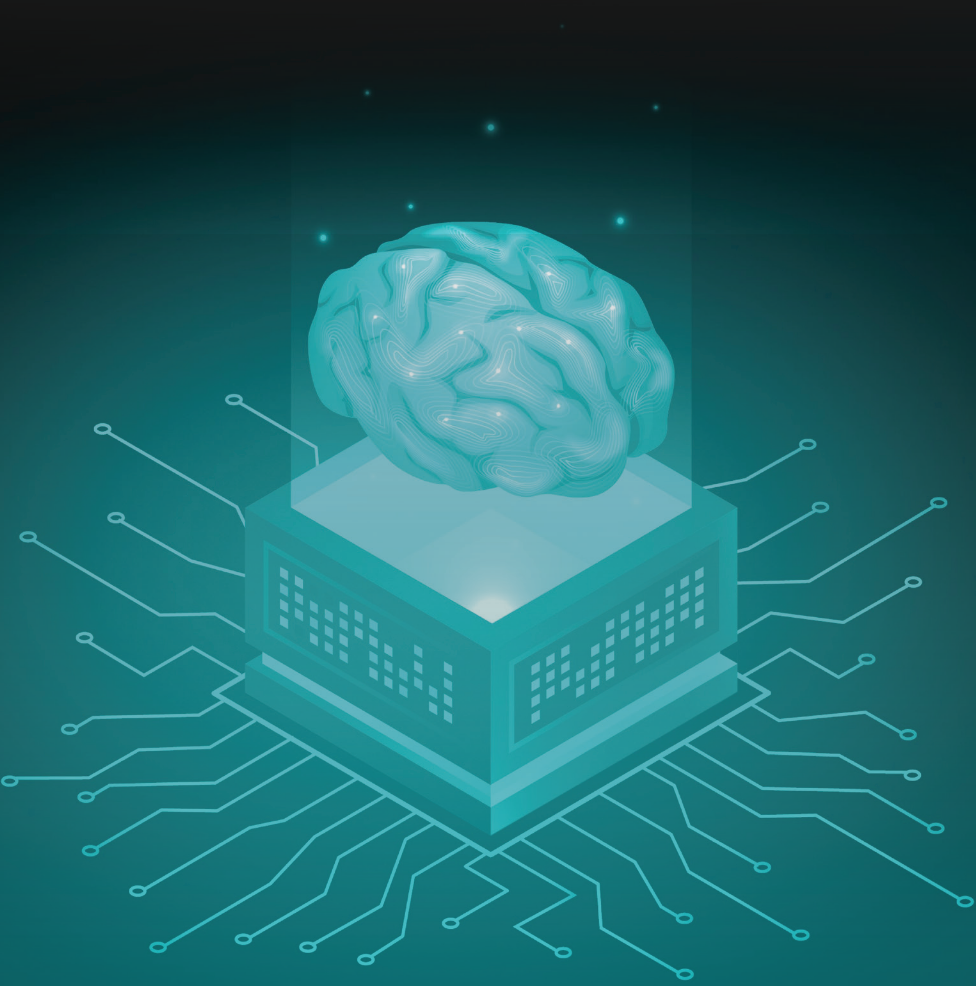
THE
AUTONOMOUS
Expert Circle Safety of Embedded AI

Hosted by: Infineon

AI and next generation MCUs enable safe and advanced autonomous driving



Embedded AI enables next level of safe, comfortable and energy efficient Automated Driving

Stephan Baumgärtner
Infineon Technologies Dresden GmbH
Dresden, Germany
stephan.baumgaertner@infineon.com

Mateusz Chmurski
Infineon Technologies Dresden GmbH
Dresden, Germany
mateusz.chmurski@infineon.com

Mathias Haberjahn
Infineon Technologies Dresden GmbH
Dresden, Germany
mathias.haberjahn@infineon.com

Lukas Marik
Infineon Technologies Dresden GmbH
Dresden, Germany
lukas.marik@infineon.com

Jürgen Schäfer
Infineon Technologies AG
Munich, Germany
juergen.schaefer@infineon.com

Daniel Scholz
Infineon Technologies Dresden GmbH
Dresden, Germany
daniel.scholz@infineon.com

***Abstract*—In 2022, The Autonomous welcomed Infineon Technologies as the Lead of the Expert Circle on Safety of Embedded Artificial Intelligence. In the ongoing Expert Circle, we have been developing dedicated solutions on how to use AI safely for trajectory planning and control. The presented approaches mainly build upon the results of EEmotion. EEmotion was a co-funded project by the German Federal Ministry for Economic Affairs and Climate Action (BMWK), coordinated by Infineon and partially implemented with partners as ZF. During the previous Main Event we had shown the advantages of an AI-Enhanced trajectory controller, namely a 50% improvement in tracking accuracy. Since then we have been extending the approach to trajectory planning by suggesting AI enhanced Model Predictive Control. This improves energy efficiency by 5%. To additionally ensure safety of the AI components we developed a safety blueprint. This includes safety of the employed AI components itself and the data used to develop them. We showcase the conceptual safety blueprint on the example of run time monitors for the AI enhanced trajectory control. Taken together, we could show that AI solutions can reduce system costs, increase energy efficiency and improve passenger comfort while ensuring system safety.**

***Index Terms*—Autonomous driving, Trajectory Control, Trajectory Planning, AI safety, Runtime Monitors, Microcontroller**

I. Introduction

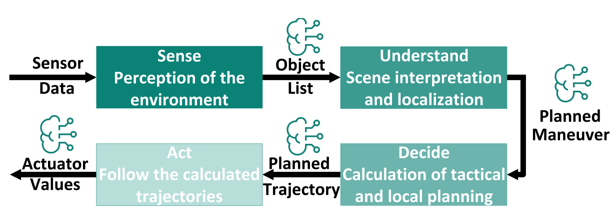


Fig. 1. Autonomous driving compute chain of the conceptual steps from environment perception to actuator values.

The autonomy of self-driving vehicles is based on four major conceptual blocks along a compute chain as depicted in Figure 1. In the first step the vehicle senses its environment, typically via different imaging techniques. In a second step the perceived environment must be interpreted to enable decision making. Here, planned maneuver can for instance be whether or not to perform a lane change. Based on the planned maneuver the vehicle plans the desired tactical driving task as the third step by calculating a trajectory. In the last step this planned trajectory is executed by a trajectory controller.

The first two blocks for environmental perception and interpretation heavily rely on high dimensional image data and AI approaches for computer vision that typically require powerful compute resources. The later blocks in the compute chain are qualitatively different in that their input data is lower dimensional and their execution is real time critical. This makes them good candidates for deployment on automotive microcontrollers such as the Infineon AURIX^{TM}.

In this paper we demonstrate our development efforts on AI enhancements of trajectory planning and control and how we ensure their safety. We describe how we made existing but fixed methods adaptable using machine learning in the next two sections. To ensure safety of the developed AI methods we describe in the following sections an AI safety assurance based on existing and new ISO standards. This is followed by the description of the implementation for run time monitors as integral part of the safety concept.

Data for training and evaluation of the presented AI approaches was based on simulations. We used a software in the loop setup consisting of a digital twin of a modern vehicle implemented in MATLAB/Simulink coupled with IPG CarMaker. This simulation based approach allowed quick iterations during development. For real world assessment the AI enhanced trajectory controller together with the reconstruction based run time monitor where successfully deployed in a test vehicle and show cased their performance improvements.

II. Energy efficient AI-enhanced trajectory planning

Model Predictive Control (MPC) has emerged as a powerful optimization-based approach for trajectory planning in complex environments. By leveraging a predictive model of the vehicle dynamics, MPC enables the computation of optimal control inputs (trajectory) that minimize a performance criterion while satisfying constraints on system states and inputs [1]. Nevertheless, complex MPC formulations often incur excessive computational overhead, rendering them infeasible for deployment on resource-constrained microcontroller units (MCUs). We propose a hybrid approach that integrates Model Predictive Control (MPC) with a Reinforcement Learning (RL) agent [2], which effectively offloads some of the computational complexity while maintaining high adaptability (see Figure 2). Additionally, by moving the energy efficiency terms out of the MPC's cost function into the reward term of the reinforcement learning agent, implementation complexity and computational costs can be reduced.

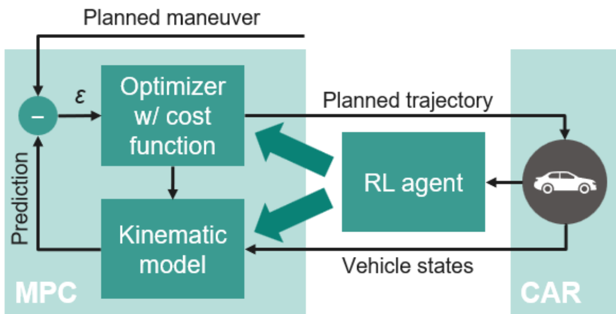


Fig. 2. RL agent running parallel to MPC planner.

The RL approach allows for semi-automated adaptation to arbitrary scenarios [3]. For the energy efficiency use case we selected a longitudinal planning scenario with heavy traffic and rapidly varying velocities, simulating a highway congestion build-up. A single training scenario instance is modelled, which is then parameterized and randomized during the agent training. The agent's reward function is designed to balance the maneuver tracking error and energy consumption.

Using this hybrid approach, we were able to gain up to 5% efficiency improvement in simulation compared to the MPC alone.

III. AI-Enhanced Trajectory Controller

To leverage the energy savings from an AI enhanced trajectory planner, the trajectory controller as the final step in the compute chain has to ensure that the vehicle follows the planned trajectory as closely as possible. Conventional control algorithms suffer from limitations such as the lack of adaptivity to dynamic environments, high computational demands and complex parameter tuning. We present an AI-Enhanced trajectory control approach that alleviates these limitations which has been developed within the German funding project EEmotion together with ZF. AI allows to make a classical PID controller adaptive. We trained a 3-layer MLP mapping information about scene, maneuver and dynamic vehicle parameters into PID gains (see Figure 3). The combination of a PID controller with a small neural network is still numerically simpler than a more complex model predictive controller and at the same time profits from AI accelerators on modern automotive MCUs such as the latest generation AURIX™.

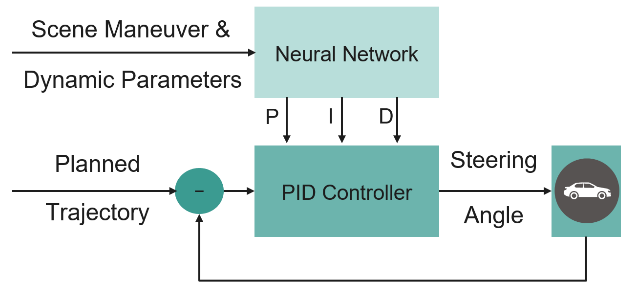


Fig. 3. AI-Enhanced PID Controller.

The proposed solution has been tested in the simulation environment as well as on the test track. As the test scenario in the simulation, we chose Nürburgring and we compared the performance of the proposed solution with conventional - fixed parameter PID. From Figure 4, one can observe that the proposed algorithm exhibits $\sim 50\%$ improvement of the root mean squared error for the cross track error in comparison to conventional PID.

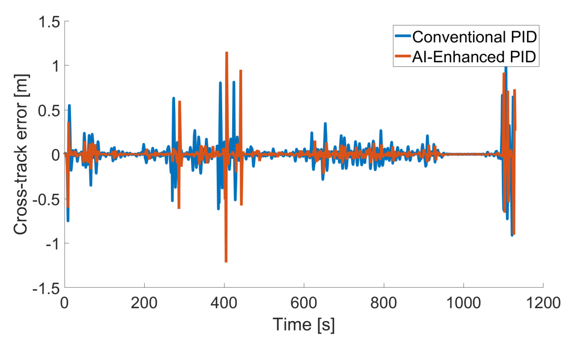


Fig. 4. Tracking accuracy for conventional and AI-enhanced PID processed on Nürburgring in simulation environment.

In the next step, the proposed AI-Enhanced PID controller has been deployed on the AURIX™- TC4x integrated into a test vehicle from ZF and compared with AI-Enhanced MPC on a real test track. Specifically, we tested our algorithm on double lane change scenarios. From the Figure 5, one can notice the proposed solution achieves similar tracking accuracy to AI-Enhanced MPC, while maintaining low-computational overhead for efficient execution on automotive MCU.

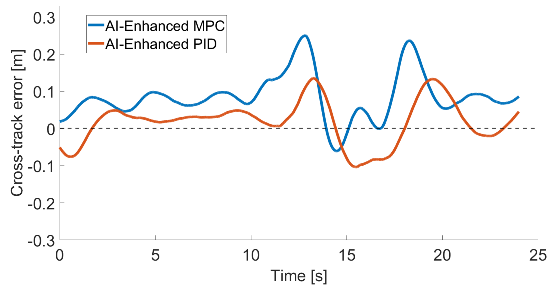


Fig. 5. Tracking accuracy for AI-enhanced MPC (blue) and AI-enhanced PID (red) of a real double lane change scenario on ZF test track.

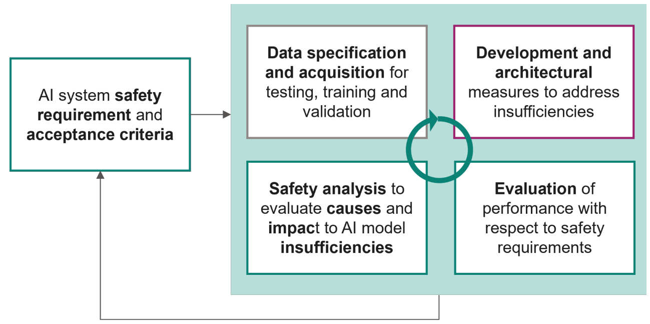


Fig. 6. Schematic of AI safety lifecycle depicting continuous safety assurance.

IV. AI Safety Assurance

The deployment of AI-based system calls for methodologies to cover the safety aspect of their components and overall behavior. The most relevant standards for AI and safety in the automotive domain interact together, namely ISOs 26262, 21448 (SOTIF), and 8800. We show a comprehensive framework, based on the established V-model outlined in ISO 26262, to ensure the reliability and safe deployment of AI technologies within the automotive industry. The activities from SOTIF and ISO 8800 are mapped into this V-model and separated into three levels of abstraction: vehicle, system, and component levels. Additionally, the recently passed EU AI Act imposes regulations for the AI system. Safety-critical systems like AI-based vehicle trajectory planning and control are categorized as high-risk AI systems and have some associated requirements. To comply with the EU AI Act, one must conduct an assessment to determine whether the development process which adheres to the mentioned standards adequately encompasses the regulatory requirements outlined in the EU AI Act for high-risk AI systems.

Since AI models are data-driven, one first must ensure the safety of the data based on standards which include: data completeness, integrity and correctness. Additionally, the correct selection of data has to be ensured by the correct definition of the Operational Design Domain (ODD) and instance scenarios. The safety assurance during the AI lifetime forms a cycle as shown in 6. We describe a procedural approach how to adopt strategies outlined in SOTIF and ISO PAS 8800 and apply them to trajectory planning and control. The latter is drawn from a framework which covers the entire lifecycle of safety-critical AI systems including: Concept phase, AI system development, and AI component development activities as well as verification and validation, and operational measures.

We discuss specific technical approaches that one can put into action like runtime monitors, redundant components for robustness and fault tolerance, and continuous monitoring including collection of data when insufficiencies are detected.

V. Reconstruction based runtime monitoring

Runtime monitors act as an essential part of the safety concept for the previously introduced AI enhanced components. We showcase this here on the example of the trajectory controller. Runtime monitors are necessary on both the input and the output of the controller 7. On the input side monitoring has to ensure that the planned trajectories lie within the specified operational design domain. As the controller is designed for the ODD trajectories outside the ODD could not be handled safely. A runtime monitor on the output side limits the impact of the controller onto actuators.

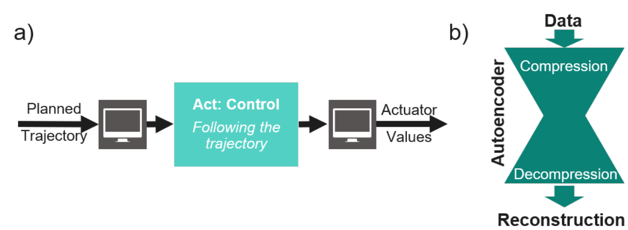


Fig. 7. a) Runtime monitors on the input and output side of the trajectory controller. b) General architecture of reconstruction based outlier detection models: the input data is compressed into a lower dimensional latent space followed by reconstruction back to the original space.

Functionally runtime monitors need to detect when the seen data deviates from the training data making them anomaly detectors. The technical implementation highly depends on the dimensionality and complexity of the data. The output of the trajectory controller is the steering angle. Due to the inertia of the vehicle and the update interval of 5 ms simple linear extrapolation is a sufficient method to detect defective steering angle values.

The planned trajectory as input for the controller has multiple dimension such as position, angle, curvature & velocity. In addition, assessment of the trajectory requires the context of the current driving situation such as the position and velocity of the ego vehicle. Thus, the data on the input side of the trajectory controller is multidimensional and complex requiring a potent monitoring approach. We employed established reconstruction based neural networks so called auto encoders [5] as runtime monitors. An autoencoder compresses the input data to a lower dimensional latent space followed by a decompression back to the original data space. Comparing the input of an autoencoder to its reconstruction on the output allows to calculate the reconstruction error. During training the autoencoder this error is minimised allowing the model to find an effective mode of the compression-decompression scheme for defect free data. We use the reconstruction error as a metric to classify unseen data. Defective input data would lead to a large reconstruction error allowing to effectively detect safety critical trajectories.

TABLE I
STATISTICAL METRICS TO COMPARE THREE DIFFERENT MODEL TYPES. F1: F1 SCORE, FPR: FALSE POSITIVE RATE, FNR: FALSE NEGATIVE RATE

Model	F1 Test	FPR Test	FNR Test
Base Line	0.934	0.184	0.07
Autoencoder	0.981	0.098	0.005
Variational Autoencoder	0.982	0.096	0.004

We tested three different reconstruction based model architectures: 1) Principle component analysis followed by its inverse transform. PCA is a common linear technique for dimensionality reduction and in the given context it acts as a linear base line model [5]. 2) An autoencoder based on a simple multilayer perceptron or fully connected architecture. 3) A variational autoencoder as an extension of the regular autoencoder. As a generative model it can generate multiple reconstructions via repeated sampling [4]. This can provide better prediction stability at the cost of higher computational demands.

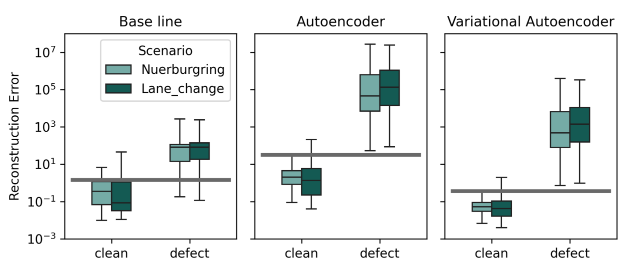


Fig. 8. Distributions of the reconstruction error for training (light green) and test scenario (dark green) for all three tested model types. Horizontal grey lines indicate the threshold values that optimally separate clean from defective data leading to statistical metrics shown in table I.

We used the Nürburging race track implemented in IPG CarMaker as training scenario and tested the models on various lane change scenarios. In both data sets we injected synthetic defects to assess the detection quality. The goal for all three models is to clearly separate clean from defective data as shown in Figure 8. Statistical assessment shows that a simple MLP based autoencoder can classify clean from defective data on better than a linear base line model (Figure 8 and Table I). Increasing model complexity to a variational autoencoder does not improve classification quality substantially.

After deployment of the trained autoencoder on the Infineon AURIX™ automotive microcontroller it has a memory footprint of $\sim$40 kb and a latency of $\sim$12 μs allowing real time execution. We showed the functionality of the autoencoder alongside the described AI enhanced trajectory controller in a test vehicle on a test track with synthetic defect injections.

VI. SUMMARY

In conclusion, the proposed approaches for trajectory planning and control for an autonomous vehicle can be leveraged by utilizing machine learning techniques in several aspects. We showed that the energy efficiency of the driving task is mainly affected by an proper planning of the trajectory.

Using Reinforcement Learning to anticipate energy savings along the longitudinal path can lead to an reduction of up to 5% for the energy consumption. However, we strongly belief that there is even a higher potential extending the approach to act on more planning parameters and complex scenarios.

We found that using a simple MLP to adjust the PID gains during runtime can improve the overall tracking accuracy by up to 50% compared to a non AI PID trajectory controller. In addition, we showed that the AI-driven PID performs in a similar magnitude of accuracy related to a nonlinear AI-enhanced MPC by consuming much less computational resources on the AURIX™- TC4x.

Regarding the safety assessment of AI-based systems, the blue print we describe can be applied to different applications when the individual details are adapted. One must bring safety elements into the context of the target application.

To ensure safe operation of our proposed AI functionalities we developed the safety assessment based on existing and new standards. Run time monitors are an integral part of the safety approach. We showed that reconstruction based neural networks can act as runtime monitors for an AI enhanced trajectory controller and reliably detect defective trajectories.

REFERENCES

[1] Pendleton, S.D.; Andersen, H.; Du, X.; Shen, X.; Meghjani, M.; Eng, Y.H.; Rus, D.; Ang, M.H. Perception, Planning, Control, and Coordination for Autonomous Vehicles. Machines 2017, 5, 6. https://doi.org/10.3390/machines5010006

[2] H. Xie, X. Xu, Y. Li, W. Hong and J. Shi, "Model Predictive Control Guided Reinforcement Learning Control Scheme," 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 2020, pp. 1-8, doi: 10.1109/IJCNN48605.2020.9207398. keywords: Training;Linear systems;Delay effects;Simulation;Training data;Reinforcement learning;Prediction algorithms;Deep reinforcement learning;Model predictive control;Time delay;Process control

[3] Kuutti, Sampo, et al. "A survey of deep learning applications to autonomous vehicle control." IEEE Transactions on Intelligent Transportation Systems 22.2 (2020): 712-733.

[4] Iqbal, Touseef, and Shaima Qureshi. 2023. "Reconstruction Probability-Based Anomaly Detection Using Variational Auto-Encoders." International Journal of Computers & Applications 45 (3): 231–37.

[5] Ruff, Lukas, Jacob R. Kauffmann, Robert A. Vandermeulen, Grégoire Montavon, Wojciech Samek, Marius Kloft, Thomas G. Dietterich, and Klaus-Robert Müller. 2020. "A Unifying Review of Deep and Shallow Anomaly Detection." arXiv [Cs.LG]. arXiv. http://arxiv.org/abs/2009.11732.

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag