Neural Networks and Deep Learning: Part 2: Implementation issues

The first lecture is scheduled on **February 22, 2022 - 9:00**
Grey Room, Retis Lab, TeCIP Institute, Via Moruzzi, 1 – Pisa

Giorgio Buttazzo and Alessandro Biondi

**Objectives:** The aim of the part is to provide practical and implementation issues useful to deploy neural networks on a variety of embedded platforms using different languages and development environments.

**Course program**

1. **Implementing Neural Networks from scratch in C.** General implementation principles. Main and auxiliary functions.
2. **Sample implementations** of common neural network models in C language.
6. **Explainable and Trustworthy AI:** methods, techniques and ethical issues.
7. **Introduction to attention mechanisms.** Application in natural language processing. Transformers and their comparison with recurrent networks.
8. **Recent models for image processing.** Applications to detection and semantic segmentation. Dataset enhancement, sensor fusion, recurrent networks in computer vision. Open challenges.
9. **Accelerating deep networks on GPGPUs.** Overview of the Nvidia TensorRT framework. Executing a DNN modelled in Caffe in TensorRT.
10. **Real-time neural vision.** How to accelerate a neural network on TensorRT to detect objects from a video camera.
11. **Accelerating deep networks on FPGA.** Common technologies and approaches for deploying deep networks on FPGA.