

# Autonomous Driving and Artificial Intelligence An Approach to Achieve Functional Safety

## Introduction

In the past few years, the automobile industry has been dramatically improving the features implemented in vehicles aimed at assisting the driver in many different traffic situations. Commonly named Advanced Driver Assistance System (ADAS), those features range from parking assistance to highway lane keeping assistance even involving more complex situations such as emergency steering and braking assistance.

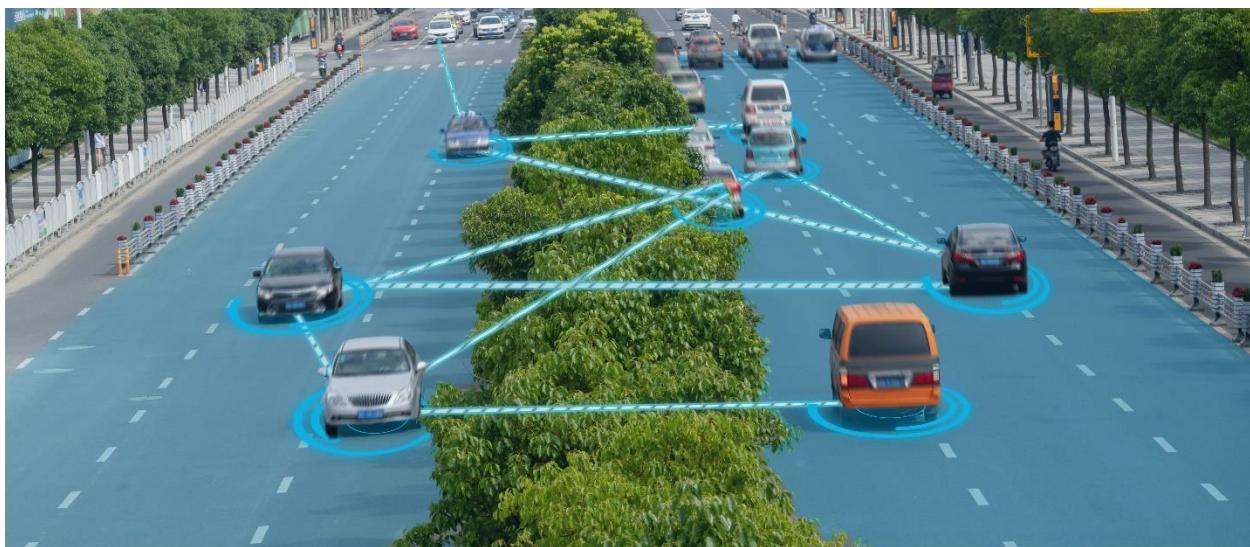


Figure 1: While Advanced Driver Assistance Systems only require a subset of those sensors, Level 4 and L5 make intensive usage

To apprehend increasingly complex AD tasks, the application needs to create more elaborate and more accurate representations of the traffic situation, through an increasing variety of sensors, in number and types (refer to Fig. 1). As a consequence, the AD system becomes more and more complex and hence is increasingly subject to various types of faults. An undetected fault in any of the components in this AD system can have dire consequences. To cope with this undesired situation, Functional Safety has been widely used in the industry to eliminate unacceptable risks that can lead to physical injuries. The objective of a safe system is to ensure that faults are detected and that the system implements enough safety mechanisms to keep or bring the whole system back into a safe state.

Nevertheless, with higher level of AD, the complexity of the AD components forces the applications to shift the algorithms from model based to Artificial Intelligence (AI) techniques such as Machine Learning (ML) or Deep Learning (DL). However much those methods are

improving, they tend to become unstable in the presence of so-called adversarial perturbations.

## The notion of confidence

The basic task of an AD application is to capture the data from the sensors, evaluate the traffic situation, compute a safe trajectory and send it to the actuators for execution. The system suffers from a certain level of uncertainty due the lack of precision of the sensors, hence influencing the confidence one can put in the detections.

The confidence can be increased by using sensor fusion techniques in the spatial domain and target tracking methods in the temporal domain. The Kalman filter is commonly used for that purpose, notably with radars.

Taking the level of confidence of the detections into account in the process of taking decision is critical in AD applications. The level of confidence must be propagated through the processing pipeline and increased or decreased depending on how much each processing step can confirm the presence of the detection. Software components based on AI can be part of the processing pipeline and hence they should not be treated differently. Taking the concept to the full system, from sensor to actuator, it becomes possible to evaluate the system as a whole and the Safety Of The Intended Function (SOTIF)

## Keeping NN components simple

One way to be able to detect mispredictions from a NN is to keep its complexity as low as possible, feed it with features (e.g. velocity and location of objects) and to check the plausibility of the inputs and outputs.

However, keeping NNs simple and small requires implementing a collection of those, each dedicated to a specific task. This approach is subject to divergences when the preprocessing steps are not stable enough in terms of development maturity. It also has the drawback that cascading NNs can cause an increasing inaccuracy and hence impact the SOTIF.

Using an end to end approach can remove this issue. Deep Convolutional NNs such as the VGG16 have demonstrated outstanding performance in object recognition. Its complexity makes it significantly more difficult to understand the cause of mispredictions and hence to design safety mechanisms to prevent them.

The situation becomes even more complicated when the application needs to track objects. Moving objects like Vulnerable Road Users (VRU, e.g. pedestrians) or surrounding vehicles need to be localized very precisely to compute a safe trajectory.

To allow a NN to also give a precise localization of the detected objects, network designs such as the U-Net[2] have been developed. The U-Net adds yet another level in the complexity of the NN. It makes the task of detecting mispredictions even more complicated. On the other hand, because the location of the objects are available, temporal tracking becomes possible not only on the type of object like in the case of the VGG16 but also on the location of those objects.

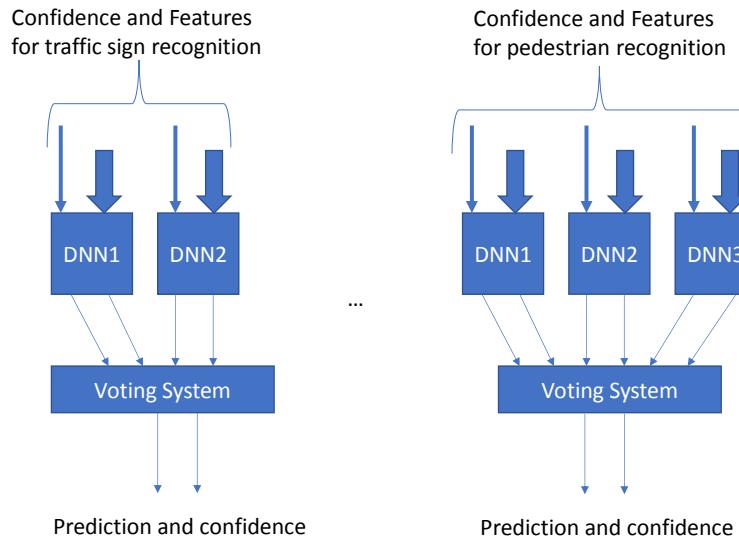


Figure 2: An example of simple redundant neural networks. All are specialized to a very specific task. There are combined in groups of different sizes. Their output is sent to voting system that provides the predictions along with the level of confidence.

One way to reduce the size of a NN based on raw images consists in using a collection of specialized NNs, dedicated to recognizing very specific objects, e.g. traffic signs (refer to Fig. 2). Having simple NNs, it makes it easier to design functional safety mechanisms to detect errors. Being smaller and dedicated to very specific tasks, those NN are usually easier to train.

## Building a failsafe system

Besides implementing safety mechanisms, it is also possible to use redundant AI components. This technique is very similar to any hardware or software redundancy. The idea is to use several implementations based on the same requirements, implemented by different teams and trained on different data sets to avoid systematic faults. The output of the different versions of NNs are sent to a voting system, along with the confidence of their predictions (refer to Fig. 2). This system can then elaborate the true output of the NN set.

## Concluding remarks

Autonomous Driving becomes more complex as it increases its level of autonomy. The number and the diversity of the sensors increase and the AD application requires more and more sophisticated algorithms to capture the traffic situation and the surroundings of the vehicle.

AI is involved in a growing number of processing stages, from perception to situation analysis and trajectory planning. Despite the difficulty of understanding their internal processing, AI component can be treated like any other hardware and software components with respect to functional safety and SOTIF. Functional safety mechanisms can be built around the AI components. They can check the consistency of the predictions both in the spatial and temporal domains. The temporal domain must be handled with caution when it comes to set timing constraints in the safety requirements.

On the other hand, through the ISO-26262 standard, functional safety offers tools to assign probabilities of faults and level of confidence to analyze the complete system involving AI. Safety mechanisms can be designed for the complete system, including the AI components.

The Functional safety and SOTIF standards are of great help to refine the complete software architecture so that all safety requirements can be met and the safety goals can be reached.

## References

- 1 *Very Deep Large Convolutional Networks for Large-Scale Image Recognition*, K. Simonyan and A. Zisserman, 10-04-2015, Published at ICLR 2015,
- 2 *Convolutional Networks for Biomedical Image Segmentation*, O. Ronneberger, P. Fischer and T. Brox, 18-05-2015,

## About KPIT

KPIT is a global technology company with software solutions that will help mobility leapfrog towards autonomous, clean, smart and connected future. With 6000+ Automobielers across the globe, specializing in embedded software, AI & Digital solutions, KPIT enables customers accelerate implementation of next generation mobility technologies. With development centers in Europe, USA, Japan, China, Thailand and India – KPIT works with leaders in mobility and is present where the ecosystem is transforming.

## Marketing Contact:

Stefanie Köhler, Head of Marketing Germany, Phone: +49 89 322 99 66 140, [Stefanie.koehler@kpit.com](mailto:Stefanie.koehler@kpit.com)