast results, but they are also "intrusive", since drivers have to wear these physiological sensors on their body, which may cause discomfort and even interfere with their driving movements. Therefore, at least at the moment, their use is limited in real and practical applications.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

The vehicle control data include – among others – steering wheel and pedal positions, assuming that, such movements, are usually different for distracted drivers than for attentive drivers [20]-[21]. However, since driving performance signals refer to the signals of the dynamics of the vehicle and the driving task itself (e.g., velocity, accelerations, engine r.p.m., etc.) they are affected by the driving style of each driver and also by changes in mental state like, for example, distractions. This can lead to misclassification of states and DD methods based only on vehicle dynamics signals can take long time to take a decision.

Finally, visual signals are those signals related to the gaze, eye movements (blinking, time/percentage of eyes closure, etc.) or head pose and they also include images or videos of the driver facial expression and body movements.

Examples can be found in [22]-[26], in which several techniques and methods are used to detect distracted driving in several projects, such as PERCLOS (Percentage of Eye Closure) and FaceLab (a computer vision system that provides real-time measurement of eye glance using head and eye tracking techniques to determine the instantaneous distraction and allowing for a classification of the distraction level). Infrared (IR) cameras are also used to detect the driver's eyes and thus to monitor driver vigilance.

3.2.2 Methods for Distraction Detection

There are two methods for developing a classifier of driver's distraction. The first one is based on the use of a threshold, in which a certain feature value is compared with a pre-set threshold (see the work [22] of Tabiti and colleagues, defining this value by using PERCLOS). Due to the nature of this problem, we can think also about a second method, based on Machine Learning (ML) techniques, which seem to be very appropriate. In fact, the data, that are usually collected, are definitely nonlinear and several studies in literature have proved that in such situations ML approaches can outperform the traditional analytical methods. Moreover, also human's driver mental and physical behaviour is non-deterministic. Therefore, since the mental state of the drivers is not observable, no simple measure can index visual and cognitive distractions precisely [27-29].

In this context, researchers have used different machine learning techniques such as Support Vector Machines (SVMs) (where control data and eye movements are used as features, such as in [15]-[20]-[30]), Hidden Markov Models (HMMs) [31], Random Forest (RF) [32], Neural Networks [38], k-nearest neighbours (KNN) [33] or AdaBoost (AB) [32]. In order to model the inherited uncertainty associated with the face features, als

o Bayesian Networks (BN) are used, which are considered to determine the probability of distraction of the driver, such as in [34]-[35]-[36]. In addition, Artificial Neural Networks (ANNs), one of the most popular machine learning approaches [37] have been applied to solve these problems, due to its robustness, ability to learn by example, and efficiency in intelligent systems (such as [38]). Other examples of the combination of dynamic signals and facial data can be found in [39] and [40]. Finally, also Fuzzy Logic (FL) based systems are widely used (like in [41]; a thorough review on fuzzy

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

methods can be found in [42]). Sometimes, FL is combined with an adaptive neuro-fuzzy inference system (ANFIS) to detect the DD [43]. The authors claim this approach showed better performance comparing to the ANN and radial basic function prediction algorithms.

3.3 Short Description of the Experiments

In this section, we describe the methodology followed to conduct the experiments, in order to collect the data for a twofold reason.

The research questions (RQs) we want to answer with this activity are the following:

1. Which are the performances of DMS based on the use of internal camera? Moreover, which is the best system to use for the goal of InterACT project, in terms of accuracy?
2. Is it possible to create a distraction classification without using the signals of internal camera as inputs, having the same – or anyway comparable – performances?
3. Which are the improvements that are possible to achieve if we combine the DMS with the behavioural and dynamic data?

RQ1 is crucial for the project and it is used mostly when the automated mode is possible and active. RQ2 and RQ3 are (more optional and) used when the driver is involved in a normal driving mode.

First, the evaluation of some Driver Monitoring Systems (DMSs), provided by different suppliers and based on camera looking at the driver's face and eyes. Second, the possibility to create also a training set, to develop new classification models and make a comparison with the aforementioned DMSs[2].

3.3.1 The Procedure

The following figure sketched the apparatus used to collect the data:

[2] It is worth to note here that these two approaches are not mutually exclusive, since the DMS-only can be used during the automated mode, while the other classifiers can be used when the driver is involved in the driving task.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

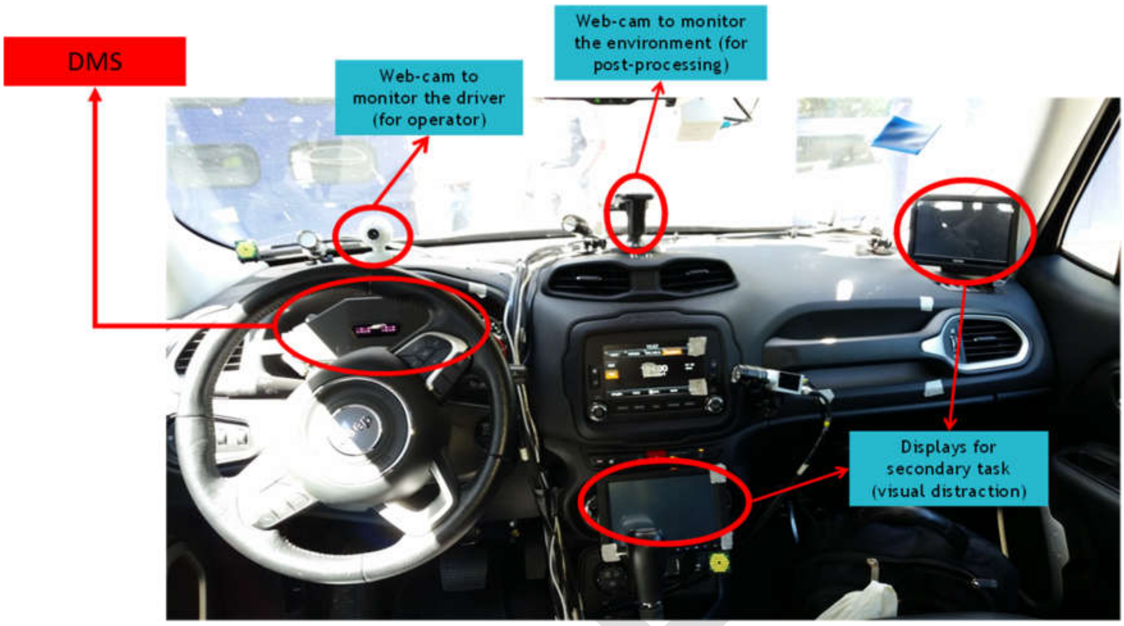

Figure 15: Cockpit Installation for the experimental set-up in test-vehicle.

The experimental set-up is illustrated in figure 1, where the main components are illustrated. First, the driver monitoring system (DMS), whose data are used both to improve the models performance and to compare the achieved results. Then, there are two web-cams, one used for the additional monitoring of the driver (to build the target-set in the post-processing phase) and the other for monitoring the external environment (in case an ambiguous situation has to be solved). Finally, there are the two displays where the secondary task is displayed to the subjects, in order to cause the distractive conditions.

Participants were asked to drive on the dedicated test-site in real-traffic situations, while completing a secondary task session. This consists in reading a sequence of random letters, displayed on one of the two secondary screen (lateral or bottom, namely at climate control and right A-pillar, selected randomly, see figure 1). The necessary time to read the letters sequence was about 2s, in such a way that subject's eyes were out of the road for this time period. The reason of this choice can be found in literature and in particular there is a study by AAA Foundation for Traffic Safety (https://aaafoundation.org/) and the University of Utah, showing that on-board infotainment systems often create unnecessary visual and cognitive demands on drivers (due to the fact that they are often complex, frustrating, and maybe even dangerous to use). In such a study, AAA wrote in a press release, that "*Removing eyes from the road for just two seconds doubles the risk for a crash*". In general, other studies claim a value in the interval (1.8÷2.2)s.

In our experiments, during the reading period (with secondary task active), three options are possible:

- Task fully completed = 1 (driver is fully distracted).

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

- Task partially completed = 2 (driver is partially dis-tracted).
- Task not completed at all = 3 (driver has not performed the task, so the status is unknown).

Total of 40 distraction tasks during the single test has been presented. In particular, the distraction task execution worked as following:

- Speaker announces the imminent task
- Display is activated and characters sequence is presented on the selected display.
- User is requested to read text aloud (drivers are instructed to complete the reading task but guaranteeing a safe driving).
- Experimenter confirms if the reading is complete.

Another operator on-board vehicle wrote down one of these labels as well, based on driver's behaviour. Then these labels are validated and confirmed in the post-processing phase.

For the creation of the target-set and for assessment purposes, we have considered only when the task was fully completed (to be sure that the driver took the eyes out of the road for 2s). In fact, the main parameters to be evaluated were the correct / missing detection during a "classified" distraction task (complete execution by the driver without glances to the route).

3.3.2 The Subjects

Data have been collected from dedicated driving session. Thirty (30) test subjects (internal users with special car license for driving prototype vehicles) drove for about 1h (76km) on extra-urban and motorway roads. Two of them hasn't provided coherent data and have been neglected, thus only twenty-eight subjects have been considered for the analysis. A minimum amount of driver experience was required, in particular:

- at least 2 years of driving license;
- at least 6000 km driven per year.

Gender has been controlled (7 female and 23 male). Overall, we obtained the following hints: about 1992km, 30h, #960 distraction tasks.

3.4 Data Analysis and Methods

Data have been collected with a frequency of 10 Hz (1 data-point each 100ms), using CANAPE software tool for data logging and synchronization. The data collected (primary variables) are reported in the following table:

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Variables	Unit of Measure
Speed	[m/s]
Lateral Position	[m]
Steering Angle	[deg]
Yaw Rate	[deg/s]
Lane Width	[m]
Road Curvature	[1/m]
Heading Angle	[deg]
Position of the accelerator pedal	[%]
Use of the brake pedal	[#] (yes/not)
X,Y coordinates of car in front	[m] (if any)
Speed of car in front	[m/s] (if any)
Combination (such as TTC, TTLC, HD, etc.)	[s]

Table 5: List of primary variables, as recorded by CAN bus through the CANAPE software tool.

Then, these variables have been combined using statistics (e.g. mean, standard deviation, derivative, percentiles, max-min, etc.), considering a mobile window of 2s (the same one used for distraction). This means that for each parameter in the list, several features have been computed to group the data. Totally, we have obtained 135 features as inputs for the different classifiers.

Following the ordinary procedure for supervised learning, the whole dataset has been divided into three subsets, as following:

- Training data (around 60% of the whole dataset), which are presented to the network during training and the network is adjusted according to its error.
- Checking data (around 15% of the whole dataset), which are used to measure network generalization and to halt training when generalization stops improving.
- Testing data (around 25% of the whole dataset), which have no effect on training and so provide an independent measure of network performance during and after training.

This dataset division has been used, considering subject 1-25, to develop and implement the distraction models. For the validation purposes, we have considered subjects 26-28.

3.5 Results

In this paragraph, we describe how we addressed the detection of distraction, using Machine Learning (ML) techniques to model the driver's state [44], as sketched in the following scheme.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

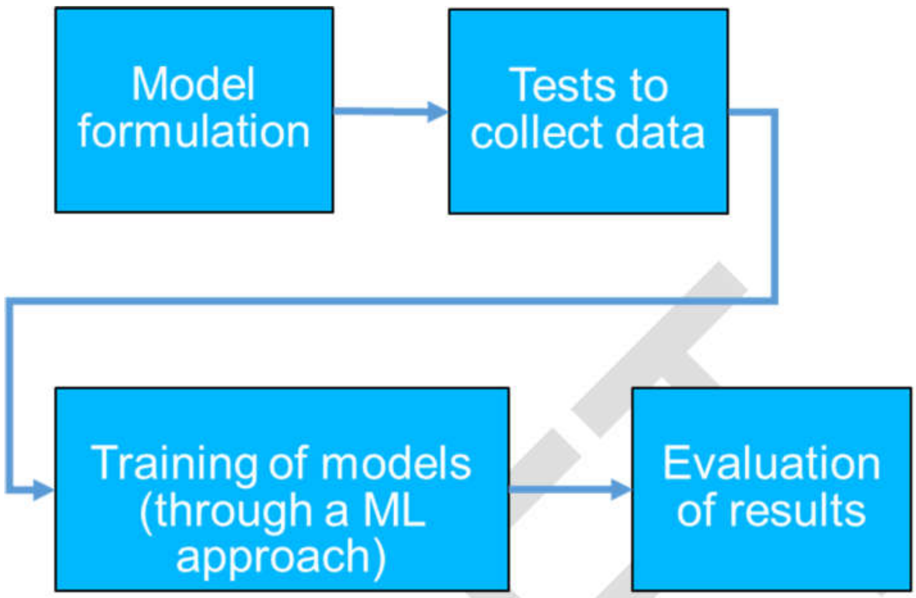

Figure 16: Traditional scheme to model data in Machine Learning approach.

ML (and Data Mining) technology may be able to provide the right algorithms to cope with the challenge of de-signing and developing a model to classify the driver's distraction. In fact, ML is the technique of searching large volumes of data for unknown patterns. In particular, this technology can be applied to build a discrimination model that captures the differences in behaviour when people drive normally and when they are distracted. We have identified the following ML algorithms: Support Vector Machines (SVMs), Ensemble Methods (EMs), Feed-Forward Neural Networks (FFNNs) and Decision Trees (DTs). In addition, we have considered two Driver Monitoring Systems (DMSs) from two different suppliers, hereafter named DMS1 and DMS2 (not provided the real names for sake of anonymity).

As performance indicator (PI), the Correct Rate (CR) of classification has been regarded as one meaningful parameter to assess the different models (for all the implemented models). In addition, we have considered also the number of true positives (TP), of false positives (FP), of false negatives (FN) and the number of true negatives (TN); they represent the confusion matrix, since it is also important not only to understand the global accuracy of the classifier, but also – and maybe above all – the actual capacity of the model to correctly identify when the driver is distracted.

3.5.1 Results of the Driver distraction Classifier alone

The first evaluation considers the DDC alone, that is without the support of any DMS (internal camera looking at the driver). We have considered different ML techniques and create different classifiers, as illustrated in this section. For the ones based on SVM, EM and DT, we adopted a Bayesian Optimization algorithm (see [58]-[60]), which attempts to minimize a scalar objective

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

function f(x) for x in a bounded domain. The function can be deterministic or stochastic, meaning it can return different results when evaluated at the same point x. This allows us to optimize these models in the selection of many characteristics, such as the kernel functions and related parameters value, the depth of the trees and number of leaves, the type of weak learners and related parameters value, and so forth.

For SVMs, the best model had the following parameters and characteristics:

- Type of Kernel = Polynomial
- Order = 2
- Box Constraints = 0.099062

For EMs, we have selected a model with:

- Method = AdaBoost
- Number of Learning Cycle = 448
- Learning Rate = 0.83507
- Minimum Leaf Size = 1

For ANN, we have obtained:

- training method = Scaled Conjugate Gradient Back-propagation
- number of layers = 2 layers topology has been chosen, with one Hidden Layer (HL – very rare the case in which more than 2 are needed; in our case, two did not provide appreciable improvement of results) and one Output Layer (OL)
- transfer function = a Sigmoid transfer function has been used for both the HL and OL.
- Hidden Neurons (number of neuron in the HL) = 50

Finally, for DTs, the best model was:

-
- Minimum Leaf Size = 1
- Maximum number of slits = 89
- Split criterion = Gini's diversity index (GDI)[3]

A Comparison of the performances is done in the following table:

[3] The Gini index of a node is defined as such: $1 - \sum_i p^2(i)$, where the sum is over the classes *i* at the node, and *p(i)* is the observed fraction of classes with class *i* that reach the node. A node with just one class (a pure node) has Gini index 0; otherwise the Gini index is positive. So the Gini index is a measure of node impurity (for more details, see also [16]-[17], or https://en.wikipedia.org/wiki/Diversity_index for a short view).

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Model	Performance					
	ACC	TP	FN	FP	TN	Tr. Time
SVM	0,5689	0,4025	0,1485	0,2826	0,1664	8024,0206
EM	0,5975	0,4902	0,0608	0,3417	0,1073	3353,7505
ANN	0,596	0,438	0,292	0,113	0,157	3s (20 epochs)
DT	0,5295	0,3775	0,1735	0,297	0,1521	246,0368

Table 6: Performances of the ML-based classifiers for each model we created, without the camera. ACC means "Accuracy", TP means "True Positive", FN means "False Negative", FP means "False Positive", TN means "True Negative" and "Tr. Time" means "Training Time". Finally, the first five performance indicators are in the proportion or fraction (parts per one), the last one in [s].

The Performance Indicator (PI) we have selected are illustrated in Table 2. Overall, the performances are not good, even if ANN and EM show better results (but the number of errors is still too big, with the risk to annoy the driver or provide wrong results that can lead to dangerous situations).

This means that the answer to the RQ1 is negative: it seems not possible to create a classifier of driver's visual distraction only based on vehicle dynamic and traffic data, since the performances are too poor to be used in a reliable way.

3.5.2 Results of the Driver distraction Classifier with DMS

Now, we use the same models, but considering also the DMS (namely, the internal camera). Thus, the RQ becomes: "is it possible to improve the performances of the classifier and at which extension"? A new training phase has been carried out and hereafter the main features of the models.

For SVMs, the best model had the following parameters and characteristics:

- Type of Kernel = Gaussian
- Scale = 0.0993
- Box Constraints = 35.063

For EMs, we have selected a model with:

- Method = AdaBoost
- Number of Learning Cycle = 476
- Learning Rate = 0.95799
- Minimum Leaf Size = 4

For ANN, we have obtained:

- training method = Scaled Conjugate Gradient Back-propagation
- number of layers = 2 layers topology has been chosen, with one Hidden Layer (HL – very rare the case in which more than 2 are needed; in our case, two did not provide appreciable improvement of results) and one Output Layer (OL)

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

- transfer function = a Sigmoid transfer function has been used for both the HL and OL.
- Hidden Neurons (number of neuron in the HL) = 50

Finally, for DTs, the best model was:

- Minimum Leaf Size = 10
- Maximum number of slits = 12
- Split criterion = Deviance

Comparing the performances, this is done in the following table:

Model	Performance					
	ACC	TP	FN	FP	TN	Tr. Time
SVM	0,8857	0,5761	0	0,1143	0,3095	1709,8496
EM	0,9666	0,5703	0,0058	0,0276	0,3963	1767,9638
ANN	0,722	0,441	0,137	0,141	0,281	2s (39 epochs)
DT	0,8141	0,5599	0,0162	0,1698	0,2541	199,3629

Table 7: Performances of the ML-based classifiers for each model we created, with the camera. ACC means "Accuracy", TP means "True Positive", FN means "False Negative", FP means "False Positive", TN means "True Negative" and "Tr. Time" means "Training Time". Finally, the first five performance indicators are in the proportion or fraction (parts per one), the last one in [s].

We have considered the same PI of the previous section. The results are much more interesting, especially for the Ensemble Method classifier which achieved a very good result overall. Also the trade-off between TN and TP is quite good (of course the dataset is originally much more unbalance towards not-distracted driver). The behaviour is good even for Support Vector Machine and Decision Trees (very similar), but the difference be-tween the percentage of TP and TN is bigger than the one of Ensemble Method. Only the Artificial Neural Network per-forms worst (at least with the architecture and topology we selected). Thus, the answer to the RQ2 is now positive: the performances are improved.

3.5.3 Comparison of Results

Finally, we have now to compare the results between the two driver monitoring systems tested and the classifiers we have developed. This is done using the following table:

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Model	Performance				
	ACC	TP	FN	FP	TN
EM	0,9666	0,5703	0,0058	0,0276	0,3963
DMS1	0,76	0,92	0,34	0,08	0,6
DMS2	0,7925	0,935	0,65	0,065	0,35

Table 8: Performances of the ML-based classifiers for each model we created, in comparison with two Driver Monitoring Systems we tested (named DMS1 and DMS2). ACC means "Accuracy", TP means "True Positive", FN means "False Negative", FP means "False Positive" and TN means "True Negative". Finally, the first five performance indicators are in the proportion or fraction (parts per one), the last one in [s].

The performance of the model given by the combination of ML and DMS outperforms the results obtained using only the DMS applications (named DMS1 and DMS2[4]), in terms of accuracy. Moreover, the balance between TP and TN percentage is much better and equilibrated. So, the final answer to our RQs is that a combination of DMS info and classifiers based on ML techniques can improve the overall result (with benefits in terms of system reliability and user acceptability, in order to reduce vehicle accidents and improve transportation safety).

In summary, we can say that the performance of the DMS are good enough to be used inside the project, in order to detect status of the driver and thus develop a more effective and efficient HMI. The system named as DMS2 has been therefore selected and its purchase is on-going for the scope of the project.

In addition, other conclusions are possible. The "virtual on-board user monitoring" (without the DMS) it not usable due to the very poor performances, while the "enhanced on-board user monitoring" (with the DMS) can provide a good improvement in driver distraction classification, for those period when the driver is involved in the driving task

[4] In order to avoid mentioning the names of the suppliers, providing the driver monitoring systems, we decide to name them anonymously.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

4. Conclusions

This deliverable has described the activities carried out in the interACT on the sensor data fusion and driver monitoring. We started from the scenarios, use cases and requirements selected in the WP1 of this project, where – in particular – three main sources of requirements are given. The requirements on the sensing layer are derived by means of the target scenarios (described in deliverable 1.1), the partner's requirements and system architecture (described in deliverable D1.2) and, finally, the basic requirements due to legal aspects (special context in a public funded project and ethical constraints, described in deliverable D9.1).

Based on that, we provided a direct association of the requirement to the different technical tasks, in order to make easier to derive an appropriate sensors setup. We focused specifically on object detection and traffic participants tracking (both static and dynamic), as well as pedestrians intention features recognition (all these topics are described in Section 2).

Besides this part, we have also described the work done for the on-board user's monitoring system, which will be installed on the CRF prototype vehicle, to provide valuable inputs to the HMI (developed inside WP4). First of all, we evaluated two Driver Monitoring Systems (DMSs) from two different suppliers, in order to assess their performances and choose the one with the better accuracy for our project. To do so, we performed a dedicated experimental phase, where a number of subjects has been "distracted" and the ground-truth data collected, for assessment purposes. Furthermore, using these data, we have also investigated two other possibilities:

- The development of a "virtual" on-board user's monitoring system, based on ML techniques, using <u>only</u> behavioral and dynamic data.
- The development of an "enhanced" on-board user's monitoring system, based on ML techniques, using <u>also</u> the DMS camera (as well as the behavioral and dynamic data).

More details are available in Section 3. The system named as DMS2 has been therefore selected and its purchase is on going for the scope of the project.

The "commercial" on-board user's monitoring system will be installed on at least one car (CRF vehicle), in order to detect the status of the driver and thus develop a more effective and efficient HMI. In particular, we use the driver monitoring data to adapt the on-board HMI strategy. Different amount of information will be displayed, depending on the driver distraction level. The ultimate goal is that the driver feels safe and well informed about the next actions of the automated vehicles and does not feel the need to intervene in situations that are well handled by the automated vehicle.

The next steps inside WP5 (task T5.1) are the complete integration of the perception platform on the prototype vehicles and the data collection, in order to perform the validation of the sensor data fusion. In addition to that, the on-board driver monitoring system will be installed and data will be available for WP4 (HMI) and possibly WP3 (for modules such as "Interaction Planning" and, possibly, "Safety Layer").

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

5. References

[1] K. Bengler. Three decades of driver assistance systems: Review and future perspectives. IEEE Intelligent Transportation Systems Magazine, (winter), 2014

[2] Treat J., et al.; Tri-level Study of the Causes of Traffic Accidents: Final Report, volume 1. Technical Report Federal Highway Administration, US DOT (1979)

[4] AAA Foundation for Traffic Safety. 2016 Traffic Safety Culture Index. Washington, DC: AAA Foundation for Traffic Safety, 2017. [Online] Available: https://www.aaafoundation.org/sites/default/files/2016TrafficSafetyCultureIndexReportandCover_0.pdf

[5] US Department of Transportation - National Highway Traffic Safety Administration. 'Traffic safety facts'. Available from: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812318?-ga=1.78055380.1104132544.1489526594

[6] US Department of Transportation - National Highway Traffic Safety Administration. 'Distracted driving'. Available from: https://www.nhtsa.gov/riskydriving/distracted-driving

[7] ust, M.A., Keller, T.A., Cynkar, J.: 'A decrease in brain activation associated with driving when listening to someone speak', Brain research, 2008, 1205, pp. 70–80

[8] Luke Ameen. 'The 25 scariest texting and driving accident statistics'. Available from: http://www.icebike.org/texting-and-driving

[9] T. Ranney. Driver distraction: A review of the current state-of- knowledge. National Highway Traffic Safety Administration, 2008, Tech. Rep. Paper No. 05-0400, 15, 2008.

[10] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama. Driver inattention monitoring system for intelligent vehicles: A review. In 2009 IEEE Intelligent Vehicles Symposium, pages 875-880, June 2009

[11] M. A. Regan and C. Hallett and C. P. Gordon, "Driver distraction and driver inattention: Definition, relationship and taxonomy", Accident Analysis & Prevention, vol. 43, no. 5, 2011.

[12] M. Westin, R. Dougherty, C. Depcik, A. Hausmann, and C. Sprouse III, "Development of an adaptive-human-machine-interface to minimize driver distraction and workload", in Proc. in the ASME 2013 Int. Mechan. Engin. Cong. and Expos., San Diego, California, USA, 2013, pp. 1 -13.

[13] S. Arun, K. Sundaraj, and M. Murugappan, "Driver inattention detection method: A review," in 2012 IEEE Cong. On Sust. Utiliz. And Develop. in Engin. and Technol. (STUDENT), Kuala Lumpur, Malaysia, 2013, pp. 1 -6.

[14] Esurance: '3 types of distracted driving', 2016, https://www.esurance.com/info/car/3-types-of-distracted-driving, accessed October 2017.

[15] Liang Y. and Reyes M.L. and Lee J.D.: "Real-time Detection of Driver Cognitive Distraction Using Support Vector Machines". In Proc. IEEE Transactions on Intelligent Transportation Systems, Vol. 8, No. 2 – June 2007.

[16] Jost, L. "Entropy and diversity". Oikos. 113: 363–375 - 2006.

[17] Hurlbert, S.H. "The nonconcept of species diversity: A critique and alternative parameters". Ecology. 52: 577–586.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

For more information:

interACT Project Coordinator

Anna Schieben

DEUTSCHES ZENTRUM FUER LUFT - UND RAUMFAHRT e.V. (DLR)

Lilienthalplatz 7

38108 Braunschweig, Germany

Anna.Schieben@dlr.de

interact-roadautomation.eu/

interACT

Designing cooperative interaction of automated vehicles with other road users in mixed traffic environments