Performance Evaluation of AI Algorithms on Heterogeneous Edge Devices for Manufacturing

Bernhard Rupprecht, Dominik Hujo, Birgit Vogel-Heuser

IEEE CASE, August 20-24th, 2022

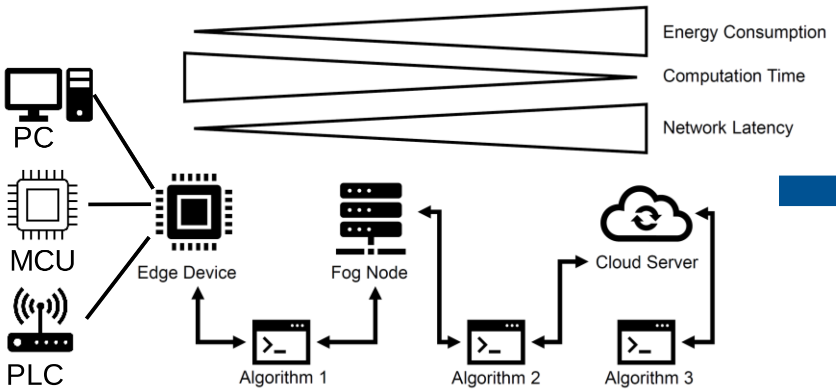
Source:
B. Rupprecht, D. Hujo and B. Vogel-Heuser, "Performance Evaluation of AI Algorithms on Heterogeneous Edge Devices for Manufacturing," 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE), 2022, pp. 2132-2139, doi: 10.1109/CASE49997.2022.9926482.
Link: https://ieeexplore.ieee.org/document/9926482

Edge Computing in Smart Manufacturing

- Limited Network Bandwidth
- Real-time Requirements

Edge Computing Approaches

- Challenge of Software/Hardware Co-Design: “Where to run which Algorithm?”



This paper examines algorithm behavior on heterogeneous hardware in order to assist SW/HW-codesign of CPPS

Contributions of this Paper

C1: Overview of properties of ML Algorithms -> well-conceived deployment

C2: Examination of ML Algorithms on heterogeneous hardware platforms -> chose appropriate ML inference hardware

C3: Introduction of metrics tailored to the manufacturing domain -> supports assessment of ML Algorithms

C4: Examination ML Algorithms on heterogeneous hardware platforms -> reveals unexpected behavior and specialties.

Concept of this Paper

Heterogeneous Hardware

Microcontroller:
PSOC63 (150 MHz)

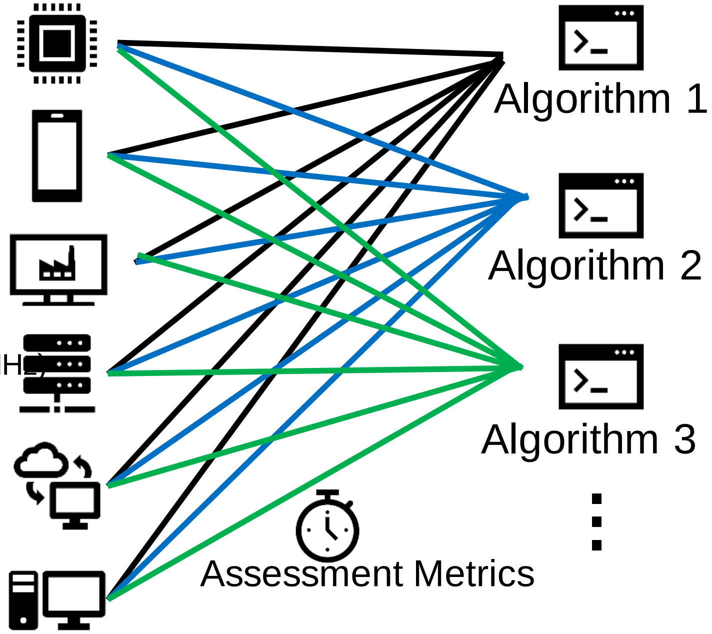
Single-board Computer:
Beaglebone black (1 GHz)

Dedicated Edge Device:
Siemens IoT2050 (1 GHz)

PLC:
STW ESX.3cm PLC (300 MHz)
Siemens S7-1500 (?)

Industrial PC:
Beckhoff CX9020 (1 GHz)
CX2040 (2.1 GHz)

Consumer PC:
HP Spectre (2.5 GHz)



- Execute various ML Algorithms on different hardware
- Examine behavior using assessment metrics:

Performance Metrics:
Execution Time,
Memory Consumption

Numerical Accuracy:
Double- vs. Single-Precision

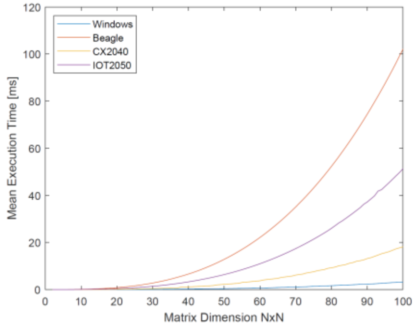
Energy Consumption

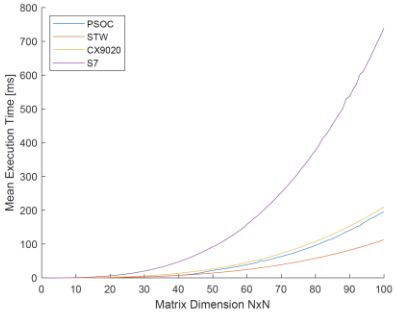
Selected Algorithms

Algorithm	Domain	Applications	Time Complexity	Space Complexity
Standard Matrix Multiplication	Basis Operation	ML, MPC	$\mathcal{O}(N^3)$	$\mathcal{O}(N^2)$
Gauss Jordan Matrix Inversion	Basis Operation	ML, MPC	$\mathcal{O}(N^3)$	$\mathcal{O}(N^2)$
Grubbs Test	Statistic	Outlier Detection	$\mathcal{O}(n)$	$\mathcal{O}(n)$
Butterworth Filter	Signal processing	Preprocessing	$\mathcal{O}(n)$	$\mathcal{O}(n)$
DBSCAN	Clustering	Outlier Detection	$\mathcal{O}(n^2)$	$\mathcal{O}(n)$
Random Forest (Prediction)	Classification	Fault detection	$\mathcal{O}(k * D(n))$	$\mathcal{O}(k * \#nodes)$
SVM (Prediction)	Classification	Fault detection	$\mathcal{O}(f)$	$\mathcal{O}(f)$

Example: Matrix Multiplication

N: Matrix Dimension
n: Number of Datapoints
k: Number of trees in the forest
D(n): Average tree depth in the forest
f: Number of features





Findings of this Paper- Summary

Conclusions-Excerpt:

- All algorithms except the RF depth show their expected theoretical time complexity on all devices
- Single-precision execution is not always faster than double-precision
- Behavior of the algorithms can differ dependent on the considered hardware, compiler and runtime
- Energy consumption directly depends on the execution time of the algorithm (except RF algorithm)

Identified relevant aspects of SW/HW Codesign for the creation of CPPS