ACM Journals

CALL FOR PAPERS

ACM Transactions on Embedded Computing Systems

Special Issue on Accelerating AI on the Edge

Guest Editors:

- **Muhammad Shafique**, NYU Abu Dhabi, United Arab Emirates, muhammad.shafique@nyu.edu
- **Theocharis Theocharides**, University of Cyprus, Cyprus, ttheocharides@ucy.ac.cy
- **Hai (Helen) Li**, Duke University, USA, hai.li@duke.edu
- **Chun Jason Xue**, City University of Hong Kong, jasonxue@cityu.edu.hk

Machine Learning is nowadays embedded in several edge computing devices, consumer electronics and cyber-physical systems. Smart sensors are deployed everywhere, in applications such as autonomous vehicles, robotics, wearables and perceptual computing devices, and intelligent algorithms power our connected world. These devices collect and aggregate volumes of data, and in doing so, they augment our society in multiple ways; from healthcare, to social networks, to consumer electronics and many more. These emerging systems require real-time inference and decision support; such scenarios therefore may use customized hardware accelerators, are typically bound by limited compute, memory and energy resources, and are restricted to limited connectivity and bandwidth. Thus, near-sensor computation and near-sensor intelligence are starting to emerge as necessities, to continue supporting the paradigm shift of our connected world.

The need for real-time intelligent data analytics for decision support near the data acquisition points, emphasizes the need of revolutionizing the way we design, build, test and verify processors, accelerators and systems that facilitate machine learning (and deep learning in particular) implemented in resource-constrained environments for use at the edge and the fog. To facilitate AI at the edge, we need to re-focus on problems such as design, verification, architecture, scheduling and allocation policies, optimization, and many more, for determining the most efficient way to implement these novel applications within a resource-constrained system, which may or may not be connected. Acceleration of AI at the edge therefore, is a fast-growing field of machine learning technologies and applications including algorithms, hardware, and software capable of performing on-device sensor (vision, audio, IMU, biomedical, etc.) data analytics at extremely low power, typically in the mW range and below, and hence enabling a variety of always-on use-cases and targeting battery-operated devices. There is growing momentum demonstrated by technical progress and ecosystem development. This special issue therefore targets research at the intersection of AI/machine learning applications, algorithms, software, and hardware in deeply embedded machine learning systems.

Topics

Topics of interest include, but are not limited to:

- Specialized hardware architectures for energy-efficient Edge AI / Tiny ML
- Accelerator design and evaluation for ML/AI inference: Circuit and architecture design
- Advanced techniques for energy-efficient inference and power management
- Reliability, security, and robustness for EdgeAI
- Ultra-low-power memory system design for Edge AI / Tiny ML
- Memory optimizations and data flow optimizations for ML/AI accelerators
- In-sensor processing, design, and implementation
- Hardware-aware software optimizations and model compression (like pruning, quantization, etc.)
- Hardware-aware Neural Architecture Search
- Hardware-software co-design methodologies for EdgeAI
- Novel applications across all fields and emerging use cases of Edge AI / Tiny ML

- Distributed learning framework for IoT-Edge
- Evaluation methods, measurement tools and techniques
- Evaluation and measurement of real production systems
- Benchmark creation, assessment and validation for EdgeAI / TinyML systems
- Open datasets and frameworks for accelerating AI at the edge
- Fast design space exploration of ML/AI accelerators
- Analytical models and performance estimation tools
- Characterization of tiny real-world embedded ML/AI systems
- Security and privacy implications in the EdgeAI / TinyML systems
- Open-source hardware and software for EdgeAI

Important Dates

- Submissions deadline: June 1st, 2021
- First-round review decisions: August 15th, 2021
- Deadline for revision submissions: September 15th, 2021
- Notification of final decisions: November 1st, 2021
- Tentative publication: Spring 2022

Submission Information

Prospective authors should follow the TECS submission guidelines at https://dl.acm.org/journal/tecs/author-guidelines. All manuscripts must be submitted electronically to the ACM Manuscript Central Web site at https://mc.manuscriptcentral.com/tecs. Select the paper type "Special Issue on Accelerating AI on the Edge". All papers will undergo the standard TECS review process with a fast review cycle.

For questions and further information, please contact **Muhammad Shafique and/or Theo Theocharides at muhammad.shafique@nyu.edu and ttheocharides@ucy.ac.cy**