# **HAORAN YOU**

Tel.: +1-281-236-1978 Email: haoran.you@gatech.edu Profile: Homepage; Linkedin; Github

# **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, CVPR, 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.
   Selected as an ICLR Spotlight! [Link]
   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.
   Selected as the Meta Faculty Research Award of 2022! [Link]
   ViTCoD: Vision Transformer Acceleration via Dedicated Algorithm and Accelerator Co-Design. HPCA 2023
- 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. *Selected as the IEEE Micro's Top Pick of 2023! [Link]*EyeCoD: Eye Tracking System Acceleration via FlatCam-Based Algorithm and Accelerator Co-Design. **ISCA 2022**
- H. You, H. Shi, Y. Guo, Y. Lin. ShiftAddViT: Mixture of Multiplication Primitives Towards Efficient Vision Transformer. **NeurIPS 2023**
- 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. CVPR 2023
- 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, B. Li, H. Shi, Y. Lin ShiftAddNAS: Hardware-Inspired Search for More Accurate and Efficient Neural Networks. ICML 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, X. Chen, Y. Zhang, C. Li, S. Li, Z. Liu, Z. Wang, Y. Lin ShiftAddNet: A Hardware-Inspired Deep Network. NeurIPS 2020

- 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**, 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, Z. Li, H. You, Y. Fu, Y. Zhang, C. Li, C. Wan, S. Wu, X. Ouyang, V. Boominathan, A. Veeraraghavan, Y. Lin. Won the first place in University Best Demonstration at DAC 2022! [Link]
   i-FlatCam: A 253 FPS, 91.49 μJ/Frame Ultra-Compact Intelligent Lensless Camera System for Eye Tracking. VLSI 2022
- 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
- Y. Fu, Z. Ye, J. Yuan, S. Zhang, S. Li, H. You, Y. Lin.
   Gen-NeRF: Efficient and Generalizable Neural Radiance Fields via Algorithm-Hardware Co-Design. ISCA 2023
- S. Li, C. Li, W. Zhu, B. Yu, Y. Zhao, C. Wan, H. You, H. Shi, Y. Lin
   Instant-3D: Instant Neural Radiance Fields Training Towards Real-Time AR/VR 3D Reconstruction. ISCA 2023
- Z. Yu, Y. Fu, S. Wu, M. Li, H. You, Y. Lin.
   LDP: Learnable Dynamic Precision for Efficient Deep Neural Network Training and Inference. TinyML 2022

# RESEARCH EXPERIENCE

# EIC Lab, Georgia Institute of Technology, Atlanta, Georgia Under the instruction of Prof. Yingyan Lin EIC Lab, Rice University, Houston, Texas Under the instruction of Prof. Yingyan Lin Meta, Burlingame, California Dec. 2022 - Present Sep. 2019 - Dec. 2022 May. 2022 - Dec. 2022

Mentors/Peers: Dr. Yunyang Xiong, Peizhao Zhang, Xiaoliang Dai, Bichen Wu

# Baidu USA, Sunnyvale, California

Under the instruction of Dr. Baopu Li

WNLO & Microsoft, Wuhan, China

Under the instruction of Dr. Yu Cheng and Prof. Pan Zhou

Jan. 2018 - Dec. 2018

May. 2021 - Dec. 2021

#### SELECTED AWARDS

TOTAL PARAMETER AND ALL AND AL	023
• ICCAD 2023 Student Travel Grant! Oct. 20	
• Haoran was selected as one of the ML and Systems Rising Stars of 2023! May. 20	023
• Haoran won the Best Poster Award at the SCS Poster Competition of GaTech! Apr. 20	023
• Haoran was selected by GaTech to receive the Outstanding Graduate Research Assistant Award! Mar. 20	023
• EyeCoD (ISCA'22) was selected for being the IEEE Micro's Top Pick of 2023! Jan. 20	023
• HPCA 2023 Student Travel Grant! 2022 - 20	023
• Haoran got PhD Growth Annual Rock Star Award in EIC Lab! 2022 - 20	023
• ViTCoD (HPCA'23) was selected for Meta Faculty Research Award of 2022! Nov. 20	)22
• Our i-FlatCam won First Place in University Demo Best Demonstration at DAC Conference! Jul. 20	022
• ISCA 2022 Student Trave Grant! Jun. 20	022
• Samsung Scholarship! 2017 - 20	)18
• International Interdisciplinary Contest in Modeling (ICM), Meritorious Winner! Apr. 20	)18
• Selected as Outstanding Undergraduates in Term of Academic Performance, (HUST 1%)! Dec. 20	017
• The Eighth Chinese Mathematics Competition (CMC), First Price (China 5%)! Nov. 20	016

# **CONFERENCE PRESENTATION**

- Semiconductor Research Corporation (SRC) TECHCON, 2023 [Link]
- The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023 [Link]
- The 50th International Symposium on Computer Architecture (ISCA), 2023 [Workshop Link]
- The 29th International Symposium on High-Performance Computer Architecture (HPCA), 2023 [Link]
- 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 University of Rochester IntelliArch Lab, 2023 [Link]
- Invited Speaker at ISCA AutoDL Workshop, 2023 [Link]
- 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

- 2024: ICLR; AAAI, CVPR
- 2023: ICLR; CVPR; MLSys; ICML; ICCV; NeurIPS; IEEE TCAS-II
- 2022: CVPR; ICML; ICLR; NeurIPS; ECCV; AAAI; JCST;
- 2021: ICML; NeurIPS; ICLR; MLSys; IEEE TNNLS;
- 2020: ICML; ICLR; CVPR;

# MENTORING EXPERIENCE

- Huihong Shi (Ph.D. Student), Published two 1st author ICCAD'22 and TCAS-I journal papers, and co-authored ICML'22, HPCA'23, and TMLR journal papers, Next move: PhD@NanJing University
- Zhanyi Sun (B.S. Student), Published two 2nd author HPCA'23 and ECCV'22 papers, Next move: MS@CMU
- Zhihan Lu (B.S. Student), Published two 2nd author AAAI'22 and TMLR journal paper, Next move: MS@CMU
- Shang Wu (M.S. Student), Published co-authored ISCA'22/TMLR, Next move: PhD@Northwestern University
- Xu Ouyang (M.S. Student), Published one co-authored ECCV'22 paper, Next move: PhD@University of Virginia
- Zijian Zhou (B.S. Student), Published one co-authored AAAI'22 paper, Next move: MS@Rice University
- Yi Sun (B.S. Student), Submitted a co-authored paper, Next move: MS@CMU
- Yutong Kou (B.S. Student), Submitted a co-authored paper, Next move: PhD@HUST

# **VOLUNTEERING SERVICE**

• Student organization helper and website developer for EECS Rising Stars Workshop of 2023.