TECHNISCHE UNIVERSITÄT DRESDEN

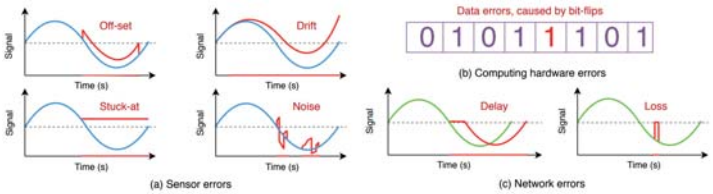
Fakultät für Elektrotechnik und Informationstechnik • Institut für Automatisierungstechnik

Reliability Engineering meets Artificial Intelligence (AI)

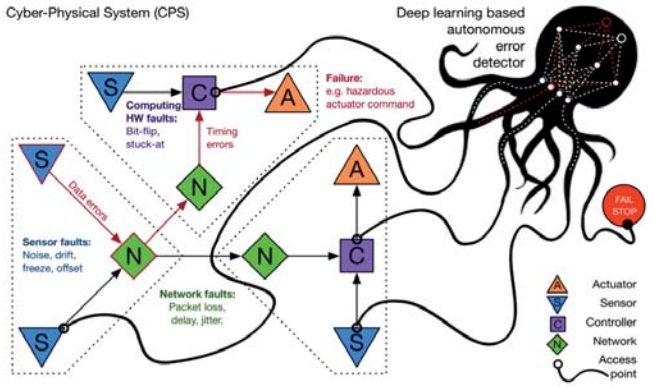
Deep learning based error detection[1]

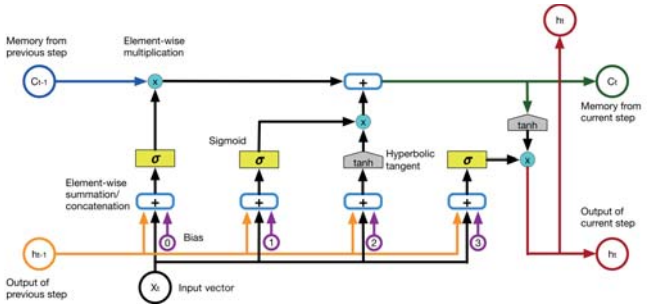
AI for error detection
- Conventional Fault Detection and Isolation methods shows low performance with the increasing of systems complexity.
- Deep learning techniques, e.g., training a neural network to distinguish between normal and erroneous system behavior, are promising solution.

Cyber-physical systems (CPS)
- Consist of numerous heterogeneous components.
- Generate big data.
- Form complex interactions and system operational profiles.
- Prone to various errors types.



Online error detection and mitigation using LSTM networks
- The signals of a CPS are mainly time series data.
- Recurrent neural networks (RNN) can use an internal state (memory) to process sequences of inputs and are applicable to error detection in time-series data.
- However, RNNs fail to capture the context as time steps increase.
- Long short-term memory networks (LSTM) overcome vanishing gradient problem.

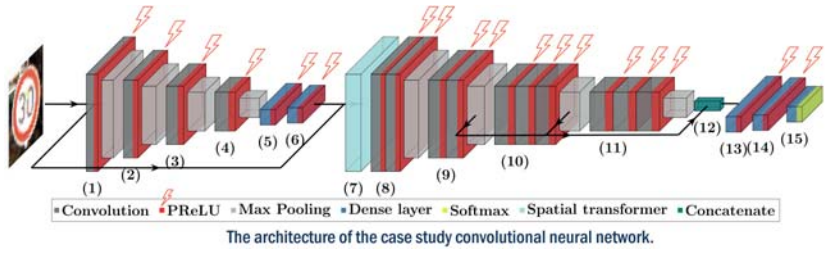




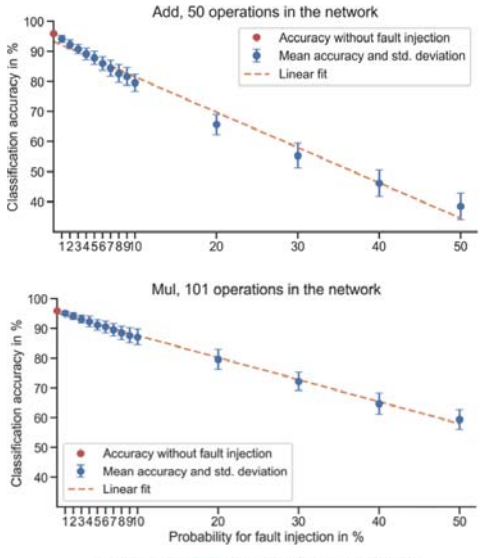
[1] Ding, K., Ding, Sh., Morozov, A., Fabarisov, T., Janschek, K.: On-line Error Detection and Mitigation for Time-series Data of Cyber-physical Systems using Deep Learning based Methods. 15th European Dependable Computing Conference (EDCC 2019), 17-20 September, 2019, Naples, Italy.

Reliability evaluation of safety-critical AI applications[2]

Safety-critical AI applications
- AI enters almost every safety-critical domain, including the automotive industry.
- AI-based applications are prone to common random hardware faults which might lead to silent data corruption.
- The next generation of functional safety standards has to define appropriate verification and validation techniques and fault tolerance mechanisms.
- It is crucial to understand how different hardware faults affect the accuracy of AI applications.
- Our fault injection framework manipulates the outputs of TensorFlow mathematical operations: For each operation type, the user can specify a fault type as well as an error injection probability.



The architecture of the case study convolutional neural network.



Results of the Bit flip fault injection experiments.

[2] Beyer, M., Morozov, A., Ding K., Ding, Sh., Janschek, K.: Quantification of the Impact of Random Hardware Faults on Safety-critical AI Applications: CNN-based Traffic Sign Recognition Case Study. ISSRE 2019.

Technische Universität Dresden
Fakultät Elektrotechnik und Informationstechnik
Lehrstuhl für Automatisierungstechnik
Prof. Dr. techn. Klaus Janschek

Tagir Fabarisov
+49 351 463-32202
tagir.fabarisov@tu-dresden.de
BAR E29