HAORAN YOU

Tel.: +1-281-236-1978   Email: haoran.you@gatech.edu   Website: Haoran's page; Linkedin

EDUCATION

Georgia Institute of Technology   Atlanta, Georgia
Transferred Ph.D. supervised by Prof. Yingyan Lin, Computer Science (CS)   Dec. 2022 - Present

Rice University   Houston, Texas
M.S. supervised by Prof. Yingyan Lin, Electrical and Computer Engineering (ECE)   Sep. 2019 - Dec. 2022

Huazhong University of Science and Technology (HUST)   Wuhan, China
Bachelor of Electronic Information Engineering   Sep. 2015 - Jul. 2019

RESEARCH INTERESTS

- Efficient and Automated Machine Learning (ICML, NeurIPS, ICLR, AAAI, ECCV, TNNLS, TMLR)
- Algorithm and Hardware Accelerator Co-Design (ISCA, HPCA, MICRO, ICCAD, ISLPED)

PUBLICATIONS

- **H. You**, C. Li, P. Xu, Y. Fu, Y. Wang, X. Chen, R.G. Baraniuk, Z. Wang, Y. Lin.
  Drawing Early-Bird Tickets: Towards More Efficient Training of Deep Networks. **ICLR 2020 Spotlight Oral**

- **H. You**, Z. Sun, H. Shi, Z. Yu, Y. Zhao, Y. Zhang, C. Li, B. Li, Y. Lin.
  ViTCoD: Vision Transformer Acceleration via Dedicated Algorithm and Accelerator Co-Design. **HPCA 2023**

- **H. You**, X. Chen, Y. Zhang, C. Li, S. Li, Z. Liu, Z. Wang, Y. Lin
  ShiftAddNet: A Hardware-Inspired Deep Network. **NeurIPS 2020**

- **H. You**, B. Li, H. Shi, Y. Lin
  ShiftAddNAS: Hardware-Inspired Search for More Accurate and Efficient Neural Networks. **ICML 2022**

- **H. You**, B. Li, Z. Sun, X. Ouyang, Y. Lin
  SuperTickets: Drawing Task-Agnostic Lottery Tickets from Supernets via Jointly Searching and Pruning. **ECCV 2022**

- **H. You**, Z. Lu, Z. Zhou, Y. Fu, Y. Lin.
  Early-Bird GCNs: Graph-Network Co-Optimization Towards More Efficient GCN Training and Inference. **AAAI 2022**

- **H. You**, T. Geng, Y. Zhang, A. Li, Y. Lin.
  GCoD: Graph Convolutional Network Accelerationvia Dedicated Algorithm and Accelerator Co-Design. **HPCA 2022**

- **H. You**, Y. Zhao, Z. Yu, C. Wan, Y. Fu, C. Li, S. Zhang, S. Wu, J. Yuan, Y. Zhang, Vivek B., Ashok V., Z. Li, Y. Lin.
  EyeCoD: Eye Tracking System Acceleration via FlatCam-Based Algorithm and Accelerator Co-Design. **ISCA 2022**

- **H. You**, R. Balestriero, Z. Lu, Y. Kou, Y. Lin, R.G. Baraniuk.
  Max-Affine Spline Insights Into Deep Network Pruning. **TMLR 2022**

- **H. You**, Y. Cheng, T. Cheng, C. Li, P. Zhou.
  Bayesian Cycle-Consistent GANs via Marginalizing Latent Sampling. **IEEE TNNLS**

- **H. You**, Y. Xiong, X. Dai, B. Wu, P. Zhang, H. Fan, P. Vajda, Y. Lin.
  Castling-ViT: Compressing Attention via Switching Towards Linear-Angular Attention at ViT Inference. **Under Review**

- **H. You**, T. Geng, C. Wan, Q. An, Y. Zhang, Z. Lu, Y. Sun, C. Li, Y. Zhao, Y. Lin.
  HA-GCN: Hardware-Aware Graph Convolutional Networks Compression. **Under Review**

- Y. Fu, **H. You**, Y. Zhao, Y. Wang, C. Li, K. Gopalakrishnan, Z. Wang, Y. Lin.
  FracTrain: Fractionally Squeezing Bit Savings Both Temporally and Spatially for Efficient DNN Training. **NeurIPS 2020**

- H. Shi, **H. You**, Y. Zhao, Z. Wang, Y. Lin.
  NASA: Neural Architecture Search and Acceleration for Hardware Inspired Hybrid Networks. **ICCAD 2022**

- Y. Zhang, **H. You**, Y. Fu, T. Geng, A. Li, Y. Lin
  G-CoS: GNN-Accelerator Co-Search Towards Both Better Accuracy and Efficiency. **ICCAD 2021**

- C. Li, T. Chen, **H. You**, Z. Wang, Y. Lin.
  HALO: Hardware-Aware Learning to Optimize. **ECCV 2020**

- Y. Zhao, X. Chen, Y. Wang, C. Li, **H. You**, Y. Fu, Y. Xie, Z. Wang, Y. Lin.
  SmartExchange: Trading Higher-cost Memory Storage/Access for Lower-cost Computation. **ISCA 2020**

- T. Geng, C. Wu, Y. Zhang, C. Tan, C. Xie, **H. You**, M. Herbordt, Y. Lin, A. Li
  I-GCN: A GCN Accelerator with Runtime Locality Enhancement through Islandization. **MICRO 2021**

- X. Chen, Y. Zhao, Y. Wang, P. Xu, **H. You**, C. Li, Y. Fu, Y. Lin, Z. Wang
  SmartDeal: Re-Modeling Deep Network Weights for Efficient Inference and Training. **IEEE TNNLS**

- 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. **ISLPED 2021**

- C. Li, Z. Yu, Y. Fu, Y. Zhang, Y. Zhao, **H. You**, Q. Yu, Y. Wang, C. Hao, Y. Lin.
  HW-NAS-Bench: Hardware-Aware Neural Architecture Search Benchmark. **ICLR 2021 Spotlight Oral**

RESEARCH EXPERIENCE

EIC Lab, Georgia Institute of Technology, Atlanta, Georgia   Dec. 202 - Present
Under the instruction of Prof. Yingyan Lin

EIC Lab, Rice University, Houston, Texas   Sep. 2019 - Dec. 2022
Under the instruction of Prof. Yingyan Lin

Meta, Burlingame, California   May. 2022 - Dec. 2022
Under the instruction of Dr. Peizhao Zhang, Dr. Xiaoliang Dai, Dr. Bichen Wu

Baidu USA, Sunnyvale, California   May. 2021 - Dec. 2021
Under the instruction of Dr. Baopu Li

WNLO & Microsoft, Wuhan, China   Jan. 2018 - Dec. 2018
Under the instruction of Dr. Yu Cheng and Prof. Pan Zhou

SELECTED AWARDS

- PhD Growth Annual Rock Star Award   2022 - 2023
- ViTCoD Work Was Selected for Meta Faculty Award   Nov. 2022
- i-FlatCam Won First Place in University Demo Best Demonstration at DAC Conference   Jul. 2022
- ISCA 2022 Student Trave Grant   Jun. 2022
- Samsung Scholarship   2017 - 2018
- International Interdisciplinary Contest In Modeling (ICM), Meritorious Winner.   Apr. 2018
- Selected as *Outstanding Undergraduates in Term of Academic Performance*, (HUST 1%).   Dec. 2017
- The Eighth Chinese Mathematics Competition (CMC), First Price (China 5%).   Nov. 2016

CONFERENCE PRESENTATION

- European Conference on Computer Vision (ECCV), 2022 [Link]
- The 49th International Symposium on Computer Architecture (ISCA), 2022 [Link]
- The 39th International Conference on Machine Learning (ICML), 2022 [Link]
- The 28th International Symposium on High-Performance Computer Architecture (HPCA), 2022 [Link]
- The 36th AAAI Conference on Artificial Intelligence (AAAI), 2022 [Link]
- Ken Kennedy AI and Data Science Conference, 2021 [Link]
- The 8th International Conference on Learning Representations (ICLR), 2020 [Link]
- The 34th Conference on Neural Information Processing Systems (NeurIPS), 2020 [Link]

INVITED TALKS

- Invited Speaker at UT Austin Energy-Aware Computing Group, 2023 [Link]
- Invited Speaker at the Tutorial in Asian Conference on Computer Vision, 2022 [Link]

PC OR REVIEWER SERVICES

- 2023: ICLR; CVPR; MLSys;
- 2022: CVPR; ICML; ICLR; NeurIPS; ECCV; AAAI; JCST;
- 2021: ICML; NeurIPS; ICLR; MLSys; IEEE TNNLS;
- 2020: ICML; ICLR; CVPR;