D3.3 - Design methodology for trustable sensing

Due date of deliverable: 30/05/2024

Responsible partner: SUPSI

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

Document Details

Title	Title
Deliverable No.	3
Work Package	3
Task	T3.3
Deliverable Type	Report
Lead Partner	SUPSI
Contributing Partner(s)	KCL, DFKI
Due date of deliverable	30/05/2024
Actual submission date	30/05/2024
Abstract	This document describes the smart sensing subsystem for a reference case-studies with the aim of providing trusted event detection. The results of a simpler fusion approach applied to the described scenario are also provided.
Keywords	Sensor fusion, Probabilistic modelling, Risk evaluation support, Performance evaluation.

Dissemination level:		
PU	Public	X
RE	Restricted to a group specified by the consortium (including Commission Services)	
CO	Confidential, only for members of the consortium (including Commission Services)	

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

Change History

Version	Date	Changed by	Changes Made
1.0	22/05/2024	SUPSI	First version sent to the partners and the reviewers, although the 'Conclusions' section needs to be completed.
1.1	23/05/2024	DFKI	Provided comments to clarify and improve the content of the document.
1.2	24/05/2024	V-RESEARCH	Provided first review
1.3	27/05/2024	USE	Provided second review
1.4	29/05/2024	SUPSI	Changes requested by DFKI, V-RESEARCH and USE have been completed.
1.5	29/05/2024	KCL	Comments provided to clarify content and fix typographical aspects.
1.6	30/05/2024	SUPSI	Changes requested by KCL have been completed. The deliverable is ready for final submission.

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

Table of Contents

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

List of Figures

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

List of Tables

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

Acronyms Table

Acronym	Expanded Form
AI	Artificial Intelligence
AS	Autonomous System
AW	Autonomous Wheelchair
BN	Bayesian Network
DF	Danger Function
LiDAR	Light Detection And Ranging
ML	Machine Learning
REXASI-PRO	REliable & eXplainable Swarm Intelligence for People with Reduced mObility
RM	RoboMaster EP
RMSE	Root Mean Square Error
ROS	Robot Operating System
RSU	Range Sensors Unit
TAS	Trustworthy Autonomous Systems
YOLO	You Only Look Once

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

1. Overview of the deliverable

1.1. Scope

In this deliverable we describe on a theoretical level an approach for a sensor fusion where uncertainties related to sensors and environmental factors are modeled with a Bayesian Network. As a case of study, a road-crossing scenario is presented and simulated in our laboratory to record a novel dataset. As indicated at the end of the document, in our future work we plan to use this dataset for development and verification of our approach. The results will be presented in deliverable D4.5.

1.2. Audience

This document is intended for everyone who is interested in gaining an understanding of the approach we intend to adopt for safer sensor fusion against uncertainty and in having a description of the scenario adopted. Initial results can also be visualised to identify the best level to apply sensor fusion on the basis of a simpler fusion.

1.3. Summary

We address the problem of multi-sensor fusion safe road-crossing by autonomous wheelchairs supported by flying drones. We have focused on the generation of a laboratory dataset using more simple devices to simulate such scenario. We also designed an analytical danger function to enable run-time risk assessment for road-crossing decision support. As a first approach, a simple mean average for sensor fusion has been applied on the generated dataset at different levels of data processing. As a result, we achieved better performance for the fusion applied at the least elaborate data level. We also presented the approach for a more sophisticated sensor fusion where environmental conditions and sensors uncertainties are modeled through a Bayesian Network's architecture. Finally, future steps are provided.

1.4. Structure

Sec. 2 gives an introduction to the approach adopted for sensor fusion and the case of study. Sec. 3 introduces the reference road-crossing scenario and the design of an analytical danger function for danger evaluation. Sec. 4 addresses dataset generation in the lab environment. Sec. 5 describes the two approaches adopted for fusion. Finally, Sec. 6 provides conclusions and hints about future developments.

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

2. Introduction

In the last decades, *Autonomous Systems* (AS) have seen a growing development in several areas, including automotive [4], navigation, aerospace, industry [19], and military [11] applications. In many cases, those systems are aimed at carrying out operations that are impossible or critical to perform for human workers. Currently, AS are mostly applied in environments where uncertain events and disturbances are either absent or largely limited, and they are supervised to some extent by human operators.

Thanks to the recent technological achievements in *Artificial Intelligence* (AI) and robotics, AS have been improved to perform increasingly complicated tasks such as driving vehicles in complex, open and uncontrolled environments, even without human supervision. However, due to the possible criticality of those applications, new vital requirements have been introduced to set next research challenges. A new vocabulary has been recently introduced to address all the necessary aspects in the design and evaluation of those systems, not only from a technical perspective, but also in terms of ethical and legal implications, including fairness and accountability. The "Ethics Guidelines for Trustworthy Artificial Intelligence" [1], presented by the High-Level Expert Group on AI set up by the European Commission, states that trustworthy AI should be:

i. lawful, to ensure that all laws and regulations are applied and respected;
ii. ethical, to adhere to moral principles and values;
iii. robust, to avoid any unintended damage and safety issues.

Trustworthy AI is fundamental for the notion of *Trustworthy Autonomous System* (TAS). TAS must be technically robust and therefore they must be evaluated according to the attributes and means of dependable, secure and resilient computing, as defined in the seminal papers [2] and [16]. More recently, in reference [7], the notions of dependability, resilience, and cyber-security have been connected as core concepts within a comprehensive TAS taxonomy.

TAS must be robust to uncertainties in the surrounding environment, secure against threats coming from cyber-space, and capable of safe human-machine interactions [12]. As addressed in existing work by Mousavi et al. [18], trustworthiness must be put in relation not only to purely technical aspects, but also to the benefits and wellness of people as a consequence of AS actions.

All the aforementioned aspects are extremely important when addressing real-world adoption of novel technologies leveraging on *Artificial Intelligence* (AI) and *Machine Learning* (ML), whose failure can have severe consequences on human health. This is the case of *Autonomous Wheelchairs* (AWs) that are meant to support motion-impaired persons in safe door-to-door navigation.

From the invention of the first powered wheelchairs in 1956, there have been several efforts to improve their technology. The realisation of AWs is a challenge that aims to give more independence to patients suffering from any kind of disability by means of multi-modal interfaces. On-board navigation systems, human-machine voice interfaces and more sophisticated components based on AI could lead to a significant improvement in the patient's quality of life. One aspect to be carefully addressed is the safety assessment and possible certification of AWs.

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

In fact, in presence of AI, together with specific regulations such as "ISO 7176-14:2022 Wheelchairs",[1] other requirements and guidelines should be considered, such as those included in the EU Artificial Intelligence Act.[2] In case AWs also monitor biomedical parameters, further certifications might be required.

As for any complex systems, AWs can be divided in simpler components following a modular approach. Therefore, holistic safety assessment can also be based on a modular approach starting from single components (i.e. sensors) and up to subsystems (general environmental sensing) , and considering the interfaces between them. In the context of social robotic navigation, trustable environmental sensing is an essential aspect that is crucial to guarantee robustness against uncertainties, internal malfunctions, and external disturbances. Sensor systems include smart sensors elaborating raw data coming from environmental measurements and transforming it into useful information such as event/threat detection.

A smart-sensing subsystem should be used to provide trusted event detection by following a model-based approach where trustworthiness is enforced during the whole system life-cycle. Common causes of failures should also be mitigated by applying the principle of "no single point of failure" together with strategies that rely on technology diversity. To assess the trustworthiness of the system, a model-based evaluation is used, in which verification for the sensing subsystem is performed at both design-time and run-time with the aim to fulfil requirements related to *Safety Integrity Levels* (SIL).

In order to address the challenges related to quantitative safety assessment, we adopt a multi-agent, multi-modal and self-adaptive sensing system to achieve trusted event detection, where sensors outputs are combined to give a common result for the measured. To quantify the uncertainty associated with the sensors and the surrounding environment, *Bayesian Networks* (BN) are adopted to provide a representation of the system.

As a first approach for *sensor fusion*, a deterministic approach based on simple mean average has been investigated by applying it at different level of data processing to identify the best method (Subsec. 5.1). This is a first step towards the experimentation and performance evaluation of the new sensor fusion approaches that we present in Subsec. 5.2 and whose results will be presented in deliverable D4.5.

For our work, we have focused on a road-crossing scenario for the swarm system composed by an AW and a drone, using multiple sensors to estimate the level of danger and to support the decision making process for crossing. Such a scenario is quite straightforward to analyse: there is only one obstacle to detect (the approaching vehicle), and there should be no physical interaction between the AW and the obstacle. If this happens, an unwanted collision has occurred. Furthermore, simpler and cheaper devices can be used to simulate this scenario. In this regard, we have proposed as a proof-of-concept a simplified scenario using wheeled robots equipped with artificial vision and distance sensors as mock-up models of the actual system. We then perform off-line danger evaluation using an analytical danger function based on kinematics considerations. Although the proposed scenario is limited in its representation of the real system, it is useful for simulating and replicating tests with good control over the various components,

[1] https://www.iso.org/standard/72408.html
[2] https://artificialintelligenceact.eu

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

without involving too many participants in the experiments (as in the case of people detection). However, this choice does not preclude the use of our approach in other scenarios by modifying certain components (e.g. the model adopted for obstacle detection if ML models are used).

We have also described and shared with the scientific community a dataset recorded in our lab and reported the results of the preliminary experimental evaluation demonstrating the advantages of using diverse and redundant sensors [10].

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

3. Case of study

This section describes the use case scenario of road-crossing and the analytical danger function adopted for danger evaluation.

3.1. Reference scenario: Road crossing

When crossing a road, humans consider a variety of factors to ensure both safety and convenience. For instance, they assess the flow of traffic, considering variables such as the speed and quantity of vehicles. They also account for the presence and state of road-crossings, traffic signals, and zebra crossings, as these demarcate safe areas for traversing. The distance to the nearest crossing point and the estimated time required to reach it are taken into consideration, as are potential traffic delays that might offer a safer crossing opportunity. Among the many aspects that can be considered, pedestrians account for distance, speed and acceleration [24, 17] as pivotal cues in their assessment of approaching vehicles before determining whether it is safe to cross the road.

Distance.

The distance of an approaching vehicle is a central indicator on which pedestrians heavily rely. Distance perception is related to the personal experience of each single pedestrian.

Speed.

People gauge speed through visual cues, such as changes in the vehicle's apparent size and the interval it takes for the vehicle to traverse a specific distance between reference points on the road.

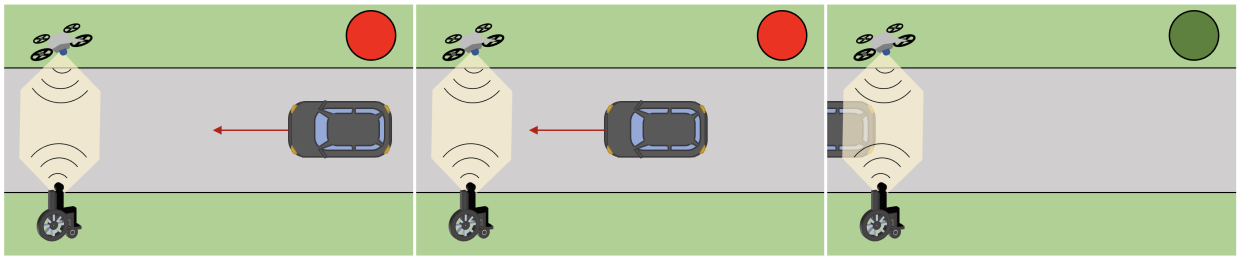
Acceleration.

Pedestrians also consider acceleration, denoting how swiftly a vehicle alters its speed over time, associating harsher/softer acceleration for cars/trucks. Acceleration contributes to the decision-making process by signaling how quickly a vehicle may reach the pedestrian's location.

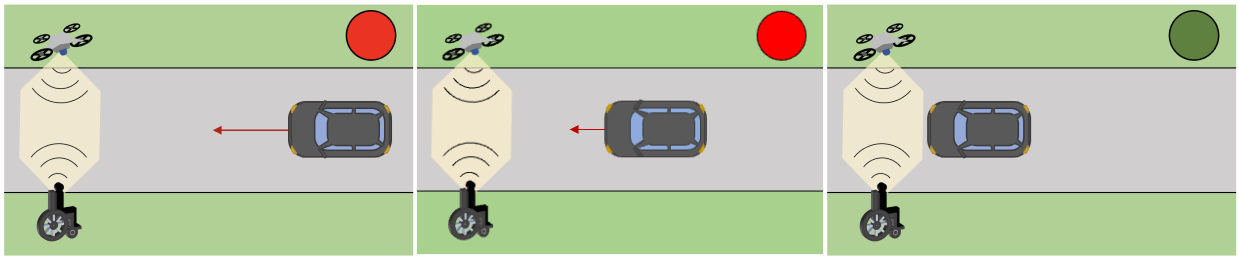
We apply a similar reasoning to the road-crossing scenario for AWs, where the system must take a safe road-crossing decision based solely on the kinematics of the obstacle. Following the principles of redundancy and diversity, a safe decision must be based on multiple sensors using different technologies whose accuracy is affected by mostly independent factors [8]. This can be achieved, for instance, by combining cameras featuring artificial vision with LiDARs or other distance sensors. Additional information can be obtained by connecting the AW with drones or specific infrastructures [14], providing information from diverse sources or different perspectives.

The road-crossing scenario we consider consists of an AW on the side of a one-way road aiming to reach the other side, an external observer (i.e. a flying drone as in REXASI-PRO or other camera system) providing further information from a third-person view, and a vehicle approaching the specific section of the road. We assume that there are no traffic lights and the driver's behaviour is not predictable;

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**



(a)



(b)

Figure 1: *Road-crossing scenario: drone and AW cooperate to support safe decisions, e.g., when a car surpasses the crossing location (a), or when a car stops before the crossing location (b).*

the vehicle might slow down to allow the crossing or move ahead, maintaining its motion. Hence, the AW must use its sensing capabilities to estimate the potential danger: the crossing is safe if the danger is estimated under a certain threshold. If the danger is too high, crossing is not recommended until the vehicle slows down or surpasses the crossing location. This fundamental notion of danger is described in the next section.

Fig. 1 provides a schematic representation of the scenario.

3.2. Risk evaluation: Danger Function design

In this section, we address the problem of assessing the risk associated with road-crossing. We consider the road-crossing scenario described in the previous Sub-sec. 3.1: the AW attempts to cross the road, the drone provides further information to the AW and an obstacle (e.g., a car) is approaching. As shown in Fig. 2, the AW and the car are assumed to have orthogonal motions, with the pedestrian crossing the road from one side to the other while the obstacle proceeds along the driveway.

In our approach, the risk evaluation is based on a continuous function g, called *Danger Function* (DF). The function is intended to reflect a kinematic analysis of the road-crossing scenario in a particular moment t. Accordingly, the value of the DF reflects the danger level associated with the kinematic configuration of the system at a specific moment. For decision-making, we summarise the assessments provided by the DF g by tagging them as potentially dangerous if the corresponding DF values exceed a given threshold, g^*.

To derive the DF, we assume that the AW can proceed with uniform linear motion and that the obstacle approaches it with a uniformly accelerated motion. Whilst both assumptions are often unrealistic, they might reflect an over-cautious

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

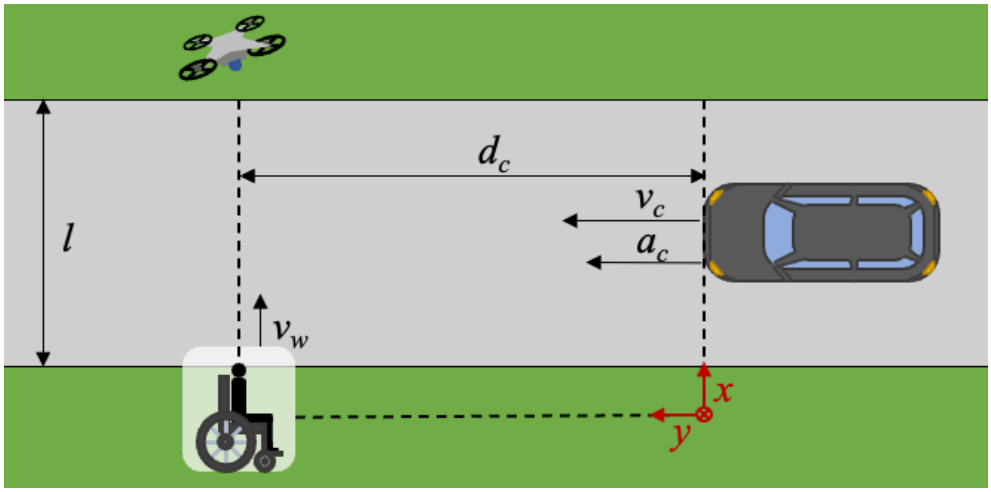


Figure 2: *Kinematics involved in the DF design.*

modelling providing a safer evaluation: in real scenarios, in case of danger, the vehicle would likely slow down and the AW would likely accelerate.

We consequently model the road-crossing scenario as follows. Assuming the AW crossing the road to move on the x-axis and starting the crossing from position $x = 0$ with constant speed v_w, we have $x_w(t) = v_w \cdot t$.

Similarly, assuming the obstacle, i.e., the car, moving on the y-axis with constant acceleration a_c and initial speed v_c: $y_c(t) = v_c \cdot t + \frac{a_c}{2} \cdot t^2$.

Then, we can compute the time t_{cross} required by the AW to cross a road of width l, i.e., $t_{\text{cross}} := l/v_w$. During the interval of time $[0, t_{\text{cross}}]$, the obstacle should be far enough to not impact with the AW, i.e., if d_c is the distance of the obstacle:

$$d_c > v_c \cdot t_{\text{cross}} + \frac{a_c}{2} \cdot t_{\text{cross}}^2 \,, \tag{1}$$

and hence:

$$\frac{\frac{l}{v_w} \cdot v_c + \frac{l^2}{2 \cdot v_w^2} \cdot a_c}{d_c} < 1 \,. \tag{2}$$

We can regard the constraint in Eq. (2) as a safety condition for a kinematics-based DF with threshold one. As expected, the danger increases for higher values of speed and acceleration of the car, while decreasing for increasing distances. Function g can now be determined on the relation between d_c, v_c and a_c as in Eq. (2).

Regarding the inverse dependence with respect to d_c, we assume that close objects are significantly more dangerous than distant objects, even if they are moving at lower speed. Consequently, to enhance the impact of the distance on the function in specific situations where the two bodies are close, a logarithm smoothing is applied to d_c. Moreover, a small threshold ϵ is also added to always obtain positive values.

Concerning the linear combination of speed and acceleration in the numerator of the left-hand side of Eq. (2), we perform linear transformations with thresholds on both v_c and a_c, and denote the output of these transformations as $\hat{v}_c$ and $\hat{a}_c$ and use a coefficients k_i to evaluate the relative contribution of the speed with respect to the acceleration. Therefore the DF and the safety condition are described as:

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

$$g(d_c, v_c, a_c) := \frac{k_v \cdot \hat{v}_c + k_a \cdot \hat{a}_c}{\log(d_c + \epsilon)}, \qquad g(d_c, v_c, a_c) < 1, \tag{3}$$

where $k_v = 1$ s, $k_a = 0.1\ s^2$ and $\epsilon = 0.6\ m$ are set following heuristics which are valid for the several tests we carried out (see Sec. 4). ϵ is chosen to ensure that the relationship ${d_c}^{min} + \epsilon \geq 1$ is always met, where ${d_c}^{min}$ is the minimum possible distance given the dimension of the devices playing the role of the AW and of the obstacle. In our tests we approximately estimated that $v_w = 0.5m/s$ and $l = 0.5m$ and, therefore, $k_v = 1s$. Regarding k_a, after some trial and error we determined that the contribution of the accelerations worsened the performance, so the k_a was reduced. Although such a choice is valid for the scene we used in our lab, the value l/v_w and ϵ might be chosen using realistic crossing time for pedestrian and more realistic measurements. Also the threshold might be changed with respect to specific scenarios or situations (e.g., bigger obstacle, wet road surface, etc...).

The piecewise linear transformations of speed and acceleration are:

$$\hat{v}_c := \begin{cases} 0 & \text{if } v_c \leq \underline{v}_c\,, \\ \frac{v_c - \underline{v}_c}{\overline{v}_c - \underline{v}_c} & \text{if } \underline{v}_c < v_c \leq \overline{v}_c\,, \\ 1 & \text{if } v_c > \overline{v}_c\,, \end{cases} \qquad \hat{a}_c := \begin{cases} -1 & \text{if } a_c \leq -\overline{a}_c \\ \frac{a_c + \underline{a}_c}{\overline{a}_c - \underline{a}_c} & \text{if } -\overline{a}_c < a_c \leq -\underline{a}_c \\ 0 & \text{if } -\underline{a}_c < a_c \leq \underline{a}_c \\ \frac{a_c - \underline{a}}{\overline{a}_c - \underline{a}_c} & \text{if } \underline{a}_c < a_c \leq \overline{a}_c \\ 1 & \text{if } a_c > \overline{a}_c\,. \end{cases} \tag{4}$$

The above transformations are intended to prevent contributions to the DF by low speeds (i.e., $v_c \leq \underline{v}_c$) and low accelerations (i.e., $|a_c| \leq \underline{a}_c$), while also putting a normalised upper bound to the contributions of high speeds ($v_c \geq \overline{v}_c$) and accelerations ($|a_c| \geq \overline{a}_c$). These values are chosen based on the maximum speeds and accelerations that the obstacle could actually reach (i.e., the devices that we will describe in Sec. 4).

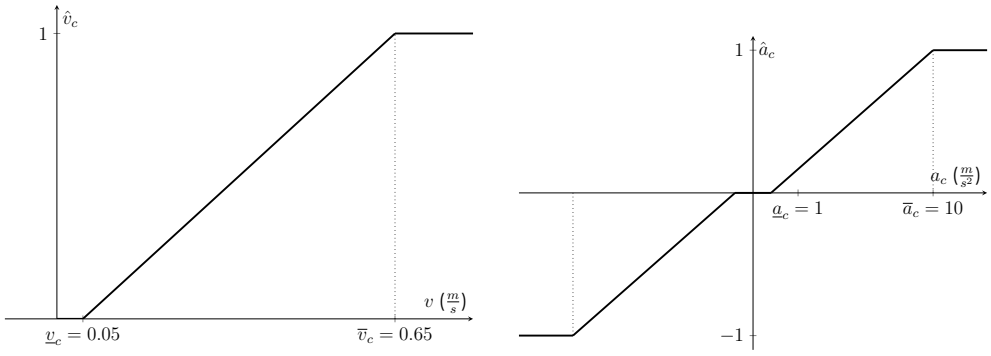


Figure 3: *Linear transformations with thresholds used for v_c (left) and a_c (right).*

On the other hand, lower limit values are included to filter out possible oscillations around zero. In addition, negative speed values are excluded as the car is never expected to go backwards, and the contribution of accelerations has been significantly reduced. In our experience, these arrangements have improved the results of the danger analysis in the case of simple fusion, which will be described in Sect 5.1. In particular, the contribution of the accelerations resulted to be quite noisy when obtained from certain types of sensors, so we significantly limited it.

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

This is the case with cameras using ML models for obstacle detection, which we will adopt in Sec. 4

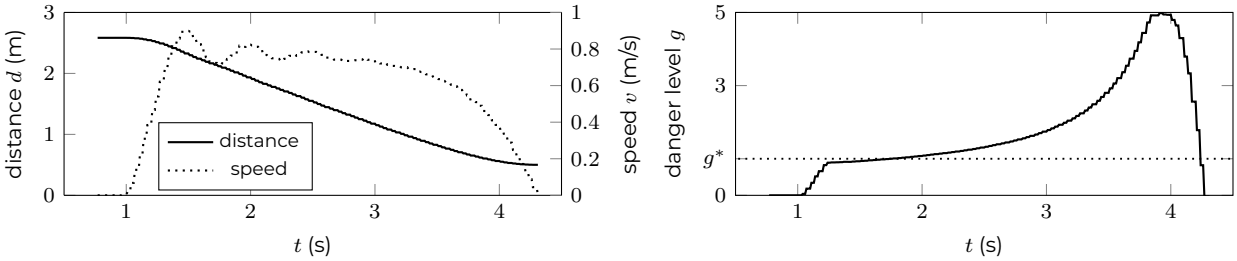


Figure 4: *Distance, speed (left) and DF values (right) on the experiment in Sec. 4.*

Fig. 3 shows the choice of the threshold parameters $\underline{v}_c$, $\overline{v}_c$, $\underline{a}_c$, and $\overline{a}_c$, for the experimental setup discussed in Sec. 4.

Fig. 4 shows an example of the kinematics (for clarity, only distance and speed on the left) and the corresponding values of the DF (right).

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

4. Dataset generation

In this section, we provide the specification of the laboratory environment — including all the equipment and hardware components — and describe the procedure of data collection, pre-processing and elaboration. In the experiments, we target a similar scenario to that described in Sec. 3.

We adopt a simplified experimental setup, where three ground robots equipped with vision and distance sensors represent the AW, drone and vehicle, as depicted in Fig. 5. In our lab, we can access ground truth poses from a high-precision motion tracking system.

4.1. Laboratory environment

The experiments have been performed in the *IDSIA Autonomous Robotics Laboratory*.[3]

Robots

For our experiments we use three wheeled omni-directional robots: each one is a *RoboMaster EP* (RM), a commercial education platform from DJI[4] that can be provided with a gimbal or a robotic arm. Their specifications are summarized in Tab. 3.

Each RM is customised for its role as "car", "AW", or "drone".

RM$_c$ simulates the vehicle and does not require additional sensors, as we are interested just on its movement. It should be noted that although the maximum speed is $3.5\,\mathrm{m/s}$, in practice we were able to apply a maximum speed of $1\,\mathrm{m/s}$.

RM$_w$ represents the AW and is equipped with a camera and four infrared range sensors. Given their narrow *Field Of View* (FOV) (see Tab. 3), range sensors are located on the RM so that their outputs can be combined to obtain information for a larger FOV. The RM$_w$ and its sensors are shown in Fig. 6(b). Conservatively, we consider the minimum distance reading returned by the four range sensors, which from now on we refer to as the single sensor *Range Sensors Unit* (RSU).

RM$_d$ has a robotic arm and acts as the drone. It has a camera positioned on top of the elbow of the robotic arm. It should be noted that the end effector of the robotic arm cannot be higher than the elbow, so we preferred a position where the camera is centred in the RM. Fig. 6(a) show the RM$_c$.

Motion Tracker

The lab used for the experiments is equipped with an OptiTrack motion tracker[5] composed of eighteen infrared cameras covering an area of $6\,\mathrm{m} \times 6\,\mathrm{m} \times 2\,\mathrm{m}$ to track the pose of the three robots at $30\,\mathrm{Hz}$ with sub-millimetre accuracy.

[3] https://idsia-robotics.github.io
[4] https://www.dji.com/ch/robomaster-ep
[5] https://docs.optitrack.com/v/v2.3

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

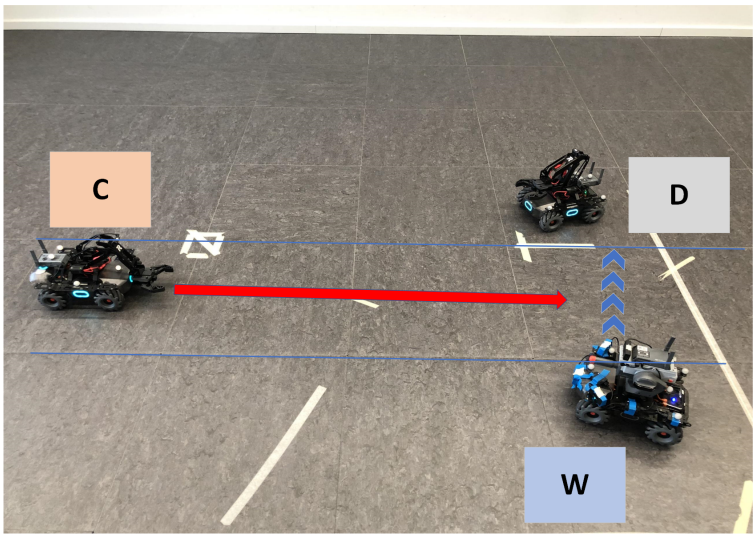
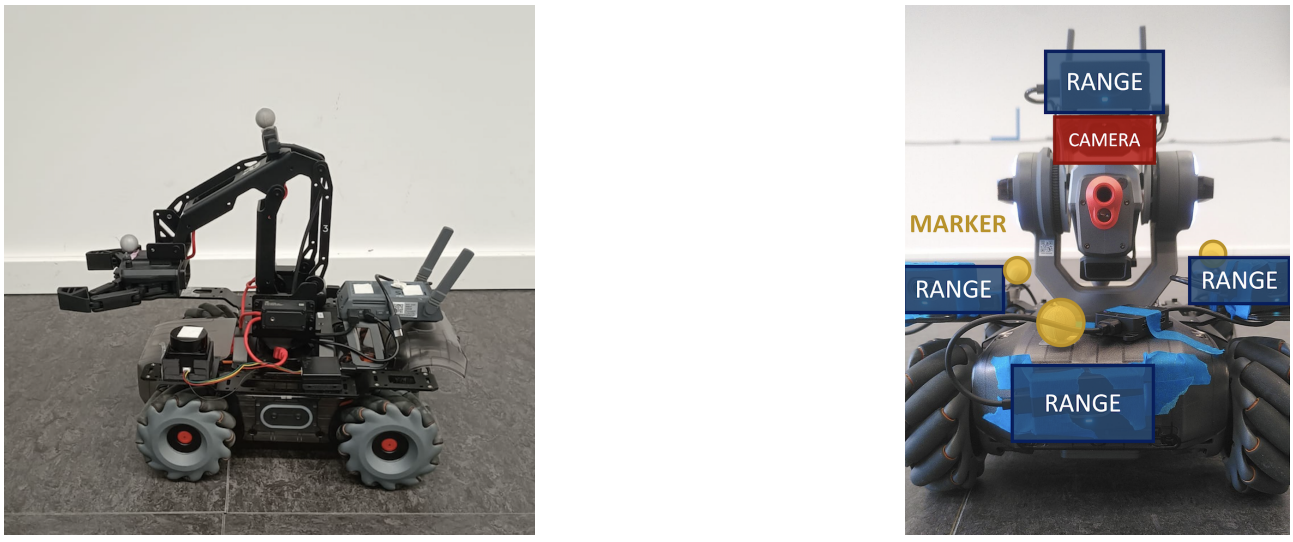


Figure 5: *Example of setup for data collection: in red is the trajectory of RM_c and in blue the crossing path of RM_w.*

Chassis	size	$32\,\text{cm} \times 24\,\text{cm} \times 27\,\text{cm}$
	maximal speed	$3.5\,\text{m/s}$
	maximal angular speed	$600\,°/\text{s}$
Camera	FOV	$120°$
	video resolution	1280×720
	video fps	$30\,\text{Hz}$
Range sensors	maximal range	$10\,\text{m}$
	FOV	$20°$
	accuracy	$5\,\%$

Table 3: *Technical specifications of RoboMaster EP.*



(a) **(b)**

Figure 6: *(a) RM_c with the camera on top of the robotic arm elbow. (b) RM_w and its components.*

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

4.2. Data collection and preprocessing

As in the scenario described in Sec. 3, RM_w and RM_d are motionless and oriented towards the approaching RM_c, which is remotely controlled.

Data collection.

To control the robots and collect data from the motion tracker, we used the *Robot Operating System* (ROS2) [20], executing a ROS2 driver for each robot.[6]

For each experimental run, we recorded data from the two cameras, the RSU and the motion tracker in bag files as detailed in Tab. 4.

Overall, we recorded 15 runs for two different setups (the first one is depicted in Fig. 5), for which runs last approximately $6\,\mathrm{s}$ and resp. $9\,\mathrm{s}$. In the second setup, the initial distance of the RM_c with respect to the RM_w is larger, allowing for more complex patterns in movement of the obstacle during experiments.

Name	Source	Type	Frequency
Poses of RM_w, RM_d and RM_c	Motion Tracker	2D poses	$30\,\mathrm{Hz}$
Video	RM_w	1080×720 RGB images	$30\,\mathrm{Hz}$
Video	RM_d	1080×720 RGB images	$30\,\mathrm{Hz}$
Range sensors	RM_w	4 distances to nearest object	$10\,\mathrm{Hz}$

Table 4: *Raw data recorded during the experiments.*

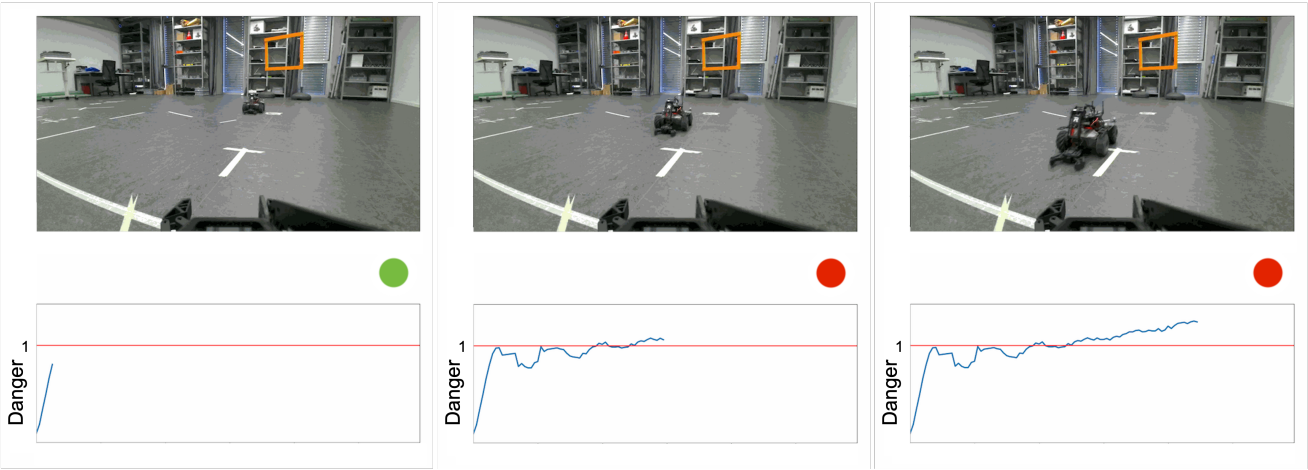


Figure 7: *Example results: RM_w camera frames during safe crossing (left) and dangerous crossing (center and right) and corresponding DF from the tracker.*

Data pre-processing.

Raw data from the bag files were processed to synchronise all streams. Relative distances between RM_w and RM_c were computed for ground truth poses and from RSU.

Raw data from cameras were split using OpenCV[7] into frames. These were

[6] https://github.com/jeguzzi/robomaster_ros
[7] https://docs.opencv.org/4.x/index.html

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

processed with *You Only Look Once* (YOLO)[8] [21], a commonly used ML model used for object detection, to detect RM_c. RMs do not have a specific class in YOLO, but are consistently recognised as *motorcycles*. YOLO was used to identify their bounding boxes in image space, from which we computed the distance using triangle similarity.[9]

The relative distance between RM_w and RM_c is chosen as comparison quantity. The Optitrack records are easily elaborated starting from the poses of RM_w and RM_c. RM_d is positioned symmetrically to RM_w with respect to the road, so its distance from RM_c can be taken directly by removing a limited bias. The range sensors raw data on RM_w already give the desired measure so their information is not futher processed.

Finally, the three distance measures from the sensors onboard, i.e. the RSU and the two cameras, were smoothed, taking an average of the previous 2 (RSU) or 5 (cameras) samples . Speed and acceleration measures were computed from the four distance measures obtained from the RSU, the two cameras and the tracker. Then, danger was estimated for each sensor and the tracker using the DF defined in Subsec. 3.2. When the condition of Eq. 3 was not met, a dangerous situation was detected and crossing was not recommended. An example of the results is shown in Fig. 7: three frames from the RM_w camera output are displayed with the corresponding DF value calculated by the tracker. A green (resp. red) circle defines a safe (resp. dangerous) situation as a result of the threshold criterion.

[8] https://github.com/ultralytics/yolov5
[9] https://pyimagesearch.com/2015/01/19/find-distance-camera-objectmarker-using-python-opencv/

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

5. Methodology for sensor fusion

5.1. First deterministic model

As first approach for *sensor fusion*, a deterministic approach based on simple average mean is applied at different level of data processing to the three values obtained from the RSU and the two cameras. This is a first step towards the experimentation and performance evaluation of new sensor fusion approaches based on the general ideas we introduced in [5] and that we describe in Subsec. 5.2. To deal with different sensor sampling times, all measures were resampled at 100 Hz before merging. We considered three different fusion architectures. A straightforward fusion procedure consists of taking the average of the distances measured by the three sensors. We call this procedure *distance fusion*. We call instead *danger fusion* the fusion procedure where the average of the DFs computed by the distances measured by the three sensors is considered. Finally, *voting fusion* was applied to the values obtained after the threshold operation on the DFs through a majority vote. All the data collected are available, together with the code used for processing and analysis, in a freely available repository.[10]

Fig. 8 depicts the comparison between the three pre-processed distance measures with the "ground truth" from the motion tracker (left), and the DF based on the *distance fusion* (right) for a particular run. From Fig. 8 (left), it can be observed that the RSU output is always available, while cameras generate signals only when RM_c is already near to the other RMs. When the output of the cameras is not available, the *distance fusion* is only based on the RSU. This is due to the computer vision system, which is unable to adequately detect the RM_c since it is confused with the lab background when it is far from cameras.

Similar considerations apply to the *danger fusion*. The DF values computed by the three sensors are displayed and compared with the tracker data in Fig. 9 (left). Those values are very noisy and unstable, while the fusion offers a smoother output that is more suitable to danger evaluation, as depicted by Fig. 9 (right).

For an evaluation, we compute a RMSE by comparing the DFs based on the tracker, with those obtained from the single sensors, and with the *distance* and *danger fusion* outputs. To evaluate the performances of the single cameras we decided to impute as default constant value a reference maximum distance ($6\ m$) when their outputs are not available. Such a descriptor cannot be computed for the *voting fusion*, which returns a binary output. When coping with such outputs, we quantify the classification performance of the different methods in terms of accuracy, precision, recall, and F1 score. To interpret those measures, note that a *false positive* is a case where a safe output is recognised as a dangerous situation.

The overall results are displayed in Tab. 5. Between all sensors, the RSU provides the most accurate road-crossing predictions in terms of accuracy, recall, precision and F1-score. This is not the case for the RMSE. We believe that this is related to the over-cautious distance measures provided by the RSU, which are not eventually affecting the road-crossing decision. However, the cameras show good performance when evaluated solely on the time intervals in which they see the obstacle. This results in a positive contribution to the fusion approach: the best fusion algorithm provide better performance than the better sensor accord-

[10] https://github.com/CarloGrigioni/safe_roadcrossing_aw

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

ing to all the performance descriptors. This clearly shows the advantages of the fusion approach. In particular, the best performance are typically achieve by the *distance fusion*, with the only exception of the recall, for which the *danger fusion* is better. Note that reaching high recall levels is especially important for our application, as we want to keep as low as possible the number of false negatives, corresponding to a crossing decision in a dangerous situation.

Sensor/Fusion	RMSE	Accuracy	Recall	Precision	F1-score
Drone camera	1.34	58.16	40.17	65.50	44.07
AW camera	1.04	68.55	52.40	71.57	55.31
RSU	2.33	76.43	68.50	76.28	68.55
Distance fusion	**0.99**	**77.25**	68.17	**76.93**	**69.59**
Danger fusion	1.09	75.16	**73.84**	72.79	69.46
Voting fusion	-	68.98	56.13	76.69	59.82

Table 5: *Mean of the performance evaluated over all tests of sensors and fusion techniques.*

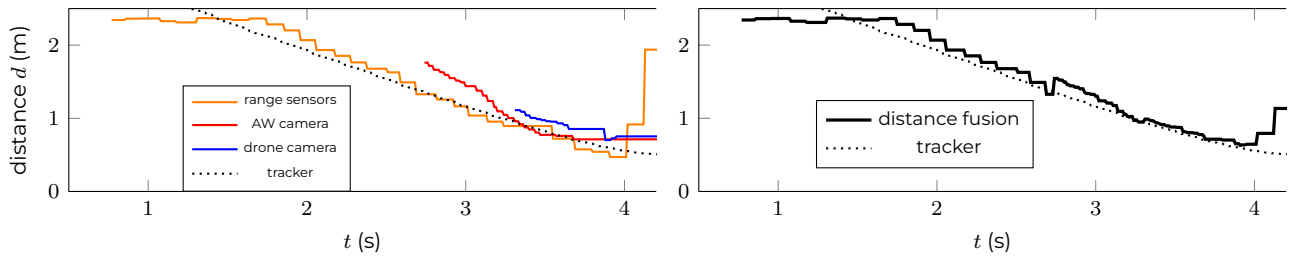


Figure 8: *Distance recorded by different sensors (left), by the ground truth (tracker) and by* distance fusion *(right).*

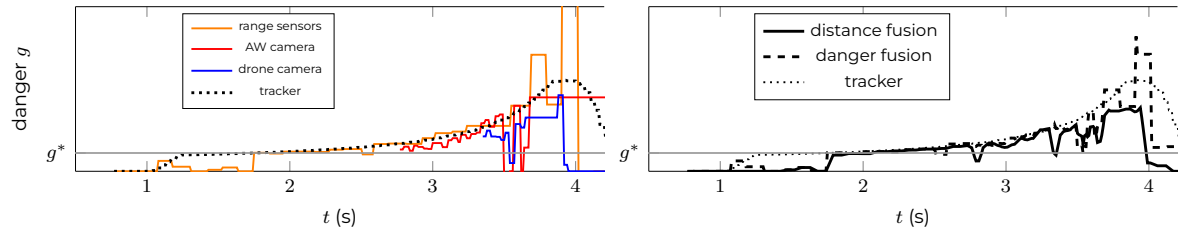


Figure 9: *DF computed from distances recorded by the sensors (left) and as obtained by* danger *and* distance fusion *(right).*

5.2. Bayesian Model

In the case of event detection, a possible approach is based on *voting*, where the output is based on the agreement of most detectors. By analysing and tracking outputs of sensors and their detection performance over time, it is possible to score their reputation, weight their contribution accordingly, and even exclude those that are no more considered *reputable*, i.e., those that could then negatively affect the outcome of the decisions. Through appropriate reconfiguration, the sensing system can consider a subset of detectors to keep the required safety level. In other words, the sensing system is able to self-adapt when internal faults

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

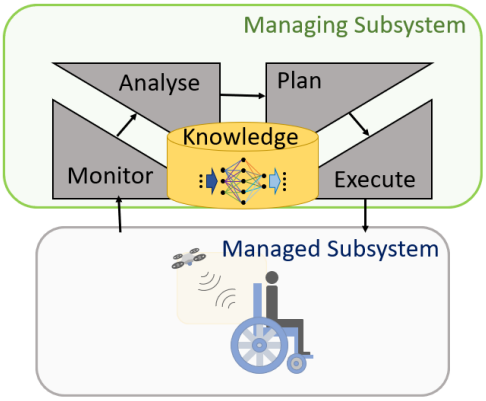


Figure 10: *Self-adaptation through MAPE-K feedback loop for safety monitoring of wheelchair-drone system.*

occur or when an exogenous environmental condition affect detection performance.

The ability of self-adaptation is achieved by combining a "Managed Subsystem", which is the system under consideration (i.e., the sensing one), with a "Managing Subsystem" based on the *Monitor, Analyse, Plan and Execute* over a shared *Knowledge* (MAPE-K) feedback loop (see Fig. 10). In AWs, the managing part implements the autonomic safety logic over the monitored sensing subsystem within the overall wheelchair-drone system. The managed subsystem is monitored along with the environment, and related data is stored in the knowledge base, which includes all relevant models representing the system from the safety perspective. Data is continuously collected and analysed to check if pre-defined safety conditions are met and no anomalies are detected. If needed, reconfiguration actions are planned at run-time to keep the system operation safe enough, including exclusion of faulty sensors, issue of alerts, application of speed limitations, switch to manual/remote control, or even fail-safe system stop.

Considering the multi-agent structure of the system, sensors characterised by different technologies are used, affected by different types of internal or external faults. The assumption about the *diversity* of sensing technology/mechanism is essential to exclude correlations between them and common-mode faults. Sensor diversity can be also achieved through different algorithms, parameters choice, and sensor displacement [6].

This multi-sensor and multi-modal approach implies enough *redundancy* to evaluate the information we are interested in. It allows to increase system robustness against the malfunction of some of its components and, to some extent, to reduce the costs by using cheaper components. Moreover, technology redundancy and diversity is a necessary feature to improve resilience against environmental disturbances. As highlighted in reference [16], "*Diversity* should be taken advantage of in order to prevent vulnerabilities to become single points of failure". More recently, the importance of diversity in machine learning systems has been highlighted to comply with functional safety requirements [3].

TAS operating in real-world environments must cope with several uncertainties such as unpredicted changes, disturbances, and the so-called "unknown unknowns". Properties such as self-adaptation allow to deal with those uncertain-

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

ties, however they also limit the possibility of employing deterministic verification approaches. One possibility to cope with uncertainties is to adopt probabilistic approaches possibly based on graphical models. BNs [15] can be used due to their suitability to represent complex causal relationships between system components, and to visually describe inter-dependencies in an easily interpretable way. The assumption about diversity of sensing technology is essential to exclude correlations between sensors and common-mode faults. BN extensions such as *Dynamic Bayesians Networks* [22] and *Time-Varying Bayesians Networks* [23] are also useful to manage time-varying and dynamic aspects of the system [9].

For the specific AW case-study, a smart-sensing subsystem is used to provide trusted event detection by following a model-based approach where trustworthiness is enforced during the whole system life-cycle. Common causes of failures are mitigated by applying the principle of "no single point of failure" together with strategies that rely on technology diversity.

In the specific case-study of vehicle detection, the presence or absence of an automobile on a specific section of the road is detected through sensors using different technologies, e.g. magnetometer sensor, video image processor or radar sensors. Unforeseen elements could interfere with the analysed scene and, depending on their nature, induce errors on one or multiple sensors.

In the described context, let X denote the actual, non-observable, value to be measured, i.e. the random variable representing the event "vehicle presence or absence". As we assume the sensors to be only partially trustable, the observation of X as returned by a sensor S_i is described by a distinct, observable, variable $O_X^{S_i}$ with the same possible values of X. Let us denote as E_i the exogenous factors possibly inducing a deterioration of sensor trustworthiness. In the case of a video image processor, it could be caused by weather condition that worsen visibility, as for instance fog or rain. We model such a correlation by setting X and E_i as *parents* of the observable variable $O_X^{S_i}$ in the BN. The quantification of BN parameters requires the quantification of conditional probabilities in $P(O_X^{S_i}|X, E_i)$. This corresponds to a confusion matrix to be assessed for each configuration of the exogenous factors in E_i. Assuming no other variables are involved in the measurement process, we have X and E_i representing root nodes of the BN graph (see an example in Fig. 11). Note that to complete the BN quantification, also the unconditional probabilities of X and E_i should be assessed. Overall, this defines a joint model of the form:

$$P(X, O_X^{S_1}, \ldots, O_X^{S_m}|E_1, \ldots, E_m) = P(X) \prod_{i=1}^{m} P(O_X^{S_i}|X, E_i)\,, \tag{5}$$

where the actual values of the exogenous factors are assumed to be observed. If this is not the case, a weighted average over the different exogenous configurations should be considered. Such a joint probabilistic model allows to infer information about the variable X in the form of the posterior probability:

$$P(X|O_X^{S_1}, O_X^{S_2}, \ldots, O_X^{S_m}, E_1, \ldots, E_m)\,. \tag{6}$$

Sensor diversity implies the conditional independence of the different sensor measurements given the latent variable as well as the unconditional independence between the exogenous factors affecting the reliability of the different

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

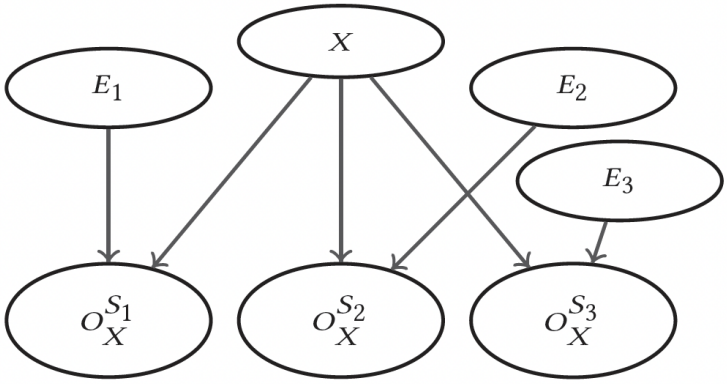


Figure 11: *A BN modelling three sensors measuring X.*

sensors. For instance, weather conditions are mostly irrelevant to radar sensors, which are subject to multi-path propagation, separability, and sensitivity of radar cross section to the aspect angle [13]. Those assumptions allow to reduce the computation of the posterior probabilities to an inference in a *naive* topology and hence derive a closed-form expression. In practical applications, extensive testing for correlation between components should be considered to evaluate the impact on safety targets. If the above independence relations are not satisfied, BN inference algorithms should be used to obtain the posterior distributions. This appears as the necessary computational counterpart of a higher expressiveness in the modelling phase. Similar considerations hold for the extension of the above static setup to a dynamic one, where the *Markovian* assumption is typically considered to reduce the modelling to two consecutive time steps only.

On a computational level, such architecture can be trained with the dataset we collected and described in Sec. 4. To include environmental factors, it is necessary to extend this dataset. One possible approach is to use filters to simulate the presence of adverse conditions such as darkness, rain and fog on the camera data. We plan to perform the computations in the continuous case by adopting the formalism of linear-Gaussian Bayesian networks. As we are collecting data in lab environments, it is possible to compute the conditional probabilities for the O and X variables statistical learning techniques, while dedicated techniques such as the expectation-maximisation algorithms should be considered for the *latent* variables E. The trained architecture and results as well as the augmented dataset will be presented in the deliverable D4.5 at the end of task T4.3.

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

6. Conclusions

We addressed the problem of multi-sensor fusion safe road-crossing by autonomous wheelchairs supported by flying drones in the context of REXASI-PRO through a simplified scenario realised in our lab. We have focused on the generation of a laboratory dataset from multiple artificial vision and distance sensors installed on RoboMasters operating with ROS in an OptiTrack enviroments. We also designed an analytical danger function to enable run-time risk assessment for road-crossing decision support. We have experimentally evaluated the danger function to provide some preliminary results as a proof-of-concept applying simple average mean for sensor fusion at different level of data processing. As a result, we achieved better performance for the fusion applied at the least elaborate data level. We also presented the approach for a more sophisticated sensor fusion where environmental condition and sensors uncertainties are modeled through a BN architecture. As a next step, we plan to train and test such architecture on our dataset. The results will be presented in deliverable D4.5 at the end of Task 4.3.

As future developments, we also plan during Task 4.3 to extend the dataset and experimentation with more complex real-world situations to be recognised and managed, such as multiple vehicles/obstacles from different directions, as well as with interference and disturbances due to vibrations, dirty camera lenses, glares, darkness, obstructions, weather conditions (rain, fog, etc.), some of which can be simulated by applying appropriate software filters. This new dataset will be used to test the model under adverse environmental conditions.

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

References

[1] HLEG AI. High-level expert group on artificial intelligence, 2019.

[2] Algirdas Avizienis, Jean-Claude Laprie, Brian Randell, and Carl Landwehr. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, 1(1):11–33, 2004.

[3] Axel Brando, Isabel Serra, Enrico Mezzetti, Francisco J. Cazorla, Jon Perez-Cerrolaza, and Jaume Abella. On neural networks redundancy and diversity for their use in safety-critical systems. *Computer*, 56(5):41–50, may 2023.

[4] Long Chen, Yuchen Li, Chao Huang, Bai Li, Yang Xing, Daxin Tian, Li Li, Zhongxu Hu, Xiaoxiang Na, Zixuan Li, Siyu Teng, Chen Lv, Jinjun Wang, Dongpu Cao, Nanning Zheng, and Fei-Yue Wang. Milestones in autonomous driving and intelligent vehicles: Survey of surveys. *IEEE Transactions on Intelligent Vehicles*, 8(2):1046–1056, feb 2023.

[5] Franca Corradini, Francesco Flammini, and Alessandro Antonucci. Probabilistic modelling for trustworthy artificial intelligence in drone-supported autonomous wheelchairs. In *Proc. 1st Int. Symp. on Trustworthy Autonomous Systems*, TAS '23, New York, US, 2023. ACM.

[6] Francesco Flammini. Model-based analysis of 'k out of m' correlation techniques for diverse redundant detectors. *International Journal of Performability Engineering*, 9(5), 2013.

[7] Francesco Flammini, Cristina Alcaraz, Emanuele Bellini, Stefano Marrone, Javier Lopez, and Andrea Bondavalli. Towards trustworthy autonomous systems: Taxonomies and future perspectives. *IEEE Transactions on Emerging Topics in Computing*, pages 1–13, 2022.

[8] Francesco Flammini, Stefano Marrone, Roberto Nardone, Mauro Caporuscio, and Mirko D'Angelo. Safety integrity through self-adaptation for multi-sensor event detection: Methodology and case-study. *Future Generation Computer Systems*, 112:965–981, 2020.

[9] Francesco Flammini, Stefano Marrone, Roberto Nardone, Mauro Caporuscio, and Mirko D'Angelo. Safety integrity through self-adaptation for multi-sensor event detection: Methodology and case-study. *Future Generation Computer Systems*, 112:965–981, 2020.

[10] Carlo Grigioni, Franca Corradini, Alessandro Antonucci, Jérôme Guzzi, and Francesco Flammini. Safe road-crossing by autonomous wheelchairs: a novel dataset and its experimental evaluation. *arXiv preprint arXiv:2403.08984*, 2024.

[11] Q.P. Ha, L. Yen, and C. Balaguer. Robotic autonomous systems for earthmoving in military applications. *Automation in Construction*, 107:102934, 2019.

[12] Hongmei He, John Gray, Angelo Cangelosi, Qinggang Meng, T. M. McGinnity, and Jorn Mehnen. The challenges and opportunities of artificial intelligence for trustworthy robots and autonomous systems. In *2020 3rd International Conference on Intelligent Robotic and Control Engineering (IRCE)*, pages 68–74, 2020.

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**

[13] Martin Holder, Philipp Rosenberger, Hermann Winner, Thomas D'hondt, Vamsi Prakash Makkapati, Michael Maier, Helmut Schreiber, Zoltan Magosi, Zora Slavik, Oliver Bringmann, and Wolfgang Rosenstiel. Measurements revealing challenges in radar sensor modeling for virtual validation of autonomous driving. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2616–2622, 2018.

[14] Mhafuzul Islam, Mizanur Rahman, Mashrur Chowdhury, Gurcan Comert, Eshaa Deepak Sood, and Amy Apon. Vision-based personal safety messages (PSMs) generation for connected vehicles. *IEEE Transactions on Vehicular Technology*, 69(9):9402–9416, 2020.

[15] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. Adaptive computation and machine learning. MIT Press, 2009.

[16] Jean-Claude Laprie. From dependability to resilience. In *38th IEEE/IFIP Int. Conf. On dependable systems and networks*, pages G8–G9, 2008.

[17] Feng Li, Wenjun Pan, and Jiali Xiang. Effect of vehicle external acceleration signal light on pedestrian-vehicle interaction. *Scientific Reports*, 13, 09 2023.

[18] Mohammad Reza Mousavi, Ana Cavalcanti, Michael Fisher, Louise Dennis, Rob Hierons, Bilal Kaddouh, Effie Lai-Chong Law, Rob Richardson, Jan Oliver Ringer, Ivan Tyukin, and Jim Woodcock. Trustworthy autonomous systems through verifiability. *Computer*, 56(2):40–47, 2023.

[19] Manuel Müller, Timo Müller, Behrang Ashtari Talkhestani, Philipp Marks, Nasser Jazdi, and Michael Weyrich. Industrial autonomous systems: a survey on definitions, characteristics and abilities. *at - Automatisierungstechnik*, 69(1):3–13, 2021.

[20] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Ng. ROS: an open-source robot operating system. volume 3, 01 2009.

[21] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You Only Look Once: Unified, real-time object detection, 2016.

[22] Pedro Shiguihara, Alneu De Andrade Lopes, and David Mauricio. Dynamic bayesian network modeling, learning, and inference: A survey. *IEEE Access*, 9:117639–117648, 2021.

[23] Zhaowen Wang, Ercan E Kuruoğlu, Xiaokang Yang, Yi Xu, and Thomas S Huang. Time varying dynamic bayesian network for nonstationary events modeling and online inference. *IEEE Transactions on Signal Processing*, 59(4):1553 – 1568, 2011.

[24] Milan Simeunović Nenad Saulić Milan Lazarević Zoran Papić, Andrijana Jović. Underestimation tendencies of vehicle speed by pedestrians when crossing unmarked roadway. *Accident Analysis & Prevention*, 143:105586, 2020.

This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement **No 101070028-2**