Zhiyuan Zhao

Personal Information

Address:	Beijing, China	Birthday:	2000.11.29
Phone (CHN):	(+86)150-5608-7685	Email:	zhiyuan-zhao@outlook.com
Research Interests:	AI Chip, Computer Architecture, FPGA Accelerator		

Educational Background

2020.9 – 2023.6 **University of Science and Technology of China** **Master of Engineering**

Integrated Circuit Engineering (Two First-Class Academic Scholarships) GPA：3.59/4.30

2016.9 – 2020.6 **Hefei University of Technology** **Bachelor of Engineering**

Integrated Circuit Design and Integration System GPA：3.29/4.30

Research Experiences

2022.5 – 2023.6 **Research and Implementation of Lightweight Neural Network Hardware Accelerator** **Master's Thesis in USTC** 70,000 Chinese characters

- To address the problem of high off-chip memory access and low computational efficiency in existing accelerators, proposed a hybrid computing engine dataflow accelerator.
- To address the problem of insufficient on-chip bandwidth supply for processing elements, proposed a pipeline-friendly adaptive bandwidth computing engine.
- To address the problem of imbalanced resource mapping and low deployment efficiency, proposed a network deployment algorithm based on storage and computing resource awareness.
- Deployed ShuffleNetV2×0.5(41MOPS, 1.4MB), ShuffleNetV2×1.0(146MOPS, 2.3MB) and MobileNetV2 (300MOPS, 3.4MB) on ZC706 at 200MHZ, achieving 6492, 1771 and 886 FPS, respectively.
- Compared to existing accelerators, the proposed accelerator achieves 70% improvement in single DSP throughput, 92% reduction in off-chip memory access, and 2.3 times improvement in energy efficiency.

2022.1 – 2022.12 **Quantum Fiber Optic Communication Project** **Institute of Semiconductors**

- Implemented SHA1 and SM3 encryption algorithm IPs, supporting password encryption and password cracking modes, with IP throughput of up to 10GB/s.
- Developed a self-designed Gigabit Ethernet hardware module based on UDP and TCP protocols, utilizing only 2k LUTs, achieving communication speeds of 990Mb/s (UDP) and 700Mb/s (TCP) with PC.

2021.7 – 2021.10 **"One Life, One Chip" Project Phase III** **Institute of Computing Technology**

- Designed a five-stage pipelined CPU based on the RV64I instruction set, with an AXI4 bus interface. The CPU supports data forwarding, instruction prefetching, dynamic branch prediction, and exception handling mechanisms. It is capable of running RT-Thread OS with IPC of 0.68.
- Evaluated and optimized critical paths using DC, taped out (top 5% quota) on SMIC 110nm process, with chip area of 0.256mm^2, and achieved a maximum frequency of 200MHz.

2019.2 – 2020.6 **Vehicle License Plate Recognition System Based on RISC-V Processor and FPGA Acceleration** **Bachelor's Thesis** 30,000 Chinese characters

- Implemented training, quantization, and weight exporting of a recognition network using Python.
- Developed a real-time vehicle license plate recognition system based on RISC-V CPU, OV5640 camera, DDR, VGA screen, and CNN module, achieving a frame rate of 30FPS and 100% recognition accuracy.

Work Experiences

2023.7 – Now **Baidu Kunlun, Beijing** **Chip Design and Architecture Engineer**

- Design the vector computing subsystem in XPU, focusing on high-performance implementation of vector computation and nonlinear function operations.
- Responsible for RTL integration of the software-defined neural network accelerator module in XPU, analyzing synthesis/backend reports, and generating directed test cases using simulator.

2019.8 – 2020.1 **Hesai Technology, Shanghai** **FPGA Engineer(Intern)**

- Developed an FPGA-based real-time computing module for rotation speed and angle, achieving high output precision of 0.01 degree, applied to the mass-produced LiDAR named Pandar40.
- Received an "Excellent" overall internship evaluation and secured an extended job offer.

Papers

- **A Low-Bit Post-Quantization Algorithm for Lightweight Network Hardware-Friendly Quantization**
  Jixing Li, Benzhe Dai, **Zhiyuan Zhao**, et al. Journal of Xidian University, 2023.(under review)
- **A Novel Skip Connection Structure in Transformer**
  Xintao Xu, Yi Liu Y, **Zhiyuan Zhao**, et al. 2022 IEEE 2nd International Conference on Software Engineering and Artificial Intelligence (SEAI). IEEE, 2022: 83-87.(EI)
- **Fire Detection System Based on Embedded Microcontroller**
  Shixing Liu, **Zhiyuan Zhao**, et al. China Science and Technology Online, 2020.

Chinese Patents

- **Adaptive Bandwidth Computing Engine Based on Pipeline Structure Accelerator**
  Zhiyuan Zhao, et al, Published July 27, 2023. (2023109121405)
- **FPGA-Based Low Resource Overhead TCP/IP Protocol Stack**
  Zhiyuan Zhao, et al, Published July 25, 2023. (202310912039X)
- **Preprocessing Device and Method for Multi-Channel Sparse Data**
  Ruixiu Qiao, Gang Chen, **Zhiyuan Zhao**, et al, Published June 27, 2023. (2023107603475)
- **A Lightweight Neural Network Accelerator**
  Zhiyuan Zhao, et al, Published March 27, 2023. (2023103029157)

Skills

Hardware Design: Proficient in Verilog and SystemVerilog, Familiar with various bus protocols such as AXI4, AXI Stream, APB, TCP/IP, I2C, SPI and UART. Experienced in timing analysis and PPA optimization.

Algorithm and Programming: Familiar with Transformer, CNN, LSTM and other deep learning algorithms, Proficient in Python, Makefile, Git, Latex and C. Knowledgeable with Tcl, Shell, CUDA, and RISC-V assembly.

Design Tools: Proficient in Vivado, Experienced with DC, VCS, Verdi, Verilator and Modelsim.

Awards & Certificates

2019	National College Students Integrated Circuit Innovation and Entrepreneurship Competition **Second Prize** in National Final and **First Prize** in Huazhong Regional Competition **(Top 3%)**	
2018	College English Test Band 6	**Passed**
2018	Hefei University of Technology Electronic Design Competition	**First Prize (Top 10%)**
2018	Anhui Provincial College Students Robot Competition	**First Prize (Top 5%)**
2018	National Computer Rank Examination Certificate of Level 3	**Passed**