Diana Wofk

dwofk@alum.mit.edu • dwofk.github.io • linkedin.com/in/dianawofk

Summary

I am a recent graduate of MIT working on energy-efficient learning-based vision for robotics applications. My experience and interests lie in algorithm-architecture-hardware co-design for efficient deployment of deep learning models on edge devices. I am excited about opportunities to work on ML/AI system design at the hardware-software interface.

Education

Massachusetts Institute of Technology (MIT) — **Cambridge, MA, USA**
Master of Engineering in Electrical Engineering and Computer Science, 5.0/5.0 — *2018–2020*
Bachelor of Science in Electrical Engineering and Computer Science, 5.0/5.0 — *2014–2018*

Research Experience

Intel Intelligent Systems Lab — **Munich, Germany**
Incoming Deep Learning Resident — *Sep 2020–Sep 2021*

Energy-Efficient Multimedia Systems Group, MIT — **Cambridge, MA, USA**
Research Assistant, Graduate Student — *Jun 2018–May 2020*

- Thesis on algorithm and hardware design for fast and energy-efficient monocular depth estimation on embedded systems.
  - Explored DNN simplification and compilation to enable real-time depth inference on Jetson TX2 (178 fps) and iPhone X (40 fps).
  - Designed and implemented a custom dataflow and accelerator to reduce data movement and ensure high spatial utilization of hardware. On Ultra96 FPGA, the depth estimation DNN accelerator achieved 1.5–2× better energy-efficiency than the TX2 CPU.

Undergraduate Researcher (UROP, SuperUROP) — *Jun 2015–May 2018*

- Investigated tradeoffs between the complexity, runtime, and energy consumption of depth estimation DNN designs.
- Designed and built communication interfaces between PCs, FPGAs, and embedded computing platforms.
  - Interfaced an FPGA board to a Jetson TX1 via PCIe to demonstrate CNN-based super-resolution applied to decoded video.
  - Developed a PC-FPGA Ethernet interface that uses raw 1GbE for bulk data transfer to and from a Xilinx VC707. Explored and tested different methods of data flow control (stop-and-wait, credit-based). Final design achieved transfer speeds of 118 MBps.

Publications

- D. Wofk*, F. Ma*, T.-J. Yang, S. Karaman, V. Sze. **FastDepth: Fast Monocular Depth Estimation on Embedded Systems**. *IEEE International Conference on Robotics and Automation (ICRA)*, May 2019. Website: `fastdepth.mit.edu`

Additional Professional Experience

Intel Corporation — **Santa Clara, CA, USA**
Pre-Silicon Validation Intern, Scalable Performance CPU Development Group — *May 2017–Aug 2017*

- Developed Python-based scripts to automate log parsing and regression testing for errors in pre-silicon validation.

International Business Machines Corporation (IBM) — **Poughkeepsie, NY, USA**
System ASIC Bring-Up Intern, IBM Systems — *May 2016–Aug 2016*

- Supported the bring-up of a chip that was part of the FICON/OSA adapter designed for the next-generation z System.
- Wrote shell scripts to automate system characterization (voltage, temperature, and frequency variation) on the test floor.

Additional Academic Experience

Courses: Hardware Architecture for Deep Learning; Digital Integrated Circuits; Digital Image Processing; Signals, Systems, and Inference; Computation Structures; Computer System Engineering; Computer Systems Security; Engineering Leadership
Teaching: MIT 6.111 FPGAs & Digital Systems TA (Fall 2018, Fall 2019); 6.033 Computer System Eng TA (Spring 2020)

Skills

Languages: English (native), Russian, German
Design Tools: Vivado Design Suite, Xilinx ISE, ModelSim
Coding: Python, C, Bash, MATLAB, Verilog
Frameworks: PyTorch, TVM