# ANDANTE

Tools and methodologies for training and mapping a neural network on ANDANTE hardware targets

ANDANTE

## Objectives

- Leverage embedded NVM for designing new accelerators for edge AI applications
- Develop dedicated tools and methodologies to facilitate the training, simulation and deployment of neural network models on the targeted HW accelerators.
- Define the software-hardware co-design methodologies to pursue according to the targeted application

## Context

- High-level design choices must be made depending on the application, the type of sensor used (event-based sensor or not) and the size of the network, which may dictate analog or digital design.
- Each of those design choices requires a dedicated tool flow (Training, generation (RTL, schematic), mapping and simulation tools).

## Main Results

- A total of **16 tools** were developed in ANDANTE and are now **completed, half** of which is available in **open source.**

## Application-dependent high-level design choices

### What coding style to use?

Classical coding is Best for still inputs (e.g. images)

Event / Spike coding is Best for always-on, low-activity inputs (e.g. surveillance)

- Rate code is to be used for "Frame-type" inputs
- Temporal code is best on Temporal series (e.g. audio)

### What network topology?

A Deep, Feed Forward topology for high accuracy and fixed inputs applications

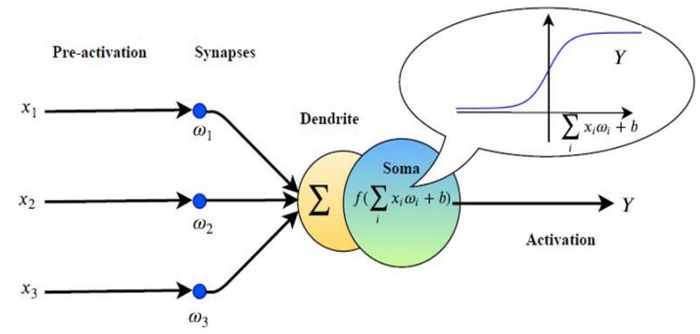
A Shallow, Recurrent topology for temporal series

### What implementation strategy?

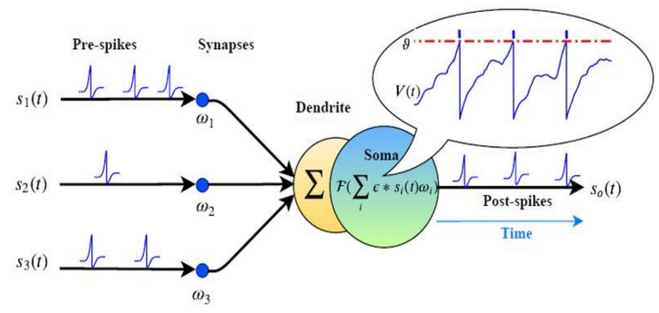
A Fully digital implementation is Mandatory for very deep, large, networks

An Analog mixed-signal implementation can be Efficient for shallower networks

Classical coding




Spike coding

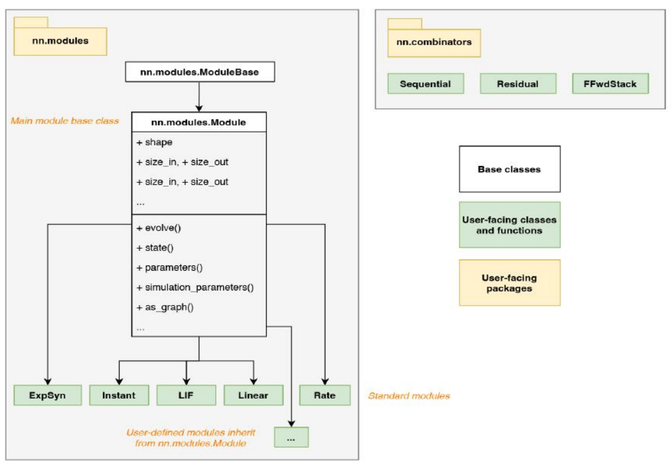



## Training tools

### Spike-coding

ANN-to-SNN conversion

Direct SNN training (Offline and On-chip learning)

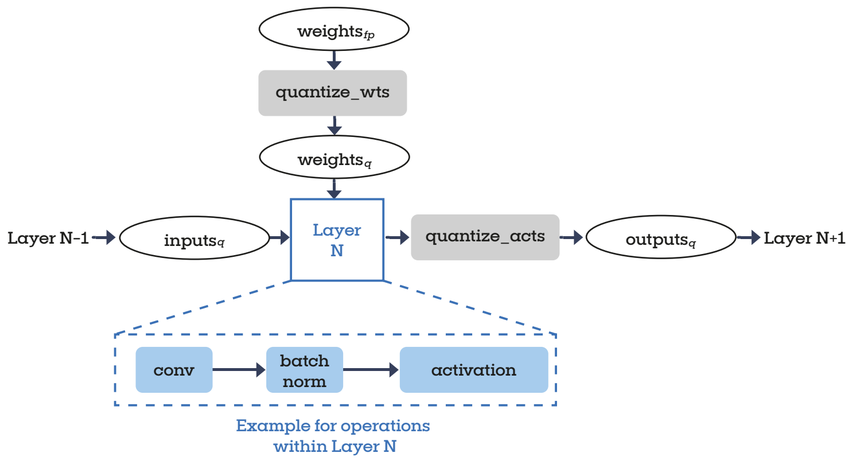



Offline Training tool for SNN

### Classical coding

Transfer learning

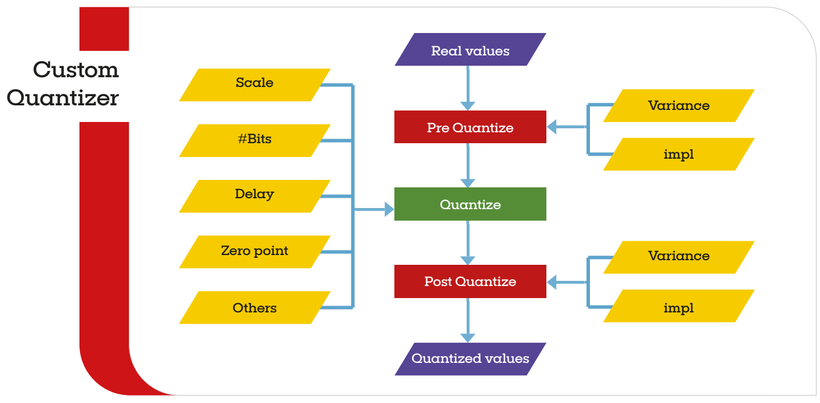
Quantization Aware Training: Reduces memory footprint and enables the use of embedded NVM




QAT training: Weights and activations quantization

### HW-aware training for mixed-signal implementations of ANN ASICs

Takes non-idealities into account during the training phase



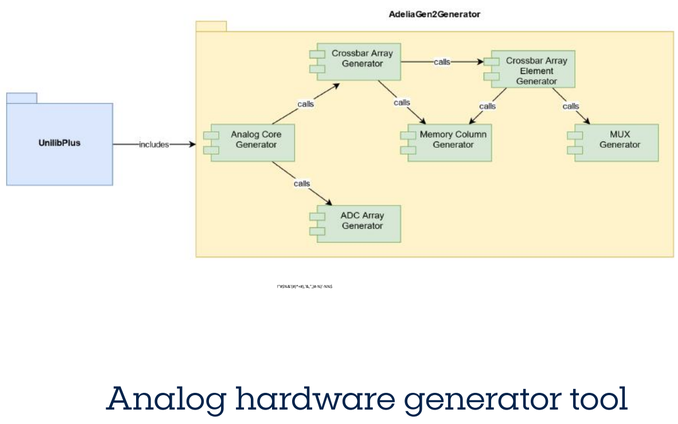

HAT (hardware-aware training) tool

## Generation, Mapping and Simulation tools

### Hardware generator

Writes RTL (for digital NN)

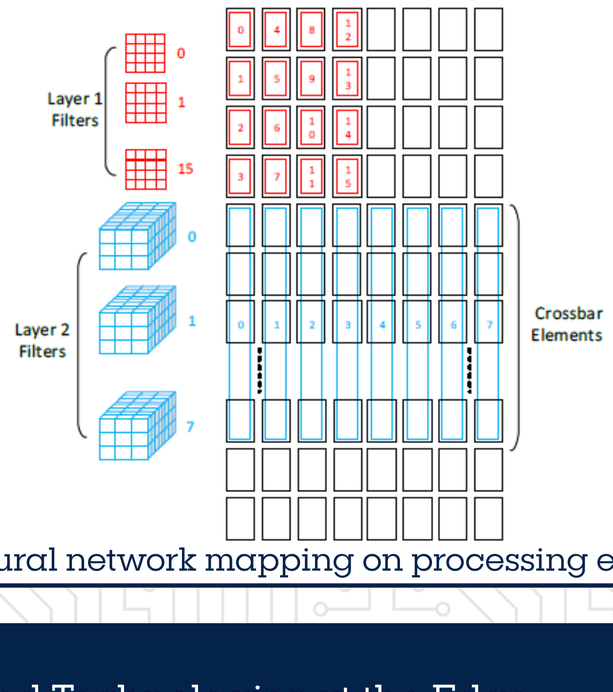
Generates schematic and layout (for analog NN)




Analog hardware generator tool

### Neural Network mapper on hardware

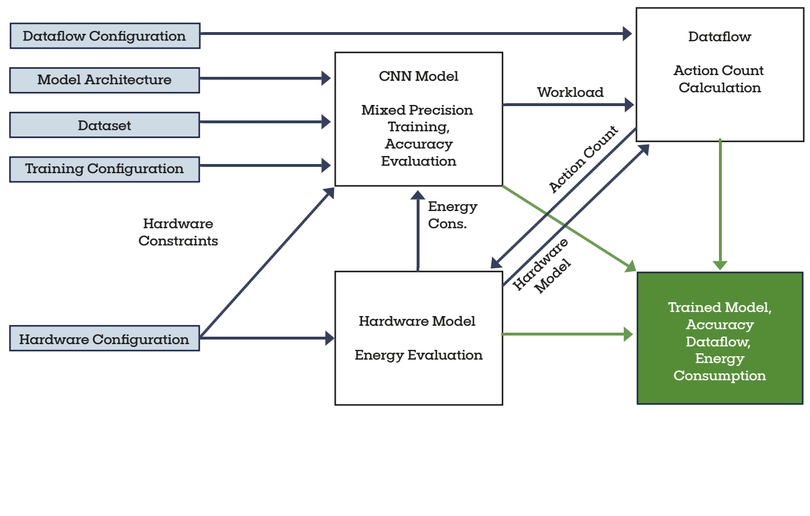
Maps onto Processing engines or NVM arrays




Neural network mapping on processing engines

### Simulation

Estimates power dissipation




Hardware power estimation tool

**ANDANTE**

AI for New Devices And Technologies at the Edge

andante-ai.edu  linkedin.com/company/andante-ai  mario.diaznava@st.com (coordinator)

"This project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement № 876925. The JU receives support from the European Union's Horizon 2020 research and innovation programme and France, Belgium, Germany, Netherlands, Portugal, Spain and Switzerland"

ECSEL Joint Undertaking
Electronic Components and Systems for European Leadership

Scan Me
to visit website