Xiaoxuan Yang

Ph.D. Candidate
Electrical and Computer Engineering Department
Duke University

Address: 534 Research Dr, Room 431
Durham, NC 27705
Email: xy92@duke.edu
Tel: +1-310-447-4619
Website: xiaoxuan-yang.github.io

Research Interests

- **Efficient Processing-in-Memory-based System Design**
- **Robust and Reliable Hardware-Software Co-Design for Non-Volatile Memory**
- **Generalized Machine Learning Algorithm for Robustness**

Education

Duke University Expected: Jun. 2023
Ph.D. Candidate in Electrical and Computer Engineering Advisors: Dr. Hai Helen Li and Dr. Yiran Chen
Thesis: Improving the efficiency and robustness of in-memory computing in emerging technologies.

University of California, Los Angeles (UCLA) Jun. 2018
M.S. in Electrical Engineering Advisor: Dr. Ramin Ramezani

Tsinghua University Jul. 2016
B.S. in Electrical Engineering Advisor: Dr. Chen Shen
Thesis: Power system transient stability evaluation method based on measurement.

Professional Experience

KLA Corporation May 2019 – Aug. 2019
Research Intern, Advanced Algorithm Group Mentor: Dr. Heng (Helen) Liu
Project: Customer-friendly wafer map inpainting with Generative Adversarial Network.

Sohu, Inc. Jun. 2017 - Aug. 2017
Technology Intern, Changyan Forum Group Mentor: Mr. Chao Chen
Project: Effective forum service system for users to offer customized service to subscribers.

Awards

- **Rising Scholars Postdoc Fellow**, University of Virginia, School of Engineering and Applied Science, 2023
- **NSF iREDEFINE Fellow**, ECE Department Heads Association Annual Conference, 2023
- **Third Place of ACM Student Research Competition SRC** at International Conference on Computer-Aided Design (ICCAD), 2022
- **Rising Star in Electrical Engineering and Computer Science (EECS)**, 2022
- **Best Research Award at ACM SIGDA Ph.D. Forum** at Design Automation Conference (DAC), 2022
- **Travel Awards** for iREDEFINE Workshop 2023, ACM SRC @ ICCAD 2022, ACM Ph.D. Forum @ DAC 2022, and IGSC 2021
- **Duke Graduate School Conference Travel Award**, 2022
- **Duke Electrical and Computer Engineering Conference Travel Fellowship**, 2022
- **Duke Electrical and Computer Engineering Diversity Award**, 2018
- **Henry Samueli Fellowship**, UCLA, 2018
- **Zheng-Geru Academic Scholarship**, Tsinghua University, 2015
- **Cai-Xiong Academic Scholarship**, Tsinghua University, 2013

Publications

Journal Articles

[1] **X. Yang**, H. Yang, J. R. Doppa, P. P. Pande, K. Chakrabarty, and H. H. Li. "ESSENCE: Exploiting Structured Stochastic Gradient Pruning for Endurance-aware ReRAM-based In-Memory Training Systems." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, 2022.

[2] C. Wu, **X. Yang**, Y. Chen, and M. Li. "Photonic Bayesian Neural Network using Programmed Optical Noises." *IEEE Journal of Selected Topics in Quantum Electronics (JSTQE)*, vol. 29, no. 2: Optical Computing, pp. 1-6, March-April 2023.

[3] **X. Yang**, C. Wu, M. Li, and Y. Chen. "Tolerating Noise Effects in Processing-in-Memory Systems for Neural Networks: A Hardware–Software Codesign Perspective". *Advanced Intelligent System*, 2200029 (2022).

[4] **X. Yang***, B. Taylor*, A. Wu, Y. Chen, and L. O. Chua. "Research Progress on Memristor: From Synapses to Computing Systems." *IEEE Transactions on Circuits and Systems I: Regular Papers (TCAS-I)*, vol. 69, no. 5, pp. 1845-1857, May 2022. [**Selected as TCAS-I Highlight**]

[5] C. Wu, **X. Yang**, H. Yu, R. Peng, I. Takeuchi, Y. Chen, and M. Li. "Harnessing Optoelectronic Noises in a Photonic Generative Network." *Science Advances* 8, no. 3 (2022): eabm2956.

Conference Proceedings

[6] **X. Yang**, S. Li, Q. Zheng, and Y. Chen. "Improving the Robustness and Efficiency of PIM-based Architecture by SW/HW Co-Design." In *Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2023.

[7] J. Henkel, H. H. Li, A. Raghunathan, M. B. Tahoori, S. Venkataramani, **X. Yang**, and G. Zervakis. "Approximate Computing and the Efficient Machine Learning Expedition." In *International Conference on Computer-Aided Design (ICCAD)*, 2022.

[8] **X. Yang**, H. Yang, J. Zhang, H. H. Li, and Y. Chen. "On Building Efficient and Robust Neural Network Designs." In *Asilomar Conference on Signals, Systems, and Computers (ASILOMAR)*, 2022.

[9] **X. Yang***, H. Yang*, N. Z. Gong, and Y. Chen. "HERO: Hessian-Enhanced Robust Optimization for Unifying and Improving Generalization and Quantization Performance" In Proceedings of *59th Design Automation Conference (DAC)*, pp. 25-30. 2022. [**Rank First in the Track**]

[10] C. Wu, **X. Yang**, H. Yu, I. Takeuchi, Y. Chen, and M. Li. "Optical Generative Adversarial Network based on Programmable Phase-change Photonics." In *CLEO: Science and Innovations*, pp. STu1G-3. Optical Society of America, 2021.

[11] **X. Yang**, S. Belakaria, B. K. Joardar, H. Yang, J. R. Doppa, P. P. Pande, K. Chakrabarty, and H. H. Li. "Multi-Objective Optimization of ReRAM Crossbars for Robust DNN Inferencing under Stochastic Noise." In *Proceedings of the 40th International Conference on Computer-Aided Design (ICCAD)*, pp. 1-9. 2021.

[12] **X. Yang**, B. Yan, H. H. Li, and Y. Chen. "ReTransformer: ReRAM-based Processing-In-Memory Architecture for Transformer Acceleration." In *Proceedings of the 39th International Conference on Computer-Aided Design (ICCAD)*, pp. 1-9. 2020. [**Rank First in the Track**]

Manuscripts

[13] G. Qian, **X. Yang**, S. Li, and Y. Chen. "Selecting the Optimal Layerwise Precisions with Iterative Post-training Quantization."

[14] S. Li, E. Hanson, Q. Zheng, **X. Yang**, and Y. Chen. "Hardware-Algorithm Joint Optimization for Sparse DNN."

[15] C. Wolters, B. Taylor, E. Hanson, **X. Yang**, U. Schlichtmann, and Y. Chen. "Biologically Plausible Learning on Neuromorphic Hardware Architectures."

Mentorship and Teaching Assistant Experience

- **Summer Intern Mentor in CEI Lab** Summer 2022
  12-Week Research Project Exploration
  - I mentor one high-school student on the neural network mixed quantization project with an iterative approach. The project is summarized in a technical report.
  - I mentor one undergraduate student on the thesis work of biologically plausible learning on neuromorphic hardware architectures, and one research paper is submitted to a conference.
- **TA for Enterprise Storage Architecture** Fall 2020
  Instructor: Dr. Tyler K Bletsch
  - I am the sole TA for this graduate course, working on the gradings of homework, code projects, and exams.
- **TA for Introduction to Signals and Systems** Spring 2020
  Instructor: Dr. Vahid Tarokh
  - As a TA for this undergraduate course, I take part in the design and grading of homework and tests.
- **TA for Neural Signal Processing** Spring 2018
  Instructor: Dr. Kao Jonathan
  - I work as TA for this undergraduate- and graduate-level course. Aside from grading, I prepare the jupyter notebook templates for homework and lead two weekly discussion sections.

Service

Journal Reviewer

- ACM Transactions on Design Automation of Electronic Systems (TODAES)
- ACM Transactions on Embedded Computing Systems (TECS)
- IEEE Design & Test (D&T)
- IEEE Embedded Systems Letters (ESL)
- IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS)
- IEEE Journal of Exploratory Solid-State Computational Devices and Circuits (JxCDC)
- IEEE Transactions on Circuits and Systems I: Regular Papers (TCAS-I)
- IEEE Transactions on Computers (TC)
- IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)
- IEEE Transactions on Emerging Topics in Computing (TETC)
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- IEEE Transactions on Very Large Scale Integration Systems (TVLSI)

Conference Reviewer

- IEEE International Conference on Artificial Intelligence Circuits & Systems (AICAS)
- Asilomar Conference on Signals, Systems, and Computers (ASILOMAR)
- AAAI Conference on Artificial Intelligence (AAAI)
- Secondary Reviewer: DAC, ICCAD, DATE, ASP-DAC, HPCA, ISCA, SEC, CODES+ISSS, MLSys, etc.

Conference Service

- Session Chair for "Repeal Murphy's Law: Avoid Errors" at DAC 2022

University Service

- Panelist for "Science & Engineering Exploration in Durham (SEED)" at 2022 First Year Students Orientation

Education Outreach

- Volunteer in "COSMOS Education Toolkit @ Inspiring Minds" at Hillside High School, Durham

Presentations

Oral Presentations

- **ASP-DAC**, Hybrid. Jan. 2023
  Improving the Robustness and Efficiency of PIM-based Architecture by SW/HW Co-design
- **ICCAD**, San Diego. Nov. 2022
  Efficient Processing-in-Memory Design for Transformer-based Models
- **ACM Student Research Contest at ICCAD (Final Round)**, San Diego. [**Third Place**] Oct. 2022
  Improving the Efficiency and Robustness of In-Memory Computing in Emerging Technologies
- **ASILOMAR**, Hybrid. Oct. 2022
  On Building Efficient and Robust Neural Network Designs
- **ICCAD**, Hybrid. Nov. 2021
  Multi-Objective Optimization of ReRAM Crossbars for Robust DNN Inferencing under Stochastic Noise
- **ICCAD**, Virtual. Nov. 2020
  ReTransformer: ReRAM-based Processing-In-Memory Architecture for Transformer Acceleration

Invited Seminars

- **Improving the Efficiency and Robustness of In-Memory Computing in Emerging Technologies**
  Indiana University Bloomington, Bloomington. Mar. 2023
  University of Texas at Dallas, Richardson. Mar. 2023
  Auburn University, Auburn. Mar. 2023
  Louisiana State University, Baton Rouge. Mar. 2023
  University of Oklahoma, Norman. Mar. 2023
  University of Virginia, Charlottesville. Mar. 2023
  Stevens Institute of Technology, Hoboken. Feb. 2023
  University of Central Florida, Orlando. Feb. 2023
  Purdue University, West Lafayette. Feb. 2023
  Colorado State University, Fort Collins. Feb. 2023
  Washington State University, Pullman. Feb. 2023
  George Mason University, Virtual. Oct. 2022

Posters

- **iREDEFINE Workshop 2023**, Santa Ana Pueblo. Mar. 2023
  Improving the Efficiency and Robustness of In-Memory Computing in Emerging Technologies
- **MEMRISYS Workshop 2022**, Cambridge. Nov. 2022
  ESSENCE: Exploiting Structured Stochastic Gradient Pruning for Endurance-aware Memristor-based In-Memory Training Systems
- **ACM Student Research Contest at ICCAD**, San Diego. Oct. 2022
  Improving the Efficiency and Robustness of In-Memory Computing in Emerging Technologies
- **Rising Stars Workshop 2022 at UT Austin**, Austin. Oct. 2022
  Improving the Efficiency and Robustness of In-Memory Computing in Emerging Technologies
- **NSF AI Institute ATHENA Annual Showcase at Duke**, Durham. Aug. 2022
  HERO: Hessian-Enhanced Robust Optimization for Unifying and Improving Generalization and Quantization Performance
- **Fifth ACSIC Symposium on Frontiers in Computing**, Chicago. Aug. 2022
  Improving the Efficiency and Robustness of In-Memory Computing in Emerging Technologies
- **ACM SIGDA Ph.D. Forum at DAC**, San Francisco. [**Best Research Award**] Jul. 2022
  Improving the Efficiency and Robustness of In-Memory Computing in Emerging Technologies
- **Intern Poster Session at KLA Corporation**, Milpitas. Aug. 2019
  Wafer Map Inpainting with Generative Neural Network