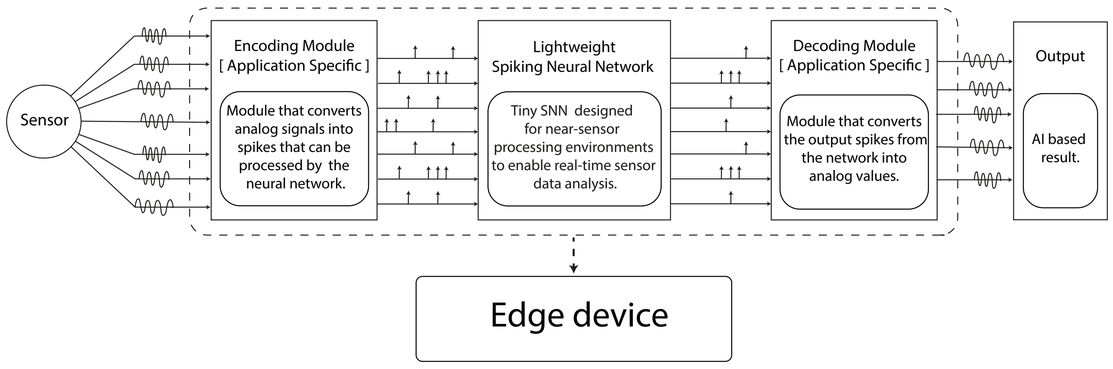
Open-Source Design of a Low-Power SNN Hardware Accelerator for Edge AI

Luca Martis, Paolo Meloni
Dipartimento di Ingegneria Elettrica ed Elettronica, University of Cagliari, Italy
luca.martis@unica.it, paolo.meloni@unica.it

Overview approach

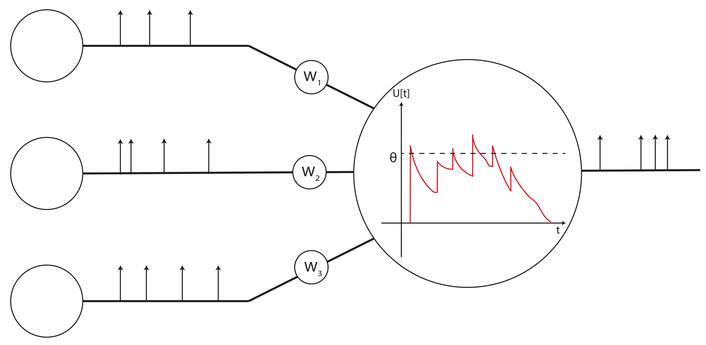


Real-time processing

Low-power consumption

Resource constrained devices

Spiking Neural Network



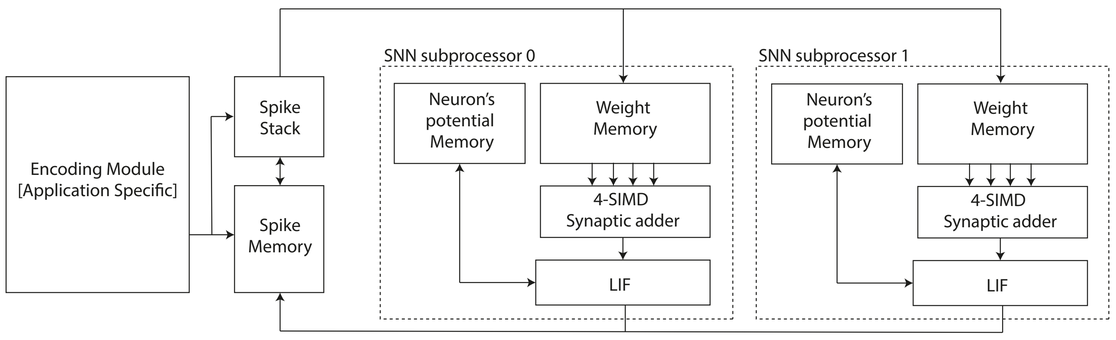
Event-based processing.

Only accumulation operations are needed, thus reducing computational complexity.

Neuromorphic hardware limited by high development costs and restricted accessibility.

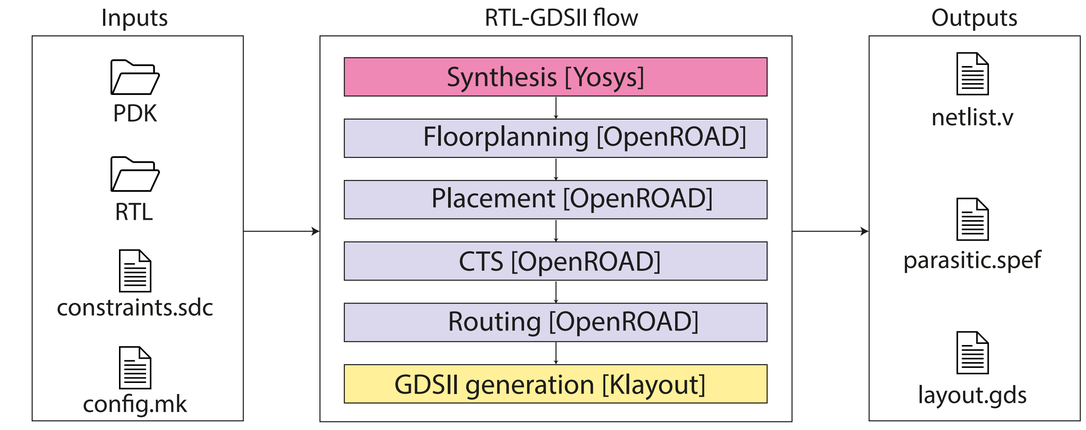
Spiking neural networks, considered as the third generation of neural networks, are characterized by neurons communicating via binary signals called spikes. This property makes them well-suited for edge devices, as it reduces computational complexity and enhances energy efficiency.

SYNtzulu architecture



The accelerator used in this work is a modified version of the open-source accelerator presented in [1], called SYNtzulu, which can be downloaded using the provided QR code. It includes an encoding module for converting continuous signals into spikes and an SNN processor that executes dense LIF neuron layers using two parallel cores.

Open-source RTL-GDSII flow

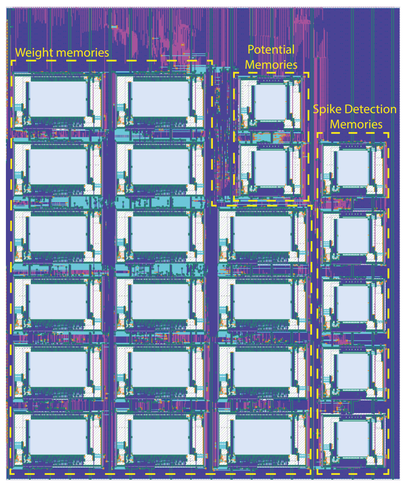


Broader accessibility

Reduced development cost

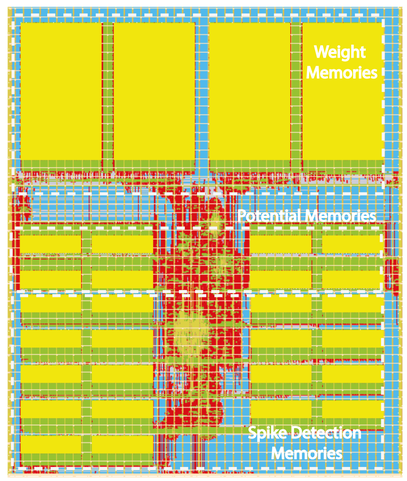
Legacy Tecnolgy node

HW accelerator with SKY130 PDK



Core Area	3000 x 3600 um
Standard Cells	8.922
Max Neurons	512
Max Synapses	32.768
F_{max}	100 MHz
Performance [100 MHz]	0,8 GSOP/s
$P_{inference}$	1,25 mW/MHz
P_{idle}	350 uW

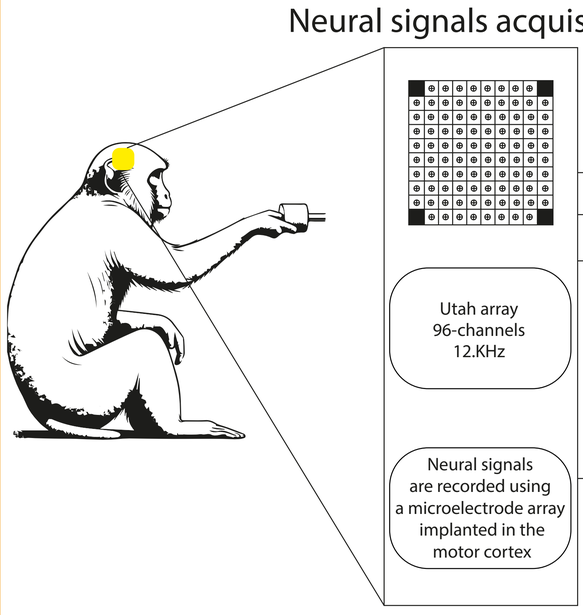
HW accelerator with IHP-SG13G2 PDK



Core Area	1750 x 2000 um
Standard Cells	15.725
Max Neurons	512
Max Synapses	32.768
F_{max}	90 MHz
Performance [90 MHz]	0,72 GSOP/s
$P_{inference}$	0,53 mW/MHz
P_{idle}	81 uW

Use case: Real time decoding of intracortical neural activity

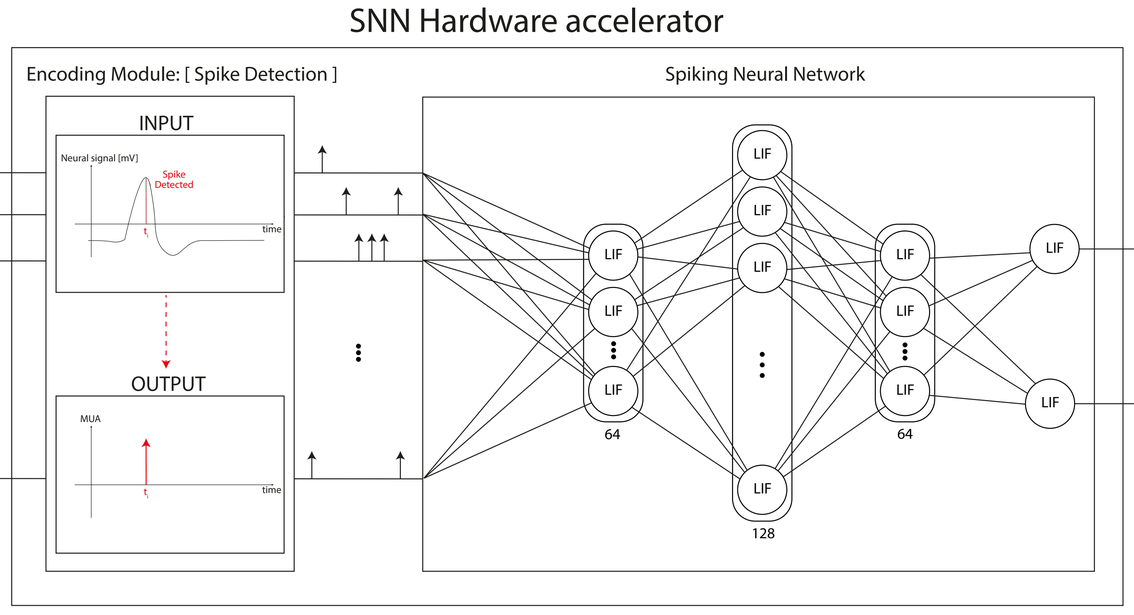
Neural signals acquisition



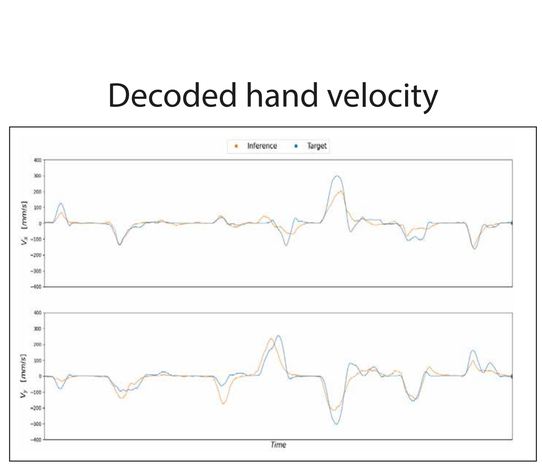
Utah array
96-channels
12.KHz

Neural signals are recorded using a microelectrode array implanted in the motor cortex

SNN Hardware accelerator



Decoded hand velocity



The dataset used to train and test the system is public and described in [3].

We extract multi-unit activity (MUA) from the neural signal using the encoding module.

We process the MUA using a four-layer dense spiking neural network trained to decode two variables representing the hand's velocity along the x-axis and y-axis

RELATED WORKS COMPARISON

Work	Decoder	Platform	Parameters	V_{CC}
This	SNN	ASIC	22.6k	0.82
[2]	LSTM	PC	341k	0.84
[2]	QRNN	PC	327k	0.85
[2]	MLP	PC	155k	0.82

References

[1] G. Leone, M. A. Scrugli, L. Badas, L. Martis, L. Raffo and P. Meloni, "SYNtzulu: A Tiny RISC-V-Controlled SNN Processor for Real-Time Sensor Data Analysis on Low-Power FPGAs," in IEEE Transactions on Circuits and Systems I:Regular Papers.
[2] N. Ahmadi, T. Adiono, A. Purwarianti, T. G. Constandinou and C. -S. Bouganis, "Improved Spike-Based Brain-Machine Interface Using Bayesian Adaptive Kernel Smoother and Deep Learning," in IEEE Access.
[3] J. E. O'Doherty, M. M. B. Cardoso, J. G. Makin, and P. N. Sabes, "Nonhuman primate reaching with multichannel sensorimotor cortex electrophysiology," Zenodo, Sabes Lab, Univ. California, San Francisco, CA, USA, Tech. Rep., May 2020. [Online]. Available: https://doi.org/10.5281/zenodo.3854034

H2TRAIN project is supported by the Chips Joint Undertaking and its members, including the top-up funding by National Funding Authorities from Austria, Finland, Germany, Italy, Poland and Spain, under grant agreement no. 101140052. H2TRAIN project is supported by the Chips Joint Undertaking (GA 101140052) and its members including top-up funding by The National Funding Agency for Austria, The Innovation Funding Agency Business Finland, Bundesministerium für Bildung und Forschung, The Ministry of Universities and Research in Italy, The Ministry of Enterprises and Made in Italy, The National Centre for Research and Innovation in Poland, The Ministry for Digital Transformation and the Civil Service in Spain, and The State Research Agency in Spain.
This work was funded by H2TRAIN project, supported by the Chips Joint Undertaking (GA 101140052) and its members including top-up funding by The National Funding Agency for Austria, The Innovation Funding Agency Business Finland, Bundesministerium für Bildung und Forschung, The Ministry of Universities and Research in Italy, The Ministry of Enterprises and Made in Italy, The National Centre for Research and Innovation in Poland, The Ministry for Digital Transformation and the Civil Service in Spain, and The State Research Agency in Spain.