Yifan Qin

Notre Dame, IN 46556, USA
+1 574-401-9049 ⋄ yqin3@nd.edu

EDUCATION

University of Notre Dame 2022 - present
Ph.D. Candidate, Computer science and engineering
Research interest: Efficient and robust deep learning, Computing-in-memory, AI accelerator, Efficient LLM and generative AI.
Honorably work with Prof. Yiyu Shi & Prof. X. Sharon Hu

Huazhong University of Science and Technology 2018 - 2021
MS, Software engineering
Research interest: quantized low-bit neural networks with RRAM

Huazhong University of Science and Technology 2013 - 2017
BS, Electronic science and technology

AWARDS AND HONORS

William J. McCalla Best Paper Award Candidate at IEEE/ACM ICCAD (10 out of 802 submissions) 2024
Young Fellow (DAC) 2024
William J. McCalla Best Paper Award at IEEE/ACM ICCAD (2 out of 750 submissions) 2023
Young Fellow (DAC) 2023
Outstanding Graduates (HUST) 2020 - 2021
Outstanding Volunteer Docent (Wuhan Museum) 2015 - 2016
National 2nd Prize (Contemporary Undergraduate Mathematical Contest in Modeling) 2015

RESEARCH EXPERIENCE

University of Notre Dame Notre Dame, IN
Doctoral Researcher August 2022 - present
Established and implemented hardware/software codesign methods to mitigate the impact of device variations and noise on inference of non-volatile compute-in-memory accelerators. Achieved high robust and efficient solutions for LLM models and AI accelerator.

Hong Kong University of Science and Technology (HKUST)
AI Chip Center for Emerging Smart Systems(ACCESS) Hong Kong
Intern May 2024 - July 2024
Designed and implemented a convolutional neural network system for ventricular arrhythmia detection with a 40nm LP TSMC CNN accelerator, delivering a deployable chip demo. Led the full-stack design, from UI to backend, achieving substantial reductions in inference latency and energy consumption through optimized quantization and pruning techniques, demonstrating high-performance real-time detection capabilities.

Huazhong University of Science and Technology Wuhan, Hubei
Master's Researcher, Research Assistant August 2018 - June 2022
Designed low-bit quantized CNNs for RRAM accelerators, addressing non-idealities of RRAM crossbars during inference. Developed a novel binary neural network RRAM accelerator with half area and maintained high accuracy.

PUBLICATION

Journal

[1] Han Bao, Yifan Qin, Jia Chen, Ling Yang, Jiancong Li, Houji Zhou, Yi Li, and Xiangshui Miao. "Quantization and sparsity-aware processing for energy-efficient NVM-based convolutional neural networks". In: *Frontiers in Electronics* 3 (2022), p. 954661.

[2] Yifan Qin, Han Bao, Feng Wang, Jia Chen, Yi Li, and Xiangshui Miao. "Recent progress on memristive convolutional neural networks for edge intelligence". In: *Advanced Intelligent Systems* 2.11 (2020), p. 2000114. (**Back Cover**).

[3] Yifan Qin, Rui Kuang, Xiaodi Huang, Yi Li, Jia Chen, and Xiangshui Miao. "Design of high robustness BNN inference accelerator based on binary memristors". In: *IEEE Transactions on Electron Devices* 67.8 (2020), pp. 3435–3441.

Conference

[1] Jianbo Liu, Zephan Enciso, Boyang Cheng, Likai Pei, Steven Davis, Yifan Qin, Zhenge Jia, Xiaobo Sharon Hu, Yiyu Shi, and Ningyuan Cao. "A 65nm Uncertainty-quantifiable Ventricular Arrhythmia Detection Engine with 1.75 µJ per Inference". In: *Proceedings of the IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE. 2025.

[2] Yifan Qin, Zhenge Jia, Zheyu Yan, Jay Mok, Manto Yung, Yu Liu, Xuejiao Liu, Wujie Wen, Luhong Liang, Kwang-Ting Tim Cheng, X. Sharon Hu, and Yiyu Shi. "A 10.60 µW 150 GOPS Mixed-Bit-Width Sparse CNN Accelerator for Life-Threatening Ventricular Arrhythmia Detection". In: *Proceedings of the Asia and South Pacific Design Automation Conference (ASP-DAC)*. ACM. 2025.

[3] Likai Pei*, Yifan Qin*, Zephan M. Enciso, Boyang Cheng, Jianbo Liu, Steven Davis, Zhenge Jia, Michael Niemier, Yiyu Shi, X. Sharon Hu, and Ningyuan Cao. "Towards Uncertainty-Quantifiable Biomedical Intelligence: Mixed-signal Compute-in-Entropy for Bayesian Neural Networks". In: *2024 IEEE/ACM International Conference on Computer Aided Design (ICCAD)*. IEEE. 2024. (* contributed equally)(acceptance rate 24%)(**2024 William J. McCalla Best Paper Award Candidate, 10 out of 802 submissions**).

[4] Yifan Qin, Zheyu Yan, Zixuan Pan, Wujie Wen, Xiaobo Sharon Hu, and Yiyu Shi. "TSB: Tiny Shared Block for Efficient DNN Deployment on NVCIM Accelerators". In: *2024 IEEE/ACM International Conference on Computer Aided Design (ICCAD)*. IEEE. 2024. (acceptance rate 24%).

[5] Yifan Qin, Zheyu Yan, Wujie Wen, Xiaobo Sharon Hu, and Yiyu Shi. "Sustainable Deployment of Deep Neural Networks on Non-Volatile Compute-in-Memory Accelerators". In: *International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)*. IEEE. 2024.

[6] Yifan Qin, Zheyu Yan, Wujie Wen, Xiaobo Sharon Hu, and Yiyu Shi. "Negative Feedback Training: A Novel Concept to Improve Robustness of NVCiM DNN Accelerators". In: *arXiv preprint arXiv:2305.14561*. 2023. (under review).

[7] Zheyu Yan, Yifan Qin, Xiaobo Sharon Hu, and Yiyu Shi. "On the viability of using LLMs for SW/HW co-design: An example in designing CiM DNN accelerators". In: *2023 IEEE 36th International System-on-Chip Conference (SOCC)*. IEEE. 2023, pp. 1–6.

[8] Zheyu Yan, Yifan Qin, Wujie Wen, Xiaobo Sharon Hu, and Yiyu Shi. "Improving realistic worst-case performance of NVCiM DNN accelerators through training with right-censored gaussian noise". In: *2023 IEEE/ACM International Conference on Computer Aided Design (ICCAD)*. IEEE. 2023, pp. 1–9. (**2023 William J. McCalla Best Paper Award, 2 out of 750 submissions**).

PRESENTATIONS & TALKS

Computer science department, Shandong University (SDU) **Aug, 2024**
Electrical engineering department, Zhejiang University (ZJU) **Aug, 2024**
University of Michigan-Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University (SJTU) **Aug, 2024**
Electrical engineering department, Southern University of Science and Technology (SUSTech) **July, 2024**
AI Chip Center for Emerging Smart Systems, Hong Kong University of Science and Technology (HKUST) **June, 2024**

TEACHING EXPERIENCE

CSE-40868 Neural Networks, TA SP23

REVIWER FOR JOURNALS/CONFERENCE

ACM/IEEE International Conference on Computer-Aided Design (ICCAD) 2024

LEADERSHIP AND SERVICE

Member, Graduate Student Association, Huazhong University of Science and Technology	2019-2020
Volunteer Docent, Wuhan Museum	2015-2016
Team Captain, College Table Tennis Team, Huazhong University of Science and Technology	2016
President, Table Tennis Association, Huazhong University of Science and Technology	2015-2016