AI accelerators for reconfigurable technologies

A. García Ríos (grios@ugr.es), and R. Padial Allué (rubenpadial@ugr.es)

Department of Electronics and Computer Technology
CITIC-University of Granada, Granada, Spain.

Calendar

	Morning session		Afternoon session	
Date	**Responsible (Schedule)**	**Unit**	**Responsible (Schedule)**	**Unit**
Monday 17/11			A. García, R. Padial (16:00-19:30)	Fundamentals of Artificial Intelligence
Tuesday 18/11			A. García, R. Padial (16:00-19:30)	Introduction to Accelerated Computing and Reconfigurable Technologies I
Wednesday 19/11			A. García, R. Padial (16:00-19:30)	Introduction to Accelerated Computing and Reconfigurable Technologies II

Calendar

	Morning session		Afternoon session	
Date	**Responsible (Schedule)**	**Unit**	**Responsible (Schedule)**	**Unit**
Thursday 20/11			A. García, R. Padial (16:00-20:00)	Introduction to Neural Network Acceleration on FPGAs
Friday 21/11			A. García, R. Padial (16:00-20:00)	Specific Design Flows for Implementing AI Algorithms on Reconfigurable Devices I
Monday 24/11			A. García, R. Padial (16:00-20:00)	Specific Design Flows for Implementing AI Algorithms on Reconfigurable Devices II

Calendar

- **Fundamentals of Artificial Intelligence**
  - Description of the main neural network architectures (MLP, CNN, etc.) and their applications.
  - Typical operations in AI models and their implications in hardware.
  - Model quantization and efficient data representation (int8, float16).
  - Trade-off between accuracy, performance, and power consumption.

- **Introduction to Accelerated Computing and Reconfigurable Technologies I**
  - Comparison between CPU, GPU, TPU, and FPGA for AI.
  - Advantages and disadvantages of FPGAs for AI acceleration.
  - Applications: edge computing, robotics, computer vision, etc.

Calendar

- **Introduction to Accelerated Computing and Reconfigurable Technologies II**
  - Workflows for AI algorithm inference in embedded systems.
  - Definition and training of an AI model.

- **Introduction to Neural Network Acceleration on FPGAs**
  - What is HLS (High-Level Synthesis) and what is it used for.
  - Advantages compared to classic RTL design.
  - Design of simple blocks (multiplication, convolution, activation functions): implementation examples.

Calendar

- **Specific Design Flows for Implementing AI Algorithms on Reconfigurable Devices I**
  - Vitis AI, OpenVINO, Intel FPGA AI Suite, MATLAB: Basic. WorkflowCompilation of a pre-trained model in MATLAB and inference.

- **Specific Design Flows for Implementing AI Algorithms on Reconfigurable Devices II**
  - Quantization, compilation, and testing of a pre-trained model.
  - Performance measurement: latency and resource usage.
  - Quantization of the pre-trained model, inference, and performance measurement.