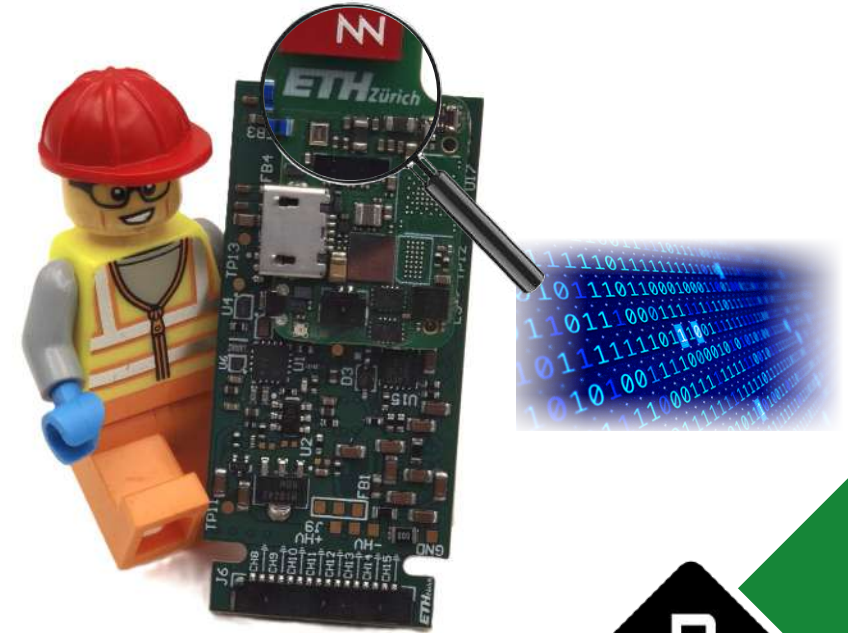


# Beyond the Transducer: A Lens on Computing for Wearable Ultrasound

ETH Future Computing Laboratory (EFCL)  
Integrated Systems Laboratory (IIS)

**Andrea Cossettini**

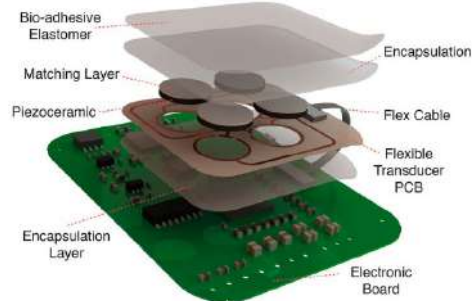
[cossettini.andrea@iis.ee.ethz.ch](mailto:cossettini.andrea@iis.ee.ethz.ch)



# Ultrasound went wearable, we need more wearable computing



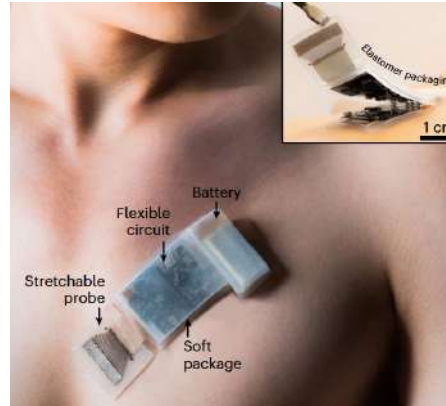
[1]



[6]



[2]



[3]



[4]



[5]



[7]

...

always-on

body-worn

untethered



Privacy



Low-latency



Long  
battery life



More bandwidth  
(compression)



Accuracy

edgeAI

- [1] Frey et al, IEEE IUS, 2022  
[2] Vostrikov et al, IEEE TUFFC, 2024  
[3] Lin et al, Nature Biotechnology, 2024  
[4] Lu et al, Nature Electronics, 2026

- [5] Giordano et al, IEEE Internet Things J, 2025  
[6] Toymus et al, Nature Communications, 2024  
[7] Flopatch, Flosionics Medical



# In today's talk I'll focus on computing for wearable US

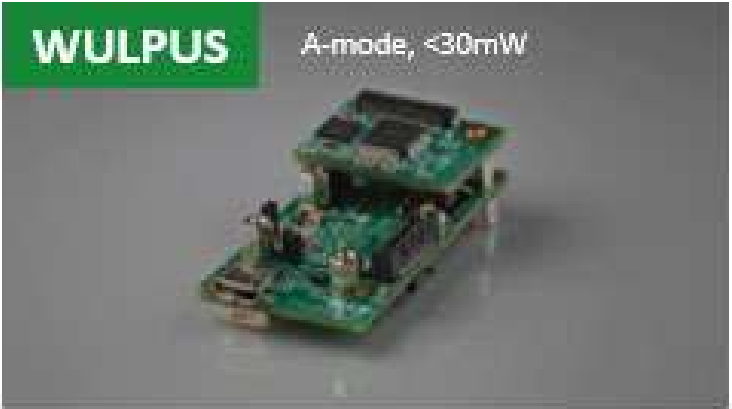
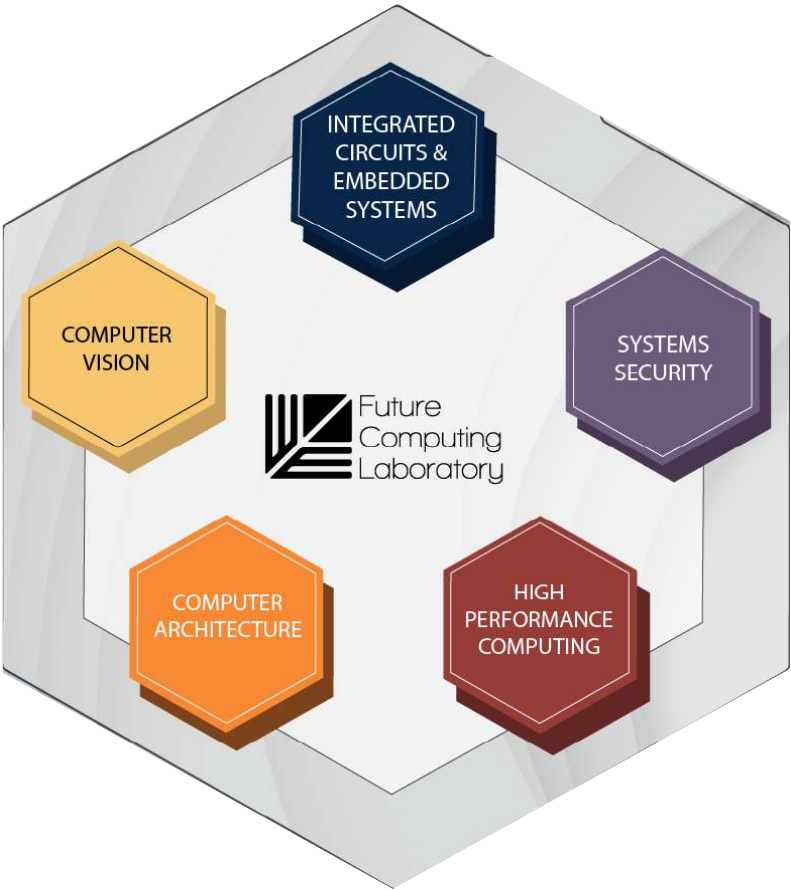
*wearable US  $\neq$  standard images*  
*wearable US  $\neq$  standard biosignal time series*

makes it interesting as a computing target  
opportunities for ultrasound and computing community together

# Introducing ourselves

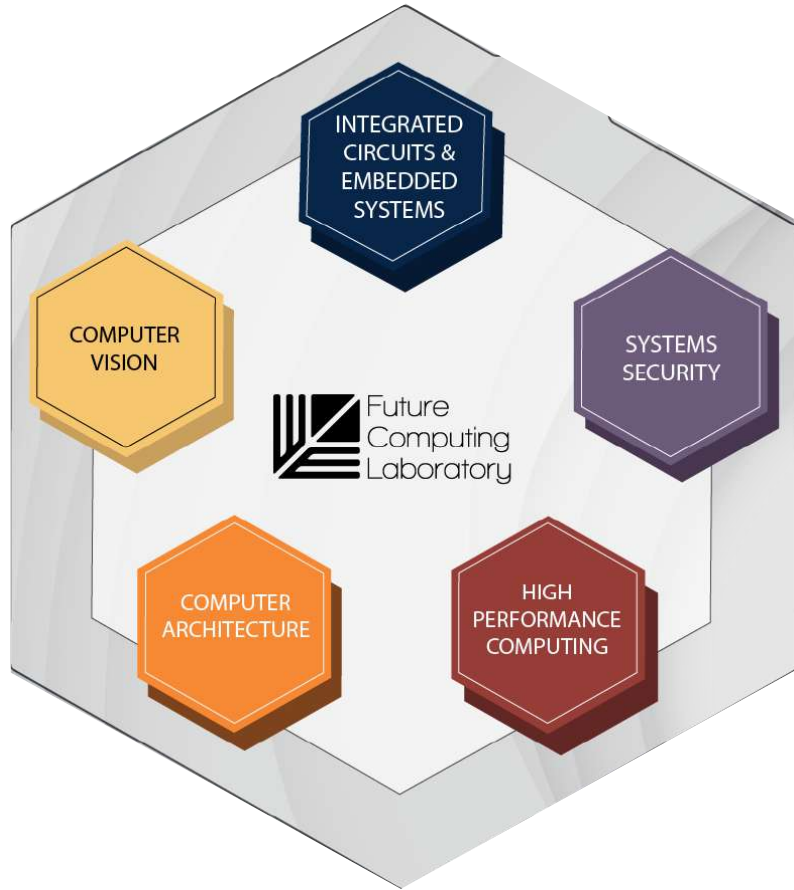


# A computing center, looking at ultrasound



you may know some of our wearables

# A computing center, looking at ultrasound



*center-wide research inspiring  
groups vertical implementations*

## Horizontal computing methods

**HW/SW co-design**

**Edge computing**

**Foundation models**

**open source**

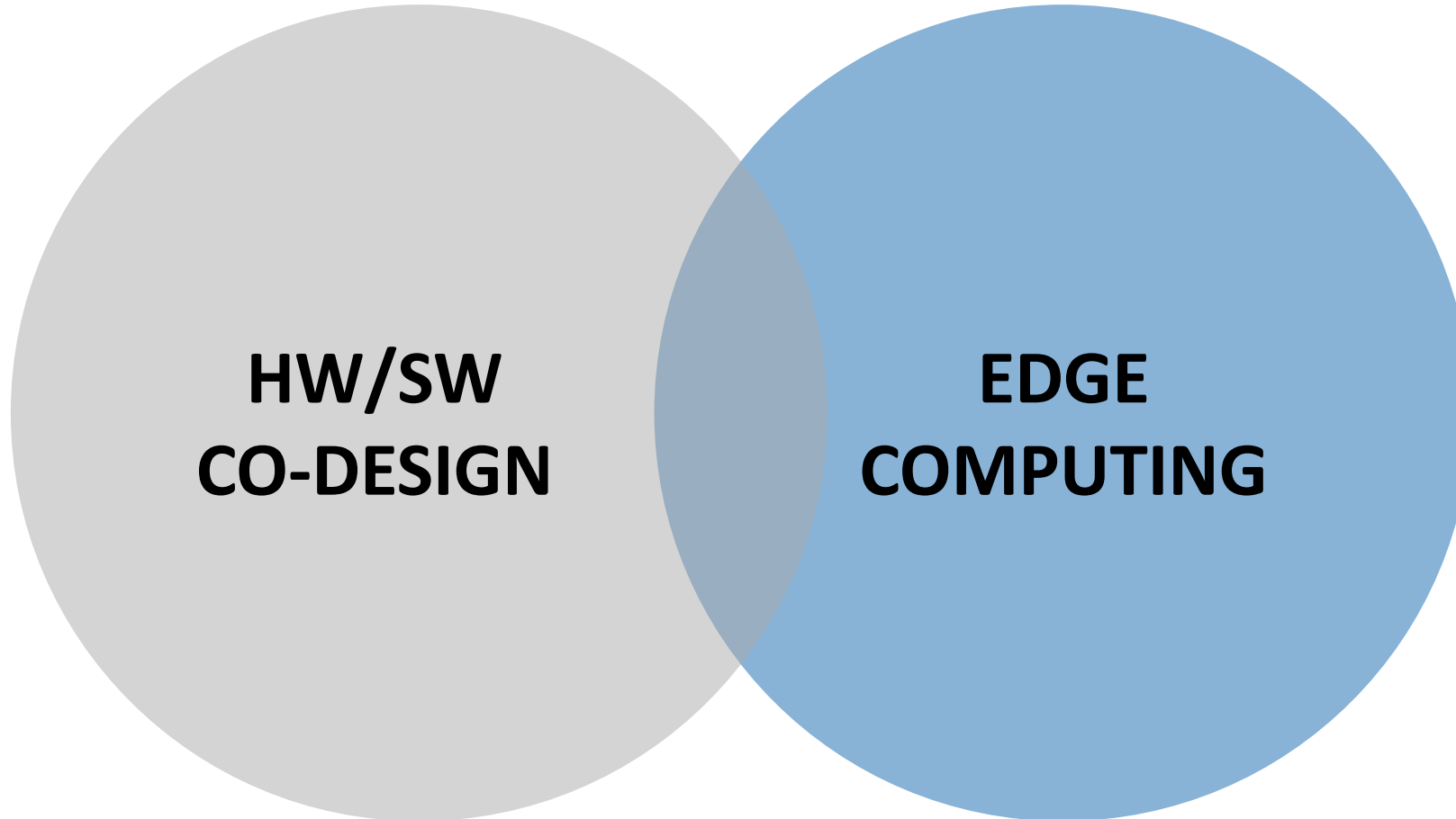
## Wearable ultrasound

Beneficiary domain

Earlier maturity w.r.t. other biosignals

Enable steeper growth

# How did we add more computing in wearable US and how did we shape our wearables around computing?



*Implementation examples*



Today's focus is on A-mode

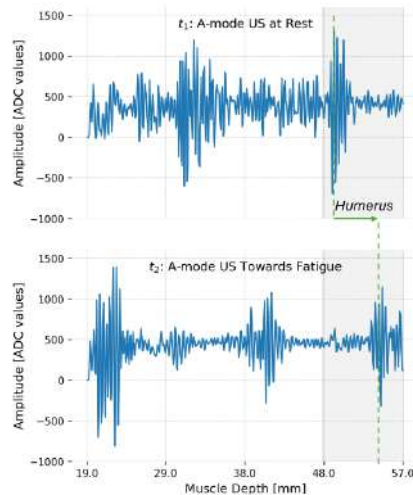
# First example: muscle fatigue on the edge w/ peak finding



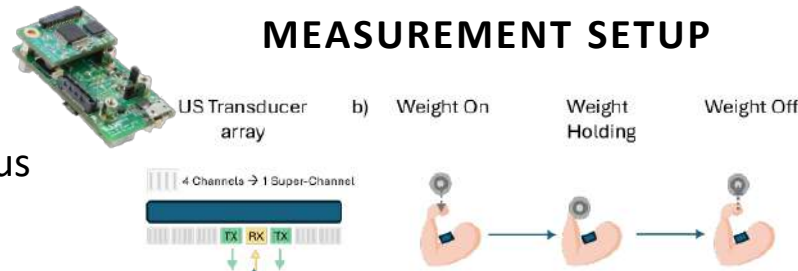
## End-to-end on-device muscle-thickness tracking for fatigue

### SIGNAL PROCESSING

Bandpass filter → Hilbert envelope  
→ select ROI (depth range of humerus bone) → peak finding

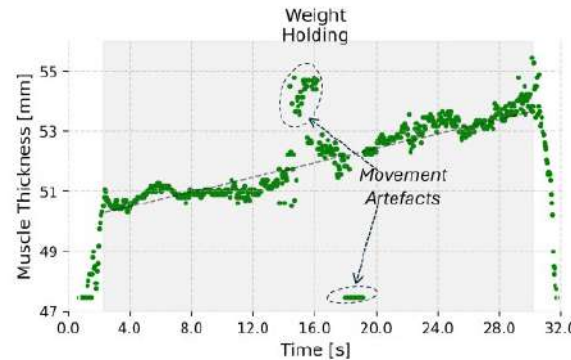


### MEASUREMENT SETUP



WULPUS + 2.25 MHz transducer - isometric 80% 1RM hold to exhaustion

### MUSCLE-THICKNESS FATIGUE PROFILE



Biceps thickness rises ~linearly under sustained load



**4.1 ms**

On-MCU processing per frame



**26%**

Total system power saved by edge proc.  
(11.7 mW total system power)



**201×**

Wireless bandwidth reduction  
(from 161 kbps to 0.8 kbps)



**4 days**

Continuous use (320 mAh)

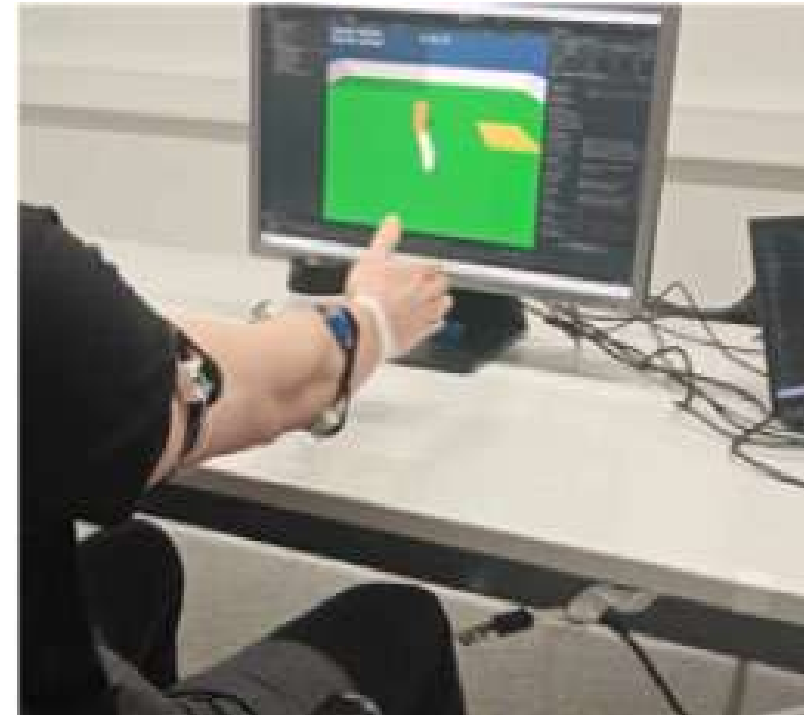
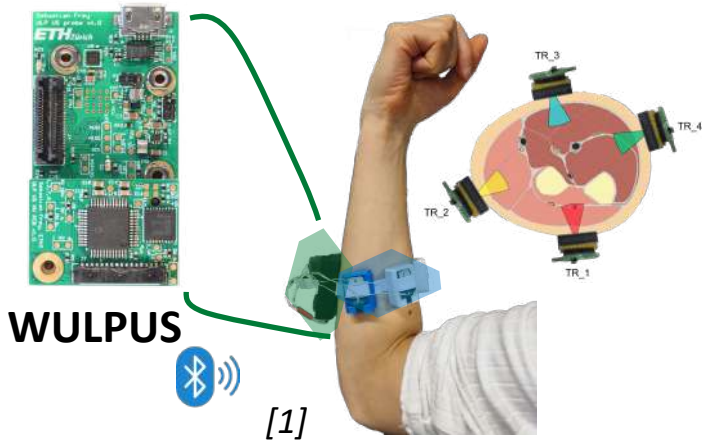
# Going more complex: hand kinematics



- More challenging problem: ultrasound armband to identify hand gestures / hand kinematics

*TODAY's DEMO → WULPUS for VR interaction*

4 transducers



[2]

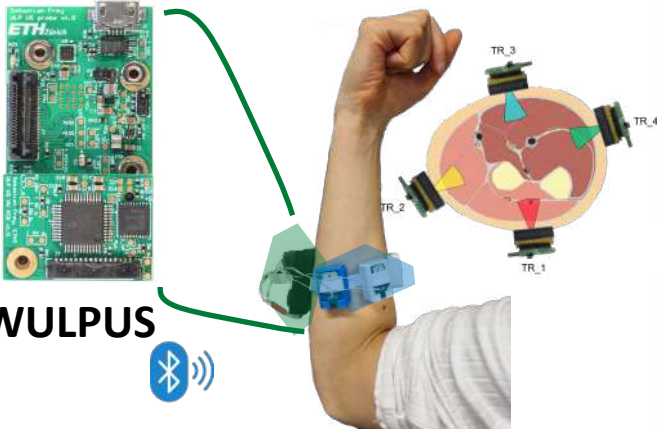
Feasible doing it on the edge?

# Going more complex: hand kinematics



- **Task:** full-hand kinematics: 20 hand + 3 wrist DoF; MANUS glove for ground truth
- **Approach:** Compact multi-output CNN: 11,285 params (very little memory available on the nRF52!)
- PoC on one subject → incremental fine-tuning: train on session 1+2, incrementally add new data:  $R^2 \sim 0.7$

4 transducers



WULPUS



nRF52: 64 kB of RAM memory  
(43 kB used for the acquisition firmware)

 **0.73 mJ**

Energy per inference

 **29.1 ms**

Latency (< 100 ms target)

[1]

 **88 %**

Wireless bandwidth saved

 **33 mW**

End-to-end system power

 **36 h**

Continuous use (320 mAh)

**Deployed: WULPUS (nRF52)**

*see how at tomorrow's Lab Secrets session on ML edge deployment*

**Better accuracy requires more complex models [2] → more capable processors**

# when it's not enough: co-processor integration



*The next generation of WULPUS brings powerful edge computing to the device*

BioGAP baseboard [1]

**Memories**  
512Mb RAM  
512Mb Flash

 **BLE**  
nRF5340

**edgeAI**  
GAP9 SoC  
(9 RV cores cluster;  
400 MHz;  
L1: 128kB;  
L2: 1.6 MB)

Connector for modules

BioGAP for  
ultrasound

new  
WULPUS  
frontend [2]

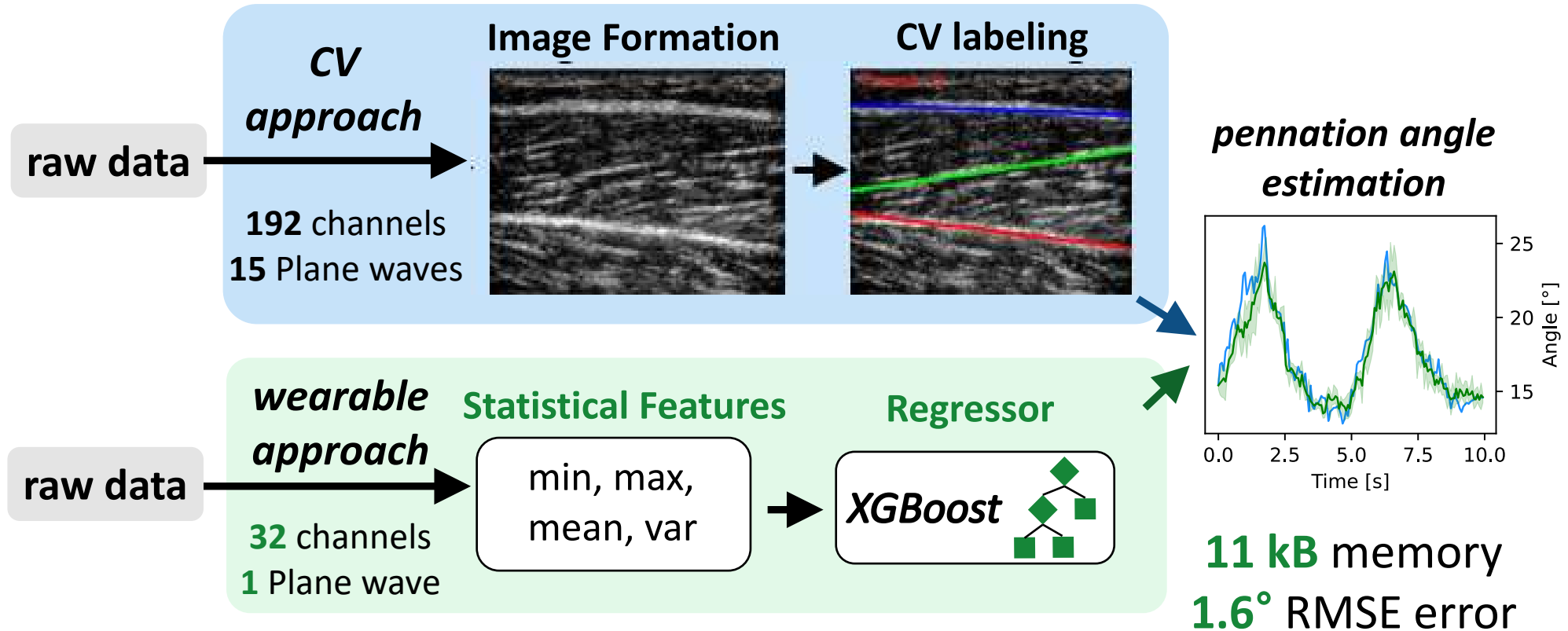
power



# More capable wearables: co-design for multi-channel US



Dataset of gastrocnemius muscle contractions (192-ch. Verasonics).  
Can we implement a processing pipeline tailored for wearables?



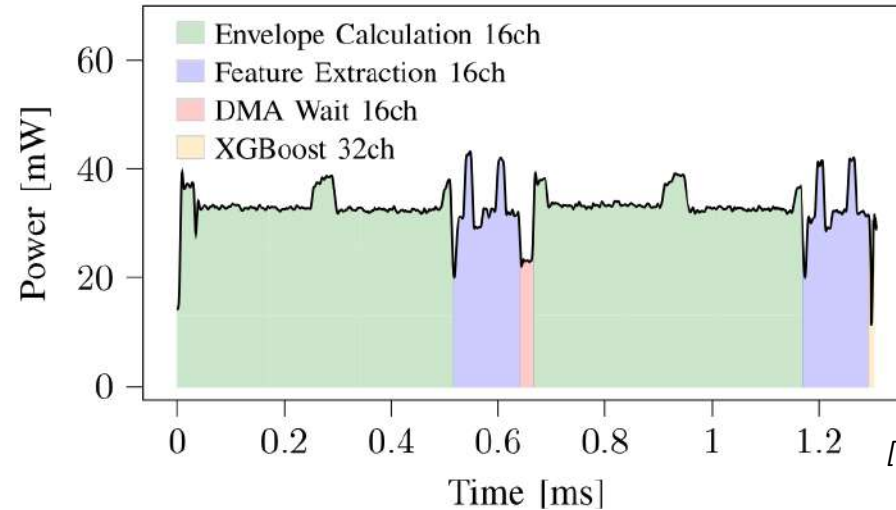
# Co-design & edge compute for multi-channel wearable US



*Model deployed on GAP9  
running at 240 MHz, 0.65V core voltage*

**33 mW** avg power

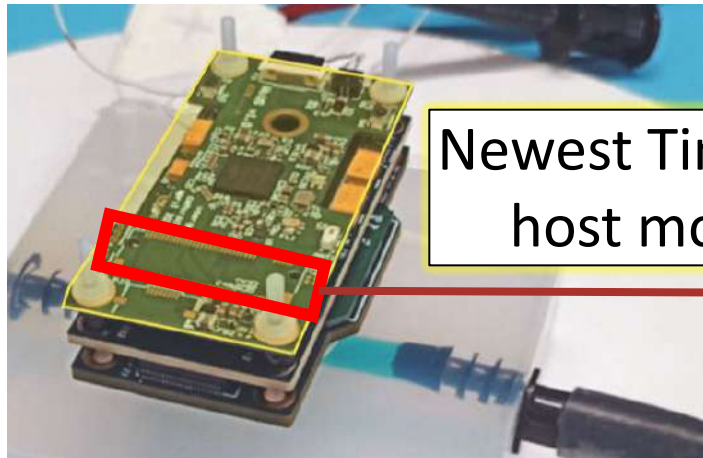
**1.3 ms** inference time



[Vostrikov et al., IEEE TIM, 2023]

Execution time allowing for frame rates of 750 Hz!

*when data are not transmitted  
continuously and processed locally*



Newest TinyProbe  
host module

Includes expansion  
for GAP9  
co-processor



# From classic ML and CNN: the next frontier is foundation models



The current key focus of our center  
involving several groups

**BIOSIGNALS**

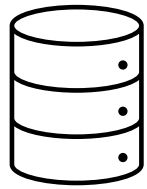
CLIMATE

...

EGO VISION

LANGUAGE

# The next frontier: foundation models



$X$ : Raw Data  
**Cheap**

Heart Rate (ECG)  
Brain Signals (EEG)  
Muscle Activations (EMG)

$Y$ : Labels  
**Expensive**

Atrial Fibrillation, Seizure, etc.

## Supervised Learning



## Self-Supervised Learning



many downstream tasks based on  
the pre-trained backbone

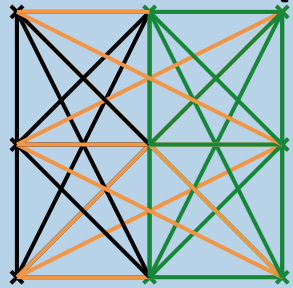


**Foundation**

# The next frontier: foundation models on the edge



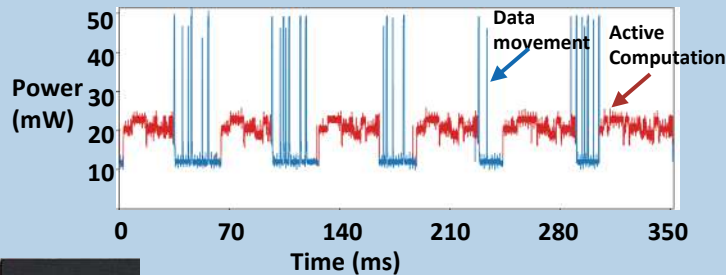
## CEREBRO (EEG)



Full Attention  
 $O(N^2)$

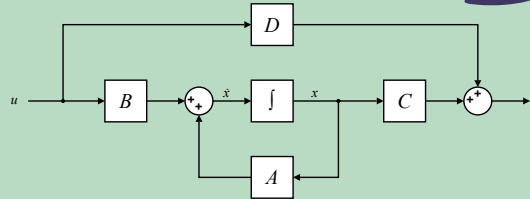
Spatial  
Temporal  
Alternating Attention

85M/40M/3.5M  
model sizes



10 mJ per inference  
0.35s runtime

## FEMBA (EEG)



Utilizing State Space Models

$O(N^2) \rightarrow O(N)$

78M/48M/7.8M  
model sizes



7.8M model (5s EEG windows)

75 mJ per inference  
1.7s latency

## TinyMYO (EMG)



Transformer-based

EMG FM spanning multiple  
diverse downstream tasks



4.5M model (on 5-s windows)

45 mJ 0.785s

2.6M model (on 1-s windows)

0.089s latency

Transfer this knowledge to wearable US?

# The future: foundation models for wearable US



- We're now attempting to do the same for ultrasound
- Main issue: data availability
  - Many public datasets of ultrasound reconstructed images (good for clinical/imaging systems)
  - Very little available for wearables (raw RF data)!!! compared to e.g. TBs of data for EEG
  - Difficult to do pre-training without sufficient data

**We need more open datasets (across different transducers, probes, use-cases...)**

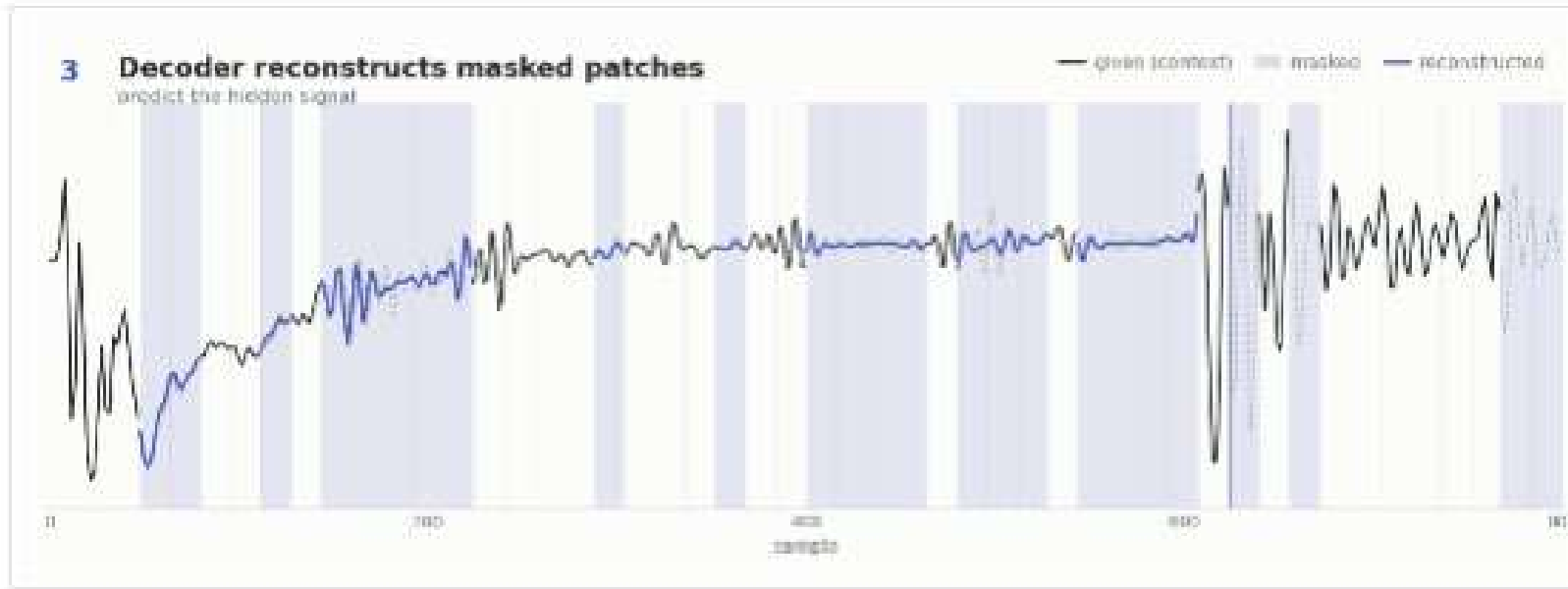
- Our preliminary explorations: possible thanks to recent open public dataset releases + multiple internal data collections (new datasets to be released soon)

# Foundation models for wearable US: WiP



- Masked autoencoder for raw A-mode ultrasound data
- Early-stage results show good signal reconstruction accuracy

*Measured on 27 held-out A-mode traces across 9 datasets.*



**0.98**

median masked-patch correlation ( $r$ )

- and promise for downstream-tasks (accuracy increase w.r.t. supervised model)
  - 56.5% → **70.5%** (6-class hand-gesture classification)
  - 76.0% → **81%** (3-class forearm pose estimation)

**WiP - now exploring different architectures & pre-training strategies**

# Thinking of wearable US from a computing perspective



- Horizontal innovation in computing transferring to wearable US

**hw/sw co-design**

**edge computing**

**Foundation models**

- and of course there's more...

**specialized architectures  
& hw acceleration**

**Computer vision**

**Computing on  
heterogeneous platforms**  
*(presentation tomorrow)*

**Our #1 lesson learned from the computing domain: openness is a boost for innovation**  
*open data, open hardware, open algorithms*

Thanks to all our past and present students and researchers  
making this possible

# ETH Future Computing Laboratory

**ETH** zürich



**See you later  
at the demos!**

