Amirhossein Esmaili

+1 (925) 434 8859
esmailid@usc.edu
www.linkedin.com/in/amirhossein-esmaili/

Resume

EDUCATION

- Ph.D. of Computer Engineering, University of Southern California, Aug 2017-Expected May 2022, GPA of 4.0/4.0
- M.Sc. of Computer Science, University of Southern California, Jan 2018-May 2020, GPA of 4.0/4.0
- M.Sc. of Electrical Eng.-Digital Electronics Systems, Sharif University of Technology, 2015-2017, GPA of 18.84/20
- B.Sc. of Electrical Eng.-Digital Systems, Sharif University of Technology, 2011-2015, GPA of 18.24/20

SKILLS

Software Programming: Python, C, C++, MATLAB
Deep Learning: Keras, TensorFlow, PyTorch, Python Numpy Package
Mathematical Programming: IBM CPLEX, OPL
Version Control: Git
Parallel Programming: Cuda, Pthread, OpenCL
Database: SQL, Gremlin
Hardware Design: Verilog Coding, Vivado HLS, Xilinx SDAccel & Vitis, Xilinx Vivado Design Suite

PROFESSIONAL EXPERIENCES

- Hardware Engineering Intern at Bigstream Solutions Inc., Summer 2020.

  * Worked on hardware acceleration of Apache Spark SQL workloads transparently (without zero code change).

  * Developed predictive models for estimating the throughput for various Apache Spark SQL workloads on different formats of input data sources such as Parquet, CSV, and JSON.

SELECTED PROJECTS

- Developed a compiler for low-latency, energy-efficient, and automatic mapping of CNNs on FPGAs. Ongoing project, ongoing research started in Spring 2020. **Some results are published here**
- Developed a hybrid, synergic machine learning model that excels at characteristics such as high accuracy, quick training, computational efficiency, and adaptability and is suitable for incremental, on-line learning on a chip. Spring 2020. **The resulted publication is here**.
- Developed a synthesis flow for mapping combinational circuits to memristive crossbar arrays (MCAs), with the goal of enabling processing in memory (PIM). Achieved considerable improvements in both latency (2.12x) and area (1.38x) compared to the state-of-the-art. Fall 2019. **The resulted publication is here**.
- Developed an online energy-aware scheduler for cluster systems with multiple machines having heterogeneous energy profiles, using the deep reinforcement learning paradigm. Achieved 40% performance improvement compared to manually tuned state-of-the-art heuristics. Spring 2019. **The resulted publication is here**.
- Developed deep neural network architectures aiming collaborative intelligence frameworks (where the inference network is partitioned between the mobile and cloud). Achieved on average 30x improvement in end-to-end latency and 40x improvement in mobile energy consumption compared to the cloud-only approach. Fall 2018. **The resulted publications are here and here**.
- Implemented a framework for energy-constrained scheduling of periodic task graphs with end-to-end deadlines on MPSoC platforms, with the aid of imprecise computations for the goal of QoS maximization. Achieved feasible scheduling of tasks while having very low energy budgets. Summer 2018. **The resulted publication is here**.
- Implemented a framework for energy-optimized scheduling of periodic task graphs with hard deadline constraints on MPSoC platforms, using the combination of DVFS and DPM. Achieved an average energy saving of 15.34% compared to the state-of-the-art. Spring 2018. **The resulted publication is here**.

SELECTED PUBLICATIONS (Google Scholar Link)

- Mahdi Nazemi, Arash Fayyazi, Amirhossein Esmaili, Atharva Khare, Soheil Nazar Shahsavani, and Massoud Pedram, "NullaNet Tiny: Ultra-low-latency DNN Inference Through Fixed-function Combinational Logic." International Symposium on Field-Programmable Custom Computing Machines (FCCM), 2021.
- M. Nazemi, A. Esmaili, A. Fayyazi, M. Pedram. "SynergicLearning: Neural Network-based Feature Extraction for Highly-Accurate Hyperdimensional Learning." International Conference on Computer Aided Design (ICCAD), 2020.
- A. Fayyazi, A. Esmaili, M. Pedram. "HIPE-MAGIC: A Technology-Aware Synthesis and Mapping Flow for HIghly Parallel Execution of Memristor-Aided LoGIC." International Symposium on Low Power Electronics and Design (ISLPED), 2020.
- A. Esmaili, M. Pedram. "Energy-aware Scheduling of Jobs in Heterogeneous Cluster Systems Using Deep Reinforcement Learning." International Symposium on Quality Electronic Design (ISQED), 2020.

- A. E. Eshratifar, A. Esmaili, M. Pedram, "BottleNet: A Deep Learning Architecture for Intelligent Mobile Cloud Computing Services." International Symposium on Low Power Electronics and Design (ISLPED), 2019.
- A. E. Eshratifar, A. Esmaili, M. Pedram, "Towards Collaborative Intelligence Friendly Architectures for Deep Learning." International Symposium on Quality Electronic Design (ISQED), 2019.
- A. Esmaili, M. Nazemi, M. Pedram. "Energy-Aware Scheduling of Task Graphs with Imprecise Computations and End-to-End Deadlines." ACM Transactions on Design Automation of Electronic Systems (TODAES), 2019.
- A. Esmaili, M. Nazemi, M. Pedram, "Modeling Processor Idle Times in MPSoC Platforms to Enable Integrated DPM, DVFS, and Task Scheduling Subject to a Hard Deadline." Asia and South Pacific Design Automation Conference (ASP-DAC), 2019.

SELECTED COURSEWORK

	at USC:	at Sharif University of Technology:
CSC1570:	*Analysis of Algorithms*, 4/4	*Parallel Programming and Architectures*, 19.6/20
CSC1561:	*Foundations of Artificial Intelligence*, 4/4	*Advanced (Objected Oriented) Programming*, 18.6/20
CSC1566:	*Deep Learning and Its Applications*, 4/4	*Advanced Computer Architecture*, 19/20
EE503:	*Probability for Electrical and Computer Engineers*, 4/4	*Microprocessors II*, 19.9/20
CSCI585:	*Database Systems*, 4/4	*ASIC/FPGA Systems Design*, 17.2/20
CSCI402:	*Operating Systems*, 4/4	*Computer Vision*, 17.5/20
CSCI544:	*Applied Natural Language Processing*, 4/4	*Embedded Systems*, 17.8/20