



Assignment of bachelor's thesis

Title: SW support for NPU accelerators for Linux-type operating systems
Student: Michal Žáček
Supervisor: doc. RNDr. Ing. Petr Zemánek, CSc.
Study program: Informatics
Branch / specialization: Computer Systems and Virtualization 2021
Department: Department of Computer Systems
Validity: until the end of summer semester 2026/2027

Instructions

NPU (Neural Processing Unit) accelerators in PCs or notebooks are becoming more and more important these days. The target platform considered in this thesis is the board TQMa8MPxL [1] but results should apply to any device with the NPU accelerator running under Linux and/or PikeOS. Neither the Linux kernel nor PikeOS (real-time embedded OS) currently support AI/ML acceleration using NPUs. The objective of the thesis is to conduct research of open source frameworks with the aim to integrate a reasonably good SW support for utilizing this accelerator to enhance the performance of machine learning (ML) algorithms on these target platforms.

- 1/ Analyze and evaluate implementation of the SW support for utilizing NPU accelerators in representative open-source frameworks.
- 2/ Investigate modules or libraries used in these frameworks for communication with NPU accelerators.
- 3/ Review ML algorithms implemented in these frameworks and select most appropriate frameworks for testing.
- 4/ Design benchmark data sets for testing the ML performance of the selected frameworks on the given Linux platform.
- 5/ Test the performance of selected frameworks on the given Linux platform.
- 6/ Based on these findings, select the most appropriate framework for the integration into the Linux platform and elaborate an integration procedure based on e.g. Yocto packaging system.



7/ Document problems encountered during the integration process.

8/ Propose a procedure for the integration of the SW support for the NPU accelerator on the PikeOS platform.

[1] <https://www.tq-group.com/en/products/tq-embedded/arm-architecture/tqma8mpxl>

