ETH zürich

CENTER FOR
PROJECT-BASED
LEARNING

RVEdge-Vision: A Fully Open, Ultra-Efficient On-Device AI Platform for Smart Eyewear

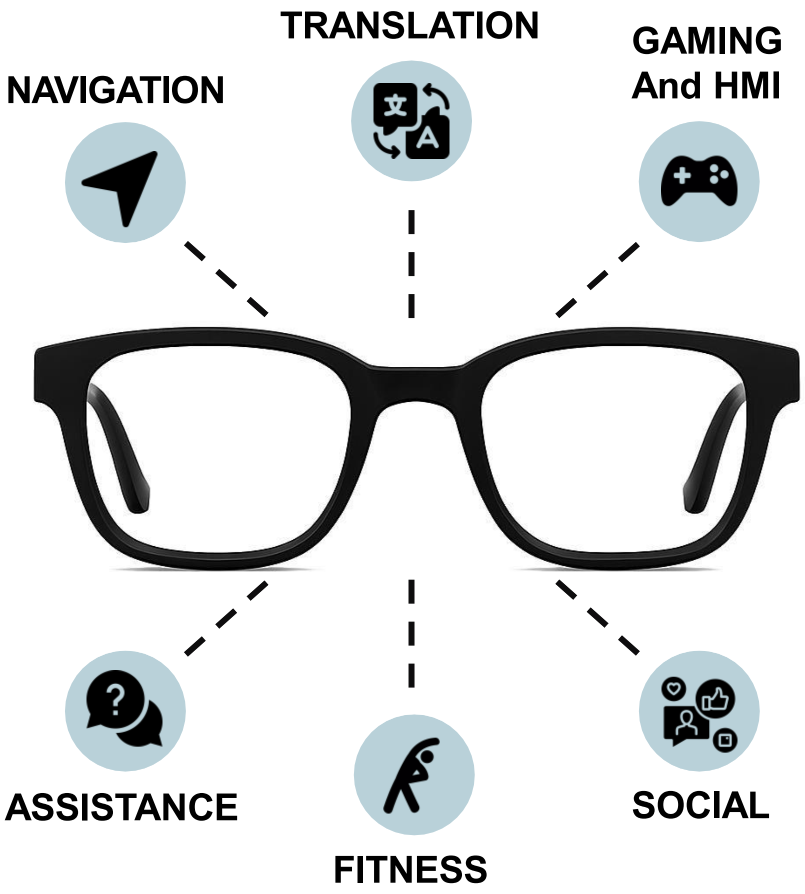
Michele Magno

2026 – RISC-V Summit Bologna

Credits: Mayer, Muller, Bonazzi, Moosmann, Celik.

Smart Glasses as an Edge Platform

CENTER FOR PROJECT-BASED LEARNING



Wearables Are Going Mainstream

Growing demand for hands-free, context-aware computing

Battery Life Is a Bottleneck

Continuous sensing & vision require ultra-low power solutions

Edge AI Is the Future

Research is shifting towards local, efficient processing

Motivation – Background

Rokid Glasses

≈ 4 h

Ray-Ban Meta

up to 4 h

XREAL Air 2 Ultra

external power

Solos AirGo V

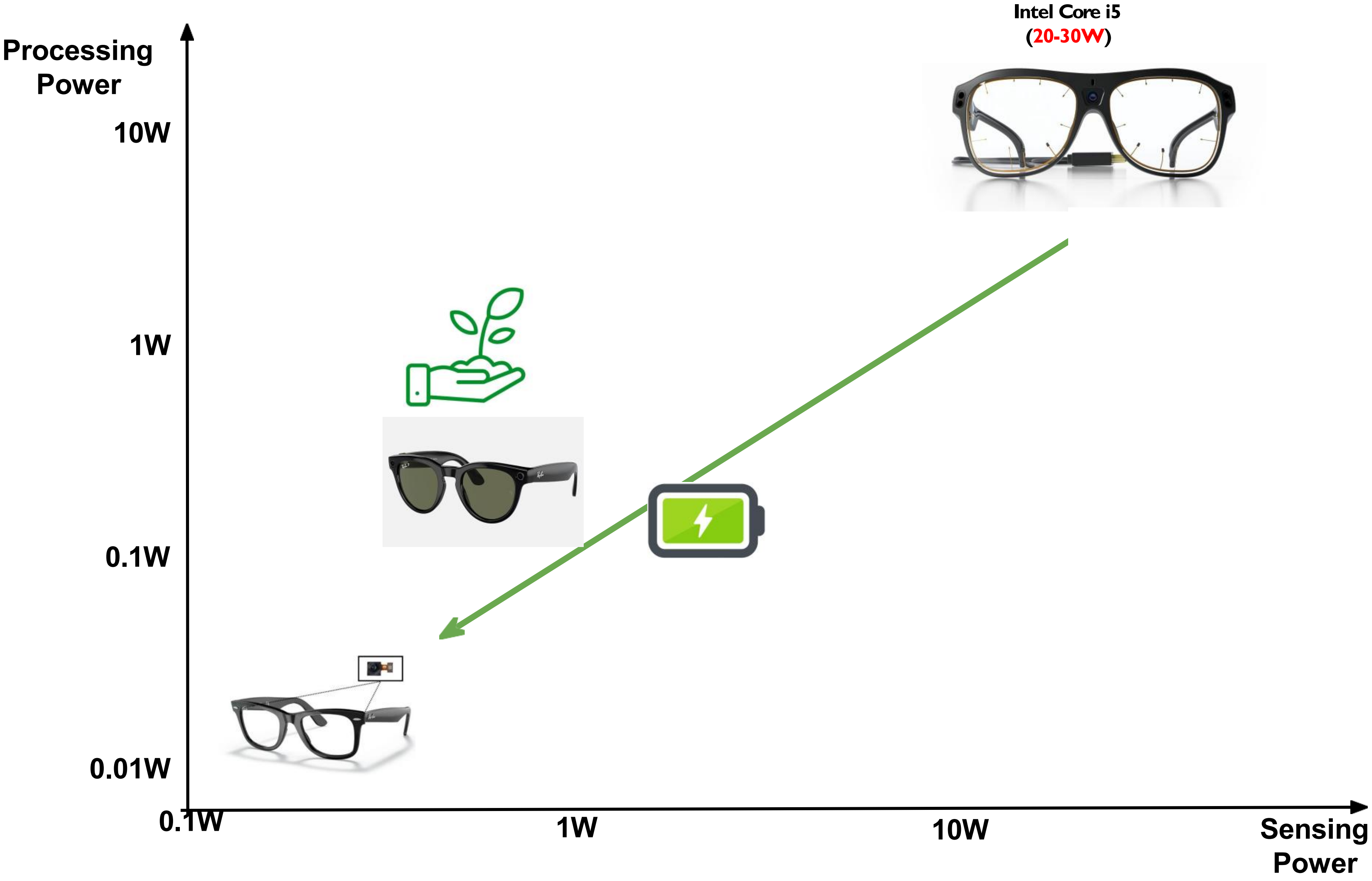
Recent Trends:

- Greater HW/SW complexity
- More functionality per area
- Better sensors

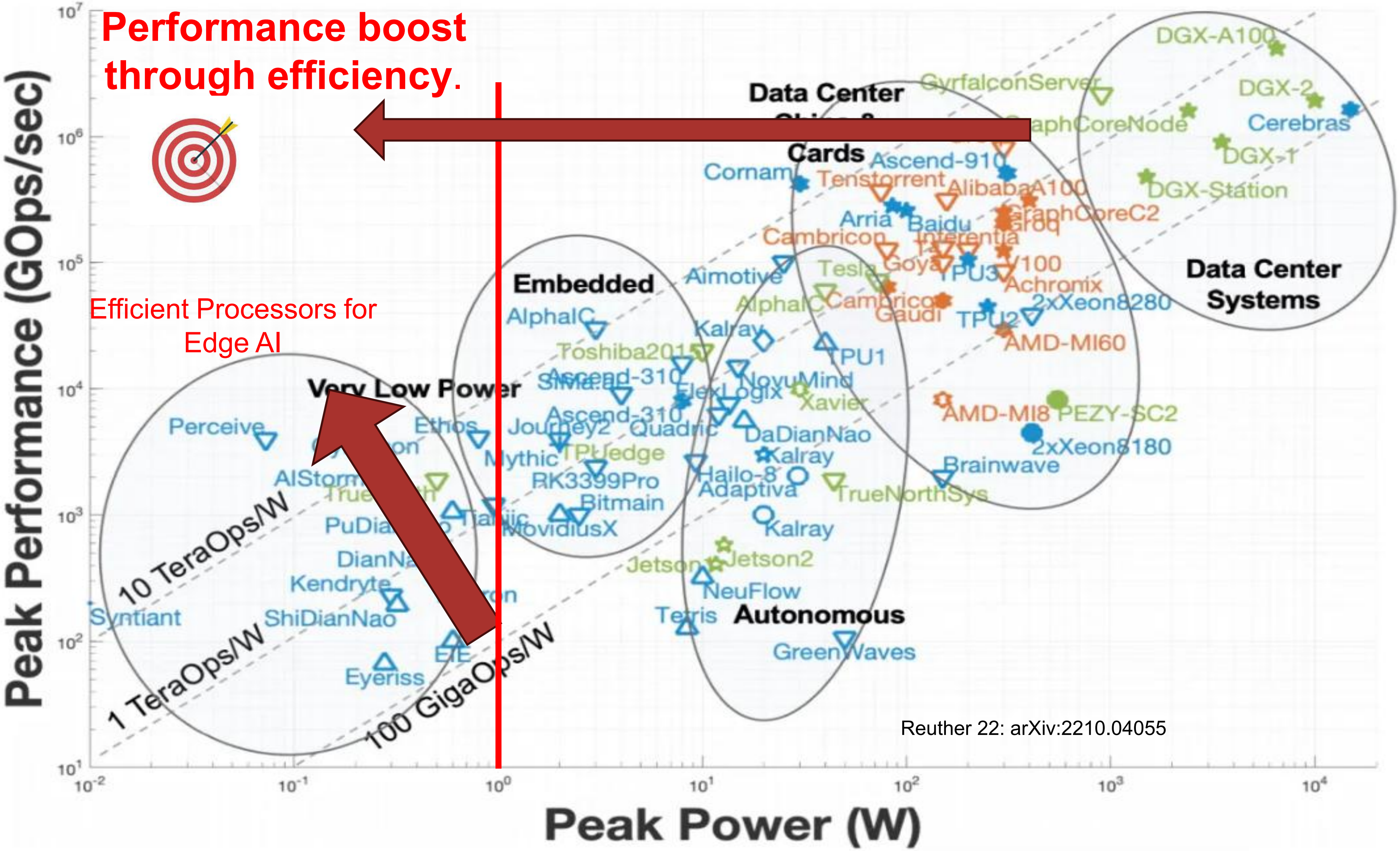
Biggest Challenges:

- Energy storage/consumption/source
- On Device processing

How to extend the lifetime?



Challenges of Bringing AI to the Edge



Reuther 22: arXiv:2210.04055

Nvidia GPUs

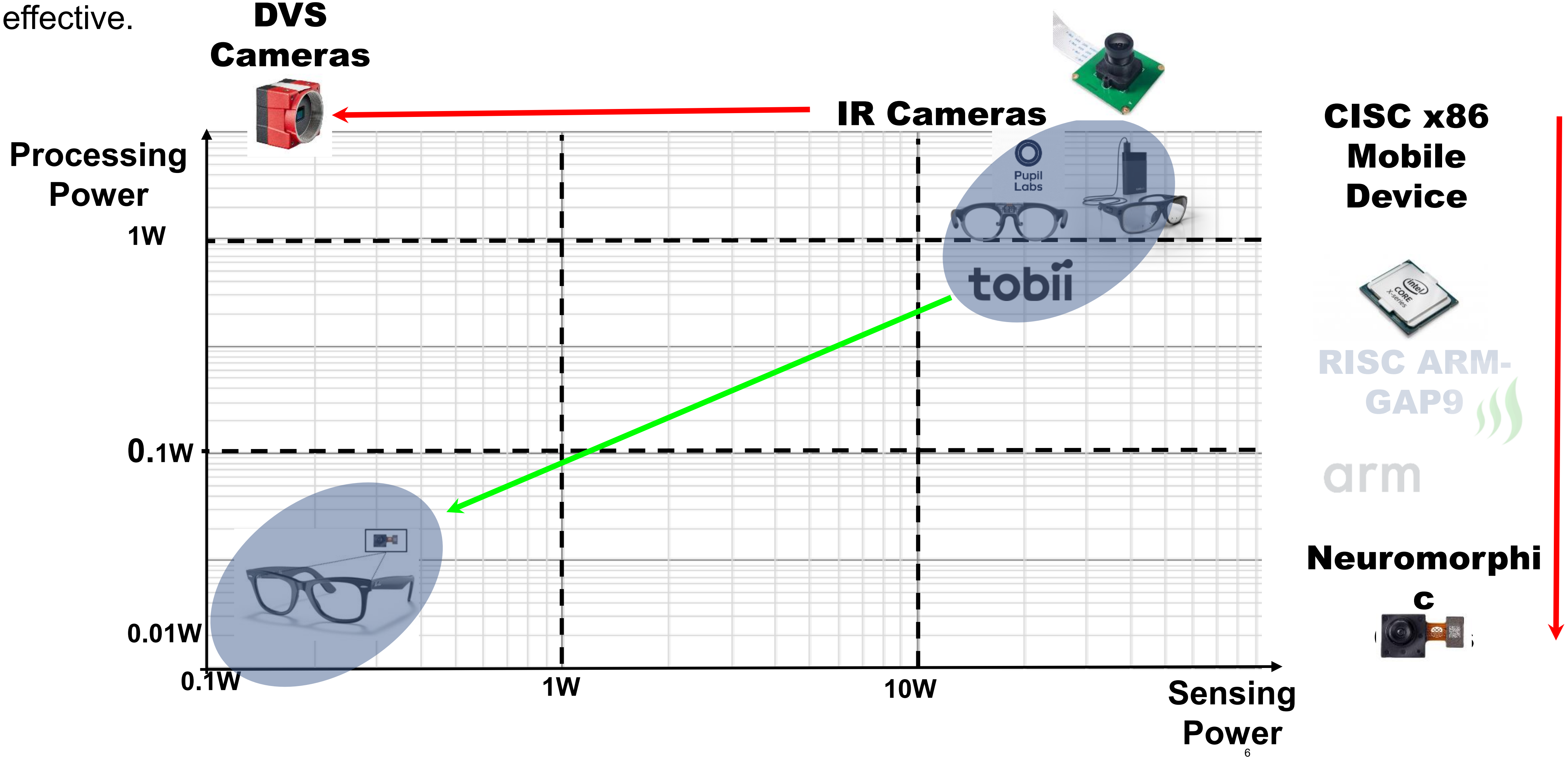
Advanced AI

Hardware alone is not enough— efficient systems, model compression and software optimization can reduce memory and compute needs by up to 90%, enabling real-time AI even on sub-Watt devices.

Application-Specific Sensors Reduce Data Load

CENTER FOR PROJECT-BASED LEARNING

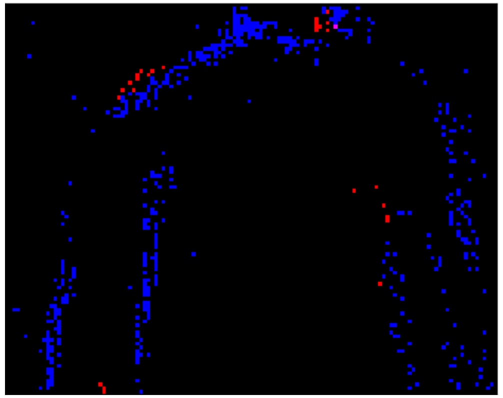
- Instead of using high-bandwidth cameras for every application, **specialized sensors** can be more effective.



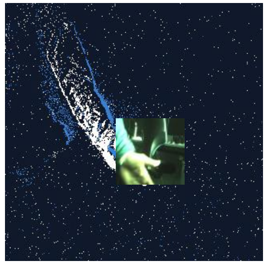
Energy Efficient and Low Latency Smart Sensors

Edge AI & Sensor Fusion
Efficient processing and AI-driven adaptation

10-1000x more intelligence
10-100x Less Latency



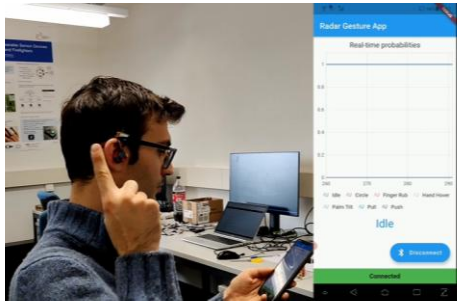
1000-10000x Less data
1-100x Lower Power



Applications Specific: Intelligent Smart glasses

Low Power Sensors & Efficient Hardware Design

100-10000x Efficiency



Eff. Wireless Communication
(Optimized data transfer, seamless interaction).

Open Smart Glasses Platform

Development Platform

Visually Appealing

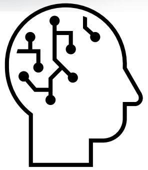
On-device AI execution

Privacy Conscious

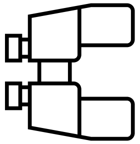
Ultra-low Power

Multi-day Battery runtime

Wake-up Sensors



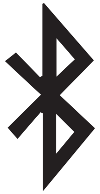
On-Board Inference



Vision Sensors

Spatial Sensors

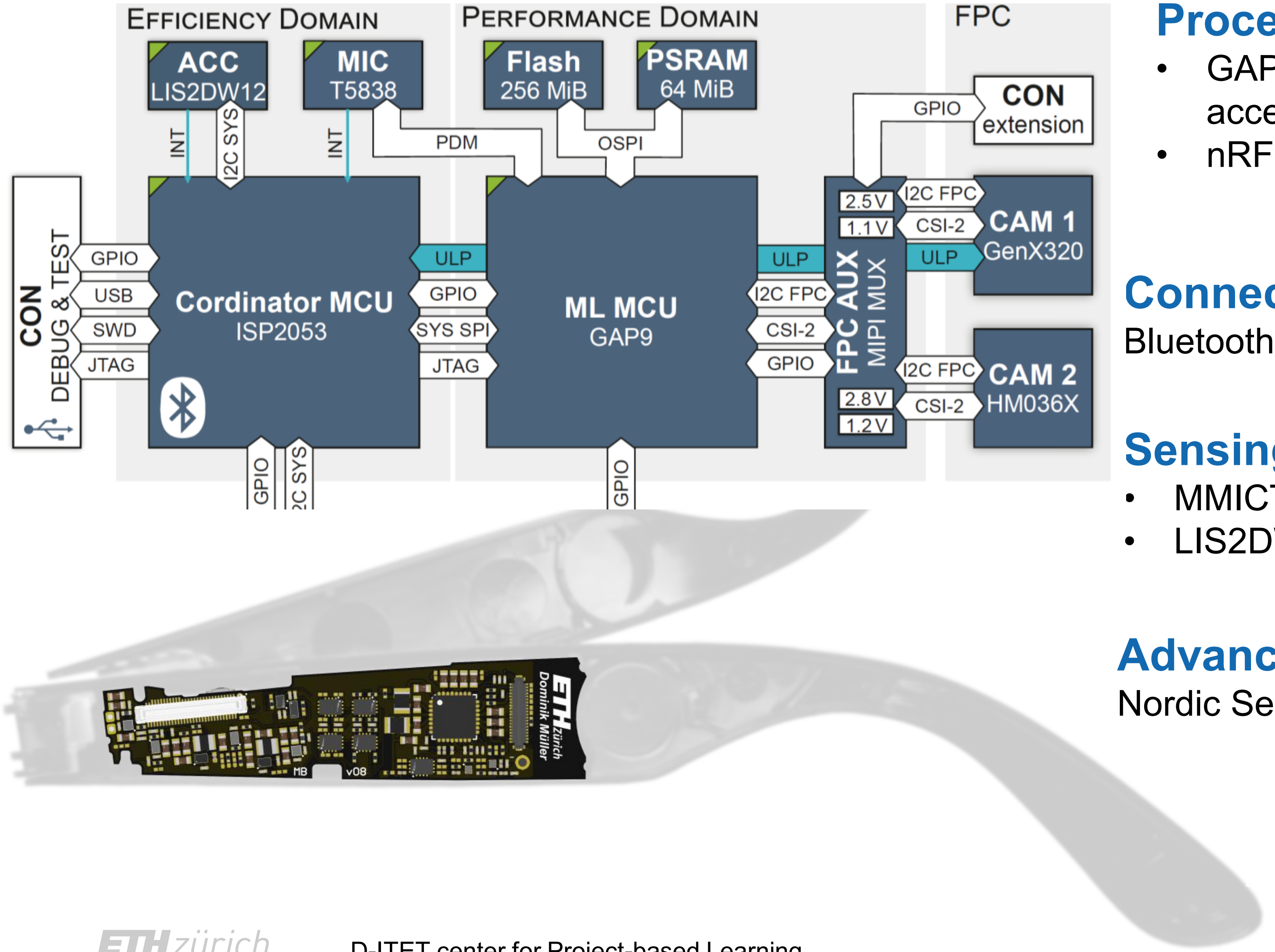
Battery Operated



BLE

Hardware Architecture and Sensors

Open Smart Glasses



Processing

- GAP9 – 9 RISC-V core compute cluster with AI accelerator - L3: 256 MB RAM & 256 MB Flash
- nRF5340 – Cortex-M33

Connectivity

Bluetooth 5.4, BLE, BT Mesh, Thread, Zigbee, Matter, ANT

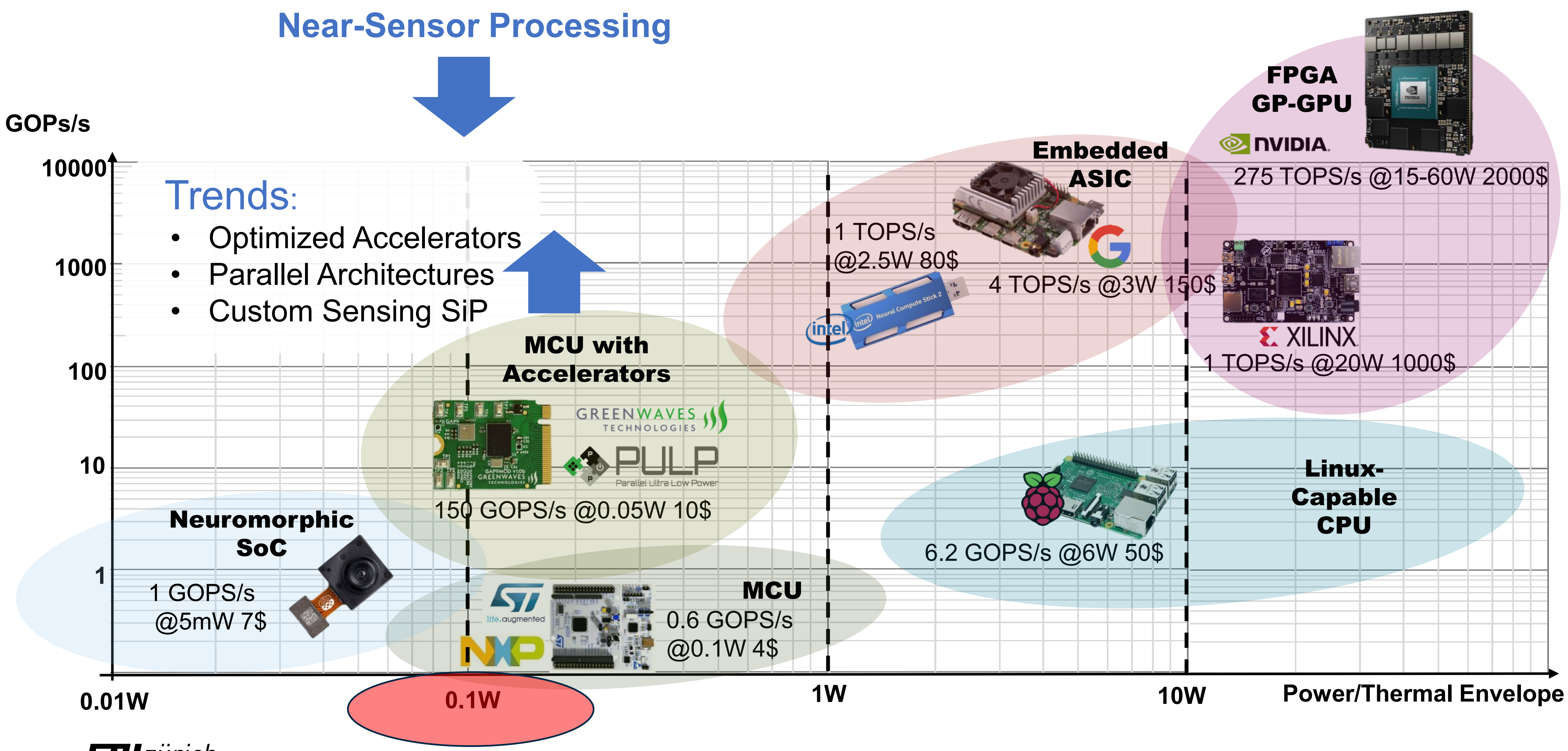
Sensing

- MMICT5838 – Microphone
- LIS2DW12 – Accelerometer

Advanced Power Management & Monitoring

Nordic Semiconductor nPM1300 – fine-grain power gating

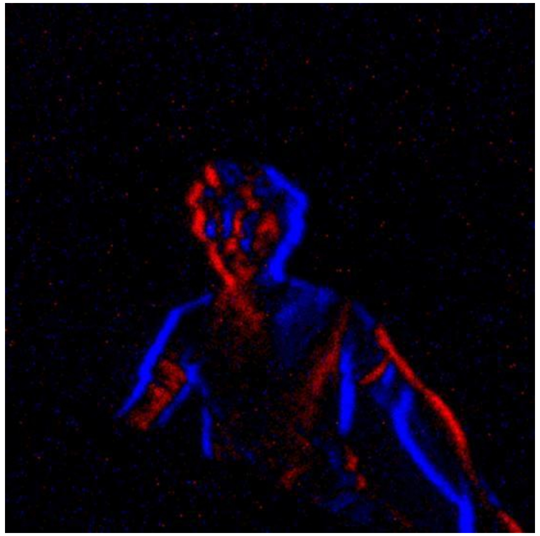
New Platforms for Edge Computing: Importance of Hardware Aware



Smart Glasses – Frame Flex

Battery

GenX320 [2]
Event-based Vision Sensor



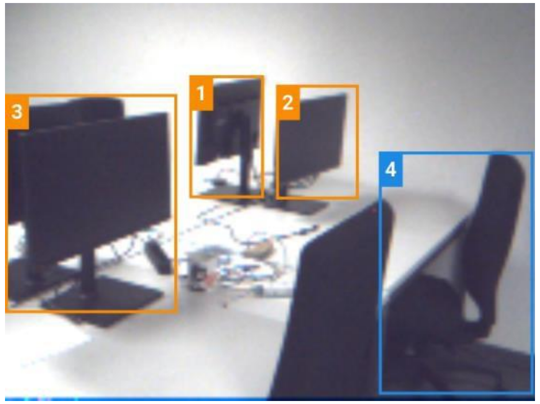
320 x 320 px
>140 dB

6.7 x 4.7 x 1.58 mm

ADAP UT-P2023 [4]
MEMS Speaker

HM0360 [3]
ULP CMOS Image Sensor

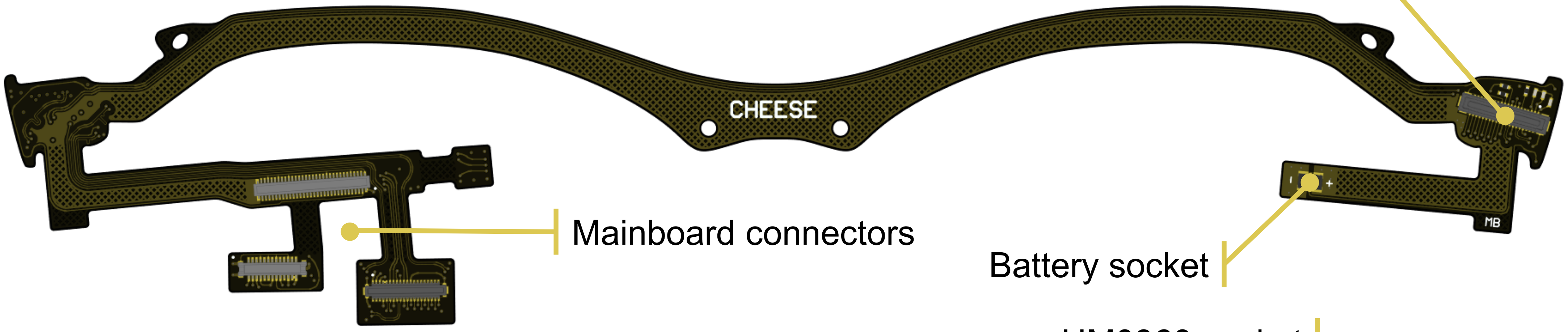
640 x 480 px
85 dB



Images: usound.com

Frame FPC

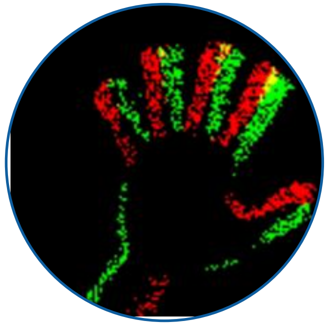
Front View



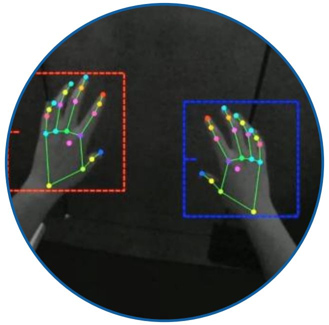
Back View

Use-Case Gesture Recognition

Event-based Gesture Recognition



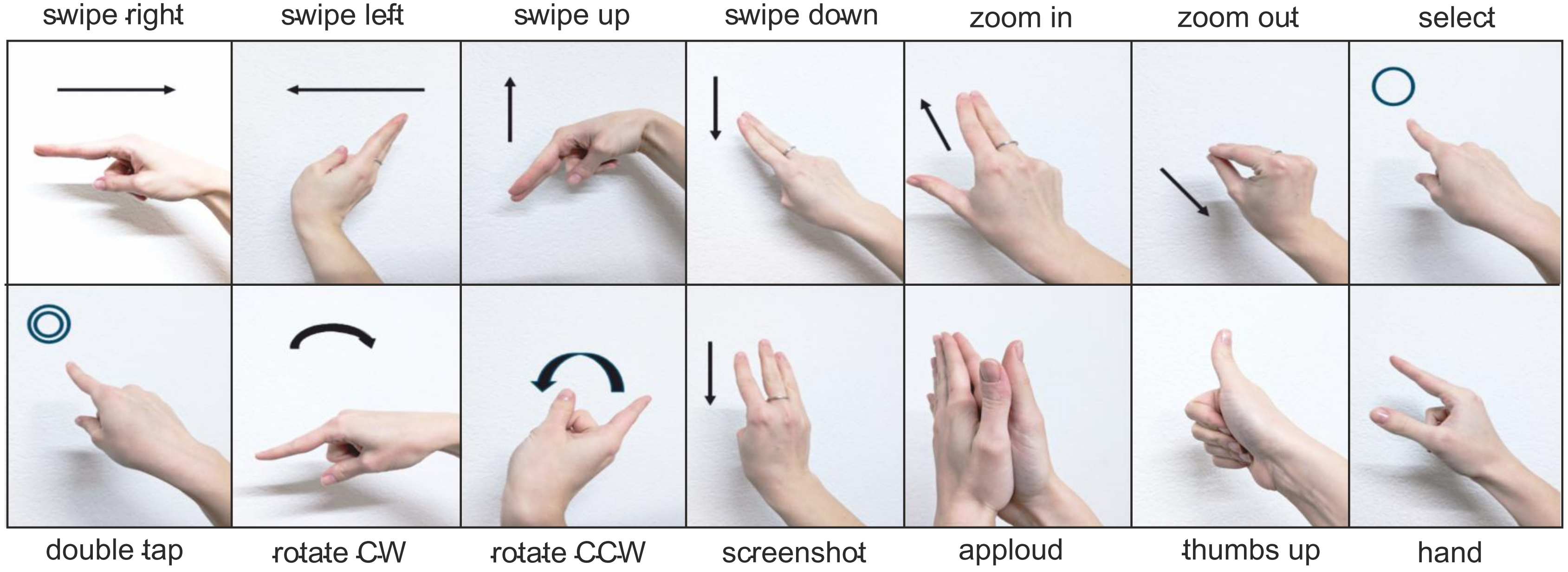
Event-based dataset for robust gesture recognition

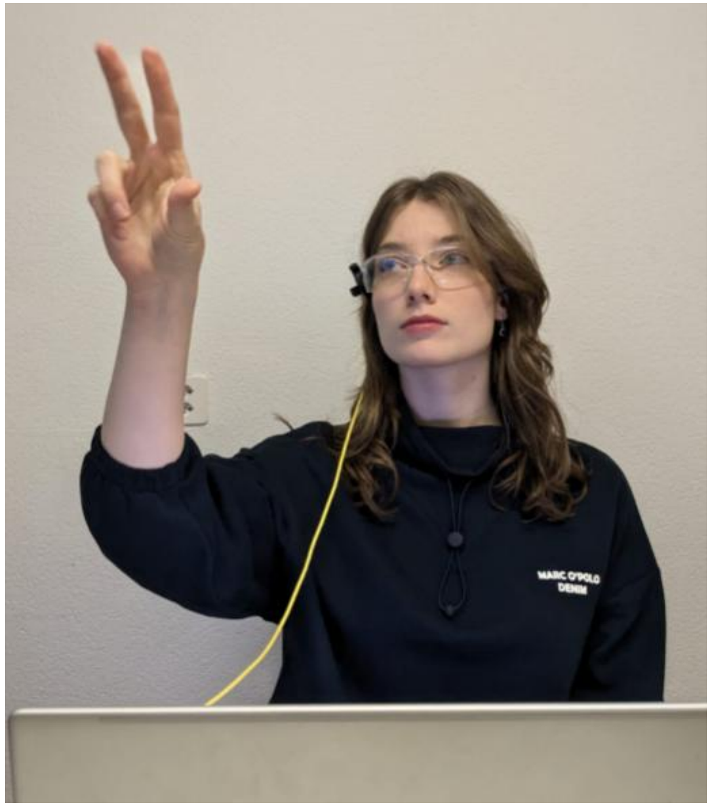


Simultaneous gesture recognition and hand tracking

Deployment on custom ULP HW

Event-based Gesture Recognition
Dataset



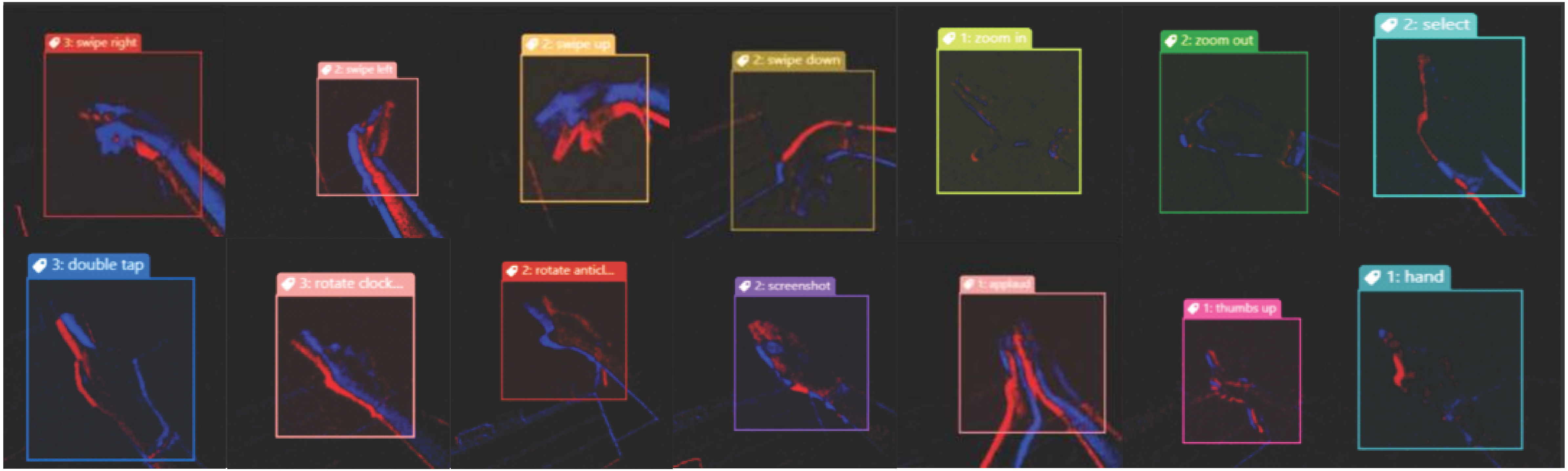


#Subjects	# Gestures	# Sequences	Avg. Sequence lenght (s)	# Frames
19	14	988	18	375,440

Event-based Gesture Recognition
Dataset

CENTER FOR PROJECT-BASED LEARNING

swipe right | swipe left | swipe up | swipe down | zoom in | zoom out | select



double tap | rotate CW | rotate CCW | screenshot | aploud | thumbs up | hand

Event-based Gesture Recognition
Evaluation of Gesture & Bounding Box Prediction

BEST TEST RESULTS PER ARCHITECTURE (STRICT SUBJECT HOLD-OUT, T=10, STRIDE 5).

Architecture	#Params	Test Acc.	Macro F1	Wtd. F1
MobileNet	3 776 855	61.79%	0.486	0.609
DD-TCN	526 545	63.79%	0.486	0.638
TCN	493 137	69.06%	0.551	0.690
TSM	589 887	79.62%	0.731	0.795
R(2+1)D	627 407	**83.94%**	**0.781**	**0.840**

R(2+1)D wins: 83.94% accuracy at ~6× fewer params than MobileNet

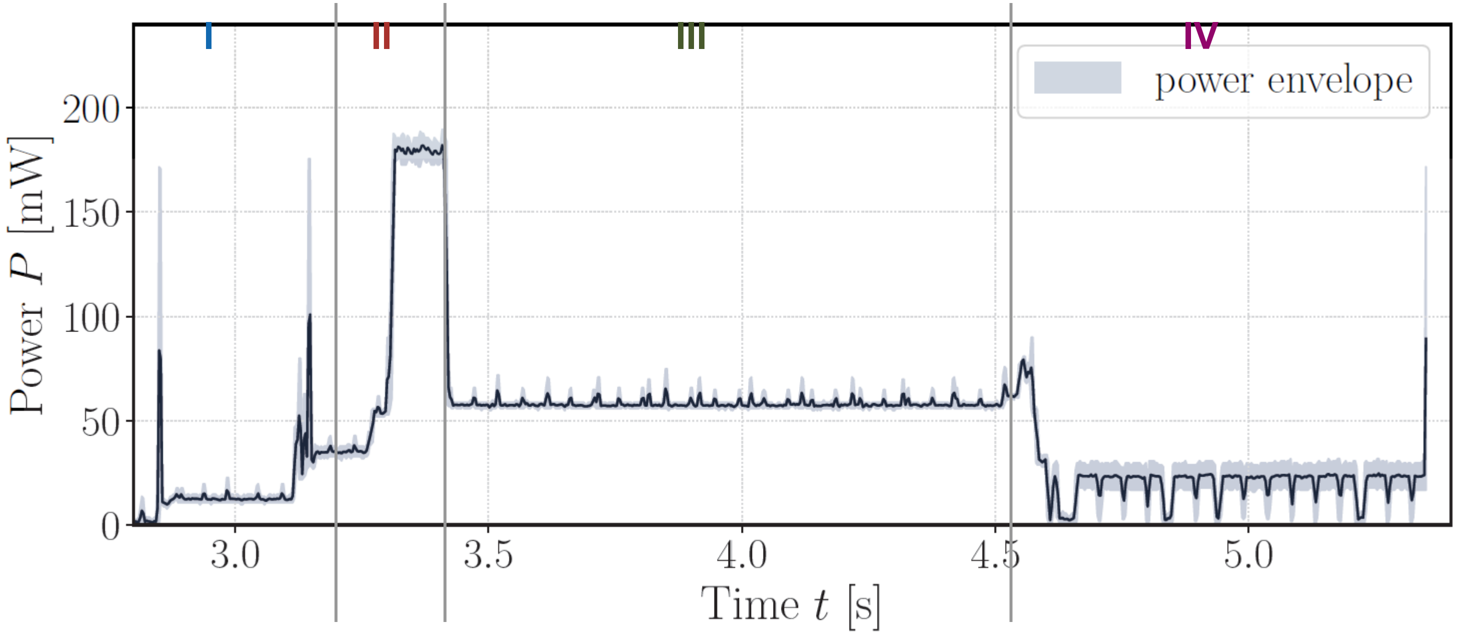
- Best accuracy and smallest viable footprint (627K params, 0.840 wtd. F1)
- Strict subject hold-out confirms generalization to unseen users
- Distinct gestures reach F1 ≥ 0.90; ambiguous ones (Select, Zoom In) lag

PER-CLASS RESULTS FOR THE BEST MODEL (R(2+1)D, DROP G12/13, AUG, BUNDLED; TEST ACCURACY 83.94%).

Gesture	Precision	Recall	F1	Support
Swipe Left	0.806	0.844	0.824	64
Swipe Right	0.970	0.955	0.962	67
Swipe Up	0.821	0.622	0.708	37
Swipe Down	0.870	0.944	0.905	71
Zoom In	0.561	0.649	0.602	57
Zoom Out	0.630	0.741	0.681	85
Select	0.400	0.286	0.333	28
Double Tap	0.833	0.938	0.882	160
Rotate CW	0.987	0.854	0.916	178
Rotate CCW	0.918	0.855	0.885	221
Screenshot	0.868	0.917	0.892	72
Weighted avg			**0.840**	1040
Macro avg			**0.781**	1040

P. Bonazzi, J. Moosmann, A. Celik, P. Mayer and M. Magno, *OpenGlass: Open-Source Smart Glasses Platform for Event-Driven Gesture Interaction*, ETH Zurich, 2026.

Power Measurements



OVERVIEW OF EXECUTION TIME AND ENERGY CONSUMPTION ACROSS THE STAGES OF THE IMAGE PROCESSING PIPELINE EXECUTED ON THE GAP9 MICROCONTROLLER (MCU).

		t [ms]	$\overline{P}$ [mW]	E [J]
	Initialization	505.180	22.942	11.590
Vision Pipeline	Image Acquisition	20.251	45.903	0.930
	Pre-Processing	0.532	80.527	0.043
	Model Inference	75.400	96.061	7.627
	Post-Processing	0.222	56.357	0.012
	Pipeline	33.920	65.612	8.613
	Deinitialization	31.070	27.736	0.862

Battery Lifetime — 300 mAh @ 3.7 V

Always-on — ≈ 17 h

continuous vision pipeline, 65.6 mW

1 inference / day — ≈ 324 days

5 s wake cycle, 50 μW deep sleep

Inference (7.63 J) dominates the active-cycle budget.

Full end-to-end cycle resolves into four power phases:

- **I Cold start / GAP9 init** – ≈22.9 mW avg, transient spikes during peripheral bring-up
- **II Camera power-up** – ≈180 mW plateau during GENX320 cold-start
- **III Vision pipeline** – ≈65.6 mW avg; one inference every 75.4 ms (Model Inference = 7.63 J, the cycle's dominant cost)
- **IV BLE TX & deinit** – ≈27.7 mW avg, duty-cycled pulses (nRF5340 advertise / connect / transfer)

Conclusion

•**Efficiency is the Key Enabler** – Power constraints demand ultra-low-power AI, requiring optimized hardware-software co-design.

•**Open Architectures Drive Innovation** – RISC-V and modular platforms enable scalable, customizable, and efficient AI processing for wearables.

•**Sensor Fusion Enhances AI Efficiency** – Event-based vision (DVS), radar, and novel sensing (e.g., electrostatic charge) significantly reduce energy consumption.

•**Application-Specific AI is the Future** – Tailored, domain-specific AI models like TinySSIMO YOLO maximize performance while keeping power usage minimal.

•**Future Outlook: Smarter, More Efficient Wearables** – Advancements in edge AI, energy-efficient sensors, and hardware accelerators will drive the next generation of intelligent eyewear.

Thank you for your attention!
Q&A