Ph.D. in Information Technology: Thesis Defense

December 21st, 2020

online by Teams – at 9.00

Ahmet ERDEM – XXXII Cycle

"Exploration and Mapping of Deep Neural Networks to Low-Power Hardware Accelerators and FPGAs"

Advisor: Prof. **Cristina Silvano**

Abstract:

Deep Neural Networks are quickly becoming a leading edge solution to classify and analyze the behavior of people and the surrounding world, paving the way to leverage the millions of sensors available in the IoT era. The increasingly higher computational demands of these new class of algorithms call for a processing power reaching into the Tera OPS while product constraints still require power and energy to be curtailed within practical and market competitive boundaries. The state of the art is aiming to develop new hybrid architectures resorting to heterogeneous computing solutions coupling both general-purpose processors and HW accelerators to achieve the intended performance and efficiency targets. A key aspect of this effort is the ability to perform fast and accurate design space explorations of the wide parametric space keeping into account both hardware and software mapping challenges.

In this thesis, an analytical model of accelerators in Orlando architecture is developed and used to explore the design space for convolutional neural network mapping on convolution accelerators. The results of the experiment showed that the configurability of accelerators is crucial for adapting the changing computation and memory access patterns within different stages of the convolutional neural network. In addition to this methodology, fused-layer convolution acceleration is studied as an emerging trend in accelerating memory-bound convolution layers and our incremental innovative improvements are detailed in this thesis. Experimental results show that even on smaller FPGAs it is possible to take advantage of consecutive convolution layers and reduce the external bandwidth traffic which is one of the main contributors to power consumption.

PhD Committee

Prof. **Gianluca Palermo**, DEIB
Prof. **Michael Huebner**, Brandenburg University of Technology
Prof. **Sotirios Xydis**, Harakopio University Athens