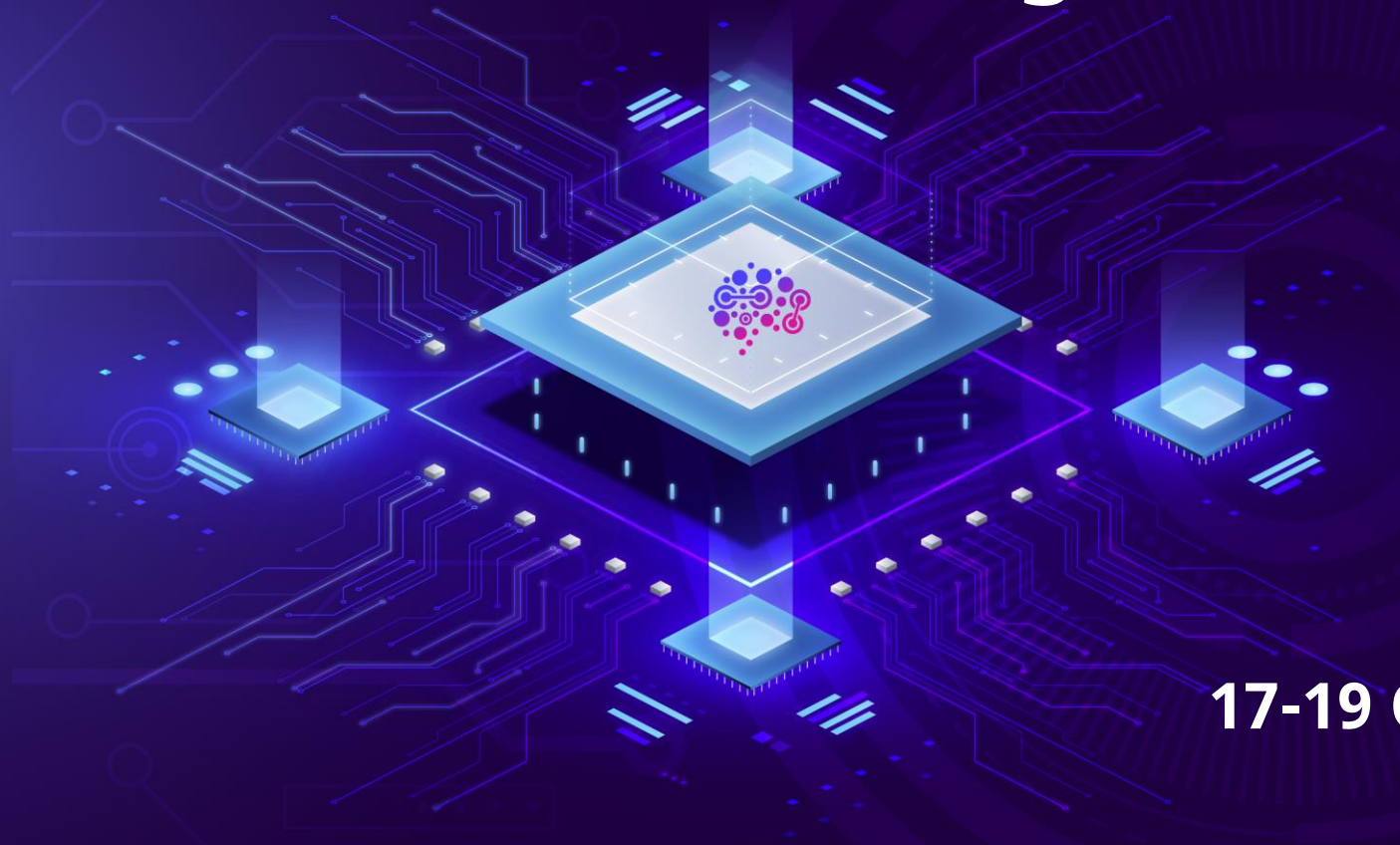
European Conference on EDGE AI Technologies and Applications - EEAI



Advancing emerging edge AI technologies and driving next generation intelligent applications.

17-19 October 2023, Athens, Greece

European Conference on EDGE AI Technologies and Applications - EEAI

SNN-based accelerators for low-latency, ultra-low-power inference at the edge

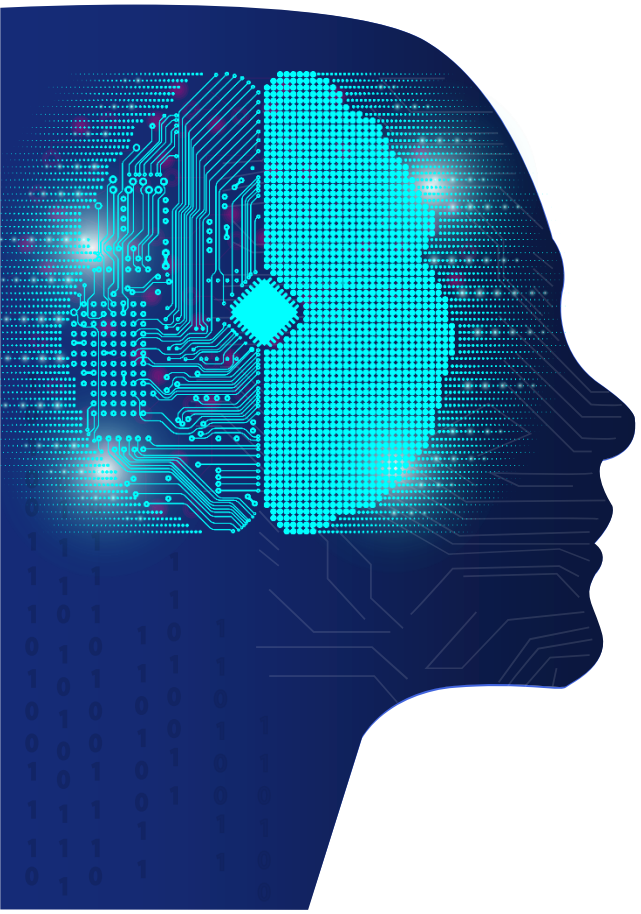
Maxim Rykunov, IMEC, Belgium (presenter)

Sumeet Kumar, Innatera, Netherlands

17-19 October 2023 Athens, Greece

Copyright © 2023

Presentation Outline



- Introduction
- Motivation
- Methodology
- INNAT’s prototype
- IMEC’s prototype
- Future activities
- Conclusion

Copyright © 2023

Introduction

A joint Neuromorphic sensor fusion demonstrator in **AI4CSM project**:

A. **Scalable neuromorphic accelerator for low-power and high-performance deep edge processing (IMEC)**

B. **Analog-mixed signal neuromorphic accelerator prototype (INNATERA)**

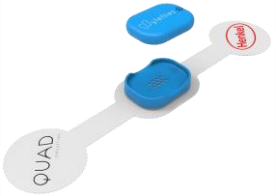
Demonstrate and evaluate neuromorphic accelerator fabric **for low-power and high-performance deep edge processing for automotive applications.**

Copyright © 2023

Motivation

Why event-based neural networks (e.g., SNNs)?

- From autonomous cars to health patches, similar challenges emerge:



✓ Time-based signals

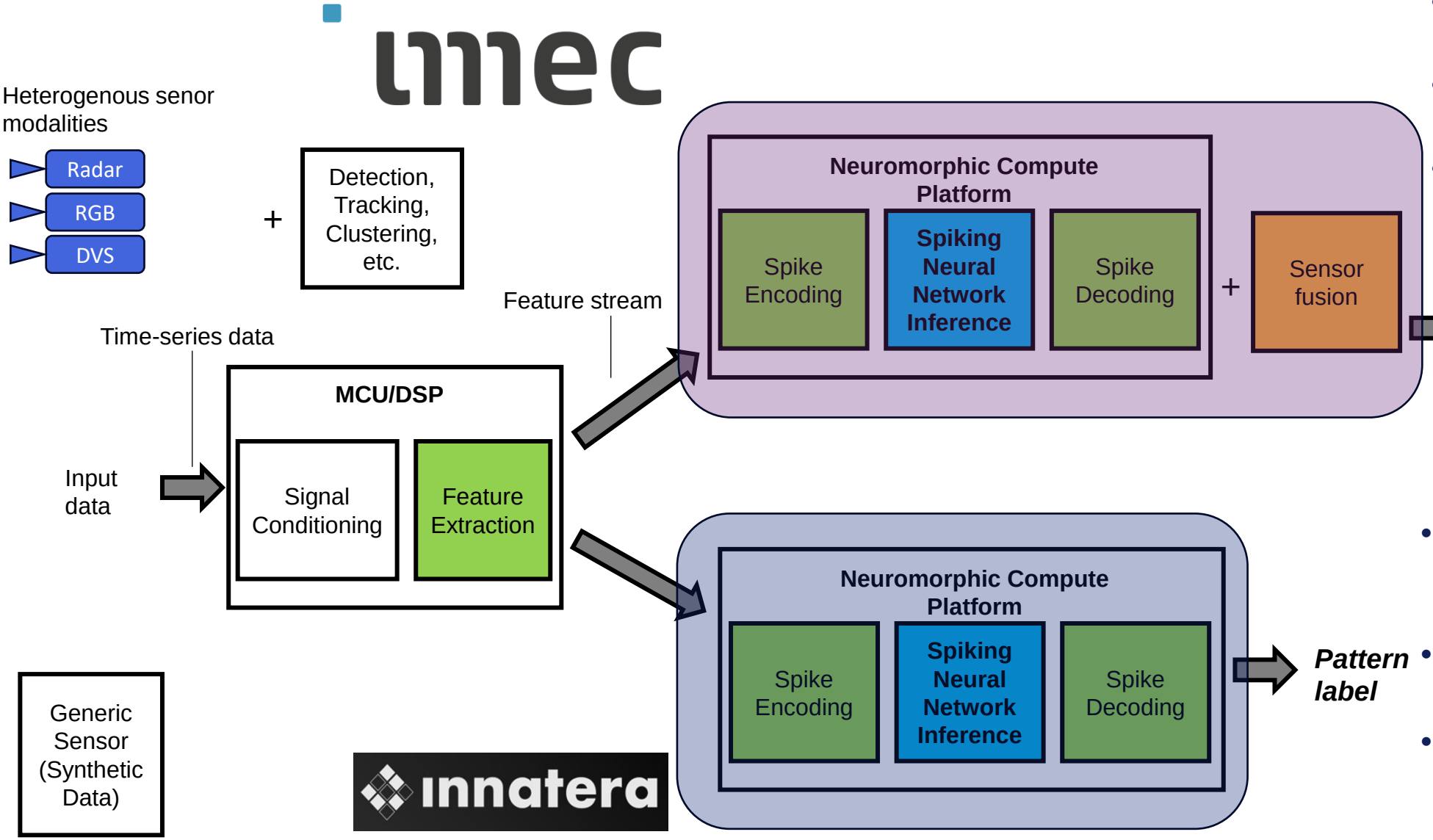
✓Low power/latency and accurate sensing to be matched with low power/latency compute architecture

✓Need for online learning and/or reduced bandwidth towards central node

✓Many (different) sensors enquires a need for event-driven sensor fusion

Copyright © 2023

Methodology



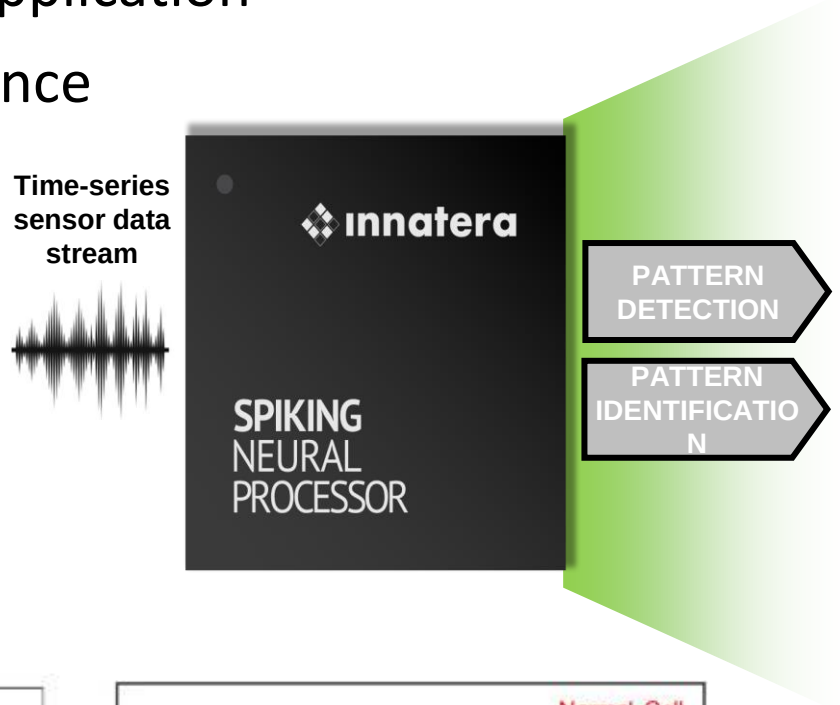
- Fully digital SNN-core implementation
- Multisensor fusion processing
- Vulnerable road users detection, tracking and classification application

- Mixed-signal circuitry for parameterized control of neuron/synapse circuits
- Burst and continues modes of data injection.
- Battery cell voltage monitoring application

Copyright © 2023

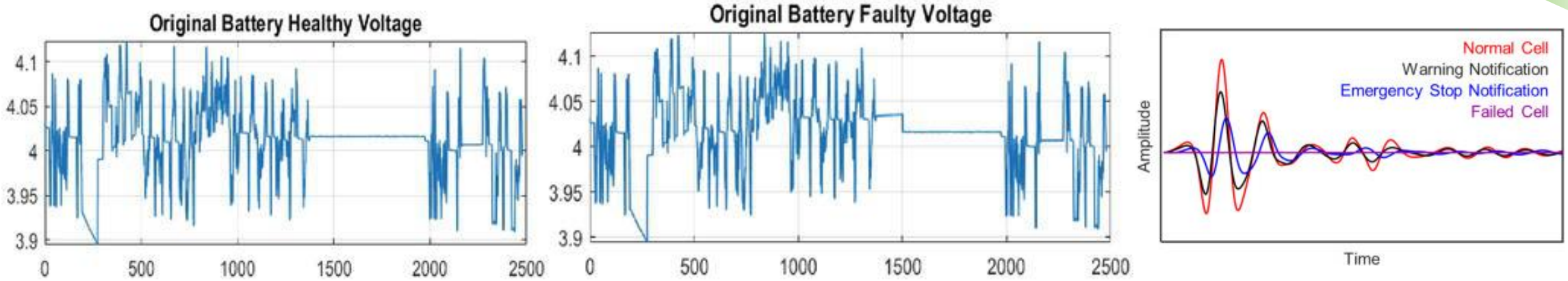
Mixed signal accelerator prototype (INNATERA)

- Power-efficient processors for pattern recognition in sensing application
- Targets a sub-mW power envelope & sub-ms latency for inference
- Uses brain-inspired SNNs for compact, sparse, event-driven processing



Targets

- Enables always-on sensing in battery-powered devices

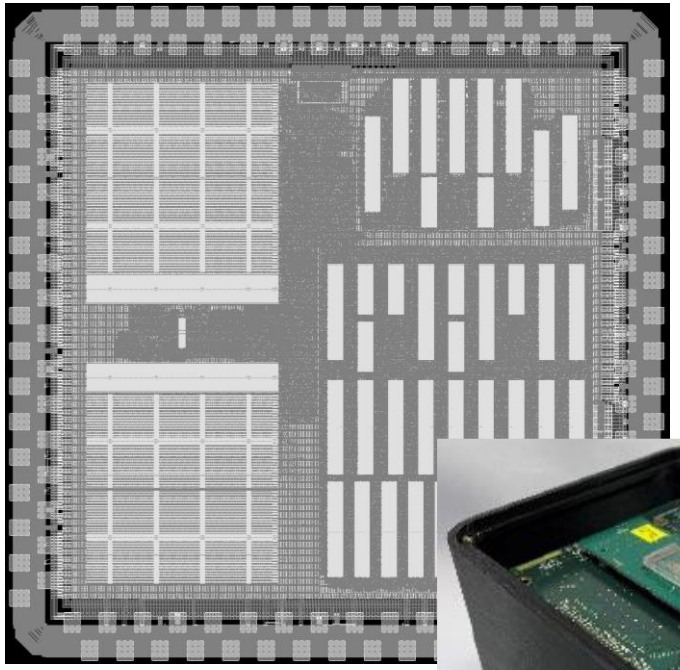


Copyright © 2023

Neuromorphic accelerator for deep edge processing (IMEC)

A highly versatile spiking neural network SoC:

- 1280 IF neurons, clustered in 4 cores of 256+64 fully connected neurons (less than 20 pJ per synaptic operation)
- Reconfigurable into different topologies, enabling optimization for different applications
- On-chip RISC-V μC and standard interfaces (SPI, I2C, I2S, SDIO, UAR, JTAG)
- Efficient toolchain for native training in event-domain (outperforming CNN-to-SNN conversion)
- Google Soli DS, IBM hand gesture DS, etc. were evaluated



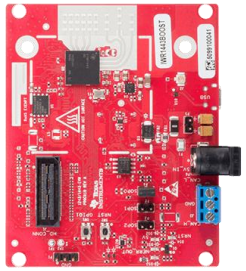
- TSMC 28 nm HPC+
- 4 mm^2 area
- P_{max} = 45 mW (100 MHz clock – 100% neuron activity)
- Typical single core P = 100 uW (sparse activity)

Copyright © 2023

Neuromorphic accelerator for deep edge processing (IMEC)

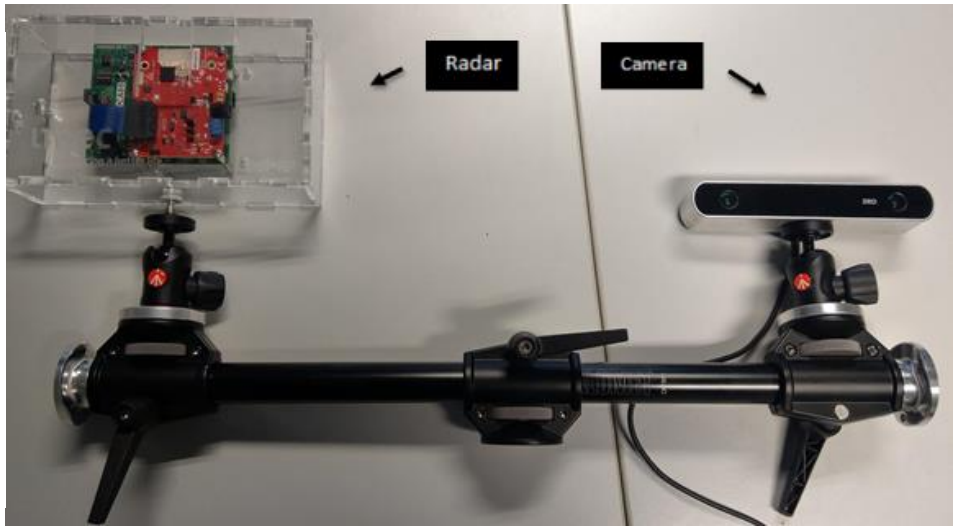
Ongoing work in data collection

- COTS mmWave 76-81GHz radar supported with a raw radar data accusation board



- ZED camera for RGBD sensing

Sequences	Training sequences	Validation sequences	Testing sequences	Total number of sequences
Pedestrian	136 selected randomly from 766	288	481	905
Cyclists	136 selected randomly from 300	42	37	215
Motorbike	136	13	41	190
Car	136 selected randomly from 465	123	95	354



Copyright © 2023

Neuromorphic accelerator for deep edge processing (IMEC)

Ongoing activities in the main application development WP:

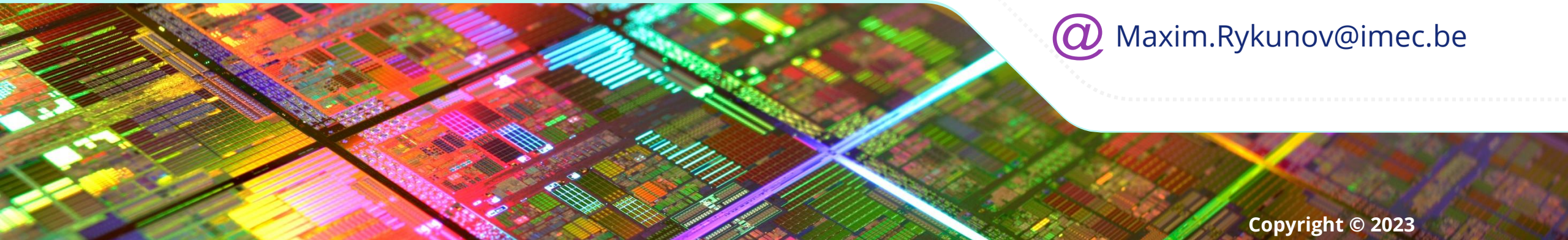
- Finalize all the pre-processing of the done measurements
- Validation of recordings in the developed SNN-based NN
- Porting and testing NN with the designed Neuromorphic accelerator
- Post-processing and heterogeneous sensor fusion
- Extensive benchmarking with SoA CNN/SNN solutions

Copyright © 2023

EDGE AI

Thank You

For your attention



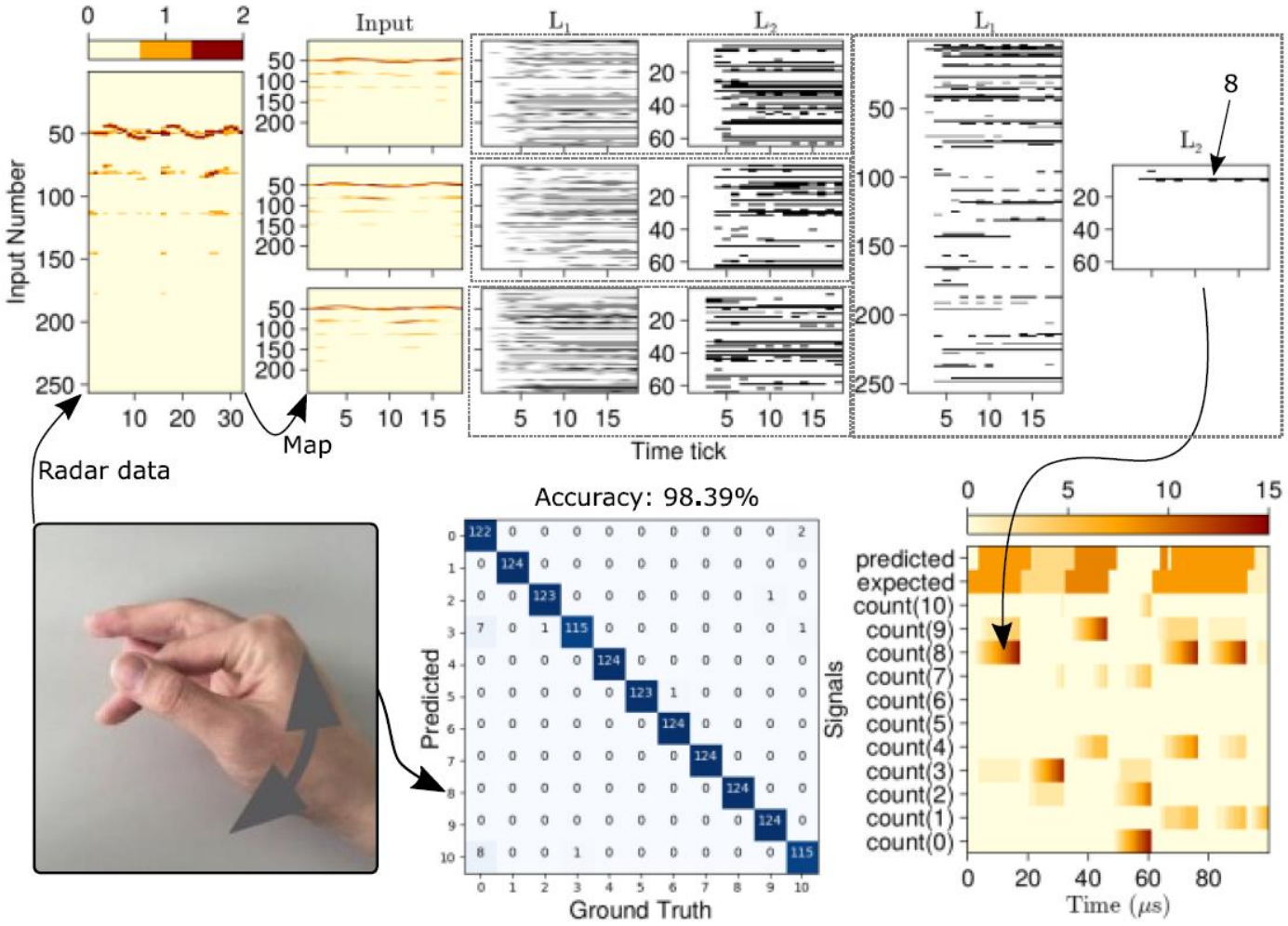
@ Maxim.Rykunov@imec.be

Copyright © 2023

Neuromorphic accelerator for deep edge processing (IMEC)

Experimental Results

- A radar-based application with the Google SOLI data set
- A vision-based application based on the IBM DVS128 gesture data set
- A keyword recognition application with the Heidelberg Digits data set



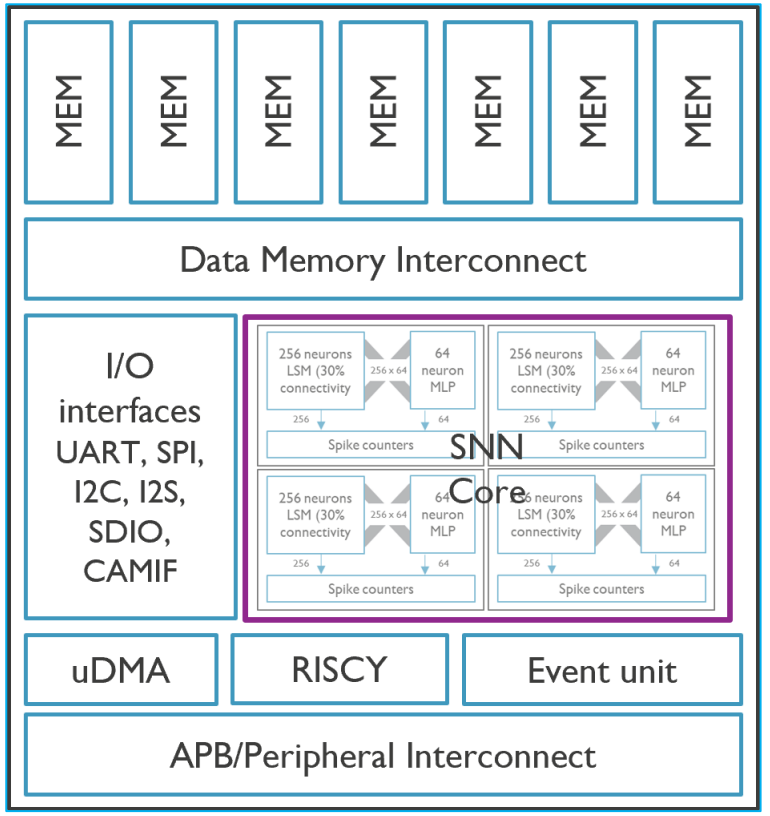
The inference of SOLI data

Copyright © 2023

Neuromorphic accelerator for deep edge processing (IMEC)

A highly versatile spiking neural network SoC:

- Fully digital architecture (self-timed logic with clocked periphery and μC)
- 1280 IF neurons, clustered in 4 cores of 256+64 fully connected neurons (less than 20 pJ per synaptic operation)
- Time multiplexing supported to scale towards networks with 100.000's neurons
- Reconfigurable into different topologies, enabling optimization for different applications

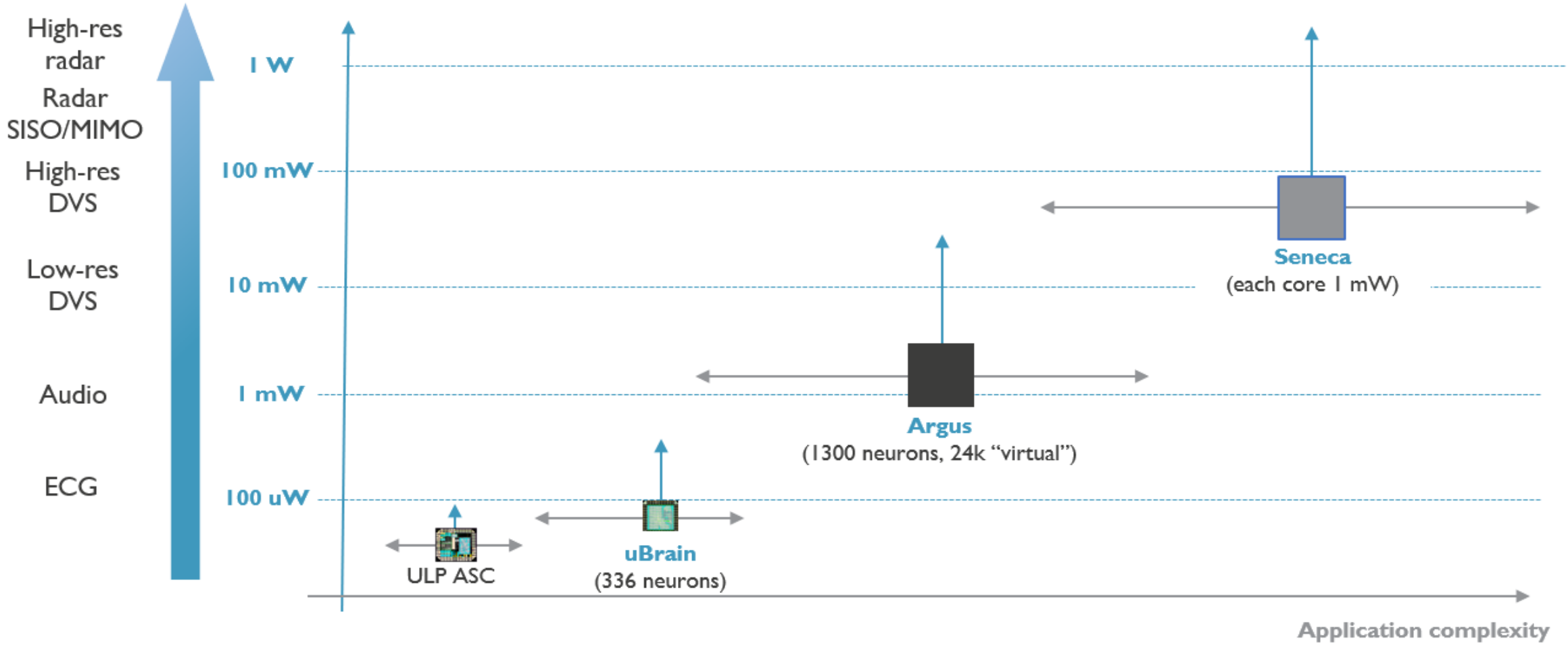


XXX put references

Copyright © 2023

Neuromorphic accelerator for deep edge processing (IMEC)

Imec's story in spiking IC developments to match processor power to sensor power



Copyright © 2023