

This project has received funding from the Electronic Components and Systems
for European Leadership Joint
Undertaking under grant agreement No 876925



ANDANTE

AI for New Devices And Technologies at the Edge

D3.5 Building blocks implementation

Deliverable No.	D3.5	Due Date	31-May-2022
Type	Report	Dissemination Level	Confidential
Version	1.0	Status	Final
Description	This deliverable provides the implementation description for the neural networks building blocks and algorithms implemented in Task 3.3 Neural Networks building blocks and algorithms.		
Work Package	WP3 – AI building blocks, Methods and Tools.		

PROPRIETARY RIGHTS STATEMENT

This document contains information, which is proprietary to the ANDANTE Consortium.

Neither this document nor the information contained herein shall be used, duplicated or communicated by any means to any party, in whole or in parts, except with prior written consent of the ANDANTE consortium.

Abstract (Published Summary)

ANDANTE has the target to design, develop and demonstrate innovative solutions and hardware accelerators for future products in the domain of edge applications, where data processing must remain on-site, power must be minimal, and latency is optimal for local interaction.

In the context of the WP3 “AI Building Blocks, Methods and Tools”, this document reports the implementation of neural network models, new training methods, optimization and benchmarking tools, building blocks for the ASICs defined in WP4 “Implementation of ICs and Platforms”, and hardware neuromorphic accelerators. It describes the implementation of the AI building blocks defined in the previous deliverable D3.4.

It represents a fundamental link between the use cases defined in WP1 “Use Case System Architectures Description and Application Requirements”, the neural networks designed in WP3, the implementation of the inference accelerators built in WP4 (ASICs) and WP3 (FPGA-based) and the final deployment on the field to execute the use case tasks in WP5 “Applications Integration, Validation and Evaluation”.

This deliverable reports the contributions and the results of the task 3.3 in the three main sections summarized below:

Section **Error! Reference source not found. - Error! Reference source not found.:**

- CSEM designed a model quantizer tool (Kuantizer), a Python library for model deployment (NeurIO) and a benchmarking tool (BenchUI).
- SynSense coded a trainable spiking softmax class.
- TUD implemented a training algorithm for recurrent neural networks (e-prop) on the SpiNNaker2 platform.
- TRT developed an algorithm and a simulator for supervised on-chip training compatible with memristive synapses.
- IFAG studied and implemented methods for parallelization and distribution of SNN models over multiple devices.
- BR&T-E, Gradient and TVES designed and implemented neural network models to address the use case UC3.4, “Robust autonomous landing”, for the following tasks:
 - runway detection,
 - object detection,
 - image registration,
 - processing the communication parameters.

Section **Error! Reference source not found. - Error! Reference source not found.:**

- FhG (IIS and EMFT): ASIC 3.1, code name Adelia
 - input and weight buffer memory,
 - analog weight emulator, AWE,
 - analog-to-digital converter, ADC,
 - neuron.

- FhG (IPMS): ASIC 3.1b, code name Felix
 - micro-controller interfaces,
 - analog-to-digital converter, ADC,
 - word-/bit-line crossbar control.
- IFAG: ASIC 3.2, code name Tiny-Accel
 - RRAM memory module,
 - multiply-and-accumulate unit, MAC,
 - word-line controller.

Section **Error! Reference source not found. - Error! Reference source not found.:**

- TRT developed a formal (non-spiking)/neuromorphic hybrid inference accelerator for neural networks evaluation on the programming logic of the Xilinx FPGA Zync UltraScale+ MPSoC.
- Imec-NL designed a neuromorphic accelerator implemented in the programming logic of the Xilinx FPGA Zync UltraScale+ RFSoc, together with an SDK to help deploying the network model on it.
- BR&T-E, Gradient and TVES accelerated the DNNs for runway detection, object detection, image recognition and process communication parameters using the DPU soft core available on the Xilinx FPGA Zync UltraScale+ MPSoC/RFSoc to evaluate their performances.