

PULP: Extreme Energy Efficiency for Extreme Edge AI Acceleration

An open platform perspective

Luca Benini,
ETH Zurich and Università di Bologna
lbenini@ethz.ch

Abstract — The next wave of pervasive AI pushes machine learning (ML) acceleration toward the extreme edge, with mW power budgets, while at the same time it raises the bar in terms of accuracy and capabilities, with new ML models being propose on a daily basis. To succeed in this balancing act, we need principled ways to walk the line between flexible and highly specialized ML acceleration architectures. In this talk I will detail on how to walk the line, drawing from the experience of the open PULP (Parallel Ultra-Low Power) platform, based on ML-enhanced RISC-V processors coupled with domain-specific acceleration engines.

References:

<https://pulp-platform.org/>

About author



Luca Benini holds the chair of digital Circuits and systems at ETHZ and is Full Professor at the Università di Bologna. Dr. Benini's research interests are in energy-efficient parallel computing systems, smart sensing micro-systems and machine learning hardware.

He a Fellow of the IEEE, of the ACM and a member of the Academia Europaea. He received various awards, including the 2016 IEEE CAS Mac Van Valkenburg award (2016) and the ACM/IEEE A. Richard Newton Award (2020).