

<b>Contact Information</b>	School of Computing KAIST 291 Daehak-ro, Yuseong-gu Daejeon, South Korea, 34141	E-mail: <a href="mailto:jspark@casys.kaist.ac.kr">jspark@casys.kaist.ac.kr</a> URL: <a href="https://jongse-park.github.io">https://jongse-park.github.io</a>
<b>Research Interests</b>	Computer Architecture, HW/SW Codesign, AI Systems, Autonomous Systems	
<b>Employment</b>	Associate Professor. <b>KAIST</b> Assistant Professor. <b>KAIST</b> Product Engineer. <b>Bigstream Solutions Inc.</b>	Mar. 2024–date Dec. 2019–Feb. 2024 Jun. 2018–Nov. 2019
<b>Education</b>	<b>Ph.D. in Computer Science. Georgia Institute of Technology</b> • Advisor: Prof. Hadi Esmaeilzadeh • Dissertation: <i>Breaking the Abstractions for Productivity and Performance in the Era of Specialization</i> <b>M.S. in Computer Science. KAIST</b> • Advisor: Prof. Seungryoul Maeng • Thesis: <i>Dynamic Resource Reconfiguration on the Cloud for Improving Data Locality</i> <b>B.E. in Computer Science and Engineering. Sogang University</b> • Graduated with Honors	Aug. 2013–Aug. 2018           Feb. 2012           Feb. 2010
<b>Honors and Awards</b>	Distinguished paper award. IEEE Symposium on High Performance Computer Architecture. “TABLA: A Unified Template-Based Framework for Accelerating Statistical Machine Learning” Honorable Mention in IEEE Micro Top Picks from 2014 Computer Architecture Conferences. “General-Purpose Code Acceleration with Limited-Precision Analog Computation” Kwanjeong Foundation Scholarship, Kwanjeong Educational Foundation (KEF) National Full Scholarship, KAIST Dean’s Honored Graduate, Ranked 3 <sup>rd</sup> among graduates of the class of 2010 DMC General Management Track Scholarship, Samsung Electronics Co., Ltd Academic Scholarship, Sogang University, 7 semesters	2016           2015           2013–2018 2010–2012 2010 2009 2004–2009
<b>Refereed Conference Papers</b>	1. M. Kim, J. Hwang, G. Heo, S. Cho, D. Mahajan, <b>J. Park</b> , “Accelerating String-key Learned Index Structures via Memoization-based Incremental Training” in <i>International Conference on Very Large Data Bases (VLDB)</i> , August 2024 (To Appear). 2. Y. Kim, C. Oh, J. Hwang, W. Kim, S. Oh, Y. Lee, H. Sharma, A. Yazdanbakhsh, <b>J. Park</b> , “DaCapo: Accelerating Continuous Learning in Autonomous Systems for Video Analytics” in <i>International Symposium on Computer Architecture (ISCA)</i> , June 2024 (To Appear). 3. G. Heo, S. Lee, J. Cho, H. Choi, S. Lee, H. Ham, G. Kim, D. Mahajan, <b>J. Park</b> , “NeuPIMs: NPU-PIM Heterogeneous Acceleration for Batched LLM Inferencing” in <i>International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)</i> , April 2024 (To Appear). 4. S. Ghodrati, S. Kinzer, H. Xu, R. Mahapatra, Y. Kim, B. H. Ahn, D. K. Wang, L. Karthikeyan, A. Yazdanbakhsh, <b>J. Park</b> , N. S. Kim, H. Esmaeilzadeh, “Tandem Processor: Grappling with Emerging Operators in Neural Networks” in <i>International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)</i> , April 2024 (To Appear).	

5. Sunho Lee, Seonjin Na, Jungwoo Kim, Jongse Park, and Jaehyuk Huh, "Tunable Memory Protection for Secure Neural Processing Units" in *The 40th IEEE International Conference on Computer Design (ICCD)*, October 2022.
6. Bokyeong Kim, Soojin Hwang, Sanghoon Cha, Chang Hyun Park, Jongse Park, and Jaehyuk Huh, "Supporting Dynamic Translation Granularity for Hybrid Memory Systems" in *The 40th IEEE International Conference on Computer Design (ICCD)*, October 2022.
7. Joon Kyung Kim, Byung Hoon Ahn, Sean Kinzer, Soroush Ghodrati, Rohan Mahapatra, Brahmen-dra Yatham, Dohee Kim, Parisa Sarikhani, Babak Mahmoudi, Divya Mahajan, Jongse Park, Hadi Esmaeilzadeh, "Yin-Yang: Programming Abstraction for Cross-Domain Multi-Acceleration", in *IEEE Micro, special issue on Compiling for Accelerators*, 2022.
8. Jinwoo Hwang, Minsu Kim, Daeun Kim, Seungho Nam, Yoonsung Kim, Dohee Kim, Hardik Sharma, Jongse Park, "CoVA: Exploiting Compressed-Domain Analysis to Accelerate Video Analytics", in *USENIX Annual Technical Conference (ATC)*, July 2022.
9. Seungbeom Choi, Sunho Lee, Yeonjae Kim, Jongse Park, Youngjin Kwon, and Jaehyuk Huh, "Serving Heterogeneous Machine Learning Models on Multi-GPU Servers with Spatio-Temporal Sharing", in *USENIX Annual Technical Conference (ATC)*, July 2022.
10. S. Lee, J. Kim, S. Na, **J. Park**, and J. Huh, "TNPU: Supporting Trusted Execution with Tree-less Integrity Protection for Neural Processing Unit" in *The 27th IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, February 2022. [To appear]
11. S. Na, S. Lee, Y. Kim, **J. Park**, and J. Huh, "Common Counters: Compressed Encryption Counters for Secure GPU Memory" in *The 27th IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, February 2021.
12. S. Ghodrati, H. Sharma, S. Kinzer, A. Yazdanbakhsh, **J. Park**, N. Kim, D. Burger, and H. Esmaeilzadeh, "Mixed-Signal Charge-Domain Acceleration of Deep Neural Networks through Inter-leaved Bit-Partitioned Arithmetic" in *The 29th International Conference on Parallel Architectures and Compilation Techniques (PACT)*, October 2020.
13. Y. Li, **J. Park**, M. Alian, Y. Yuan, Q. Zheng, P. Pan, R. Wang, A. Schwing, H. Esmaeilzadeh, N. Kim, "A Network-Centric Hardware/Algorithm Co-Design to Accelerate Distributed Training of Deep Neural Networks," *The 50th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, October 2018.
14. H. Sharma, **J. Park**, B. Samynathan, B. Robatmili, S. Mirkhani, H. Esmaeilzadeh, "From Tensors to FPGAs: Accelerating Deep Learning," *A Symposium on High Performance Chips (Hot Chips)*, August 2018.
15. H. Sharma, **J. Park**, N. Suda, L. Lai, B. Chau, J. Kim, V. Chandra, H. Esmaeilzadeh, "Bit Fusion: Bit-Level Dynamically Composable Architecture for Accelerating Deep Neural Networks," *International Symposium on Computer Architecture (ISCA)*, June 2018.
16. **J. Park**, H. Sharma, D. Mahajan, J. Kim, P. Olds, H. Esmaeilzadeh, "Scale-Out Acceleration for Machine Learning," in *The 50th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, October 2017.
17. **J. Park**, E. Amaro, D. Mahajan, B. Thwaites, H. Esmaeilzadeh, "AXGAMES: Towards Crowdsourcing Quality Target Determination in Approximate Computing," in *International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, April 2016.
18. H. Sharma, **J. Park**, D. Mahajan, E. Amaro, J. Kim, C. Shao, A. Mishra, H. Esmaeilzadeh "From High-Level Deep Neural Models to FPGAs," in *The 49th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, October 2016.
19. D. Mahajan, **J. Park**, E. Amaro, H. Sharma, A. Yazdanbakhsh, J. Kim, H. Esmaeilzadeh, "TABLA: A Unified Template-based Framework for Accelerating Statistical Machine Learning," in *The 22nd IEEE Symposium on High Performance Computer Architecture (HPCA)*, March 2016.

**(Distinguished Paper Award)**

20. D. Mahajan, A. Yazdanbakhsh, **J. Park**, B. Thwaites, H. Esmaeilzadeh, "Towards Statistical Guarantees in Controlling Quality Tradeoffs in Approximate Acceleration," in *International Symposium on Computer Architecture (ISCA)*, June 2016.
21. A. Yazdanbakhsh, **J. Park**, H. Sharma, P. Lotfi-Kamran, H. Esmaeilzadeh, "Neural Acceleration for GPU Throughput Processors," in *The 48th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, December 2015.
22. **J. Park**, H. Esmaeilzadeh, X. Zhang, M. Naik, W. Harris, "FLEXJAVA: Language Support for Safe and Modular Approximate Programming," in *The 10th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering (FSE)*, September 2015.
23. A. Yazdanbakhsh, D. Mahajan, B. Thwaites, **J. Park**, A. Nagendrakumar, S. Sethuraman, K. Ramkrishnan, N. Ravindran, R. Jariwala, A. Rahimi, H. Esmaeilzadeh, K. Bazargan, "AXILOG: Language Support for Approximate Hardware Design," in *Design Automation and Test in Europe (DATE)*, March 2015.
24. R. S. Amant, A. Yazdanbakhsh, **J. Park**, B. Thwaites, H. Esmaeilzadeh, A. Hassibi, L. Ceze, D. Burger, "General-Purpose Code Acceleration with Limited-Precision Analog Computation," in *The 41th International Symposium on Computer Architecture (ISCA)*, June 2014.  
(Nominated for CACM Research Highlights; Honorable Mention in IEEE Micro Top Picks)
25. B. Thwaites, G. Pekhimenko, A. Yazdanbakhsh, **J. Park**, G. Mururu, H. Esmaeilzadeh, O. Mutlu, T. Mowry, "Rollback-Free Value Prediction with Approximate Loads," in *The 24th International Conference on Parallel Architectures and Compilation Techniques (PACT)*, August 2014.
26. J. Choi, **J. Park**, J. Seol, and S. Maeng, "Isolated Mini-domain for Trusted Cloud Computing," in *The 13th International Symposium on Cluster, Cloud, and Grid Computing (CCGrid)*, May 2013.
27. **J. Park**, D. Lee, B. Kim, J. Huh, S. Maeng, "Locality-aware Dynamic VM Reconfiguration on MapReduce Clouds," in *The 21st International ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, June 2012.

#### Refereed Journal Articles

1. J. Park, S. Kang, S. Lee, T. Kim, **J. Park**, Y. Kwon, and J. Huh, "Hardware Hardened Sandbox Enclaves for Trusted Serverless Computing" in *IEEE Transactions on Architecture and Code Optimization (TACO)*, 2023.
2. S. Noh, J. Koo, S. Lee, **J. Park**, and J. Kung, "FlexBlock: A Flexible DNN Training Accelerator with Multi-Mode Block Floating Point Support" in *IEEE Transactions on Computers (TC)*, 2023.
3. S. Lee, R. Hwang, **J. Park**, and M. Rhu, "HAMMER: Hardware-friendly Approximate Computing for Self-attention with Mean-redistribution and Linearization" in *IEEE Computer Architecture Letters (CAL)*, 2023.
4. W. Seo, S. Cha, Y. Kim, J. Huh, and **J. Park**, "SLO-aware Inference Scheduler for Heterogeneous Processors in Edge Platforms" in *Transactions on Architecture and Code Optimization (TACO)*, 2021.
5. D. Mahajan, K. Ramkrishnan, R. Jariwala, A. Yazdanbakhsh, **J. Park**, B. Thwaites, A. Nagendrakumar, A. Rahimi, H. Esmaeilzadeh, K. Bazargan, "AXILOG: Abstractions for Approximate Hardware Design and Reuse," in *IEEE Micro, special issue on Alternative Computing Designs and Technologies*, October 2015.

#### Refereed Workshop Papers

1. Y. Lee, **J. Park**, "LVS: A Learned Video Storage for Fast and Efficient Video Understanding" in *Efficient Deep Learning for Computer Vision (ECV) in conjunction with CVPR*, June 2024 (To Appear).
2. H. Sharma, **J. Park**, E. Amaro, B. Thwaites, P. Kotha, A. Gupta, J. Kim, A. Mishra, H. Esmaeilzadeh, "DNNWEAVER: From High-Level Deep Network Models to FPGA Acceleration," in *The Second Workshop on Cognitive Architectures (CogArch) in conjunction with ASPLOS*, April 2016.

3. D. Mahajan, A. Yazdanbakhsh, **J. Park**, B. Thwaites, H. Esmaeilzadeh, "Prediction-Based Quality Control for Approximate Accelerators," in *The Second Workshop on Approximate Computing Across the System Stack (WACAS) in conjunction with ASPLOS*, March 2015.
4. **J. Park**, K. Ni, X. Zhang, H. Esmaeilzadeh, M. Naik, "Expectation-Oriented Framework for Automating Approximate Programming," in *The First Workshop on Approximate Computing Across the System Stack (WACAS) in conjunction with ASPLOS*, March 2014.
5. A. Yazdanbakhsh, B. Thwaites, **J. Park**, H. Esmaeilzadeh, "Methodical Approximate Hardware Design and Reuse," in *The First Workshop on Approximate Computing Across the System Stack (WACAS) in conjunction with ASPLOS*, March 2014.
6. A. Yazdanbakhsh, R. Amant, B. Thwaites, **J. Park**, H. Esmaeilzadeh, A. Hassibi, L. Ceze, D. Burger, "Toward General-Purpose Code Acceleration with Analog Computation," in *The First Workshop on Approximate Computing Across the System Stack (WACAS) in conjunction with ASPLOS*, March 2014.
7. B. Thwaites, A. Yazdanbakhsh, **J. Park**, H. Esmaeilzadeh, "Bio-Accelerators: Bridging Biology and Silicon for General-Purpose Computing," in *Wild and Crazy Ideas (WACI) in conjunction with ASPLOS*, March 2014.

## Research Experience

- Research Assistant.** Alternative Computing Technology (ACT) Lab Aug. 2013–Aug. 2018
- Georgia Institute of Technology
  - Advisor: Prof. Hadi Esmaeilzadeh
- Visiting Researcher.** Alternative Computing Technology (ACT) Lab Jan. 2018–Aug. 2018
- University of California, San Diego
  - Advisor: Prof. Hadi Esmaeilzadeh
- Research Intern.** Architecture Research Group (ARG) May 2017–Aug. 2017
- NVIDIA Research
  - Mentors: Dr. Arslan Zulfiqar and Dr. Eiman Ebrahimi
  - Manager: Dr. Stephen Keckler
- Research Intern.** Catapult team Jan. 2016–May 2016
- Microsoft Research
  - Mentor: Dr. Eric Chung
  - Manager: Dr. Doug Burger
- Research Assistant.** Computer Architecture (CA) Lab Feb. 2010–Jul. 2013
- Korea Advanced Institute of Science and Technology (KAIST)
  - Advisor: Prof. Seungryoul Maeng

## Teaching Experience

- Instructor.**
- |  |             |
|--|-------------|
| • CS610: Parallel Processing   | Spring 2024 |
| • CS311: Computer Organization   | Spring 2024 |
| • CS411: System for AI   | Fall 2023   |
| • CS510: Computer Architecture   | Spring 2023 |
| • CS230: System Programming  | Fall 2022   |
| • CS311: Computer Organization   | Spring 2022 |
| • CS230: System Programming  | Fall 2021   |
| • CS492: Special Topic in Computer Science: System for Artificial Intelligence | Spring 2021 |
| • CS230: System Programming  | Fall 2020   |
| • CS492: Special Topic in Computer Science: System for Machine Learning        | Spring 2020 |
- Teaching Assistant.**

- CS3220: Processor Design Georgia Institute of Technology Fall 2016
- CS3220: Processor Design Georgia Institute of Technology Fall 2014
- CS8803: Alternative Computing Technology Georgia Institute of Technology Spring 2014
- CS211: Digital System and Lab KAIST Spring 2011
- CS311: Embedded Computer Systems KAIST Fall 2010

**References  
Available to  
Contact**

**Hadi Esmaeilzadeh.** Professor, UCSD

[hadi@eng.ucsd.edu](mailto:hadi@eng.ucsd.edu)

**Nam Sung Kim** Professor, UIUC

[nskim@illinois.edu](mailto:nskim@illinois.edu)

**Doug Burger.** Technical Fellow and Corporate VP, Microsoft Research

[dburger@microsoft.com](mailto:dburger@microsoft.com)

**Stephen W. Keckler.** VP of Architecture Research, NVIDIA Research

[skeckler@nvidia.com](mailto:skeckler@nvidia.com)

**Eric Chung.** VP of AI Computing, NVIDIA

[eschung@gmail.com](mailto:eschung@gmail.com)