Tool Ecosystem for testing and increase robustness of AI

Continental Partners involved

- Continental Automotive GmbH, Siemensstr. 12, 93055 Regensburg
- Continental Automotive Hungary Kft., Házgyári út 6-8, 8200 Veszprém, Ungarn
- Automotive Distance Control Systems GmbH, Peter-Dornier-Str. 10, 88131 Lindau

Objective of the project proposal

This project defines and implements a tool ecosystem for testing and increase robustness of automotive AI applications. The tool ecosystem shall help to establish an automotive industry-wide accepted way of validating AI modules in automotive products. The tool ecosystem shall be available for use by any interested party.

Description of the approach to reach the objective

Terminology:

- The term “AI module” refers to individual modules of safety-critical automotive products/systems that use some kind “AI”, i.e., Deep Learning or any other Machine Learning methodology. The common characteristic of such “AI” methods is the use of complex models which consist of a large number of parameters that are determined using some algorithmic approach. Neither the model itself nor its parameters’ correctness can be verified using traditional Software Engineering methods, which creates the need for alternative validation approaches. The tool ecosystem shall enable such validation.
- The term “tool ecosystem” refers to a set of tools which enable validation of AI modules. For reduction of complexity and easier development, individual aspects of AI module validation shall be covered by separate tools. Redundancies shall be avoided.

Scope: The tool ecosystem covers at least the following aspects of AI module validation:

- KPIs (e.g., OD-metrics like mAP, …)
- Resiliency/robustness with regards to adversarial attacks
- Processes (data relevance, data selection, proper definition of datasets, proper training methodology, …)
- Statistics of parameter distribution (weight distribution)
- Compilation checking
- Methods for explainable and interpretable AI

Testing approaches include, depending on the individual problem:

- **white box testing**, i.e., testing with full insight into the internals of an AI module
    - Robustness with respect to adversarial attacks/manipulated inputs
    - Validation of continuity of cost function in case of Deep Learning
    - “Coverage analysis”, i.e., testing neural networks on a large amount of training data in order to test if certain activations are always zero – potential error sources
    - …
- **black box testing**, i.e., testing of an AI module without any knowledge of its internals
    - KPI computation (implementation very problem specific, not to be part of this project)
    - Performance loss of deployed models
    - Verification of deterministic behavior

    - …

Testing levels include:

- **unit testing**, i.e., testing of an AI module on it's own, for example within it's Deep Learning training framework's context
- **integration testing**, i.e., testing of the AI model after it has been transformed to be used in the final product. This includes among other things…
    - compilation to target a certain hardware accelerator for Machine Learning, which introduces the possibility of having an incorrect compilation of some aspects of an AI module
    - fixed-point conversion, which introduces error sources like quantization noise and potential overflow/saturation errors
    - loss of precision
    - change of data formats, e.g., channels first vs. channels last

Some of these methods can be regarded as the transfer of best practices from Software Engineering to AI Development.

Expected Contributions

- For well-specified problems, implementations shall be developed (e.g., running tests with different permutations of subtests and data)
- For potentially new suitable evaluation methods, research shall be done
- Requirements for suppliers shall be created (e.g., specification of guidelines for compiler suppliers or HW accelerator suppliers)

Work model:

The idea of this project is to collaborate with partners in the industry, including, but not limited to:

- OEMs
- Suppliers of AI-based automotive products/systems
- Suppliers of relevant development tools, e.g. suppliers of Neural Network compilers or Neural Network accelerator hardware

Such partners are expected to represent their requirements on the tool ecosystem, and to contribute to the implementation of tools.