Yongan Zhang

832-566-7124 | yz87@rice.edu | Personal Site

Research Interest

- Hardware acceleration for AI algorithms
- Software/Hardware co-design for efficient AI implementation

Education

- **Rice University** — Houston, TX, USA
  M.S., Electrical and Computer Engineering, Advisor: Prof. Yingyan Lin — *Jan. 2020 – May 2023*
- **Rice University** — Houston, TX, USA
  B.S., Electrical and Computer Engineering — *Aug. 2015 – May 2019*

Experiences

- **Research Assistant, Rice University** — Mentor: Dr. Joseph Cavallaro
  Parallel Hardware Applications in Science and Technology (PHAST) — *Jan 2017 – May 2018*
- **System Engineer Intern, PHAZR, Inc.** — Mentor: Dr. Robert Daniels
  5G Millimeter wave systems for the licensed-bands in the 24-40 GHz — *May 2018 – Aug 2018*
- **Ph.D. Intern, PNNL** — Mentor: Dr. Ang Li
  Multi-FPGA acceleration for scalable Graph Neural Networks implementation — *Jan 2022 – May 2022*
  - Designed the multi-FPGA architecture for large GNN acceleration
  - Implemented from arch design to final board deployment (fixed model-to-hardware mapping)
  - Worked with Xilinx HLS and Vivado for arch, and Pynq for deployment
- **Research Intern, Meta** — Mentor: Dr. Yuecheng Li
  Reconfigurable hardware acceleration for VR mobile telepresence pipeline — *May 2022 – Dec 2022*
  - Designed the run-time reconfigurable architecture for improved hardware resource reuse
  - Designed the fine-grained operation scheduling for model-to-hardware mapping
  - Designed RTL-verified performance modeling for flexible DSE
  - Constructed design automation flow to auto generate the arch design and scheduling given Pytorch models
  - Worked with a hybrid of Catapult HLS, Vivado, RTL, C++ and Python for the whole flow

Skills

C/C++, Verilog, Python; Vivado/Vitis, Catapult HLS, Pytorch, Tensorflow, Slurm

Publications

1. Y. Zhao, **Y. Zhang**, Y. Fu, X. Ouyang, C. Wan, S. Wu, A. Banta, M. John, A. Post, M. Razavi, J. Cavallaro, B. Aazhang, Y. Lin, "e-G2C: A 0.14-to-8.31 uJ/Inference NN-based Processor with Continuous On-chip Adaptation for Anomaly Detection and ECG Conversion from EGM", *IEEE Symposium on VLSI Technology and Circuits* (**VLSI**), 2022.
2. H. You, Y. Zhao, Z. Yu, C. Wang, Y. Fu, J. Yuan, S. Wu, S. Zhang, **Y. Zhang**, C. Li, V. Boominathan, A. Veeraraghavan, Z. Li, Y. Lin, "EyeCoD: Eye Tracking System Acceleration via FlatCam-Based Algorithm and Accelerator Co-Design", *IEEE/ACM International Symposium on Computer Architecture* (**ISCA**), 2022.
3. H. You, T. Geng, **Y. Zhang**, A. Li, Y. Lin, "GCoD: Graph Convolutional Network Acceleration via Dedicated Algorithm and Accelerator Co-Design", *IEEE International Symposium on High-Performance Computer Architecture* (**HPCA**), 2022.
4. **Y. Zhang**, H. You, Y. Fu, T. Geng, A. Li, Y. Lin, "G-CoS: GNN-Accelerator Co-Search Towards Both Better Accuracy and Efficiency", *IEEE/ACM International Conference on Computer-Aided Design* (**ICCAD**), 2021.
5. **Y. Zhang**, Y. Fu, W. Jiang, C. Li, H. You, M. Li, V. Chandra, Y. Lin, DIAN: "Differentiable Accelerator-Network Co-Search Towards Maximal DNN Efficiency", *ACM/IEEE International Symposium on Low Power Electronics and Design* (**ISLPED**), 2021.

6. **Y. Zhang**, A. Benta, Y. Fu, M. John, A. Post, M. Razavi, J. Cavallaro, B. Aazhang, Y. Lin, "RT-RCG: Neural Network and Accelerator Search Towards Effective and Real-time ECG Reconstruction from Intracardiac Electrograms", *The ACM Journal on Emerging Technologies in Computing Systems* (**JETC**), 2021.
7. Y. Fu, **Y. Zhang**, H. You, Y. Lin, "Auto-NBA: Efficient and Effective Search Over The Joint Space of Networks, Bitwidths, and Accelerators", *The International Conference on Machine Learning* (**ICML**), 2021.
8. Y. Fu, **Y. Zhang**, C Li, Z Yu, Y Lin, "A3C-S: Automated Agent Accelerator Co-Search towards Efficient Deep Reinforcement Learning", *The 58th Design Automation Conference* (**DAC**), 2021.
9. Y. Fu, Z. Yu, **Y Zhang**, Y Jiang, C Li, Y Liang, M Jiang, Z Wang, Y Lin, "InstantNet: Automated Generation and Deployment of Instantaneously Switchable-Precision Networks", *The 58th Design Automation Conference* (**DAC**), 2021.
10. T. Geng, C. Wu, **Y. Zhang**, C. Tang, C. Xie, H. You, M. Herbordt, Y. Lin, A. Li, "I-GCN: A Graph Convolutional Network Accelerator with Runtime Locality Enhancement through Islandization", *IEEE/ACM International Symposium on Microarchitecture* (**MICRO**), 2021.
11. M. Li, Z. Yu, **Y. Zhang**, Y. Fu, Y. Lin, "O-HAS: Optical Hardware Accelerator Search for Boosting Both Acceleration Performance and Development Speed", *IEEE/ACM International Conference on Computer-Aided Design* (**ICCAD**), 2021.
12. C. Li, Z. Yu, Y. Fu, **Y. Zhang**, Y. Zhao, H. You, Q. Yu, Y. Wang, Y. Lin, "HW-NAS-Bench: Hardware-Aware Neural Architecture Search Benchmark", *The International Conference on Learning Representations* (**ICLR**), 2021.
13. H. You, X. Chen, **Y. Zhang**, C. Li, S. Li, Z. Liu, Z. Wang, Y. Lin, "ShiftAddNet: A Hardware-Inspired Deep Network", *Conference on Neural Information Processing Systems* (**NeurIPS**), 2020.
14. Y. Zhao, C. Li, Y. Wang, P. Xu, **Y. Zhang**, Y. Lin, "DNN-Chip Predictor: A Multi-grained Graph-based Performance Simulator for DNN Accelerators", *International Conference on Acoustics, Speech, and Signal Processing* (**ICASSP**), 2020.
15. P. Xu, Y. Zhao, C. Hao, X. Zhang, Z. Guan, **Y. Zhang**, Y. Wang, D. Chen, Y. Lin, "AutoDNNchip: An Automated DNN Chip Predictor and Builder for Both FPGAs and ASICs", *ACM/SIGDA International Symposium on Field-Programmable Gate Arrays* (**FPGA**), 2020.

Awards

- Distinction in Research and Creative Work — May. 2019
- President's Honor Roll — Dec. 2017/ May. 2018

Teaching

- ELEC 327: Implementation of Digital Systems (*Teaching Assistant*) — 2018 Spring
- ELEC 539: Introduction to Communication Networks (*Teaching Assistant*) — 2020 Fall
- ELEC 515: Embedded Machine Learning (*Teaching Assistant*) — 2020 Fall
- ELEC 515: Embedded Machine Learning (*Teaching Assistant*) — 2021 Fall
- ELEC 526: High Performance Computer Architecture (*Teaching Assistant*) — 2022 Spring