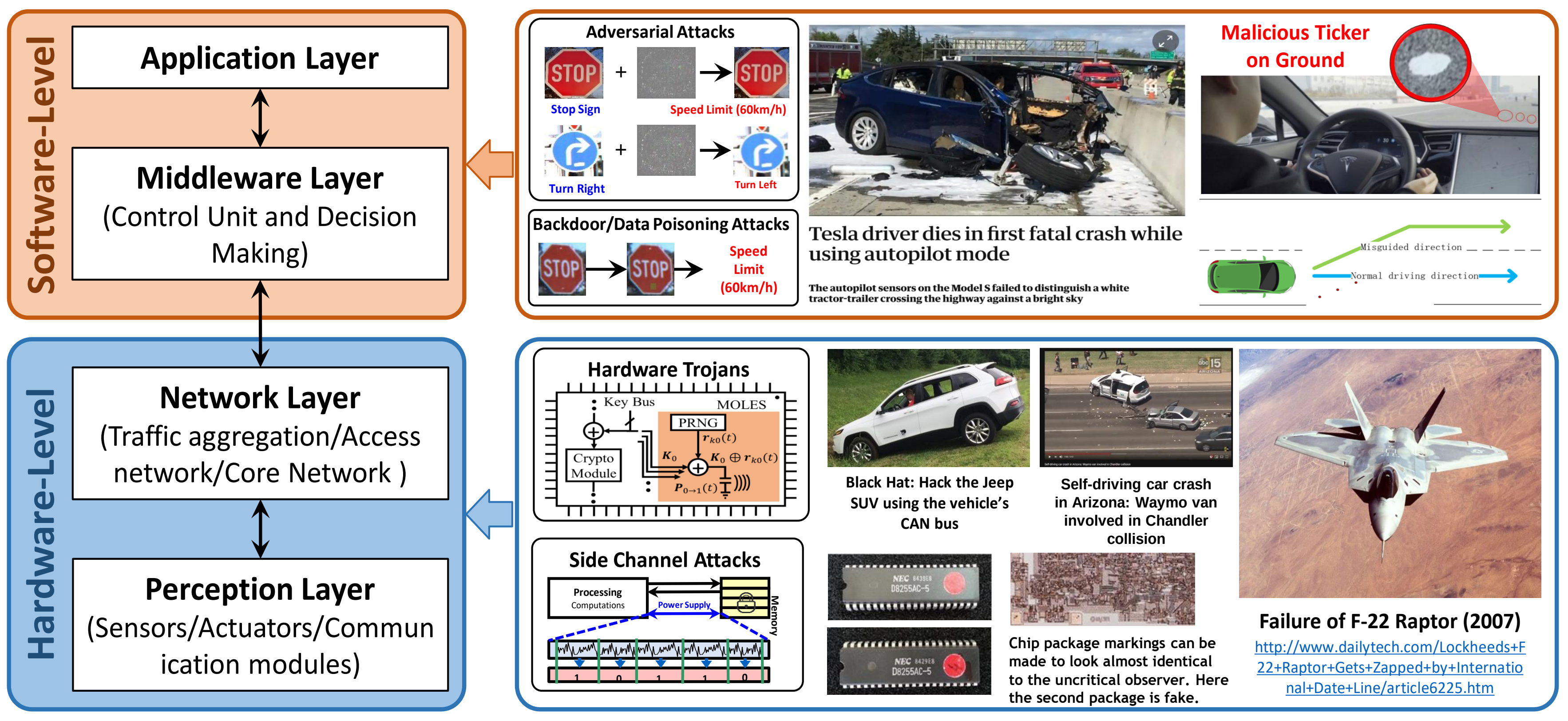
PHD FORUM

Hardware and Software Techniques for Securing Intelligent Cyber-Physical Systems

Faiq Khalid[1] (Ph.D. Candidate), Muhammad Shafique[2] (Advisor)
[1]Technische Universität Wien (TU Wien), Vienna, Austria
[2]Division of Engineering, New York University Abu Dhabi (NYUAD), Abu Dhabi, UAE

Problems and Motivation



Tesla driver dies in first fatal crash while using autopilot mode

The autopilot sensors on the Model S failed to distinguish a white tractor-trailer crossing the highway against a bright sky

Black Hat: Hack the Jeep SUV using the vehicle's CAN bus

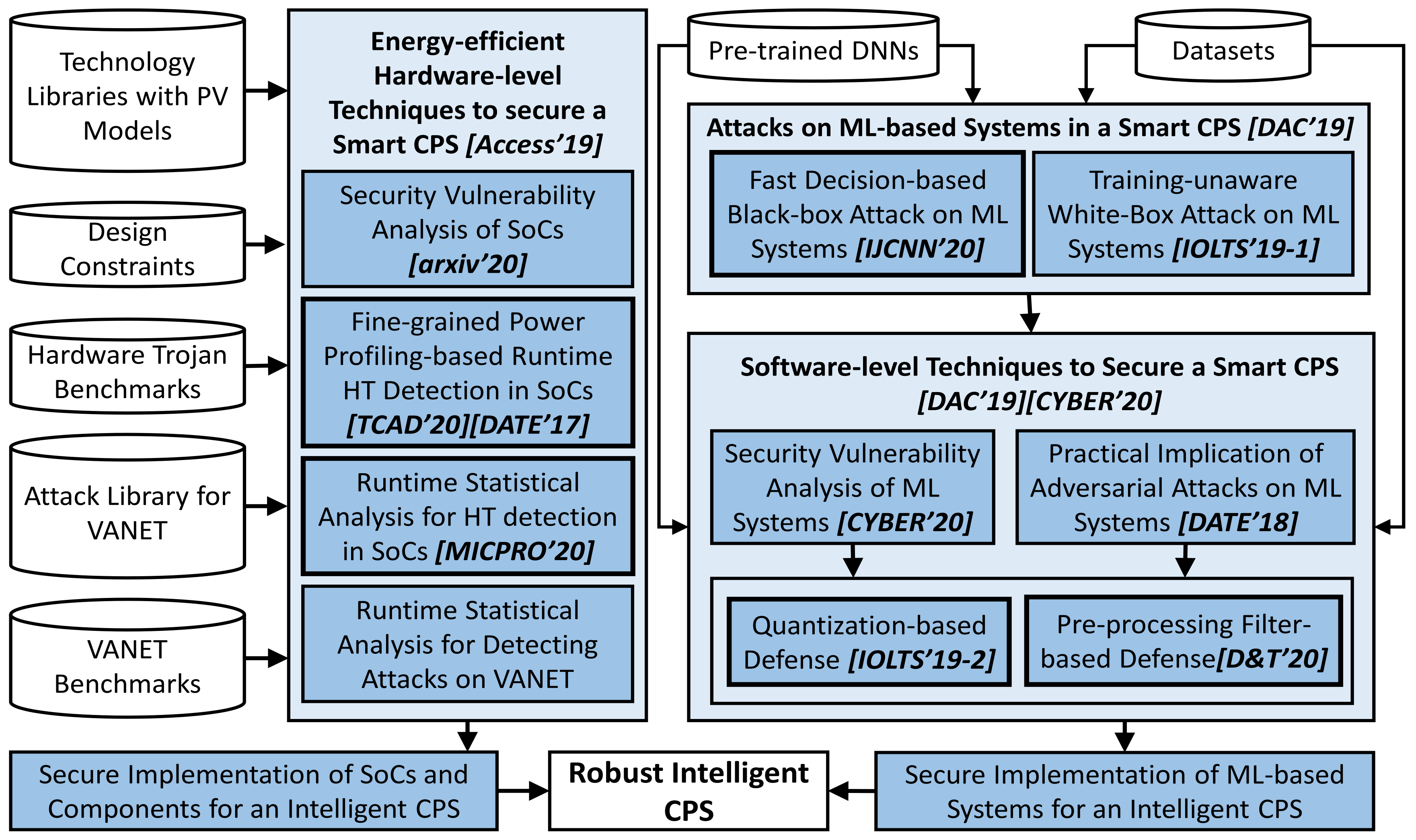
Self-driving car crash in Arizona: Waymo van involved in Chandler collision

Chip package markings can be made to look almost identical to the uncritical observer. Here the second package is fake.

Failure of F-22 Raptor (2007)
http://www.dailytech.com/Lockheeds+F22+Raptor+Gets+Zapped+by+International+Date+Line/article6225.htm

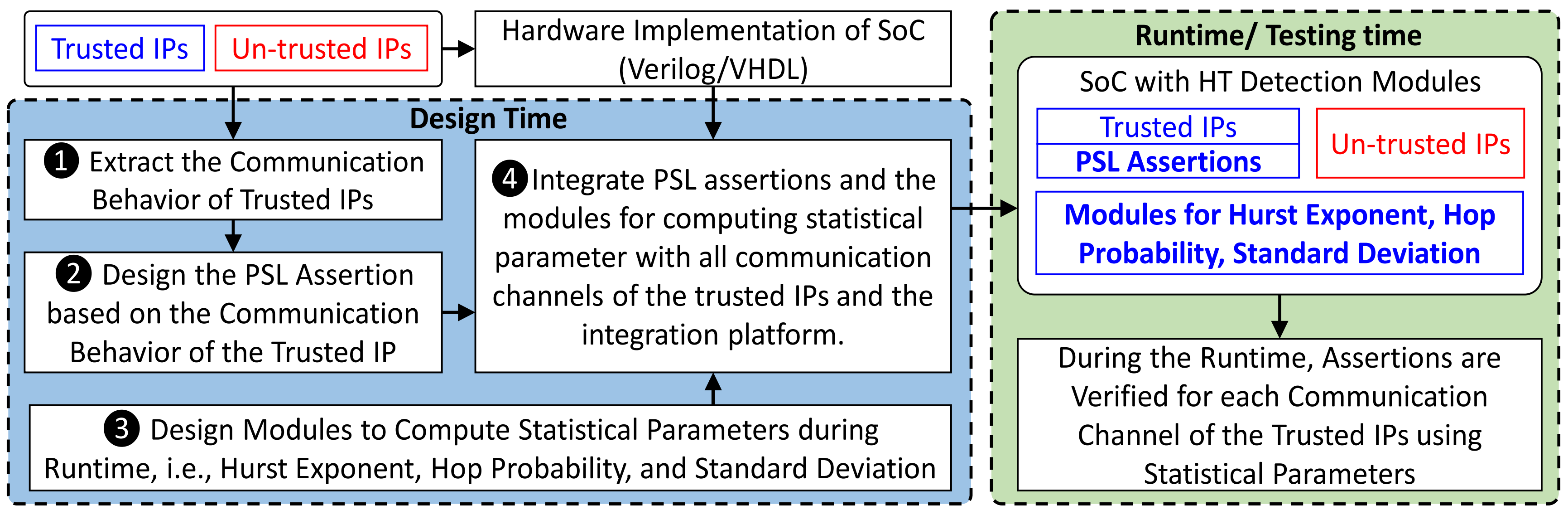
- ❑ **Design a Cost-Effective Secure Intelligent CPS**
  - ❑ Customized hardware/software solutions at appropriate system layers
  - ❑ Adapting to application properties and user requirements

Overview of Our Methodology


Technology Libraries with PV Models
Design Constraints
Hardware Trojan Benchmarks
Attack Library for VANET
VANET Benchmarks
Energy-efficient Hardware-level Techniques to secure a Smart CPS [Access'19]
Security Vulnerability Analysis of SoCs [arxiv'20]
Fine-grained Power Profiling-based Runtime HT Detection in SoCs [TCAD'20][DATE'17]
Runtime Statistical Analysis for HT detection in SoCs [MICPRO'20]
Runtime Statistical Analysis for Detecting Attacks on VANET
Pre-trained DNNs
Datasets
Attacks on ML-based Systems in a Smart CPS [DAC'19]
Fast Decision-based Black-box Attack on ML Systems [IJCNN'20]
Training-unaware White-Box Attack on ML Systems [IOLTS'19-1]
Software-level Techniques to Secure a Smart CPS [DAC'19][CYBER'20]
Security Vulnerability Analysis of ML Systems [CYBER'20]
Practical Implication of Adversarial Attacks on ML Systems [DATE'18]
Quantization-based Defense [IOLTS'19-2]
Pre-processing Filter-based Defense[D&T'20]
Secure Implementation of SoCs and Components for an Intelligent CPS
Robust Intelligent CPS
Secure Implementation of ML-based Systems for an Intelligent CPS

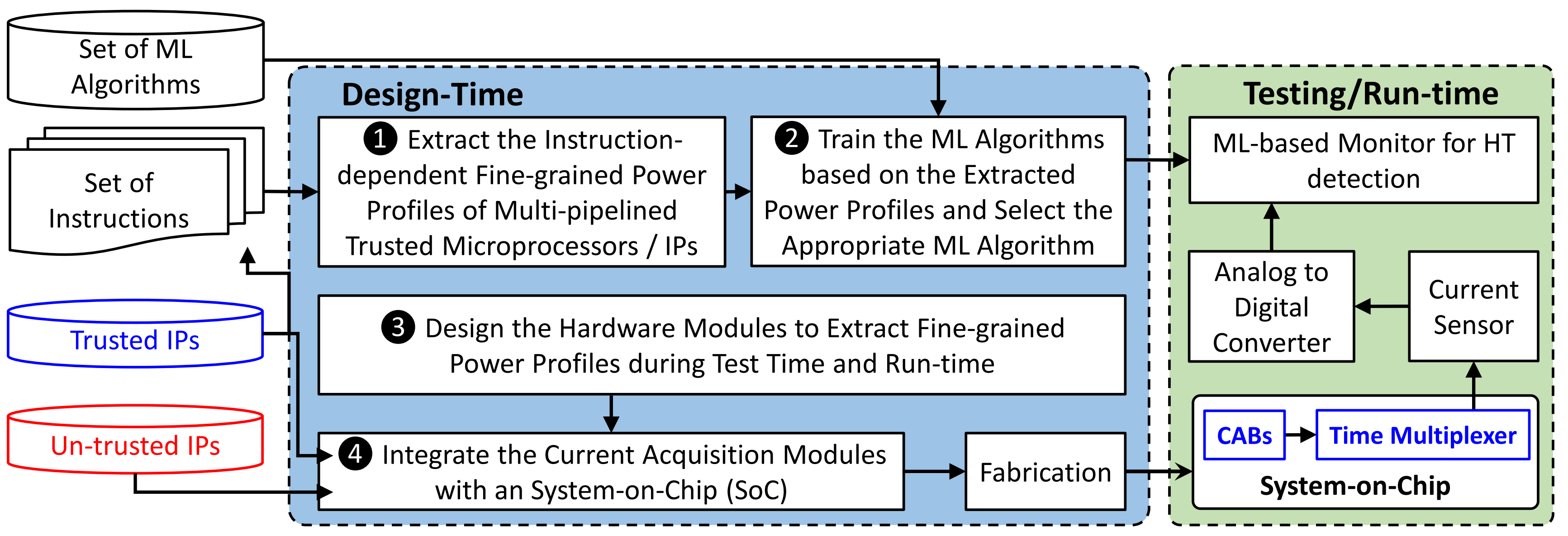
Hardware-level Techniques

❑ Communication-based Runtime HT Detection [MICPRO'20]


Trusted IPs
Un-trusted IPs
Hardware Implementation of SoC (Verilog/VHDL)
Design Time
1 Extract the Communication Behavior of Trusted IPs
2 Design the PSL Assertion based on the Communication Behavior of the Trusted IP
3 Design Modules to Compute Statistical Parameters during Runtime, i.e., Hurst Exponent, Hop Probability, and Standard Deviation
4 Integrate PSL assertions and the modules for computing statistical parameter with all communication channels of the trusted IPs and the integration platform.
Runtime/ Testing time
SoC with HT Detection Modules
Trusted IPs
PSL Assertions
Un-trusted IPs
Modules for Hurst Exponent, Hop Probability, Standard Deviation
During the Runtime, Assertions are Verified for each Communication Channel of the Trusted IPs using Statistical Parameters

On average, our approach (**SIMCom)** achieves **99% HT detection accuracy** with a **1.5% drop** due to process variations (PV) and exhibits **less than 1%** area overhead and **≈1%** power overhead.

❑ Fine-grained Power-profiling-based Runtime HT Detection [TCAD'20]

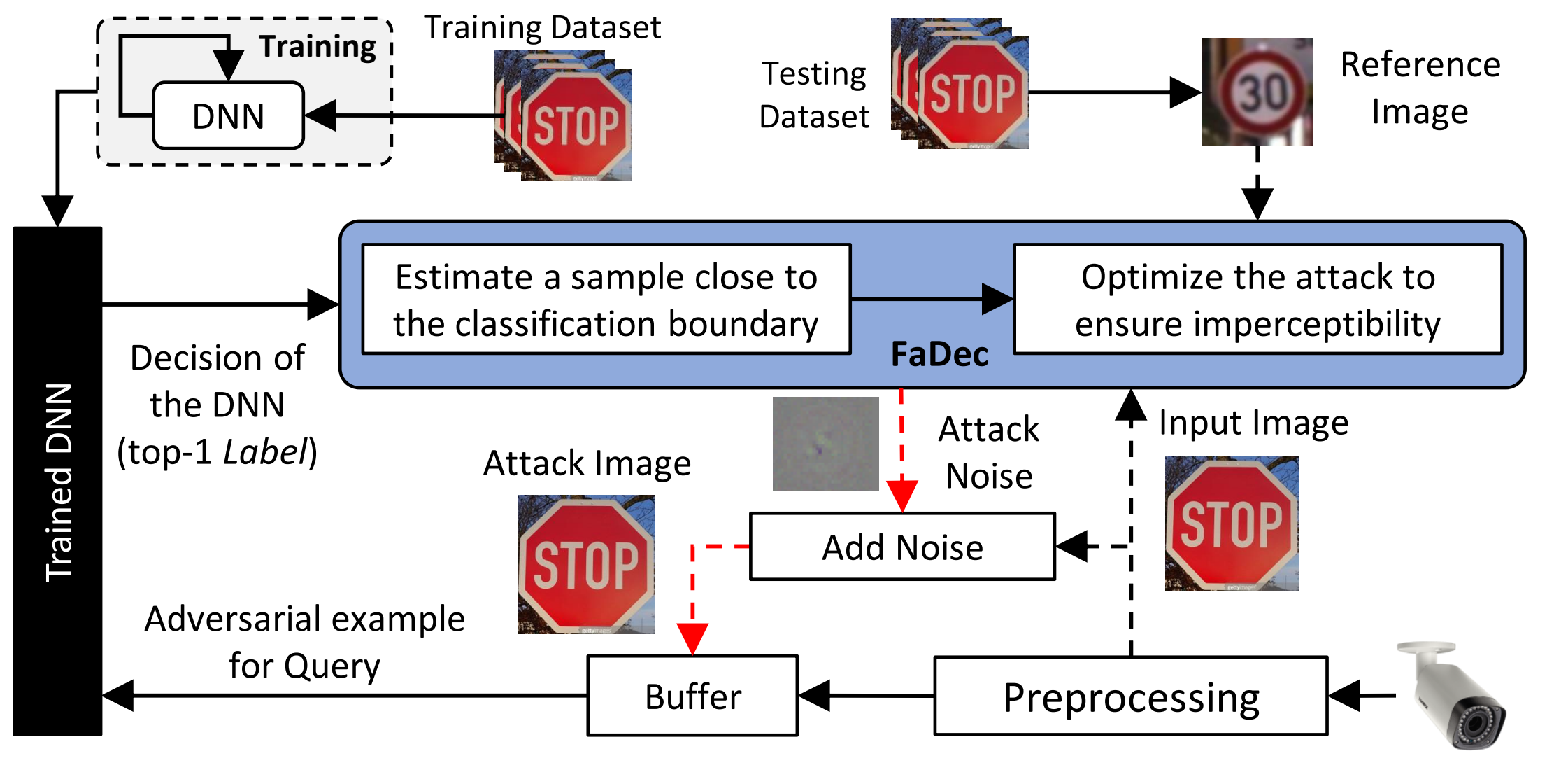

Set of ML Algorithms
Set of Instructions
Trusted IPs
Un-trusted IPs
Design-Time
1 Extract the Instruction-dependent Fine-grained Power Profiles of Multi-pipelined Trusted Microprocessors / IPs
2 Train the ML Algorithms based on the Extracted Power Profiles and Select the Appropriate ML Algorithm
3 Design the Hardware Modules to Extract Fine-grained Power Profiles during Test Time and Run-time
4 Integrate the Current Acquisition Modules with an System-on-Chip (SoC)
Fabrication
Testing/Run-time
ML-based Monitor for HT detection
Analog to Digital Converter
Current Sensor
CABs
Time Multiplexer
System-on-Chip

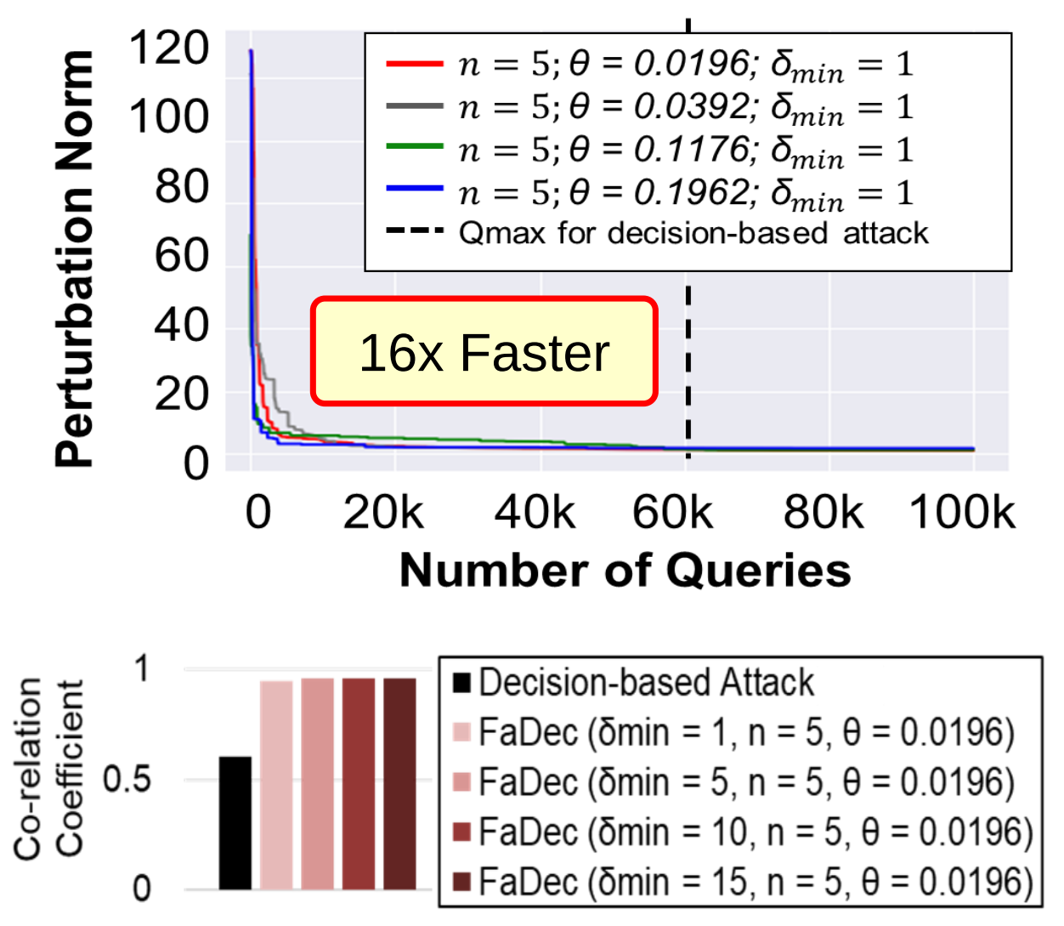
On average, our proposed approach (**MacLeR)** achieves **95% HT detection accuracy** with a **0.6% drop** due to PV, **3% drop** due to workload and aging variation, and exhibits **less then 0.5%** area and power overheads.

Software-level Techniques

❑ Security Attacks on ML-based Systems [IJCNN'20]

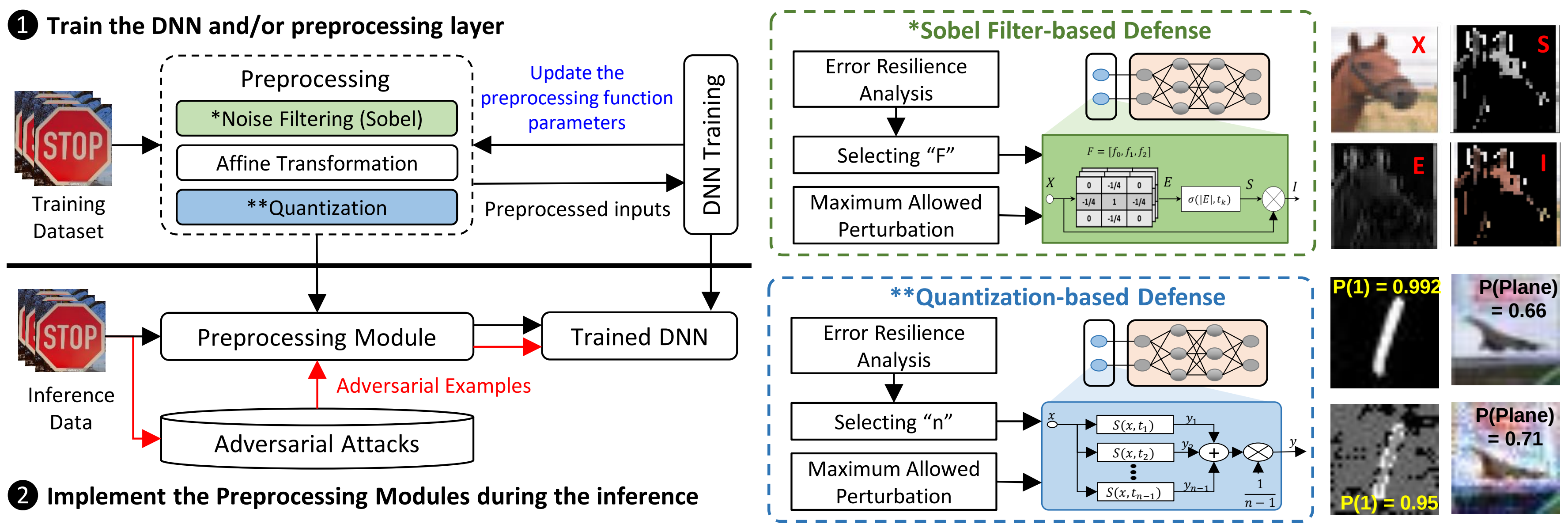
***Attack:** Stop sign is mapped to 30km/h*


Training
DNN
Training Dataset
Testing Dataset
Reference Image
Trained DNN
Estimate a sample close to the classification boundary
Optimize the attack to ensure imperceptibility
FaDec
Decision of the DNN (top-1 Label)
Attack Image
Attack Noise
Input Image
Add Noise
Adversarial example for Query
Buffer
Preprocessing



The proposed attack (**FaDec**), with appropriate attack parameters values, **converges16x faster** and generates the attack image with **≈20% better imperceptibility** than the state-of-the-art decision-based attack. **Open-source:** https://github.com/fklodhi/FaDec

❑ Software-level Defenses for ML-based Systems [IOLTS'19][D&T'20]



On average, **QuSecNets** increases classification accuracy up to **50%-96% (MNIST)** and **10%-50% (CIFAR10). SSCNets** increases classification accuracy up to **16%-30% (White-box scenario)** and **28% to 48% (Black-box Scenario)**.

Selected Publications

[TCAD'20] F. Khalid, S. R. Hasan, S. Zia, O. Hasan, F. Awwad and M. Shafique, "MacLeR: Machine Learning-Based Runtime Hardware Trojan Detection in Resource-Constrained IoT Edge Devices," in IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**), vol. 39, no. 11, pp. 3748-3761, 2020.

[MICPRO'20] F. Khalid, S. R. Hasan, O. Hasan, M. Shafique, "SIMCom: Statistical Sniffing of Inter-Module Communications for Runtime Hardware Trojan Detection," Elsevier Microprocessors and Microsystems (**MICPRO**), pp. 103-122, 2020.

[Access'19] D. Ratasich, F. Khalid, F. Geissler, R. Grosu, M. Shafique and E. Bartocci, "A Roadmap Toward the Resilient Internet of Things for Cyber-Physical Systems," in **IEEE Access**, vol. 7, pp. 13260-13283, 2019.

[IJCNN'20] F. Khalid, H. Ali, M. Abdullah Hanif, S. Rehman, R. Ahmed and M. Shafique, "FaDec: A Fast Decision-based Attack for Adversarial Machine Learning," in International Joint Conference on Neural Networks (**IJCNN**), 2020, pp. 1-8. **Received IEEE CIS Young Professional Grant**

[D&T'20] H. Ali, F. Khalid, H. A. Tariq, M. A. Hanif, R. Ahmed and S. Rehman, "SSCNets: Robustifying DNNs using Secure Selective Convolutional Filters," in IEEE Design & Test (**D&T**), vol. 37, no. 2, pp. 58-65.

[CYBER'20] F. Khalid, M. A. Hanif and M. Shafique, "Exploiting Vulnerabilities in Deep Neural Networks: Adversarial and Fault-Injection Attacks," in Conference on Cyber-Technologies and Cyber-Systems (**CYBER**), pp. 24-29, 2020.

[DATE'18] F. Khalid, M. A. Hanif, S. Rehman, J. Qadir and M. Shafique, "FAdeML: Understanding the Impact of Pre-Processing Noise Filtering on Adversarial Machine Learning," in Design, Automation & Test in Europe Conference & Exhibition (**DATE**), 2019, pp. 902-907.

[arxiv'20] F. Khalid, I. H. Abbasi, S. Rehman, O. Hasan, A. M. Kamboh, M. Shafique, "ForASec: Formal Analysis of Security Vulnerabilities in Sequential Circuits," arXiv preprint arXiv:1812.05446, (Under Review **IEEE TCAD**)

[IOLTS'19-1] F. Khalid, M. A. Hanif, S. Rehman, R. Ahmed and M. Shafique, "TrISec: Training Data-Unaware Imperceptible Security Attacks on Deep Neural Networks," in in International Symposium on On-Line Testing and Robust System Design (**IOLTS**), 2019, pp. 188-193.

[IOLTS'19-2] F. Khalid, H. Ali, H. Tariq, M. A. Hanif, S. Rehman, R. Ahmed, M. Shafique, "QuSecNets: Quantization-based Defense Mechanism for Securing Deep Neural Network against Adversarial Attacks," in **IOLTS**, 2019, pp. 182-187.

[DATE'17] F. Khalid, S. R. Hasan, O. Hasan and F. Awwad, "Power profiling of microcontroller's instruction set for runtime hardware Trojans detection without golden circuit models," in IEEE **DATE**, 2017, pp. 294-297.

[DAC'19] J. J. Zhang, K. Liu, F. Khalid, M. A. Hanif, S. Rehman, T. Theocharides, A. Artussi, M. Shafique, S. Grag, "Building Robust Machine Learning Systems: Current Progress, Research Challenges, and Opportunities," in ACM/IEEE Design Automation Conference (**DAC**), 2019, pp. 1-4. **Received a HIPEAC Paper Award**

TECHNISCHE UNIVERSITÄT WIEN

جامعة نيويورك أبوظبي
NYU ABU DHABI

This Ph.D. is supported in parts by the **FFG** and the **BMVIT,** Austria, under the "ICT of the Future" project, **IoT4CPS: Trustworthy IoT for Cyber-Physical Systems**

IoT4CPS Trustworthy IoT for CPS

FFG Österreichische Forschungsförderungsgesellschaft

Bundesministerium Verkehr, Innovation und Technologie