LOCA SERIES: "Mixed-Precision Quantization techniques for Energy-efficient DNN Inference"

Abstract:

In this project, we aimed to enhance the computational efficiency and deployment feasibility of neural networks through mixed precision quantization. We implemented two quantization-aware training (QAT) methods. Our results demonstrated significant reductions in model bit-width assignments while maintaining accuracy comparable to full-precision models.

Speaker: Omar Lahyani

Short bio:
Omar Lahyani is a fifth-year engineering student at Ecole Polytechnique de Tunisie. During 2024, he worked as a research intern at Barcelona Supercomputing Center (BSC) to develop his final thesis and obtain his diploma with a project focused on efficient AI acceleration.

Speakers

Speaker: Omar Lahyani. Synthesis and Physical design of ICs, Computer Sciences, BSC.
Host: Francesc Moll. Synthesis and Physical design of ICs Group Manager, Computer Sciences, BSC.
Barcelona Supercomputing Center - Centro Nacional de Supercomputación