Chair of Computer Architecture and Parallel Systems
TUM Department of Informatics
Technische Universität München

Master thesis: Acceleration of CNN Training using FPGAs

For multiple years now, FPGAs have been proposed as an approach to accelerate highly parallel computing problems. The computation of big DNNs consume massive amounts of computational resources and energy, accelerating these workloads is a key hurdle in achieving further advancements in the field of AI. The massive parallelism present in Convolutional Neural Networks (CNNs) has led to implementations both on FPGAs and in ASIC designs. The majority of these focus on inference instead of training, as it is much simpler to implement and requires less memory bandwidth. This thesis aims at implementing a CNN training framework for FPGAs, allowing for flexibility in the supported Layers and a possible future integration in more complex AI algorithms.

Tasks:

- Explore existing FPGA implementation of CNN training
- Evaluate different approaches in memory organization and exploiting parallelism
- Implement CNN training on a FPGA using Vivado HLS and/or SystemVerilog
- Evaluate the results and compare them with other FPGA implementations
- Analyze Bottlenecks of Training on FPGAs
- Integrate with an existing ML framework

Recommended knowledge and experience:

- Experience in programming with `C/C++`
- Experience in programming with `VHDL/Verilog`
- Experience in programming with `Python`
- Experience in deep learning

Benefits:

- Involve in the academic environment of chair of Computer Architecture and Parallel Systems
- Access to powerful Xilinx FPGAs

Application:

If you are interested in this topic, get in contact with Dirk Stober and Dai Liu (find the contact details below) through email.

Technische Universität München
Chair of Computer Architecture and Parallel Systems (Prof. Schulz)
dirk.stober(at)tum.de
dai.liu(at)tum.de
www.caps.in.tum.de