Seminar: Software-Hardware Co-Design for Edge AI Acceleration

06/06/2025, 10.00, DAUIN, Sala Riunioni 1

Abstract

In the realm of Artificial Intelligence (AI), there exist two distinct ends of the spectrum: cloud AI, which relies on servers in the cloud, and edge AI, which operates on distributed devices that are situate at the edge of a network. With the rapid proliferation of wearables, drones, and other Internet-of-Things devices, edge AI has become ubiquitous and plays a pivotal role in reshaping our daily lives.

In this talk, we will delve into various layers of the design stack of edge AI and look into the challenges it confronts. To mitigate these challenges, we will introduce state-of-the-art methodologies to enable software-hardware co-design across software layer, middleware layer and hardware layer. Specifically, we will discuss Neural Architecture Search for deep learning and automatic hardware mapping for Compute-In-Memory. We will also explore how innovative computing architectures and circuits can further accelerate AI execution and enhance energy efficiency, delivering promising solutions that are highly sustainable to enable green edge AI.

Speaker Bio

Bo Wang received the Ph.D. degree in electrical and electronic engineering from Nanyang Technological University, Singapore. In 2020, she joined Singapore University of Technology and Design, Singapore as an Assistant Professor. Her research interests span various aspects of energy-efficient architecture and circuit design, including deep neural network based hardware acceleration, neuromorphic computing processors, in-memory computing, and design automation tools for hardware-software co-optimization. She authored and co-authored many papers published at international journals and conference proceedings, including TNNLS, JSSC, TCAS-I, TVLSI, TCAS-II, A-SSCC, DAC, DATE, ISLPED, MobiSys, etc. Dr. Wang received the Distinguished Design Award at the IEEE Asian Solid-State Circuits Conference (A-SSCC) in 2023. She was also a recipient of the Best Paper Award at the ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED) in 2023 and the International SoC Design Conference (ISOCC) in 2024 and 2014. She is serving as an Associate Editor for the IEEE Open Journal of Circuits and Systems and has served as a Guest Editor for the Frontiers in Neuroscience. She is an IEEE senior member.