"The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723205"

# Deliverable D3.5

## Report on Smart Detection System (SDS)

**Due date of deliverable: 31/10/2019**

**Actual submission date: 29/102019**

**Authors:**
CEREMA: Mohamed Chafik Bakey, Jean-Paul Garrigos, Louahdi Khoudour, Christian Françoise
UTBM: Jocelyn Buisson, Zhi Yan, Yassine Ruichek
NeoGLS : André Perpey, Clément Ruffin, Anthony Roy

## Project details

| | |
|---|---|
| Project acronym | SAFER-LC |
| Project full title | SAFER Level Crossing by integrating and optimizing road-rail infrastructure management and design |
| Grant Agreement no. | 723205 |
| Call ID and Topic | H2020-MG-2016-2017, Topic MG-3.4-2016 |
| Project Timeframe | 01/05/2017 – 30/04/2020 |
| Duration | 36 Months |
| Coordinator | UIC – Marie-Hélène Bonneau (bonneau@uic.org) |

## Document details

| | |
|---|---|
| Title | Report on Smart Detection (SDS) |
| Workpackage | WP3 |
| Date of the document | 29/10/2019 |
| Version of the document | 03 |
| Responsible partner | CEREMA |
| Reviewing partner | NeoGLS |
| Status of the document | Final |
| Dissemination level | Public |

## Document history:

| Revision | Date | Description |
|---|---|---|
| 01 | 17/10/2019 | First Draft (read and corrected by UTBM team) |
| 02 | 24/10/2019 | Second draft (read and corrected by NeoGLS team) |
| 03 | 29/10/2019 | Ethical and final review |

# Consortium - List of partners

| Partner No | Short name | Name | Country |
|---|---|---|---|
| 1 | UIC | International Union of Railways | France |
| 2 | VTT | VTT Technical research centre of Finland Ltd | Finland |
| 3 | NTNU | Norwegian University of Science and Technology | Norway |
| 4 | IFSTTAR | French institute of science and technology for transport, development and networks | France |
| 5 | FFE | Fundación Ferrocarriles Españoles | Spain |
| 6 | CERTH-HIT | Centre for Research and Technology Hellas - Hellenic Institute of Transport | Greece |
| 7 | TRAINOSE | Trainose Transport – Passenger and Freight Transportation Services SA | Greece |
| 8 | INTADER | Intermodal Transportation and Logistics Research Association | Turkey |
| 9 | CEREMA | Centre for Studies and Expertise on Risks, Environment, Mobility, and Urban and Country planning | France |
| 10 | GLS | Geoloc Systems | France |
| 11 | RWTH | Rheinisch-Westfaelische Technische Hochschule Aachen University | Germany |
| 12 | UNIROMA3 | University of Roma Tre | Italy |
| 13 | COMM | Commsignia Ltd | Hungary |
| 14 | IRU | International Road Transport Union - Projects ASBL | Belgium |
| 15 | SNCF | SNCF | France |
| 16 | DLR | German Aerospace Center | Germany |
| 17 | UTBM | University of Technology of Belfort-Montbéliard | France |

# Executive summary

The deliverable D3.5 is mainly dedicated to the description of the smart detection system (SDS) developed in collaboration with UTBM and within workpackage 3 and task 3.2 of this WP.

This deliverable is a logical continuation of the deliverable D3.1 called “Proof-of-concept on data acquisition platform for risk evaluation and SDS systems”. The deliverable D3.1 was the first deliverable of the work package 3. The main aim of this work was to show the proof of concept on data acquisition platform for risk evaluation (task 3.1) and SDS system (composing task 3.2). For this purpose, a global architecture (hardware and software) was built in order to receive, manage, process the data collected in several different test sites. Different datasets were collected to test the proof of concept. The main aim during the period of work was to try to collect video data as much as possible for the development of several systems and particularly the risk evaluation (leader: UTBM) and SDS. The main objective was to verify that these data, coming from different places, with very different formats, with different acquisition rates, different image sizes are suitable for the detection and recognition of potentially dangerous situations at level crossings.

Once the proof of concepts verified, the work is divided into two different parts : one is called “Report on risk evaluation system and use cases for pilot test” led by UTBM (D3.4) and one is called “report on smart detection system” led by Cerema (D3.5).

This document provides an overall description of the Smart detection system (Hardware + software), the data gathered by the system, the test sites used to evaluate the SDS, the evaluation plan and the main results. A specific section is dedicated to the safety-related aspects when using the SDS.

## Table of content

# 1. INTRODUCTION

## 1.1. Objectives of SAFER-LC project

Over the past few years, there has been one death and close to one serious injury every day on level crossings in Europe. Therefore, SAFER-LC aims to improve safety and minimize risk by developing a fully integrated cross-modal set of innovative solutions and tools for the proactive management and design of level-crossing infrastructure.

These tools will enable road and rail decision makers to find effective ways to detect potentially dangerous collision situations at level crossings, prevent incidents at level crossings by innovative design and predictive maintenance methods and mitigate the consequences of incidents/disruptions due to accidents or other critical events.

The project will focus both on technical solutions, such as smart detection services and advanced infrastructure-to-vehicle communication systems, and on human practices, to adapt infrastructure design to end-users and to enhance coordination and cooperation between different stakeholders from different transportation modes.

The project will first identify the needs and requirements of rail-road infrastructure managers and LC users and then seek to develop innovative smart detection and communication systems, and to adapt them for use by all types of level crossing users.

A series of pilot tests across Europe will be rolled out to demonstrate how these new technological and non-technological solutions can be integrated, validated for their feasibility and evaluated in terms of their performance.

The project will deliver a bundle of recommended technical specifications (for standardisation), human practices and organizational and legal frameworks for implementation.

Finally, SAFER-LC will develop a toolbox accessible through a user-friendly interface which will integrate all the project results and solutions to help both rail and road managers to improve safety at level crossings.

## 1.2. General description of Task 3.2

The objective of the **smart detection system** (SDS), developed within task 3.2, is to develop a warning system based on intelligent detection of potentially dangerous situations occurring at LCs and some hazard situations in the larger surrounding of the LC. An optimized Automatic Incident Detection dedicated to level crossings is specified, and implemented, and evaluated. This subtask will use of recommendations and a list of requirements that are outlined as part of WP1. The SDS allows for the accurate detection of hazardous events and localization of obstacles which are motionless or in motion at the LC that could jeopardize the safety of LC users especially vulnerable users. Possible events to detect include vehicles stopped on the tracks, objects left on the tracks, trespassing and pedestrians stopping or crossing the LC.

Indeed, this task 3.2 included in WP3, shares common data with task 3.1. For the risk evaluation (task 3.1), the team develops some simulations scenarios representing potentially dangerous situations at LC. These simulations are fed and modeled by real data coming from the test sites (real level crossing or mock-up built for this purpose). Smart Detection System is planned to detect and recognize in real time dangerous situations from the collected real time data. The work performed in the purpose of this deliverable is to try to show the capability of the SDS to detect the potentially dangerous situations implemented in two different test sites : Cerema and Aachen test sites. Another very important fact is to analyse the usefulness of the SDS in case it is implemented in the daily management system of a railway operator. Some open discussions and recommendations are included in this deliverable.

## 1.3. Interactions with other tasks and evolution within the project

The work included in this deliverable, concerning tasks 3.1 and 3.2, is strongly linked to task 3.4 for the communication system. The main aim of the task 3.4 : Communication systems for cross-modal information sharing is to transmit the information generated at the level crossing by a video system to send alarms to different actors around the LC depending on the event detected : risk , dangerous situations should be sent to the approaching cars, the approaching train or to a control center. The nature of information and to whom sending it will be developed between tasks 3.1, 3.2 and 3.4. The SDS system will be connected to the communication system thanks to a specific interface. This work is described here.

## 1.4. Purpose and structure of the document

The outline of this document is the following:

- Introduction
- Global hardware architecture of the Smart Detection System
- Global software architecture of the Smart Detection System
- Datasets used for the training and the evaluation of the SDS
- Evaluation plan and global results obtained
- Discussion on safety related aspects
- Conclusion and perspectives
- Bibliography

## 2. GLOBAL HARDWARE ARCHITECTURE OF THE SDS

The global architecture of the system used for the final test is represented in Figure 1. It includes the smart detection system and the smart Roadside Unit with the NeoGLS interface which is able to send information to surroundings cars or to the train. The NeoGLS interface is connected also to Ifsttar communication system:

1. The SDS is implemented on a personal computer with Linux as operating system connected to an IP camera. The SDS processes data flows coming from the video sensor in order to detect events occurring in the field of view of the camera.
2. The video flow is stored in a video dataset.
3. The events detected by the SDS are registered using Linux Syslog standard process. This process is configured for using documents-oriented dataset, mongoDB.
4. The process (Event Proxy process) developed allows to send events stored in the database, via NeoGLS Roadside Unit (RSU) network.
5. The process (Video Proxy process) allows to send video flows stored in the video database via NeoGLS RSU network
6. The NeoGLS RSU receives all the information: events detected by the SDS, the corresponding video flow, the state of the traffic lights, the state of the barriers. Then the principle is the following. According to the status of the lights and the status of the barriers, the RSU is choosing the adequate alerts to send to the control room or to the Onboard Unit installed in the Ifsttar cars. Every alert sent to the control room is accompanied by the related piece of video.

### 2.1. Hardware components

The main component used in the SDS shown in figures 1 and 2, is the acquisition and processing units. This equipment is based on a mini ATX (Advanced Technology Extended) Personal Computer platform. This choice was essentially justified by the fact of being able to integrate a GPU (Graphics Processing Unit) card, allowing to execute process using the CUDA (Computer Unified Device Architecture) cores that these cards are equipped with, such as Deep Learning.

This architecture is therefore generic enough to test various technological solutions. The acquisition and processing unit is contained in a box whose dimensions are 26x26x20 cm. This type of form factor makes it easy to move equipment in the field, to different test sites. The second essential component of the hardware architecture is of course the acquisition camera. The choice fell on an IP (Internet Protocol) camera powered by POE (Power Over Ethernet) technology.

The criteria adopted at this point, were the universal and standardized aspects of these technologies. Indeed, as we will see later, the idea was to develop processing and analysis tools that are compatible with different operating systems (such as Windows or Linux). However, the connectivity by Ethernet is the one that offers the most ease on this field.

The selected camera comes with an optical system which is able to detect objects situated in close proximity to the pole. However, the optical field is large enough in order to observe at a glance, the entire crossing plank and the nearest roadways.

The camera sensitive assets parts are protected by a protective case compliant with the EN 60529 IP65 level.

The NeoGLS RSU is also powered by PoE (Power Over Ethernet) with a waterproof RJ45 connector. This connector allows the RSU to be IP65 compliant, to be provided with power and to communicate with the control room by Ethernet. All hardware modules are embedded in the IP65 case: that means the electronic PCB with a GPS module and ITS-G5 communication capabilities.

In addition to the RSU, a Raspberry Pi module is installed to operate the barrier. The Raspberry Pi was chosen because this is one of the most compact ways to have an operating system running with many inputs and outputs (GPIO) that can be connected to sensors or actuators. In that case, one input is used to know the state of the barrier, and one output is used to toggle the barrier state. This module is also connected to the RJ45 Gigabit link so that the RSU is capable to communicate with it and operate the barrier.

A NAS (Network Access Storage) unit was added to the proof of concept architecture. This unit is used to store video and frame data in order to retain background information of the ground truth if needed. This equipment could host 5 Hard Disk Drives. We choose some 10 TB Disks. So the total NAS capacity is about 50 TB. The data speed capacity between the processing and acquisition unit and the NAS unit can be improved, thanks to an incorporated optical 10 Gb SPF+ (Small form-factor pluggable) card.

Finally, As some components need a standard 230 volt electric power in order to operate, a cells stack was added in order to power all the other equipments in case of test on field without an AC Power Supply outlet. The batteries 12-volt power are converted to 230 volts by means of a power inverter. An UPS (Uninterruptible Power Supply) unit is added for the purposes of stop correctly the acquisition and processing unit, in case of lack of power in the cells stack. The battery pack is designed and sized to allow the entire smart detection system proof of concept architecture, proper functioning for at least one day.

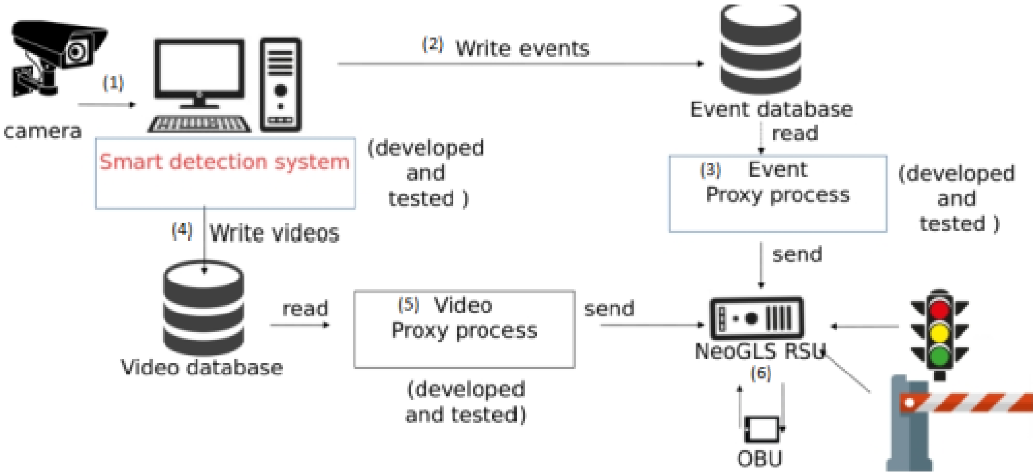


Figure 1: Interaction of SDS with NeoGLS and Ifsttar systems

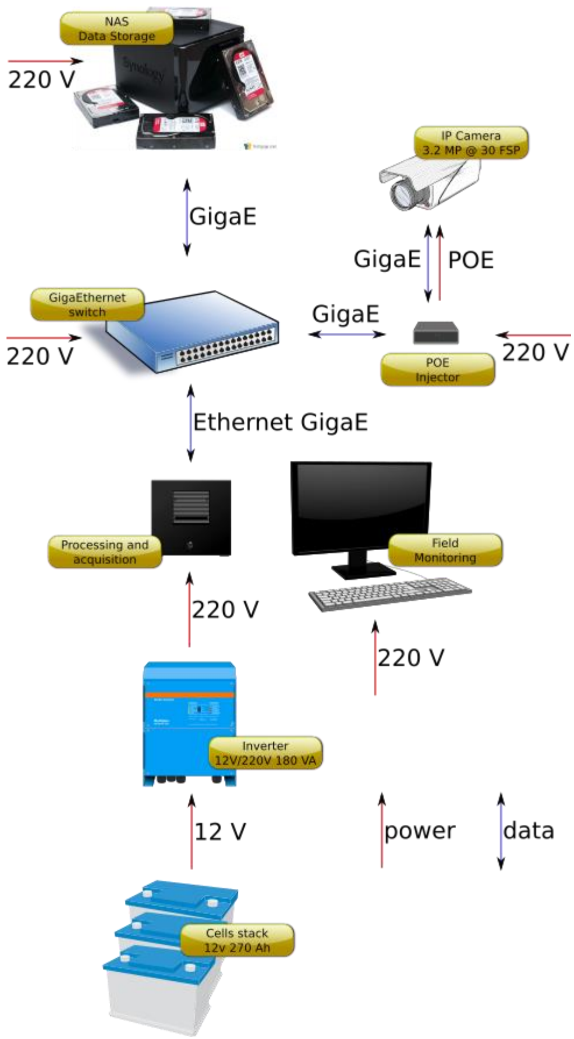



Figure 2: Hardware components and data flows

## 2.2. Intercommunication between the components and external environment

As previously seen, the intercommunication between the components, is ensured by a gigabit ethernet link. The acquisition and processing unit is equipped with two gigabit ethernet socket. So it’s possible to separate the communication between communication general standard equipments (like communication system with outside of smart detection system for example) and the specific gigabit ethernet link dedicated for the communication with the camera (the acquisition link). This kind of hardware solution allows to improve the streaming video flow quality.

The communication outside of the smart detection system is done by two ways, as shown below (Figure 3).

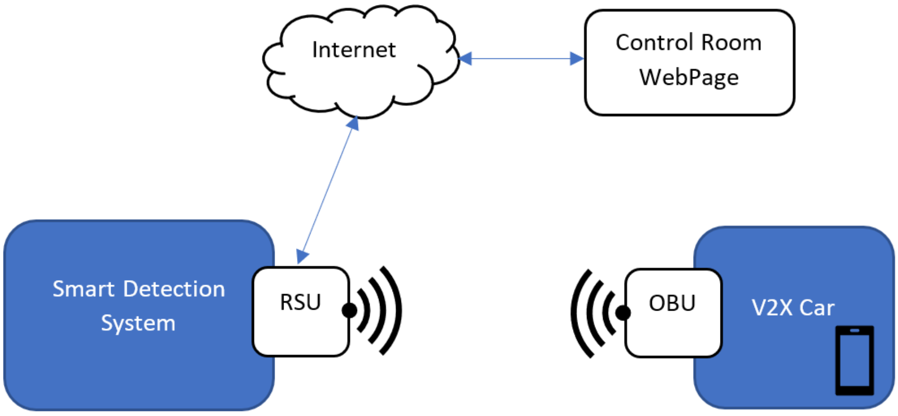



Figure 3: Communications between the SDS and external environment

The SDS POE switch, which powers the Camera and the NeoGLS RSU is connected to the NeoGLS router designed to create a bridge between the DLR partner local network (which provides Internet access for all the equipments) and the RSU local subnet. All the SDS equipments and the Raspberry Pi are directly connected to the NeoGLS RSU local subnet.

This internet access is used to establish an OpenVPN connection between the RSU and the NeoGLS web platform (Control Room). This VPN secured connection establishes an uptstream (RSU to control room) and downstream (control room to RSU) link.

Detected events and video captures received from the Camera are forwarded to the control room using the upstream link. The downstream link is used to fetch the status of the LC (barrier status, detected vehicles, transmitted events, …) and present it to the agent on a dynamic map.

The second communication way outside of the SDS is ITS-G5. This communication is done between the NeoGLS RSU and a NeoGLS OBU present in a car. Both have an antenna (internal for the RSU, and a magnetic external one for the OBU which is placed on the car's roof). The two ITS stations communicate with ITS-G5 using the CCH (Control Channel). The messages are exchanged using the GeoNetworking protocol (ETSI EN 302636-4-1 v1.2.1) which will be detailed in2.5 section.2.3 Link with the software application (Libraries)

The software architecture includes a specific layer dedicated to the communication between the acquisition and processing unit and the camera, through the gigabit Ethernet acquisition link.
This layer uses two kind of software pieces:

- Smartek proprietary camera manufacturer SDK (Software Development Kit) for Linux and Windows operating systems;
- Aravis open source library for video acquisition using Genicam cameras.

Both solutions implement the gigabit ethernet protocols used by the selected camera. Aravis software offers the capability to reduce the dependency with a specific manufacturer camera but is only Linux operating system compliant. This downside can, however, be offset by the fact that Linux supports directly tools commonly used in data sciences, making it the reason why we finally choose this operating system for the acquisition and processing unit.

From the RSU side, G5 messages are sent using the LLC (Link Layer Control) library. The above layers of the ITS stack are proprietary software from NeoGLS, whereas the VPN connection is established using OpenVPN.

## 2.3. Events storage

The SDS uses the linux standard daemon Syslog-ng, in order to store the detected events. The daemon is configured to save data in a mongoDB document-oriented database instead of a classic log text file. This allows to prepare and launch database requests easier.

Once the events are forwarded to the control room, they are also stored in a database so that past events are always available to the agent. However, there is no storage directly in the RSU. Upon restart, the current situation is rebuilt in the LDM (Local Dynamic Map) using the communication from the SDS and the OBU.

## 2.4. Communication protocol used between SDS and RSU

The SDS uses a specific JSON-RPC Protocol, in order to transmit the detected events. JSON-RPC is a remote procedure call protocol encoded in JSON (JavaScript Object Notation). It's a very simple protocol, with few data types and commands.

The version used is the version 2 specified on 2009-05-24.

This protocol works by sending a request to a server (for SAFER LC Project, the server is the NeoGLS RSU). The protocol client is a specific software process (designed by Events proxy process). Multiple input parameters can be passed to the remote method as an array or object.
The protocol specification, provides several properties:

- **method**: a string with the name of the method to be invoked;
- **params**: an object or Array of values to be passed as parameters to the defined method;
- **id**: A value of any type used to match the response with the request that is replying to.
- The receiver of the request must replay with a valid response to all received requests:
- **result**: The data returned by the invoked method. Eventually null if an error occurs;
- **error**: A specific error code if there was an error;
- **id:** The id of the request it is responding to.

One can see below the version 2 of Smart camera JSON-RPC API specification.
In the case of SAFER LC project, the id parameter is the syslog mongoDB database ObjectId value which consists of :

- a 4-byte value representing the seconds since the Unix epoch of the database saving action timestamp, and not the event timestamp;
- a 5-byte random value, and;
- a 3-byte counter, starting with a random value.

## 2.5. Communication protocol used between RSU and external environment

**RSU to OBU:** To communicate with the OBU over ITS-G5, the RSU uses the GeoNet protocol (ETSI EN 302636-4-1 v1.2.1). Both stations implement an ITS stack (Facilities) as described in ISO TS 103-301 which allows them to exchange messages. The following table shows which type of messages are sent by the RSU and the OBU:

| | CAM | DENM | MAP/SPaT |
|---|---|---|---|
| **RSU** | Sent but not used by OBU | For each event detected by the camera (and by ITS-G5 OBU position), the RSU generates a DENM which is broadcasted to alert any V2X car incoming. The causes used are various and depend on what is detected (pedestrian, traffic jam, stationary vehicle, …) | The RSU generates SPaT/MAP data depending on the status of the LC (open or closed). In that way, an arriving V2X car can know the status and adapt its speed accordingly. |
| **OBU** | The OBU sends CAM messages containing its speed and position. The RSU receives it and decides to create an alert under certain conditions (ex: the OBU is stopped on the LC). If this is the case, the alert is broadcasted using DENM messages, and is also forwarded to the control room. | Not sent | Not sent |

Table 1: Summary of broadcasted messages between RSU and OBU

**RSU to Control room:** To communicate with the control room, several protocols are used:

- JSON RPC: As the RSU is already a JSON RPC server for the Camera Events, the same server is used by the control room to request the RSU status and display it on a cartography. Only the remote method called is different from the one for event detection.
- SCP: Secure Copy is used to transfer video extracts to the control room. This protocol was used for several reasons: first, it allows authentication by key algorithm so that the control room is sure that provided data comes effectively from the RSU. Second, it allows data encryption so that videos cannot be intercepted and viewed by other people. Last, it ensures that data has been transmitted correctly using checksums.
- HTTP: Is used at RSU startup to download its configuration and fetch the last state of the LC.

# 3. GLOBAL SOFTWARE ARCHITECTURE OF THE SDS

## 3.1. Processing modules

Figure 4 reports the global processing pipeline implemented based only on the smart detection system. The different processing modules are mentioned on the figure with a few predefined scenarios to detect and recognize like the presence of an obstacle or a traffic jam at LC, atypical behavior and pedestrian(s) at LC. The SDS is composed of 6 modules.

**Data acquisition module:** The main aim of this module is to carry out data acquisition from the video sensor (all kind of IP cameras could be taken into account). The frame rate of the acquisition is 30 frames/sec. This module includes three extensions: acquisition from a webcam, acquisition from a video file (avi for example), online acquisition directly from the camera flow. These extensions were tested during the training of the software tests.

**Video compression module:** The aim of this module is to compress and store all the video data acquired. The duration of each video footage is 30 seconds (further changed into 5 seconds, for control center needs). These videos are used in the control center in order for the operator to verify the veracity of the event detected and received.

**Objects detection module:** This module allows to detect all the objects present on the scene analysed by the camera. This detection is based on background subtraction. Several algorithms based on background subtraction are developed in this module, some of them were published in ICIP conference [1]. We have also developed an improved version of Codebook algorithm [2]. The choice of the algorithm to use is proposed to the user in a dropdown list. The user can choose the algorithm desired. The detection result for each image received from the camera, is a black and white output called mask. In this image, the white pixels represent the objects detected while the black pixels represent the background. This module suggests also 4 algorithms dedicated to the shadows removal in order to improve the object tracking procedure [3].

**Objects tracking module:** This module ensures the tracking of the objects detected. The main aim is to track the objects detected in the scene. In this module a specific tracking algorithm is developed. This algorithm allocates a label to each object detected (car, pedestrian, etc.…).

**Objects classification module:** This module aims to recognize the object detected and find its class (pedestrian, car, bicycle, truck, …). The algorithm developed is based on deep learning techniques. In the algorithm, we have a neural network with an SSD-based architecture (Single Shot multibox Detector) [4].

**Events detection module**: This module is fed by the two previous algorithms (tracking and classification modules). Depending on the results of this module, the main aim is to detect the event whatever its nature. This event module is the most important of the system since it is important to not miss obstacles present on the LC. Then, another part of the module is dedicated to the possibility to recognize a scenario.

« Obstacle on the level crossing»: an obstacle (vehicle, pedestrian…) is stopped on the Level Crossing (LC);

« Pedestrian on the level crossing »: pedestrians present in the region of interest of the LC, or outside the region of interest.;

« Atypical behaviour »: an object is moving forward and then backward in the LC zone.

« Jam »: several objects (mainly cars, but also cars + pedestrians, several pedestrians, etc.…) are creating a traffic jam on the LC.

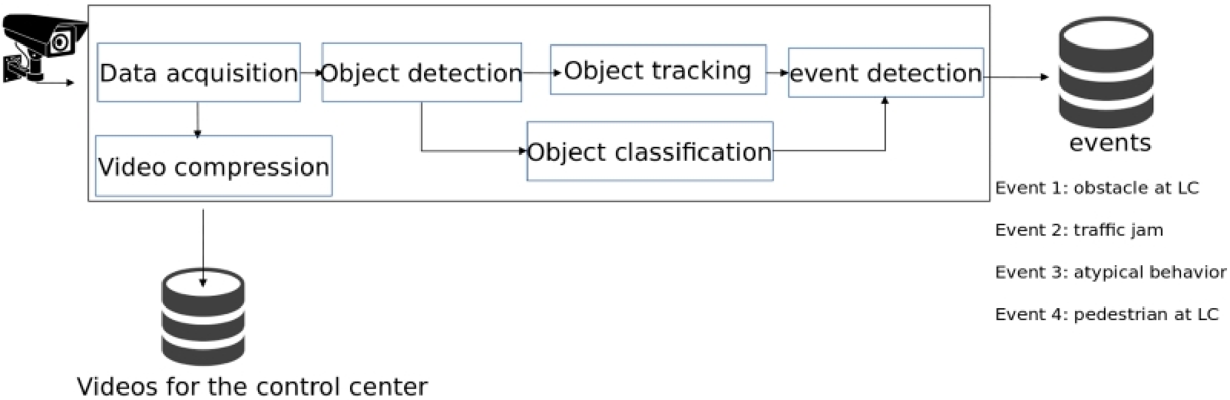


Figure 4: Software architecture of the SDS, global processing pipeline

## 3.2. User interface

Figure 5 illustrates the SDS user interface. This interface is composed of four main representations (4 rows on the figure).

**In the first row** one can find four widgets:

- The « Algorithms » widget allows the user to choose the algorithm to be applied for the three following modules:
  - Object detection module (background subtraction + shadows elimination).
  - Objects tracking module
  - Classification module

- The « Input » widget allows the user to choose the acquisition mode (online, offline, ….)
  - For the « offline » mode, it is possible for the user to read from a folder containing files (avi, mov, ….)
  - For the « online » mode, the user can choose among acquisition from an IP camera (Gig camera) or from a webcam.

The « output » widget allows the user to choose the folder in which video data is stored;

The « Region Of Interest » widget allows the user to define the region of the LC to be taken into account

**In the second row**, one can find 3 video widgets which are able to display their content in real-time:

- The first widget displays the input video (data flow coming from the camera in the case of "online version" or data coming from video files in the "offline" version);
- The second widget displays the objects detection results (binary image with masks);
- The third one displays the result of the final processing (result of the modules: tracking, classification and very important, **the detection**)

**In the third row**, one can find the images list of the detected objects.

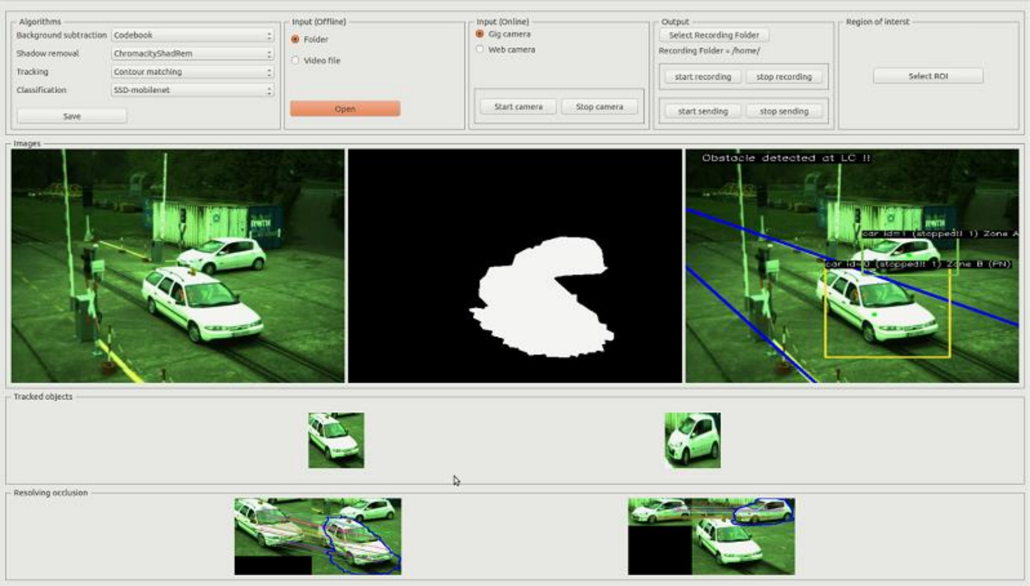



Figure 5: User interface of the SDS

In the fourth row, one can find the occlusion management algorithm. In the displayed example, the background subtraction algorithm considers two objets as a single object. The dedicated algorithm will separate the two objects thanks to knowledge in previous images.

# 4. DATASETS USED FOR THE TRAINING AND THE EVALUATION OF THE SDS

## 4.1. Data used for the training of the software

For the development and training of the SDS algorithms, there are several sources of data which were used:

- Data coming from the Internet with ground truth [5] (as seen in photo1). These datasets are used by researchers to compare their video algorithms. They represent many different scenes with scenarios including cars, pedestrian and trucks during very different weather conditions periods. These datasets are very useful to test the performances of some video algorithms. Furthermore, these datasets were built with an existing ground truth. This dataset contains 11 video categories with 4 to 6 videos sequences in each category. The categories are : "Bad Weather", "Low Framerate", "Night Videos", "PTZ", "Thermal", "Shadow", "Intermittent Object Motion", "Camera Jitter", "Dynamic Background", "Baseline" and "Turbulence".

- Data coming from a past project, belonging to a French workprogramme, called PANsafer: the datasets are coming from Mouzon in the north of France which is a rural area and from Lausans with several urban level crossings (Photo 2 for Mouzon).
- Data coming from a real level crossing in the Toulouse region (called Montaudran) (as seen in photos 3 and 4). The Montaudran LC is a semi-urban one, with a curved trajectory with a very heavy traffic including cars, two-wheels, pedestrians and a high frequency for trains.
- Data coming from a mock-up built at Cerema Toulouse premises (as seen in photos 5, 6 and 7) in which the different scenarios are played (see further in this document).

Photo1: Intersection roads/trams

Photo 2: Mouzon test site


Photo 3: A first view of the Montaudran level crossing


Photo 4: Another picture shot at Montaudran test site

Photo 5: Mockup of a level crossing installed at Cerema Toulouse

Photo 6: Mockup of a level crossing installed at Cerema Toulouse with a pedestrian and bicycle crossing the LC at the same time

Photo 7: Mockup of a level crossing installed at Cerema Toulouse a big truck zigzagging while the barriers are closed.

## 4.2. Data used for the evaluation of the SDS

### 4.2.1. Data collected at Aachen test site for the evaluation of the SDS

Three acquisitions sessions were carried out on the Aachen test site:

The main aim of **the first session** was to test the different modulesdeveloped: acquisition, events detection, tracking. Thus, this first phase (25-27 January 2019) was dedicated to test the main functionalities of the smart detection system as a standalone demonstrator. For that the video system was installed on the Aachen site and the different functionalities were tested one after another : recognition of the sensor by the software, acquisition module, processing module. To test these funtionalities, some scenarios including vehicles, pedestrians, etc.…. were simulated on the level crossing. During this first session the smart detection system was also connected to the NeoGLS interface to verify the messages exchange and verify that everything is working correctly. These tests were carried out based on events played at Aachen level crossing. Then preliminary connections between NeoGLS and Ifsttar systems were also carried out. 8 videos of events are collected (total of 30 min of recording).

During the **second phase** (26-28 march 2019), from smart detection system side, the main aim of the second phase was to test the classification of objects, the connections between NeoGLS and Ifsttar systems and also the connection with the control room and the possibility to provide video data at the same time than the alert messages. This task was partially demonstrated. We demonstrated the good connection with NeoGLS and NeoGLS demonstrated the good connection with Ifsttar system. Another important task during this second phase was to simulate events planned in the project. The four scenarios: obstacle on the tracks, pedestrians, atypical behaviour and traffic jam were simulated with repetitions to be statistically representative and in different configurations of the barriers. The scenarios were simulated during closed barriers periods and during open barriers periods. For each scenario, video data is stored, detection is performed and information exchange between smart system and NeoGLS interface is carried out. Thanks to this second phase evaluation a big amount of video data is available for further testing. 9 videos are collected (total of 58 min of recording).

The main aim of the **third session** was to test the global processing chain: SDS, RSU and communication system. Third trials period took place between 25 and 27 May 2019. The main aim of this third session for the smart detection system was to test two different things:

- In case of a detected event, send an alarm to the control room with the life video data accordingly. For this functionality and with a specific application in an internet platform, the smart detection system sends continuously 5 seconds video data buffers to the control room which could be far from the LC. When an event is detected, a visual alarm is sent to the control room with the corresponding video data. For testing these functionalities, also in this case a large set of different scenarios including cars, pedestrians, has been carried out;
- The second thing about the functionality was to test the global chain: smart detection (Cerema and Utbm) + smart interface (NeoGLS) + communication system (Ifsttar). 7 videos are collected for this testing (total of 57 min of recording).

### 4.2.2. Pre-selection of the datasets used for the evaluation

For the evaluation, four different datasets were used: the global datasets recorded at Cerema (2 sessions) and the two datasets recorded during the two first sessions at Aachen. Quantitatively this represents more than 4 hours of events to detect, 41 videos including 1038 events whatever the scenario. We have kept 24 videos coming from Cerema datasets (523 events), 8 videos from the first session of Aachen (118 events) and 9 videos of the Aachen second session (397 events).

# 5. EVALUATION PLAN AND GLOBAL RESULTS OBTAINED

## 5.1. Performance indicators for the SDS

If one day the SDS is installed at a level crossing, with a high security context, what is very important is the ability to detect events, whatever their nature. We must also keep in mind, that every time something is detected, an alarm is sent at least to the control room with the corresponding video. So, it is very important that the system is able to detect an event from the beginning to the end, to send alarm accompanied with the related video. In this case, the operator in his control room is warned, the video is displayed and that is up to him to make his decision on what to do according to the data received.

The first performance indicator is called **Perf_Detect**:
**Perf_Detect** = (number of events detected by the SDS)/(number of the events correctly detected by the SDS + the number of events non detected by the SDS)

This indicator is calculated for every video (41 videos) but we have decided to group the evaluation by distinguishing Cerema Datasets and Aachen datasets, the contexts are quite different. Table 2 provides results for the indicator Perf_Detect for global datasets, Cerema and Aachen

| | **Perf_Detect** |
|---|---|
| Global datasets | **83,7%** |
| Cerema datasets | **78,83%** |
| Aachen datasets | **88%** |

Table 2: Performance of events detection

Thus, the global **performance of detection** of the SDS (from the data processed) is around 84%. This figure seems to be low if we take into account the security context. This performance could be close to 100% if we do not multiply the number of events in a given period of time. Indeed, on a real level crossing the frequency of events is not so high. Of course, a more intensive evaluation on a real level crossing could lead to more precise results.

If we consider the test site, the performance of the system in terms of detection is a bit better for the Aachen datasets.

As we know, video sensing and image processing could be sensitive to weather conditions. So, we have decided to carry out the performance of detection of the SDS according to a classification of the atmospheric conditions.

Among the 41 videos collected we have managed to define 7 different weather conditions, which is a lucky evaluation situation. Table 3 includes the weather conditions encountered and the number of videos related to them.

| Weather | Number of videos | Number of frames | ¨% of the total frames |
|---|---|---|---|
| High sun with shadows created by objects, wind | 13 | 108572 | 35,75% |
| Sun and shadow on the LC | 1 | 3907 | 1,29% |
| Cloudy and low illumination | 10 | 55380 | 18,23% |
| Snow and low illumination | 6 | 45971 | 15,14% |
| Cloudy with low average illumination with small rain | 7 | 72293 | 23,80% |
| Snow with very low illumination | 2 | 8348 | 2,75% |
| Cloudy with higher illumination | 2 | 4882 | 1,61% |

Table 3: classification of weather conditions

The indicator **Perf_Detect_Weather** is used to calculate the ability of the SDS to detect events according to the weather conditions.

| Weather | **Perf_Detect_Weather** |
|---|---|
| High sun with shadows created by objects, wind | **80,17%** |
| Sun and shadow on the LC | **100%** |
| Cloudy and low illumination | **73,54%** |
| Snow and low illumination | **93,65%** |
| Cloudy with low average illumination with small rain | **87,78%** |
| Snow with very low illumination | **100%** |

| | |
|---|---|
| Cloudy with higher illumination | **88,89%** |

Table 4: events detection by type of weather

From the figures of table 4, it is difficult to conclude on the performance of the SDS according to the weather conditions. We have noticed that the software has a very short delay on detection when the level of illumination is low. That is why the performance of detection is lower (73.54%) for the cloudy and low illumination situation.

In principle, once the event is detected and sent to the control room with the corresponding video, it is enough for the railway operators. But in our case, it is also possible to detect and recognize the event. We recall that our scenarios are including cars, pedestrians, bicycles, etc…. So in this case it could be useful to calculate a second indicator that we call Perf_Detect_Recog. It is calculated like that:

**Perf_Detect_recog** = (number of events recognized by the SDS)/(number of the events correctly recognized by the SDS + the number of events not correctly recognized)

Recognize means that the system is recognizing correctly one of the predefined scenarios.

This indicator is calculated for every video (41 videos) but we have decided to group the evaluation by distinguishing Cerema Datasets and Aachen datasets, the contexts are quite different. Table 5 provides results for the indicator Perf_Detect_Recog for global datasets, Cerema and Aachen.

| | **Perf_Detect_Recog** |
|---|---|
| Global datasets | **71,84%** |
| Cerema datasets | **75,46%** |
| Aachen datasets | **68,56%** |

Table 5 : Performance of events recognition

The capability of the system to detect and recognize a scenario is around 72%. There is no big difference between Cerema and Aachen datasets.

As for the previous indicator (performance of detection) we have decided to carry out the performance of recognition of the SDS according to the classification of the atmospheric conditions.

The indicator Perf_Detect_Recog_Weather is used to calculate the ability of the SDS to recognize events according to the weather conditions.

Table 6 shows the performances of recognition by the SDS according to the weather.

| Weather | **Perf_Detect_Recog_Weather** |
|---|---|
| High sun with shadows created by objects, wind | **68,56%** |
| Sun and shadow on the LC | **100%** |
| Cloudy and low illumination | **72%** |
| Snow and low illumination | **77,54%** |
| Cloudy with low average illumination with small rain | **70,21%** |
| Snow with very low illumination | **38,82%** |
| Cloudy with higher illumination | **100%** |

Table 6: events recognition by type of weather

As said before, the worse situation is when the illumination is low. Indeed, the performance of recognition during “Snow with very low illumination” is 38.82%. However, this dataset is very small. The performance of **detection** for this weather situation is 93,65%. This means that the SDS is quite powerful for the detection performance (which is the most important) and could meet some difficulties for the recognition task during bad weather conditions.

## 5.2. Processing time

The processing time of the SDS is between 25 to 30 frames/sec. This processing time is compatible with a real-time implementation. This means that every object, people crossing the LC will be detected whatever its speed.

## 5.3. Connection between smart detection system and NesGLS Roadside Unit

During the tests, each scenario detected by the “Smart Detection System” is sent to the “Roadside Unit” (RSU) of NeoGLS. This exchange of data has been done on several hundred events (including pedestrians, vehicles stopped, detection of traffic jam, …). There was no anomaly in the exchanges. Figure 6 reports the detection system interface and the process of detection of a bicycle.

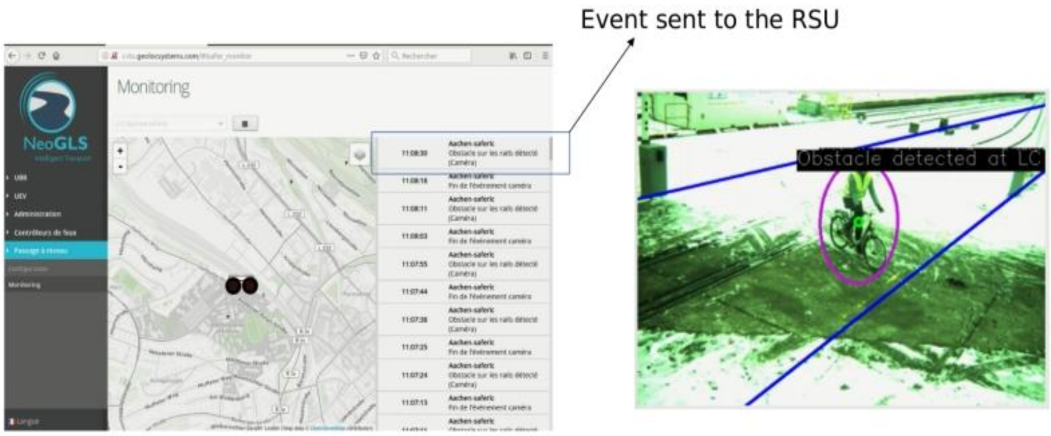



Figure 6: Example of a detection of a pedestrian crossing the LC: left: events' interface, right : illustration of the detection

The figure on the left shows the monitoring webpage of the Smart RSU. It is divided in two parts. The map on the left-hand side represents the LC and the events graphically on an OpenStreetMap cartography. The state of the barrier is also represented, as the icon blinks when the barrier is closed.

The right-hand side is a scrolling list of every detected event. When the SDS notifies the RSU of a detected event, a line is added in the list and the agent is able to view details about the detected event, and a video snippet.

## 5.4. Link between the SDS and the control center

As said before, it is very important that an event detected by the SDS on the field could be sent to the control centre to warn the railway operator that something is happening on a given LC. The application developed (with NeoGLS) for this purpose is a web platform. The application is run from SDS side and also from the control centre side.

In terms of safety, the connection with the control room and the possibility to provide up-to-date video data together with the alert messages in case of a detected event is very important.

During the last test session in Aachen, we tested the communication between the SDS and the control room. The Smart Detection System continuously sends image buffers of 5s to the RSU. When an event is detected by the SDS, the RSU starts forwarding the video samples to the control room. The video is then aggregated, and the operator in his room receives a pictorial or audible warning and the video which corresponds to the event. The tests are conclusive. The duration of the buffer is configurable. The three following images are shot at the same time (Figure 7).

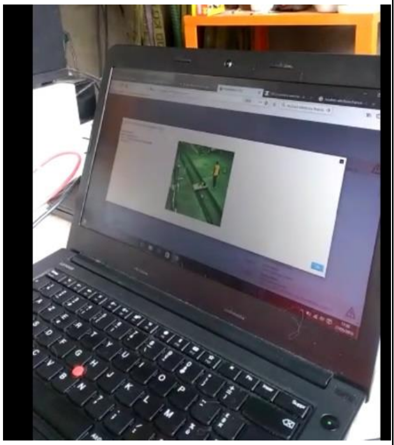
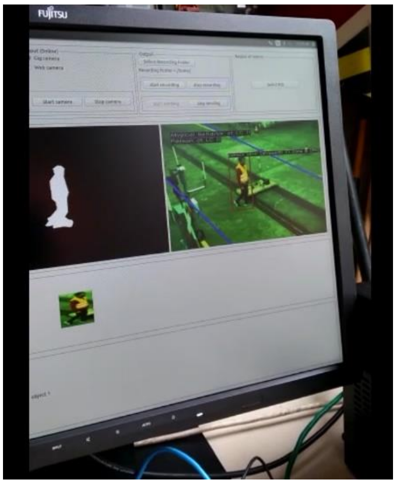
| | | |
|---|---|---|
|  |  |  |
| Control room | Smart detection system | Level crossing (Aachen) |

Figure 7 : methodology of connection with the control room

## 5.5. Detection of vulnerable users

One of the functionalities of the SDS could be the detection of vulnerable users in the level crossing premises. Of course, every object (pedestrian, car, etc...) crossing an LC is a vulnerable user. But this last functionality aims to detect workers for example present on the rails without being aware on the arrival of a train. During the last session at Aachen, we have decided to test this functionality and to see what the sensitivity of the SDS software is when detecting very small objects far from the sensor. Photos 8 and 9 illustrate this situation. Each photo is divided into three parts: left part displaying the initial video, center part displaying the masks and right part displaying the detections.

The distance between the video sensor and the objects is around 100 meters. Several video sensors could be installed to have a better coverage.


Photo 8: example of rail users on the tracks


Photo 9: example of rail users on the tracks

# 6. SAFETY ASPECTS LINKED TO THE USE OF SDS

This section is fed by an analysis of the state of the art on safety at level crossings, by an analysis of past projects [6] on the subject, by discussions with experts.

## 6.1. Intelligent Level-Crossings: Definition

An Intelligent Level-Crossing is a system which integrates functions of advanced sensors, communications and information technologies in order to improve safety and operational efficiency at rail-road crossings.

The main benefits offered by a well designed Intelligent Level-Crossing system are (i) increased security and safety of the road users, train passengers and rail staff, (ii) improved efficiency of the rail and road traffic management by provision of real-time information to rail and road users on the status of the traffic network (for example, possible route alterations due to traffic jams at level-crossings).

Such system has the capability to detect the conditions at the level-crossing, identify potentially hazardous situations, notify the local traffic management system, trigger the system response accordingly, and provide advanced warnings to the vehicle users and train drivers.

## 6.2. Obstacle/activity Detection at level-crossings

The purpose of obstacle detection systems is to prevent vehicle and train collisions at level-crossings. They allow for the timely detection of objects caught within the level-crossing area and provide means for automated alert notification and activation of an appropriate network response (such as setting off the warning signs, lifting/closing the barriers, train route alterations, or partial closing down of the rail network). The main benefit of these systems is that they are potentially capable of fully automated performance, remote control and system failure diagnostic.

Technologies deployed for obstacle detection at level-crossings generally fall into two broad categories: intrusive and non-intrusive [7]. Intrusive technologies typically require installation below the ground surface and hence cause traffic distractions during installation and maintenance; in general, they cost less to purchase but incur higher installation and maintenance costs. The intrusive technologies include inductive loops, pneumatic road tubes, magnetometers and piezoelectric cables. The non-intrusive technologies are typically installed above the road/rail surface, they are in general more expensive but easier to install and maintain. Typical non-intrusive technologies are radar, infrared, acoustic, ultrasonic and video.

The selection of the most appropriate technology for obstacle/activity detection depends on various factors and the final choice will inevitably be a result of a trade-off between the cost and performance. The findings reported in the relevant literature so far, rely on different evaluation conditions and criteria. Hence, it is not surprising that the conclusions of their performance evaluation vary. Some researchers recommend video image processing and active infrared methods [8], whereas others find that video surveillance and inductive loops [6] offer best performance.

Although a reliable methodology for evaluation of obstacle/activity detection in the level-crossing context is yet to be established, most of the researchers find fusion of the video processing and another complementary technology as most promising. The inherent limitations of the video techniques, such as inability to perform under reduced visibility due to adverse weather conditions and time of the day, may be overcome by integration with another type of technology such as radar or active infrared.

## 6.3. Video based obstacle detection

Video image processing is the most recent technology to be applied in rail transport safety applications. Although it suffers from significant drawbacks like susceptibility to weather and lighting conditions, its performance in obstacle detection applications may be significantly improved by using advanced image processing and scene understanding algorithms. Furthermore, the integration of multiple cameras allows for wide area coverage and recovering of the three-dimensional data. The multiple-view installations together with scene understanding algorithms and powerful processors create an intelligent system capable of detecting small size objects, accurate object classification and tracking. This technology allows for the capture and storage of large amounts of data and real-time processing. The data collected by video cameras may be stored and used for further research in the area of safety and security, such as human behavior modelling and traffic flow modelling in the level-crossing context. In Japan for example, a system of stereo cameras is used to achieve a large coverage area without the need for mechanical scanning [11]. Stereo cameras enable the system to obtain the three-dimensional shape information which significantly reduces the number of false detections due to shadows and vehicle headlights.

## 6.4. Recommendations

These recommendations are coming from FP7 European project SELCAT [2006-2008]. In this project, a work package was called “Level Crossing Technology” and was dedicated to the analysis of the potential use of technology for safety at level crossings. At that moment, the recommendations concerning the potential use of new technologies for existing level crossings were as follows:

- Design of any new safety system should be based on prior human factors impact analysis
- Low cost technology will allow its broader application and will contribute to a direct decrease in the number of accidents at an affordable cost
- Substitution of obsolescent technology with new technology where this can be justified on cost / benefit grounds.
- Substitution of human actions by new technology
- Assessment of safety and cost using a harmonized methodology
- Supervision of danger zones: technological improvements to level crossing safety infrastructure, such as deployment of various types of sensors (audio, video, radar, lasers) for the timely detection of potentially hazardous situations

- Use of fast, reliable, wireless links to enable a seamless communication between the train, the level crossing, and the main control centre/s.
- Optimisation and unification of crossing warning and closing times (for example, by rail board site positioning systems)
- Definition of main, basic level crossing types at a European level, based on modular designs
- The pilot development of advanced technologies for each of the basic European LC types, for the purpose of validating the proposed approaches
- Development of the harmonised level crossing information system on a European level
- Elaboration of the harmonised methodology for level crossing assessments, based on the findings of SELCAT
- Organization of a European campaign for car drivers, as well as for broader use in education.

These recommendations are intended to instigate an initiative at the European level, the aims of which are: to develop methods and tools for reduction of safety risk at level crossings; demonstrated and tested on software simulations, level crossing prototypes and in-field tests based on the selected safety scenarios.

We think that beside all these recommendations, a very important follow-up to the work on the possibility to use video sensing is to ask to the rail operators to open their test sites in order to have a very deep evaluation. We know that it is not always possible for security purposes, but we want to conclude on the usefulness of an imaging system, we should do that.

# 7. CONCLUSIONS AND NEXT STEPS

In this task 3.2 we have developed what we have called a Smart Detection System (SDS). The main aim of this work is to show a technical “proof of concept” when using video sensing for safety at level crossings. The main aim was to demonstrate that thanks to a video-based system and image processing software, we can detect potentially dangerous situations at level crossings. The SDS is interfaced with a Roadside Unit for information exchanges and the possibility to send alarms and warnings to approaching cars, the train, the control centre. All this work was developed in close collaboration with NeoGLS for the RSU and Ifsttar for the communication and data exchanges.

The main performance of the SDS is its ability to **detect events** and to transmit them. It was measured on Cerema and aachen datasets.

Thus, we have measured that the global events detection performance is around 84% and the scenarios recognition performance is around 72%. Of course, the detection performance is the most important.

There is no real influence of the weather conditions on the ability to detect or recognize except when the illumination is very low. In this case the detection is a bit delayed.

Of course, the results achieved are valid according to the datasets processed; these datasets being limited. The video-based systems need to be evaluated in a long-term manner. This is an interesting perspective.

A very good perspective could be to obtain from the railway companies the authorization to test a video imaging and communication systems during real exploitation periods. If we manage to do that, we will be able to measure the impact of these technologies on the improvements of safety at level crossings.

# 8. BIBLIOGRAPHY

[1] Bakkay, M. C., Rashwan, H. A., Salmane, H., Khoudour, L., Puigtt, D., & Ruichek, Y. (2018, October). BSCGAN: deep background subtraction with conditional generative adversarial networks. In 2018 25th IEEE International Conference on Image Processing (ICIP) (pp. 4018-4022). IEEE.
[2] Kim, K., Chalidabhongse, T. H., Harwood, D., & Davis, L. (2005). Real-time foreground–background segmentation using codebook model. Real-time imaging, 11(3), 172-185.
[3] Sanin, A., Sanderson, C., & Lovell, B. C. (2012). Shadow detection: A survey and comparative evaluation of recent methods. Pattern recognition, 45(4), 1684-1695.
[4] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham.
[5] Y. Wang, P.-M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, CDnet 2014: An Expanded Change Detection Benchmark Dataset, in Proc. IEEE Workshop on Change Detection (CDW-2014) at CVPR-2014, pp. 387-394. 2014
[6]https://trimis.ec.europa.eu/sites/default/files/project/documents/20120925_150540_41826_870-870-26_SELCAT-D13.pdf
[7] Federal Highway Administration, US Department of Transportation, "A Summary of Vehicle Detection and Surveillance Technologies Used in Intelligent Transportation Systems", Washington DC, USA, 2000.
[8] Noyce, D. A. and Dharmaraju, R., "An Evaluation of Technologies for Automated Detection and Classification of Pedestrians and Bicyclists", University of Massachusetts Transportation Center, Massachusetts, USA, 2002
[9] Erickson, J. K., "Vehicle Detection at Highway-Rail Intersections", Indian Railways Institute of Civil Engineering, 1996
[10] Salmane, H., Ruichek, Y., & Khoudour, L. (2011, October). Object tracking using Harris corner points based optical flow propagation and Kalman filter. In Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on (pp. 67-73). IEEE.

# ANNEXE 1: DETAILED RESULTS OF SDS ON CEREMA AND AACHEN DATASETS FOR DETECTION AND RECOGNITION

| Video name | Detection and recognition performance (Perf_Detect_Recog) | Detection performance (Perf_Detect) |
|---|---|---|
| video1 | 37,50 % | 100,00 % |
| video2 | 67,07 % | 84,95 % |
| video3 | 83,72 % | 93,02 % |
| video4 | 74,29 % | 94,29 % |
| video5 | 94,29 % | 97,14 % |
| video6 | 68,75 % | 100,00 % |
| video7 | 85,71 % | 82,35 % |
| video9 | 33,33 % | 25,00 % |
| video10 | 57,50 % | 59,26 % |
| video11 | 31,25 % | 66,67 % |
| video12 | 94,12 % | 100,00 % |
| video13 | 83,33 % | 75,00 % |
| video14 | 80,39 % | 64,47 % |
| video15 | 100,00 % | 100,00 % |
| video16 | 100,00 % | 100,00 % |
| video17 | 100,00 % | 36,36 % |
| video18 | 36,67 % | 65,85 % |
| video19 | 100,00 % | 100,00 % |
| video20 | 100,00 % | 77,78 % |
| video21 | 100,00 % | 100,00 % |
| video22 | 60,00 % | 70,00 % |
| video23 | 100,00 % | 100,00 % |
| video24 | 85,71 % | 82,35 % |
| video25 | 75,00 % | 17,65 % |
| video26 | 63,64 % | 77,27 % |
| video27 | 46,15 % | 84,62 % |
| video28 | 92,31 % | 100,00 % |
| video29 | 86,67 % | 100,00 % |
| video30 | 52,63 % | 100,00 % |
| video31 | 25,00 % | 100,00 % |
| video32 | 100,00 % | 100,00 % |
| video33 | 76,47 % | 100,00 % |

| Video name | Detection and recognition performance (Perf_Detect_Recog) | Detection performance (Perf_Detect) |
|---|---|---|
| video34 | 0,00 % | 55,56 % |
| video35 | 65,08 % | 96,77 % |
| video36 | 57,50 % | 80,85 % |
| video37 | 69,05 % | 85,71 % |
| video38 | 48,48 % | 100,00 % |
| video39 | 59,32 % | 92,06 % |
| video40 | 72,60 % | 86,75 % |
| video41 | 85,71 % | 92,86 % |
| video42 | 96,30 % | 87,10 % |
| **Total** | **71,84 %** | **83,70 %** |