Ahmet Inci

Email: ainci@andrew.cmu.edu • Website: https://inciaf.github.io • GitHub: https://github.com/inciaf

RESEARCH INTERESTS

Computer Architecture, Machine Learning, Hardware-Aware Machine Learning

EDUCATION

Carnegie Mellon University, Pittsburgh, PA

Aug 2017 - Present

- Ph.D. in Electrical and Computer Engineering
 - Advisors: Prof. Diana Marculescu & Prof. Gauri Joshi

Sabanci University, Istanbul, Turkey

Sep 2012 – Jul 2017

- Bachelor of Science (B.Sc.) in Electronics Engineering
 - **GPA:** 3.84 / 4.00, Salutatorian, Summa Cum Laude

RESEARCH EXPERIENCE

Energy-Aware Computing Lab, Carnegie Mellon University

• Advisor: Prof. Diana Marculescu

Aug 2017 – Present

 Designing scalable and efficient systems and ML models using HW/ML model co-design techniques to achieve the best of both worlds. Currently, I have been working on quantization-aware DNN accelerator and model co-exploration through architecture-level modeling and efficient design space exploration. Recently, I have been working on scalable and efficient RL training on CPU-GPU systems. Additionally, my previous work has explored how to utilize emerging non-volatile memories in GPU architectures for DL workloads.

Performance and Energy-Aware Computing Lab, Boston University

Advisor: Prof. Ayse Coskun

Jun 2016 - Sep 2016

- Project: Temperature Dependent DRAM Power and Performance Model
- Modeling 3D-stacked DRAM power consumption under various temperatures and embedding this temperature dependent power model into already existing DRAM power consumption simulators to optimize overall performance of 3D-stacked systems.

Signal Processing and Information Systems Lab, Sabanci University

• Advisor: Prof. Mujdat Cetin

Jan 2015 – Jul 2017

• I had multiple projects within the common theme of signal processing and machine learning. In my junior year, I worked on error-related potentials (ErrP) in brain-computer interfaces applications to better understand the relation between ErrP and error severity.

Neuroelectronics Lab, University of California, San Diego

Advisor: Prof. Duygu Kuzum

Jun 2015 - Sep 2015

• Calculating local field potentials (LFP) by using a network and performing simulations on NEURON simulator. Understanding the contributions of spikes and synaptic potentials to sharp wave-ripple complexes.

WORK EXPERIENCE

NVIDIA

• Research Intern, Architecture Research Group (ARG)

May 2021 – Aug 2021

- Optimizing Power Management of Deep Learning Systems with Reinforcement Learning
- Research Intern, Architecture Research Group (ARG)

May 2020 – Aug 2020

• Towards Scalable and Efficient Reinforcement Learning on CPU-GPU Systems

ARM

Research Intern, ML Technology Group

May 2019 – Aug 2019

• Implementing hardware-aware neural architecture search (NAS) methods for mobile platforms

Cadence Design Systems

• Research Intern, Virtuoso ML Team

May 2018 – Aug 2018

 Creating a machine learning based recommendation system for EDA tools, particularly for Virtuoso in order to alleviate the designer's workload, reduce design time, and improve productivity.

PUBLICATIONS

CONFERENCES

[1] <u>Inci, A.</u>, Isgenc, M., Marculescu, D., "DeepNVM: A Framework for Modeling and Analysis of Non-Volatile Memory Technologies for Deep Learning Applications" *DATE* '20

WORKSHOPS

Full Scholarship

- [1] Inci, A., Virupaksha, S., Jain, A., Thallam, V., Ding, R., Marculescu, D., "QADAM: Quantization-Aware DNN Accelerator Modeling for Pareto-Optimality" ML for Computer Architecture and Systems Workshop (ISCA'21)
- [2] <u>Inci, A.</u>, Virupaksha, S., Jain, A., Thallam, V., Ding, R., Marculescu, D., "QAPPA: Quantization-Aware Power, Performance, and Area Modeling of DNN Accelerators" 2nd On-Device Intelligence Workshop (MLSys'21)
- [3] <u>Inci, A.</u>, Isgenc, M., Marculescu, D., "Cross-Layer Design Space Exploration of NVM-based Caches for Deep Learning" 12th Non-Volatile Memories Workshop (NVMW'21)
- [4] <u>Inci, A.</u>, Bolotin, E., Fu, Y., Dalal, G., Mannor, S., Nellans, D., Marculescu, D., "The Architectural Implications of Distributed Reinforcement Learning on CPU-GPU Systems" 6th Workshop on Energy Efficient Machine Learning and Cognitive Computing (EMC2) 2020
- [5] Inci, A., Marculescu, D., "Solving the Non-Volatile Memory Conundrum for Deep Learning Workloads" 8th Workshop on Architectures and Systems for Big Data (ASBD), ISCA'18 JOURNALS
- [1] <u>Inci, A.</u>, Isgenc, M., Marculescu, D., "DeepNVM++: Cross-Layer Modeling and Optimization Framework of Non-Volatile Memories for Deep Learning" under review for IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems
- [2] Canakci, S., Toy, M. F., <u>Inci, A.</u>, Liu X., and Kuzum, D., "Computational Analysis of Network Activity and Spatial Reach of Sharp Wave-Ripples" *PLoS One, Sep 2017*PATENTS
- [1] <u>Inci, A.</u>, Loh, D., Meng, L., Suda, N., Kunze, E. "Specializing Neural Networks for Heterogeneous Systems" *US Patent Application* 16/724,849, *Filed: December* 2019

HONORS AND AWARDS

 Finalist for Qualcomm Innovation Fellowship Hardware-Aware Multimodal 3D Object Detection for On-Device Augmented Reality Application 	2020 tions
■ Bob Lee Gregory Fellowship, Carnegie Mellon University	2019
 Best Project Award for <i>Hardware Architectures for Machine Learning</i> MAGNETO: Evaluation of Non-Volatile Memory Technologies for Deep Learning Workloads 	2018
 Best Project Runner-Up Award for <i>Energy-Aware Computing</i> Power/Performance Analysis and Optimization for Deep Learning on a CPU-GPU Platform 	2017
 Best Project Award for <i>Networks in the Real World</i> Who Speaks to Whom? Spatiotemporal Analysis of Phone Call Networks 	2017
■ Carnegie Institute of Technology Dean's Fellow	2017
• Graduated as Salutatorian (2nd highest ranking) student in Electronics Engineering Depart	rtment 2017
 Dean's High Honor List for all semesters 	2013 – 2017
 M.I.T Sabanci University Freshman Scholars Program Chosen for Massachusetts Institute of Technology - Sabanci University Freshman Scholars Program fo success in freshman courses. 	Apr 2015 or outstanding
 Dilek Sabanci Scholarship, Sabanci University Full-tuition scholarship with stipend for undergraduate studies. It is only given to 5 students each year. 	2015
 Sakip Sabanci Encouragement Scholarship, Sabanci University Full-tuition scholarship with stipend for undergraduate studies. 	2014
 Merit Scholarship, Sabanci University Awarded for ranking in top 0.15 percent among 1.8 Million participants in the Nationwide University En 	2012 – 2017 ntrance Exam

Awarded for ranking in top 0.34 percent among 1 Million participants in the Nationwide High School Entrance Exam.

2008 - 2012

SKILLS

- Programming Languages: C / C++, C#, Python, Verilog, Assembly, MATLAB, Java, SKILL
- Tools: TensorFlow, Caffe, PyTorch, gem5, GPGPU-Sim, HotSpot, DRAMSim2, McPAT, Sniper
- CAD Tools: Xilinx ISE, Cadence Virtuoso, Mentor Graphics ModelSim, Synopsys Design Compiler, Cadence SoC Encounter, Agilent ADS

COURSEWORK

Carnegie Mellon University, Pittsburgh, PA

 Hardware Architectures for Machine Learning, Energy-Aware Computing, Machine Learning, Computer Architecture and Systems, System-on-Chip Design, Networks in the Real World

Sabanci University, Istanbul, Turkey

 Computer Architectures, VLSI Systems Design, Data Structures, Operating Systems, Digital IC, Microcomputer Based System Design

TEACHING EXPERIENCE

Carnegie Mellon University, Pittsburgh, PA

■ TA for Energy-Aware Computing (18-743)

Fall 2018

- Instructor: Prof. Diana Marculescu
- Designed and evaluated research projects, graded reports, presentations, and homeworks, and held weekly
 office hours.
- TA for ULSI Technology Status and Roadmap for SoC and SiP (18-664)

Fall 2020

- Instructor: Prof. Andrzej Strojwas
- Gave tutorials on several architectural tools, evaluated research projects and presentations.

Sabanci University, Istanbul, Turkey

■ TA for Introduction to Computing (CS-201)

Spring 2015

- Instructor: Gulsen Demiroz
- Held weekly office hours and helped students to overcome their problems on programming concepts.
- TA for Logic and Digital System Design (CS-303)

Fall 2016

- Instructor: Prof. Ilker Hamzaoglu
- Held weekly office hours, supervised students in laboratory sessions, and evaluated their performances.

PROJECTS

Senior Graduation Project

Advisor: Prof. Muidat Cetin

Sep 2016 - Jun 2017

In this project, I investigated applying DNNs for brain-computer interfaces that I implemented previously.
 Our results show that testing accuracy significantly increased by using DNNs.

Error Related Potentials in BCI Applications

■ Advisor: Prof. Mujdat Cetin

Sep 2015 – Jun 2016

• In this project, I investigated error-related potentials (ErrP) in electroencephalography (EEG) data by using two brain-computer interfaces which stimulate subjects. I investigated the relation between ErrP and error severity for different tasks by performing experiments with 8 subjects. I implemented interfaces by using C#. It is accessible in my GitHub profile. I used machine learning algorithms to analyze EEG data.

Sozlukus

Co-founder & Developer

Sep 2014 – Nov 2015

• An interactive social network with ID management and a database for Sabanci University students. It was coded in Python by using Django. It was an open lexicon created by the users. We had more than 150 members who actively used it. There were more than 500 topics. We reached 5% of the population of Sabanci University without using any digital advertisements.

Social Awareness About Street Art and Performers

■ Project Owner

Apr 2012 – Jan 2013

Sabanci University was the sponsor of this project. I have made a documentary interviewing with street
performers in Amsterdam, Paris, and Brussels. Goal of this project was to raise social awareness to the
problems and life conditions of street performers, their expectations from the society and vice versa.

Recycling in Campus (Civic Involvement Project)

Volunteer

Sep 2012 – Jun 2013

We had weekly meetings for a year to raise social awareness about significance of recycling in campus. We
organized discussions about global and local problems of environment.

[CV compiled on 2021-09-03]