

Shien Zhu

Institute for Computing Platforms - Systems Group, ETH Zurich
shien.zhu@inf.ethz.ch [Google Scholar](#) [LinkedIn](#)



EDUCATION

JULY 2018 – DECEMBER 2022

PH. D, Nanyang Technological University, Singapore

Parallel and Distributed Computing Lab, School of Computer Science and Engineering

Supervisor: Dr. Weichen Liu

Thesis: Deep Learning Acceleration: From Quantization to In-Memory Computing

AUGUST 2014 – JUNE 2018

BACHELOR, University of Science and Technology of China, China

Department of Automation, School of Information Science and Technology

Supervisor: Dr. Weiwei Shang

Thesis: Dynamic Identification of a 6-DOF Industrial Robot

AWARDS

Best Research Award

Intentional Student Conference on Artificial Intelligence, June 2022

NTU Research Scholarship

Nanyang Technological University, Singapore, July 2018 – July 2022

DAC Young Research Fellow

The 58th Design Automation Conference (DAC), December 2021

EDAthlon 3rd Place

Programming Competition on Electronic Design Automation, IEEE CEDA HK, Aug. 2020

Outstanding Student Scholarship

University of Science and Technology of China (USTC), 2014, 2015, 2016, 2017

WORK EXPERIENCE

- Part-Time Student Research Assistant, Nanyang Technological University, Aug-Dec, 2022
- **Research Fellow**, Nanyang Technological University, Feb-Apr, 2023
- **Research Fellow**, ETH Zurich, July 2023 - Now

RESEARCH INTERESTS

- Deep Learning Acceleration
- Neural Network Quantization
- Software-Hardware Co-Design
- Edge Computing and Embedded Systems
- Computer Architecture
- In-Memory Computing
- Electronic Design Automation
- Hardware Accelerator and FPGA

PUBLICATIONS

- **Shien Zhu**, Luan H. K. Duong, and Weichen Liu, "**TAB: Unified and Optimized Ternary, Binary and Mixed-Precision Neural Network Inference on the Edge**," in *ACM Transaction on Embedded Computing Systems*, 2022
- **Shien Zhu**, Luan H. K. Duong, and Weichen Liu, "**XOR-Net: An Efficient Computation Pipeline for Binary Neural Network Inference on Edge Devices**," in *Proceedings of the IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS)*, 2020
- **Shien Zhu**, Luan H. K. Duong, Hui Chen, Di Liu, and Weichen Liu, "**FAT: An In-Memory Accelerator with Fast Addition for Ternary Weight Neural Networks**," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2022
- **Shien Zhu**, Shiqing Li, and Weichen Liu, "**iMAD: An In-Memory Accelerator for AdderNet with Efficient 8-bit Addition and Subtraction Operations**," in *Proceedings of the 2022 Great Lakes Symposium on VLSI (GLSVLSI)*, 2022 (**Best Paper Candidate, 7 out of 43 papers**)
- **Shien Zhu**, Shuo Huai, Guochu Xiong, and Weichen Liu, "**iMAT: Energy-Efficient In-Memory Acceleration of Ternary Neural Networks With Sparse Dot Product**," in *Proceedings of the 2023 ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED)*, 2023 (**Best Paper Candidate, ranked 1st in the track**)
- Shiqing Li, **Shien Zhu**, Xiangzhong Luo, Tao Luo, and Weichen Liu, "**An Efficient Sparse LSTM Accelerator on Embedded FPGAs with Bandwidth-Oriented Pruning**," in *Proceedings of the 33rd International Conference on Field-Programmable Logic and Applications (FPL)*, 2023
- Hao Kong, Shuo Huai, Di Liu, Lei Zhang, Hui Chen, **Shien Zhu**, Shiqing Li, Weichen Liu, Manu Rastogi, Ravi Subramaniam, Madhu Athreya, and M. Anthony Lewis, "**EDLAB: A Benchmark for Edge Deep Learning Accelerators**," in *IEEE Design & Test*, 2022
- Hui Chen, Zihao Zhang, Peng Chen, **Shien Zhu**, and Weichen Liu, "**Parallel Multipath Transmission for Burst Traffic Optimization in Point-to-Point NoCs**," in *Proceedings of the 2021 Great Lakes Symposium on VLSI (GLSVLSI)*, 2021

SKILLS

- Training deep learning models using **PyTorch**
- **C/C++ programming on ARM CPU, x86 CPU, and Nvidia GPU**
- **Python**
- Matlab
- Java, HTML, and Spring Framework
- Circuit simulation using **Virtuoso**
- **Verilog and vHDL programming**
- Graphic design: draw.io and Inkscape
- Piano Accompaniment
- Guitar, Violin, and Hulusi
- Music theory and composition

ACTIVITIES

Academic Service

- Paper Reviewer, Future Generation Computer Systems (FGCS), 2023-Now
- Paper Reviewer, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD), 2022-Now
- Paper Reviewer, ACM Transactions on Embedded Computing Systems (TECS), 2022-Now
- Paper Reviewer, Journal of System Architecture (JSA), 2022-Now

Teaching Assistant Experience

- Sensor, Interfacing, and Control (Lab). Autumn 2021
- Multidisciplinary Design Project (Lab). Spring 2020
- **Software Engineering** (Lab). Spring & Autumn 2019, and Spring 2020
- **Advanced Computer Architecture** (Tutorial). Autumn 2018, and Spring 2023
- **Fundamentals of Electronic Design** (Tutorial + Lab), Spring 2018
- **Digital Logic Circuits** (Tutorial). Autumn 2017

Supervised Undergraduate Research Projects

- Final Year Project, Marcus Wei Jie Tham, "**Training Binary Neural Networks**", 2021-2022
- Final Year Project, Muthukrishnan Subhiksha, "Deploying AI Applications on Smartphones with Neural Network Accelerators: 'AI Benchmark Application for Android Devices'", 2021-2022
- Final Year Project, Jingwei Chew, "**Ultra-low Power Real-time Object Detection Based on Quantized CNNs**", 2020-2021
- Exchange Student Research Project, Chunyun Chen, "**A Deep Learning Accelerator Based on Eyeriss V2 Implemented in Chisel**", 2019-2020
- Exchange Student Research Project, Kaivalya Swami, "Different Quantization Techniques for Neural Networks", 2019
- URECA Project, Nguyen Le Hoang, "**Object Detection for Drone Images on Embedded Systems**", 2018-2019
- URECA Project, Tajudeen Safeek Ahmed, "Real-time Object Detection for Drones Using Deep Learning", 2018-2019

Exchange Experience

- University of Newcastle, Australia. July to September, 2017.
Undergraduate Student Research Program:
Modelling and Analysis of Cancer Drugs' IC50s Using Memetic Algorithm
- Shanghai Jiao Tong University, China. Spring Semester, 2016.

Volunteer Experience

- Secretariat, Student Research Forum @ ASP-DAC, 2021
- Liaison Officer, Global Young Scientist Summit, Singapore (GYSS). January 2020
- Volunteer, Swim for Hope, Singapore. October 2018
- Instructor, Science and Technology Week, USTC, China. May 2015 and May 2018