

Cao Gao

caogao@umich.edu

• web.eecs.umich.edu/~caogao

• (734) 834-3274

SUMMARY

Interested in computer architecture, performance analysis, mobile systems and machine learning in general.
Current work includes designing mobile architecture for future machine learning algorithms.

EDUCATION

University of Michigan, Ann Arbor, MI

September.2012 – May.2017 (expected)

PhD, Computer Science and Engineering (Hardware)

Advisor: Prof. Trevor Mudge

M.S., Computer Science and Engineering

Overall GPA: 4.0/4.0

Zhejiang University, Hangzhou, Zhejiang, China

September.2008 – June.2012

BEng, Major: Electronic and Information Engineering

Minor: English

Member of Chu Kochen Honors College

Overall GPA: 3.91/4.0

PROFESSIONAL EXPERIENCE

ARM Ltd., Austin, TX

June.2014 – August.2014

R&D Intern at the Mobile System Group. Worked on Android workload characterization.

MAJOR PROJECTS

An ultra-low power non-uniform memory accelerator for wearable devices

- Developed the architecture, ISA, and compiler for the accelerator, which has been fabricated as a chip
- Designed a framework that automatically generates optimal memory layout based on target applications

Accelerating deep learning algorithms on mobile platforms

- Analyzed the characteristics of Deep Neural Network workloads on mobile / server GPUs
- Participated in designing a framework which intelligently partitions workloads between mobile and server

Graph analytics processing accelerator

- Proposed an accelerator architecture for billion-edge scale graph applications
- Led four grad students to characterize applications, explore algorithm and architecture choices

User quality-of-experience metrics for Android applications (at ARM)

- Identified a set of metrics that measures the user experience of Android applications
- Implemented a framework to automate workload execution and metrics collection

A study of mobile device utilization

- Evaluated the CPU and GPU utilization of a wide range of common mobile applications on real hardware
- Identified the diminishing returns of increasing core counts and suggested a more flexible system

SELECTED PUBLICATIONS

- Y. Kang, J. Hauswald, C. Gao, A. Rovinski, et.al. *The Neural Edge: Collaborative Intelligence Between Cloud and the Mobile Edge*. Architectural Support for Programming Languages and Operating Systems (ASPLOS), April 2017.
- S. Bang, J. Wang, Z. Li, C. Gao, Y. Kim, et.al. *A 288 μ W Programmable Deep Learning Processor with 270kB On-chip Weight Storage Using Non-Uniform Memory Hierarchy for Mobile Intelligence*. ISSCC, February 2017.
- C. Gao, A. Gutierrez, M. Rajan, R.G. Dreslinski, T. Mudge, and C.J. Wu. *A Study of Mobile Device Utilization*. 2015 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), March 2015.

SKILLS

Programming: experienced in C/C++, Python, CUDA, familiar with Verilog, Matlab

Environments: Linux, Android, git, ARM streamline, Keil uVision, nvprof, Caffe