Periodica Polytechnica
Transportation Engineering

x(y) pp. x-x, (year)
https://doi.org/10.3311/PPtr.XXXX
Creative Commons Attribution b

RESEARCH ARTICLE

System of Systems Safety Analysis and Evaluation in ZalaZONE

Hans Tschürtz[1*], Florian Wagner[2], Wilko Schröter[3], Zsolt Szalay[4], and Árpád Török[4]-

Received xx Xxxx 201x, Online xx Xxxx 201x

Abstract

Safe autonomous operation is a major challenge for today's technologies. In order to be able to define and evaluate the requirements of these technologies, a systematic and methodical approach is required.

The VISSE has developed such an approach over several years and is now to be evaluated on the basis of various use cases.

Students of the course of studies "Safety and Systems Engineering" have applied these procedures to a defined use case in a student project of a master study course.

Driving scenarios for a road intersection were defined and safety critical situations identified, analyzed and evaluated in ZalaZONE.

The analysis and test results have shown the possibility to improve a used sensor concept in beforehand. This offers the opportunity to reduce the complexity of the driving scenarios respectively to avoid unknown situations.

Keywords

ZalaZONE, System of Systems, Inherent System Safety, Functional Safety, Hazard, Critical Situation, Scenario, Scene

[1]Vienna Institute for Safety and Systems Engineering
FH Campus Wien,
University of Applied Sciences
Favoritenstraße 226
A-1100 Vienna, Austria
(e-mail: hans.tschuertz@fh-campuswien.ac.at)

[2]TeLo GmbH,
Gersdorf an der Feistritz 158,
A-8212 Gersdorf an der Feistritz, Austria
(e-mail: florian.wagner@telo.at)

[3]Student of the Master Study Program Safety and Systems Engineering
FH Campus Wien,
University of Applied Sciences
Favoritenstraße 226
A-1100 Vienna, Austria
(e-mail: wilko.schroeter@stud.fh-campuswien.ac.at)

[1]Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering,
Department of Automotive Technologies
Stoczek street 6.
H-1111 Budapest, Hungary
(e-mail: zsolt.szalay@auto.bme.hu
arpad.torok@auto.bme.hu)

Introduction

The Vienna Institute for Safety and Systems Engineering (VISSE) has been working on the topic "Inherent System Safety for Autonomous Systems" for several years. Students of the master study course "Safety and Systems Engineering" are partly integrated into these research activities. In the lecture "Interdisciplinary Safety Project" of this master study program, students were assigned to conduct a “Systems (of Systems) Safety Analysis” for a road intersection with autonomous vehicles.

Several critical situations in a defined driving scenario had to be identified, analyzed, and evaluated according to a new systematic-methodical approach, developed by VISSE. The evaluation part was executed in ZalaZONE. Our research partners Budapest University of Technology and Economics and ZalaZONE Proving Ground from Hungary, as well as our Austrian cooperation partner TeLo GmbH in Gersdorf, supported this study project.

ZalaZONE, the Hungarian test track, is designed according to a complex criteria set to fulfill the requirements of testing autonomous vehicle systems [1]. Following this, the test track has numerous test facilities that can ensure a wide range of test conditions. The proving ground has been designed based on the specifications of the most important industrial actors and scientific organizations in Hungary. The proving ground is constructed on a 265-hectare area and has the following testing facilities: High-speed oval, Dynamic surface, Braking surfaces, Handling courses, Motorway section, Rural road, Smart City Zone [2].

The applied approach was originally developed by VISSE for a system (of systems) safety analyses in the railway sector for safe autonomous people mover operations. This safety approach will now be transferred to autonomous operation in the road traffic. This student project is an important building block for demonstrating that this approach also works in the more complex area of road traffic. This paper gives an overview of the approach and important results of the tests in ZalaZONE.

Chapter 2 describes the safety analysis approach to identify critical situations and the resulting safety goals for avoiding accidents and incidents. Chapter 3 refers to the exact calculations of the critical situation with respect to the existing sensor technology and the derivation of safety requirements.

Chapter 4 describes the identification and classification procedure with the Yolo software. The test scenarios performed in ZalaZONE and the evaluation of the test data are described in Chapter 5. Chapter 6 concludes the results and shows improvement possibilities.

1 SoS Safety Analysis Approach

The challenges for an autonomous world lie on the one hand in defining the necessary legal aspects and, on the other hand, in coping with the high degree of technical complexity [3]. Autonomous systems can be a concrete and immediate threat to human life. Therefore, the state is obliged to set up a legal framework for the development and evaluation of these systems to ensure the human right to life [EGMR 08.11.2005 34056/02, §164].

Current functional safety standards will remain important, but they do not cover the overall-safety aspects. The ISO 26262 safety standard for road vehicles [4] defines functional safety as an absence of unreasonable risks due to hazards caused by malfunctioning behaviour of E/E systems. It provides requirements and guidance to control random hardware failure and to avoid systematic faults. Functional safety refers to that part of safety that depends on the correct functioning of the safety-related system. Functional safety only covers a small part of the overall safety. That means, in the context of autonomous driving, not only endogenous hazards have to be considered, also exogenous hazards have to be identified. Safety must be considered beyond the system boundaries of an E/E system, respectively, of a vehicle. The safety of autonomous systems is more than a component and/or a system property; it is a System of Systems property.

Accordingly, the ISO/PAS 21448 Safety Of The Intended Function (SOTIF) [5] standard answers some of the questions related to the safety of automated vehicle functions. It proposes to divide the space of the possible operation scenarios of the investigated system into known, unknown, safe, and unsafe scenarios. Based on this, the main objective of the approach recommended by the standard is to extend the set of known and safe scenarios as much as required. However, we have to mention that the infinite spread of autonomy of cooperative systems results in an infinite number of possible scenarios indeed. Therefore, further model development efforts are needed to combine both the bottom-up, inductive, and top-down, deductive analytical safety approaches [6], [7].

The concept of systems used in functional safety standards has to be extended to larger systems by analytically considering all existing systems in a kind of a pre-defined super-system, their interdependencies, and emergences. A System of Systems (SoS) is a set of heterogeneous entities that operate independently to each other but related to each other on an undefined level. The identification of the SoS and all its possible entities, e.g. the EGO system, cars, cyclists, pedestrians and so forth, is the basis for safety analysis. The following picture shows the process.

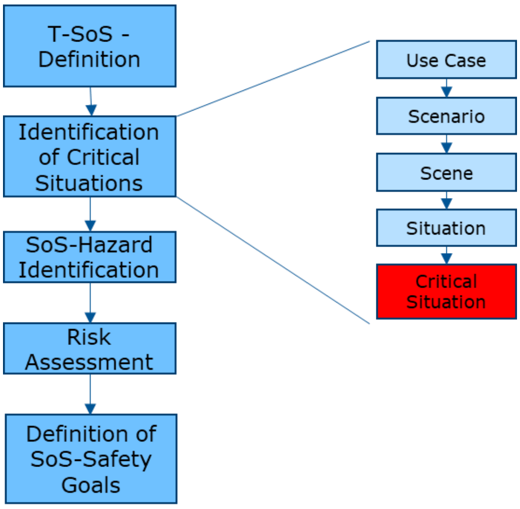
T-SoS - Definition
Identification of Critical Situations
SoS-Hazard Identification
Risk Assessment
Definition of SoS-Safety Goals
Use Case
Scenario
Scene
Situation
Critical Situation

Figure 1: SoS Safety Analysis Approach

After the SoS definition, all the possible use cases in the SoS environment will be identified. A use case describes the service, which has to be provided by the EGO-System in its defined application (e.g., road intersection). It covers the complete driving scenario with all the necessary scenes. In each scene, possible situations are identified. Each situation is transferred into a critical situation trigger by possible events that can occur. Such critical situations are the basis for a SoS Hazard identification. An SoS-Hazard is a critical situation that can cause injury or death to humans and/or loss of the EGO-System and/or other systems and/or damage to the environment.

Based on a subsequent risk assessment, necessary SoS-Safety Goals will be derived. The SoS-Safety Goals provide the basis for safety requirements of all available entities, especially for the required technologies that are necessary for safe traffic. One of the most important technologies in the context of autonomous driving is the sensor concept.

2 Sensor Evaluation

In the following sections, the investigation focuses on the requirement of “environment perception related sensor systems” (hereinafter mentioned as “sensor”). Safety requirements for the sensor technology must be included in the appropriate specification according to a mature requirements engineering process. Additional requirements for the sensor system result from the required ability of the measured values and data. This results from the requirements of the neighboring "sister" systems for the sensors. The measured values and data must contribute to the fact that the environment can be captured by the functional chain: sensor, sensor fusion, and semantic understanding (see Figure 2).

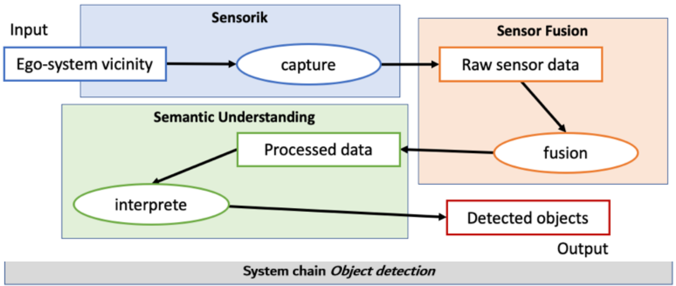

Figure 2: Functional chain “Object detection”

There are four fundamental questions to be answered when defining the sensor requirements:

1. Which objects have to be recognized?
2. How big are these objects?
3. What are the smallest relevant dimensions of these objects?
4. At what distance from the EGO vehicle must these objects be recognizable?

Beyond the above-mentioned four aspects, the system should also fulfill the requirements related to the real-time image processing in the case of transportation-related object detection processes. In accordance with this, the requirement specification needs to put considerable emphasis on the computational complexity and on the time demand of the applied object detection system model.

We note that beyond the response time of the system, the synchronization of the sensors is also a crucial issue in evaluating the safety characteristics of a sensor fusion based environment perception system. It is an outstandingly important issue how we merge signals with different frequency and how we utilize the intermediate data of the denser signals.

However, this study primarily focuses on the spatial perception aspects of the detection process, especially considering the spatial extent and position of the detected object.

Accordingly, the dimensions of the objects are relevant to define the horizontal and vertical boundaries of the area that must be covered by the sensors in order to detect the object in the identified critical situation. In addition to the requirements for horizontal and vertical coverage, requirements for the resolution of the sensors are also important. The smallest relevant dimensions are those that must be recorded by the sensors in order to enable the classification of the objects. An example to illustrate this: a bicycle is approx. 2m long and 1.5m high. An algorithm dealing with object classification will not be able to recognize the bicycle as such unless the rods of the bicycle frame can also be recognized from the measured values and data. The smallest relevant dimension would, therefore, be 3 cm.

One way to answer these questions is described in detail in [8] in the context of self-driving railway vehicles.

This method was adapted and used for analyzing the identified critical situations in detail. Figure 3 shows one of the analyzed critical situations in an intersection.

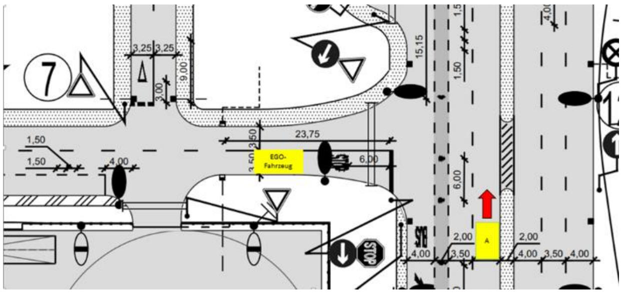

Figure 3: Example of an identified critical situation in ZalaZONE

The chosen situation involves only one object (right-hand vehicle in yellow, abbreviated A) that must be detected by the sensor concept. The following calculation, therefore, only reveals a part of the necessary sensor requirements. The exhaustive application of the situation analysis method with the subsequent determination of the respective sensor requirements gives the complete requirements.

The method is briefly described by using this example. The EGO vehicle at the intersection must be able to either safely cross the intersection or stop at the stop line in good time. For this purpose, the vehicle A on the right must be recognized so far from the intersection that the EGO vehicle can make a qualified decision. The range and opening angle of the sensors are thus largely determined by the speeds of the vehicles involved and the braking distance of the EGO vehicle.

The dimensions - length l, width w and height h - of the vehicles in the present situation must be roughly defined. While the dimensions of the EGO vehicle (l_{EGO}, w_{EGO}, h_{EGO}) can be precisely determined, the dimensions of the vehicle approaching the intersection from the right-hand side (l_A, w_A, h_A) must be determined based on a worst-case estimate to cover all possible variants of the object class *car*.

The trajectories of the EGO vehicle must be defined. This determines a critical timespan, which affects the further calculation. Firstly, the speed of the EGO vehicle approaching the intersection is set. Based on the speed, the braking distance $s_{Brake,EGO}$ can be calculated. If the vehicle is to stop in front of the stop line in good time to avoid a crash, the EGO vehicle must decide at a distance $s_{Brake,EGO}$ ahead of the intersection whether it should stop or drive. The distance $s_{Int,EGO}$ is the length of the trajectory of the EGO vehicle through the intersection area including the distance $s_{Brake,EGO}$ before the intersection. At the end of $s_{Int,EGO}$ the EGO vehicle has reached the target speed. Then a driving profile has to be created, which defines the speed at which the EGO vehicle drives through the immediate intersection area. This gives the timespan $t_{Int,EGO}$. This is the time that the EGO vehicle needs to travel the distance $s_{Int,EGO} + l_{EGO}$. The length of the vehicle must be taken into account here since the EGO vehicle enters the dangerous intersection area with the front and leaves it with the rear. Now it can be calculated from which distance $s_{Int,A}$ vehicle A, which is approaching the intersection from the right-hand side, could dangerously verge onto the EGO vehicle. Based on $t_{Int,EGO}$ and the speed v_A of the vehicle A coming from the right, $s_{int,A}$ can be calculated: $s_{int,A} = t_{int,EGO} \times v_A$. For the brake or drive decision, the EGO vehicle must fully capture the vehicle A at a distance of $s_{int,A}$ from the potential collision point, by the sensors. Before one can determine the sensor requirements, a measuring time $t_{Meas,EGO}$ must be taken into account. During this time, the EGO vehicle determines the speed of the vehicle on the right. The approach speed of the EGO vehicle $v_{EGO,0}$ influences the distance $s_{Meas,EGO}$ that the EGO vehicle travels during the measurement time. Similarly to this, v_A sets the travelled distance $s_{Meas,A}$ of the vehicle A within the investigated time frame $t_{Meas,EGO}$. The required range r_{sens} of the sensors thus results from

$$r_{sens} = \sqrt{(\mathrm{s_{Break,EGO}} + \mathrm{s_{Meas,A}} + \mathrm{d_{buf}})^2 + (\mathrm{s_{Int,A}} + \mathrm{s_{Meas,A}})^2}$$

where d_{buffer} is the distance between the stopping line and the potential collision point. The minimum vertical opening angle $\alpha_{Sens,vert}$ results from the sensor range r_{Sens} and the height of the right-hand vehicle h_{RV1}: $\alpha_{Sens,vert} = 2 * \arctan [h_{RV1} / (2 * r_{Sens})]$.

The determination of the minimum sensor resolution is not trivial. It is determined by the smallest relevant dimension of interest. This depends on the algorithms that process the measured values and data. The discussion in the following section shows that a few pixels are sufficient for image object classification algorithms. [9] dealt in their work with the processing of LiDAR data and empirically defined a formula that defines resolution requirements. If one wants to resolve a 15cm detail at a distance of 100m, the LiDAR must have a horizontal and vertical resolution of 0.009 °. LiDAR sensors for driving applications currently do not offer this. Further research efforts will still be necessary here to improve object recognition and resolution in this area. Subsequently, the calculation method for determining the requirements in this area must be improved and expanded.

3 Object Detection and Classification

Object detection is the process of detecting objects and their bounding box in an image. A bounding box is the smallest rectangle of an image that contains an object completely.

A common input for an object detection algorithm is an image. A common output is a list of bounding boxes and object classes. For each bounding box, the model outputs the corresponding predicted class and its probability.

In object recognition, precision means the probability that the predicted bounding boxes match the detected boxes. Precision is also referred to as the positive predicted value and is calculated as the percentage of diagnosed objects out of all detected bounding boxes (including the boxes included as 'false positive'). It measures how accurate the predictions are.

Recall is the true positive rate, also called sensitivity, which measures the probability that all objects will be detected. It is the percentage of all existing bounding boxes (including those not included as 'false negative') that are diagnosed. Recall measures how well all objects are found.

Choosing a threshold of accepted objects for precision and recall is always a compromise, as both parameters are in a trade-off relationship. When a model detects pedestrians, a high recall rate is usually chosen so as not to miss a passer-by, even if this means stopping the car from time to time without a valid reason. When a model detects investment opportunities, high precision is chosen to avoid choosing the wrong opportunities, even if this means missing some.

YOLO (You only look once) was used for object recognition on the test track. This object detection system, based on Convolutional Neural Networks, divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted with the predicted probabilities.

In YOLO, class possibilities are learned for individual images using bounding box coordinates, and recognition is performed by regression. YOLO divides each input image into a grid of size S x S. Each grid cell has the task of locating the object if the center of this object falls into a grid cell. YOLO simultaneously performs a classification and localization problem for each of the grid cells. Each grid cell predicts N bounding boxes and their confidence. The confidence reflects the precision of the bounding box and whether the bounding box actually contains an object despite the defined class. YOLO predicts the classification value for each box and class.

This predicts the total number of S x S x N boxes. If the confidence of the boxes is below a threshold value, the boxes are discarded.

The version 3 of YOLO used here works with 3 scales for the division into grid cells, where the size of the input image is reduced by 32, 16, 8 to detect small, medium and large objects.

This model has several advantages over classifier-based systems (such as R-CNN [10], Fast R-CNN [11], Faster R-CNN [12]), where only the generated region suggestions are considered in the first step. This context-related information helps to avoid false-positive results.

Classifier-based systems analyze an input image in two steps:
1. identification of promising regions of interest (ROI) in an input image, which probably contains foreground objects. This is done with a Region Proposal Network (RPN).
2. calculating the object class probability distribution of each ROI using a Classification Network - i.e., calculating the probability that the ROI contains an object of a particular class. Then the object class with the highest probability is selected as the classification result.

In contrast, YOLO considers the entire image at test time, so its predictions are influenced by the global context in the image. It also makes predictions with a single network evaluation, unlike systems like R-CNN, which require thousands for a single image. This makes it extremely fast.

However, YOLO has to struggle with smaller objects because of the way it detects objects. For example, it would have difficulty detecting individual birds from a flock. As with most deep learning models, it has difficulty correctly recognizing objects that are too different from the training set (unusual aspect ratios or appearance). The latest version used here, YOLOv3, can run at more than 170 frames per second (FPS) on a modern GPU at a frame size of 256 × 256.

YOLO was first released in 2015 ([13]) and surpassed almost all other object recognition architectures, both in speed and accuracy. Since then, the architecture has been improved several times (YOLOv2 [14] and YOLOv3 [15]). The latest version, YOLOv3, has, among other things, predictions of boxes in various dimensions. In addition, the network has been greatly enlarged with 53 layers. The smallest bounding box size is 13 x 13 pixels.

As a platform for object detection, "TensorFlow 2" with the model "YoloV3" was used. The 80 object classifications were used with the Microsoft COCO dataset ([16]). The weights of the pre-trained network were taken from the official "DarkNet GitHub Repository":

4 Test-Scenarios and Data-Evaluation

According to the SoS Safety Analysis Approach, the entire intersection was demarcated, the driving scenarios defined, and the critical situations identified beforehand. The sensor technology used on the EGO vehicle is an important basic element for identifying critical situations that are not addressed in detail in this paper.

Two critical situations were chosen for the tests in the ZalaZONE proving ground.

Situation 1 (see figure 3): A vehicle on the main road on the right-hand side (viewed from the EGO vehicle) had to be identified in time to avoid an accident.

Situation 2 (see figure 4): A rolling football in front of the EGO vehicle had to be identified and classified in order to be able to draw conclusions about any following persons, to avoid an accident.

The test results were evaluated on the basis of the sensor data with the YOLOv3 software. In preparation for YOLOv3, the camera data was rendered from 960 x 604 pixels to 256 x 256 pixels, and an image brightener was used for the camera shots due to the relatively dark environment. Microsoft COCO was used to classify the detected objects using a pre-trained network.

The camera data was first recorded using the GPU (Graphics Processing Unit) during the test drives and evaluated after the end of the test using YOLOv3.

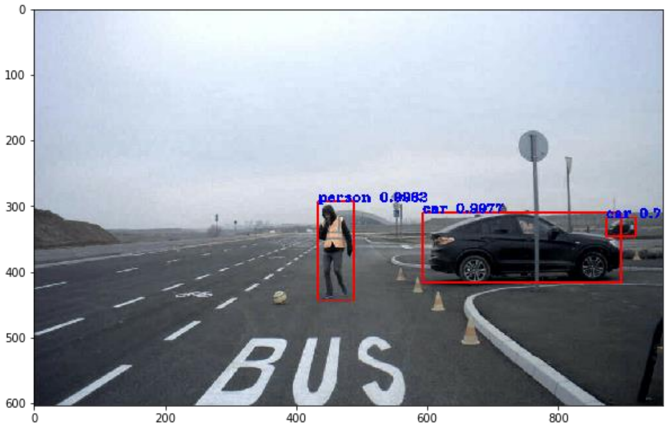

Figure 4: Frame from the test series including recognized persons and cars with corresponding probability

YOLOv3 was applied to the input images in a frame-wise manner. Only the pre-trained network was used for the evaluation. There was no supervised learning by hand-mapped bounding boxes from the input images to achieve a comprehensible result. The frames were then compared regarding the prediction probabilities and the detected objects from the COCO classification. YOLOv3 is able to evaluate the individual frames of the cameras in real-time. Cars and persons were detected very well and for the critical situation with very high probability.

Due to the triple scaling of YOLOv3, even relatively small objects could be detected if they were cars or persons, although with a low predictive probability (e.g., in figure 4, the car at the right edge of the image with a prediction probability of 0.7).

The football, appearing on the road in front of the EGO-vehicle, was not detected regardless of their size. Although Microsoft COCO has a corresponding category for these objects ("football").

5 Conclusion

Both the test scenarios and the test results in ZalaZONE have shown that the SoS Safety Analysis Approach is also working in the more complex road traffic. On the one hand, critical situations could be evaluated in advance. On the other hand, it provides the possibility to improve a used sensor concept beforehand.

This offers the opportunity to reduce the complexity of the driving scenarios because such an approach can avoid most of the unknown situations. Unknown situations in such complex SoS area can leave individual systems uncontrolled, leading to a completely new, formally non-existing type of accidents, which we call SoS-Accidents. A certain part of these are related to autonomous systems, and therefore would probably never occur in case of human control.

Many new findings were also gained during the tests in ZalaZONE. For example, an improvement of the results, especially for predictions in dynamic tests, can be expected by localizing the vehicle and sensor fusion of camera, GPS, and LiDAR data using Kalman filters.

The results of object recognition using YOLOv3 show the ambivalence of these methods based on convolutional neural networks (YOLO, Faster R-CNN): The convolution layers reduce spatial dimension and resolution. Therefore, the models can only detect relatively large objects. For small objects, the pre-trained networks have to concentrate on certain objects to achieve sufficient accuracy; on the other hand, the networks should work as universally as possible. A possible way out of this dilemma would be a situation analysis carried out in real time during autonomous driving and then the use of pre-trained networks for the respective detected situation; in the test cases, for example, a concentration (= weighting) on objects such as "car", "bicycle", "person", "bus" would be conceivable for the detected situation "road traffic", while other categories such as "tvmonitor", "sofa", "toilet" would have a lower weight.

The results have also shown that autonomous systems can be an immediate threat to human life. Therefore, a legal framework for developing and evaluating systems for autonomous driving will be necessary to ensure the human right to life. The SoS Safety Analysis Approach is an important tool for the development of complex systems for autonomous driving and can also be extended to a safety case.

Reference List

[1] Steinmetz, E., Emardson, R., Eriksson, H., Hérard, J., & Jacobson, J. (2011). High Precision Control of Active Safety Test Scenarios.

[2] Szalay, Z., Hamar, Z., & Simon, P. (2018, June). A Multi-layer Autonomous Vehicle and Simulation Validation Ecosystem Axis: ZalaZONE. In International Conference on Intelligent Autonomous Systems (pp. 954-963). Springer, Cham.

[3] Bogya, P., Vass, S., & Tihanyi, V. LONGITUDINAL CONTROL WITH FUZZY SYSTEMS FOR AN AUTOMATED VEHICLE.

[4] ISO 26262: Road vehicles - Functional safety, International Standard, ISO International Organisation for Standardisation, Geneva, Switzerland, 2018

[5] Road vehicles – Safety of the intended functionality. ISO/PAS 21448:2019.

[6] ben Othmane, L., Al-Fuqaha, A., ben Hamida, E., & van den Brand, M. (2013, October). Towards extended safety in connected vehicles. In 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013) (pp. 652-657). IEEE.

[7] Zöldy, M. (2018). Investigation of autonomous vehicles fit into traditional type approval process. Proceedings of ICTTE, 428-432.

[8] Wagner, F., 2019. Sicheres Sensorkonzept für autonome Schienenfahrzeuge.

[9] Pesci, A., Teza, G., & Bonali, E. (2011). Terrestrial Laser Scanner Resolution: Numerical Simulations and Experiments on Spatial Sampling Optimization. Remote Sensing, 3, 167-184.
DOI: 10.3390/rs3010167

[10] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 2014, pp. 580-587.

[11] Ross Girshick: Fast R-CNN: The IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440-1448.

[12] Ren Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems, 2015, pp. 91-99.

[13] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi: You Only Look Once: Unified, Real-Time Object Detection, 2015.

[14] Joseph Redmon, Ali Farhadi: YOLO9000: Better, Faster, Stronger, 2016.

[15] Joseph Redmon Ali Farhadi: YOLOv3: An Incremental Improvement, 2018.

[16] Tsung-Yi Lin et al., Microsoft COCO: Common Objects in Context, 2015.