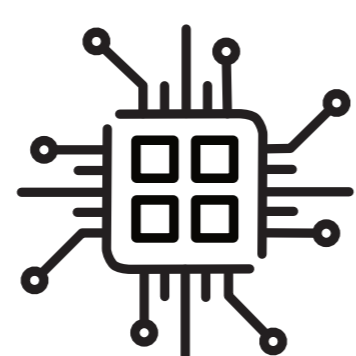


ARCHITECTURES BASED ON UNCONVENTIONAL ACCELERATORS FOR DEPENDABLE/ENERGY EFFICIENT AI SYSTEMS

Brahmaiah Gandham, Jacopo Raffi, Francesco Pio Monaco, Pedro Foletto Pimenta, Davide Gardenal, Hikaru Okamoto, Bahman Derikvandi, Bohao Zhang, Soheil Fallah Bagherinezhad, Elisa Ricci, Giovanni Iacca, Kasim Sinan Yildirim, Paolo Rech, Philippe Velha, Roberto Passerone, Flavio Vella
University of Trento, Italy

Better energy efficiency and performance of computer vision tasks for autonomous vehicles

Hardware



Heterogeneous design

On-chip, chiplet, or system-in-package integration of CMOS with non-conventional AI accelerators



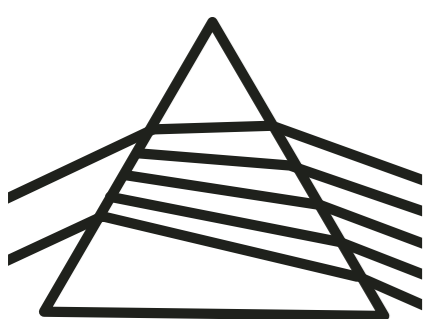
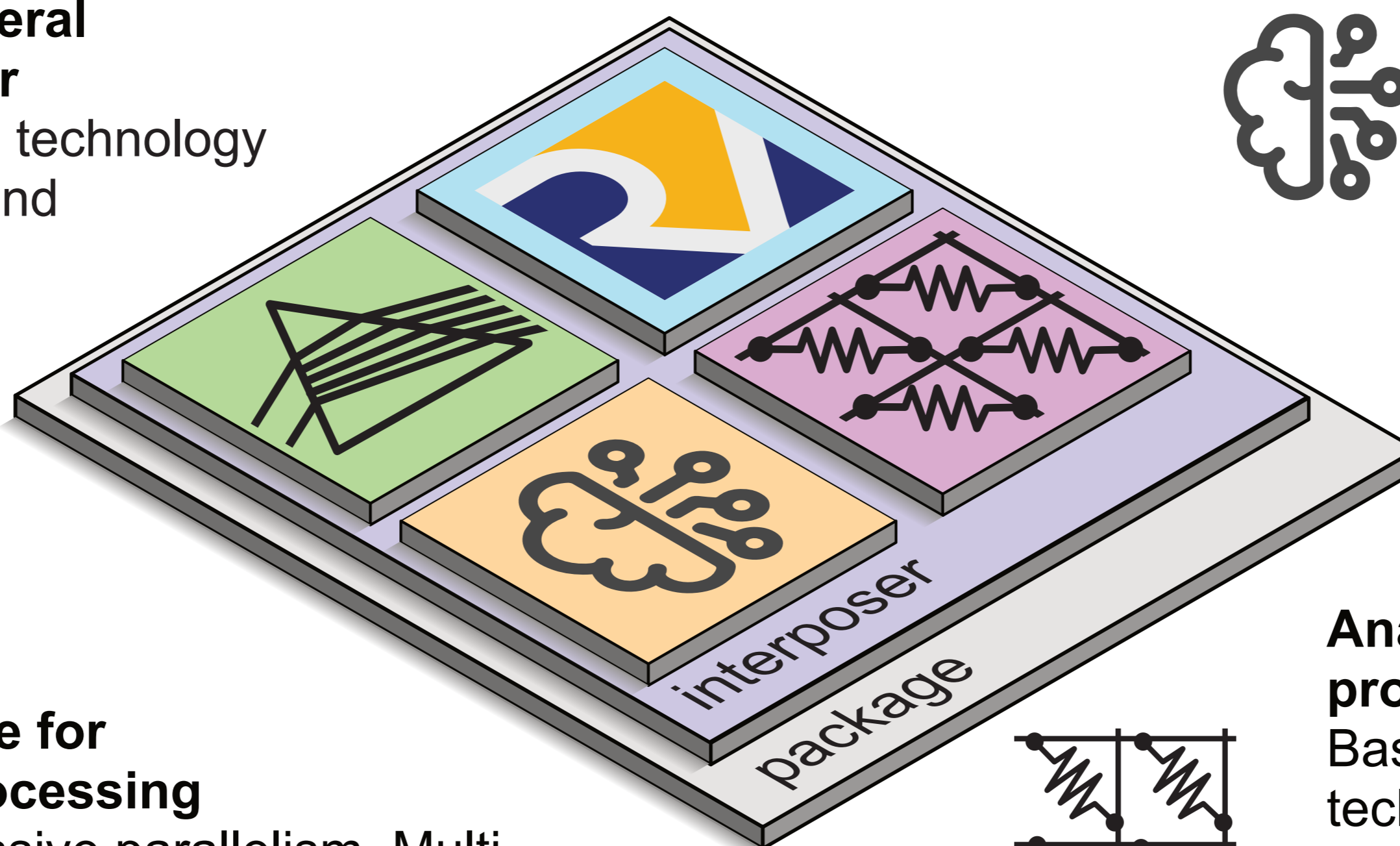
RISC-V-based general purpose processor

Conventional digital technology for system control and orchestration



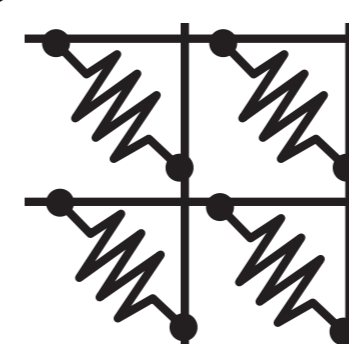
Neuromorphic accelerator

Spiking neural networks on ferroelectric/PCM devices
Ultra-low power



Photonic engine for analogue AI processing

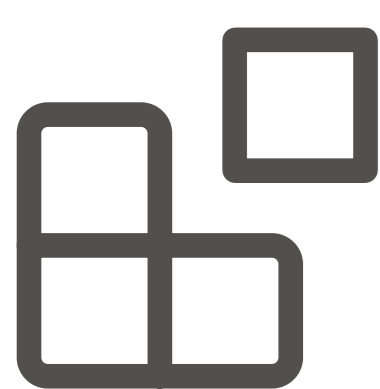
High speed, massive parallelism. Multi-channel DAC and ADC connecting the accelerator to memory



Analogue and digital processing-in-memory

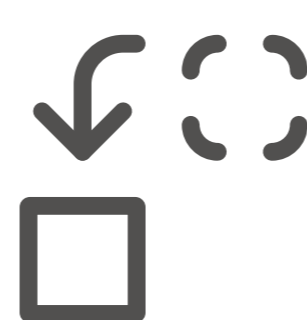
Based on DRAM or non-volatile technologies (PCM, ferroelectric, RRAM). Reduced data movement, low power, deterministic processing time

System and software



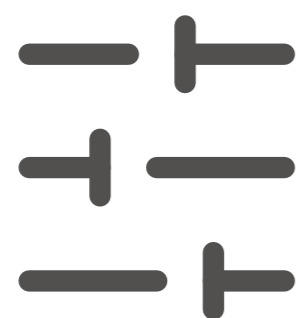
Scalable compute fabric

Modular design with templateable hardware blocks sharing common interface, even in presence of different programming models



Compilers and runtime libraries

Mapping AI kernels to accelerators and automatic precision tuning, pruning, and sparsification of NNs



Design space exploration and HW-SW co-design

Use-case specific configuration of accelerator architecture, allocating hardware blocks to achieve optimal performance



Simulation and reliability assessment

Characterizing performance, scalability, energy consumption, fault tolerance, vulnerability to attacks, and thermal properties of novel AI accelerators

Funding and Disclaimer

This work has been funded by the European Union via the European Defence Fund project ARCHYTAS under grant number 101167870. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

