ETH zürich

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

IEEE SENSORS 2025

A Parallel Ultra-Low Power Silent Speech Interface based on a Wearable, Fully-dry EMG Neckband

Integrated Systems Laboratory (ETH Zürich)

F. Meier, **G.Spacone**, S.Frey, L.Benini, A. Cossettini

gspacone@iis.ee.ethz.ch

@pulp_platform

pulp-platform.org

youtube.com/pulp_platform

PULP Platform

Open Source Hardware, the way it should be!

Introduction to Silent Speech Interfaces (SSIs)

*Novel form of assistive technology to detect **speech** with **vocalization***

Applications

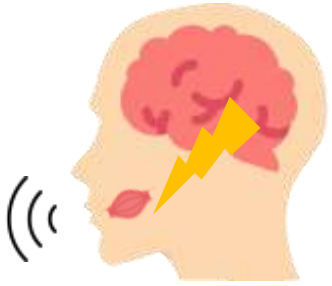
Human –machine interfaces

Speech Impairments

Noisy environments

Privacy

Sensing technology: surface electromyography (sEMG)



Monitoring of **electrical signals** transmitted to **muscles** involved in speech production

- Non-invasive
- Early-onset speech detection

[1, 2]

Previous Works and Objectives

Kapur et al.
Alter Ego [3]

- 92% (10 classes)

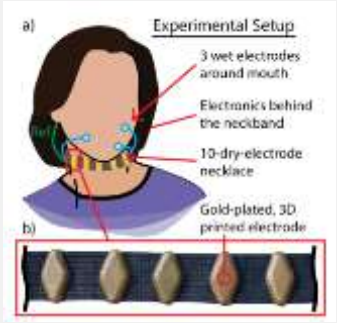
- ✗ Wet electrodes
- ✗ Obtrusive

Liu et al.
EMG Tattoo Patches [4]

- 92% accuracy (6 classes)
- Dry

- ✗ Patches on the face
- ✗ Disposable

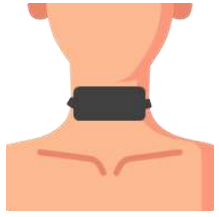
Wu et al.
Neckband form factor [5]



- 93% (11 class)
- Neck placement

- ✗ Not fully dry
- ✗ Lacks on-board processing capabilities

Goals



Wearable
Fully-dry
Minimally obtrusive
EMG-based SSI

Ultra-low power
→ multi-day use

Vocalized and Silent speech detection

Interface Design

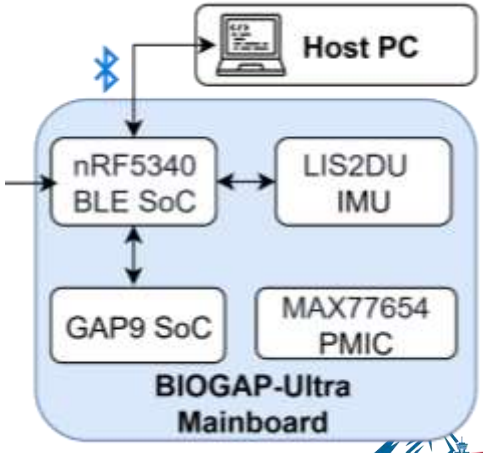
- Wearable **NeckBand**
  - 14 differential EMG channels
  - 4 bias electrodes
  - SoftPulse Dry Electrodes - Datwyler

- Acquisition Platform: **BIOGAP Ultra** [6]

Interface Design

- Wearable **NeckBand**
  - 14 differential EMG channels
  - 4 bias electrodes
  - SoftPulse Dry Electrodes - Datwyler

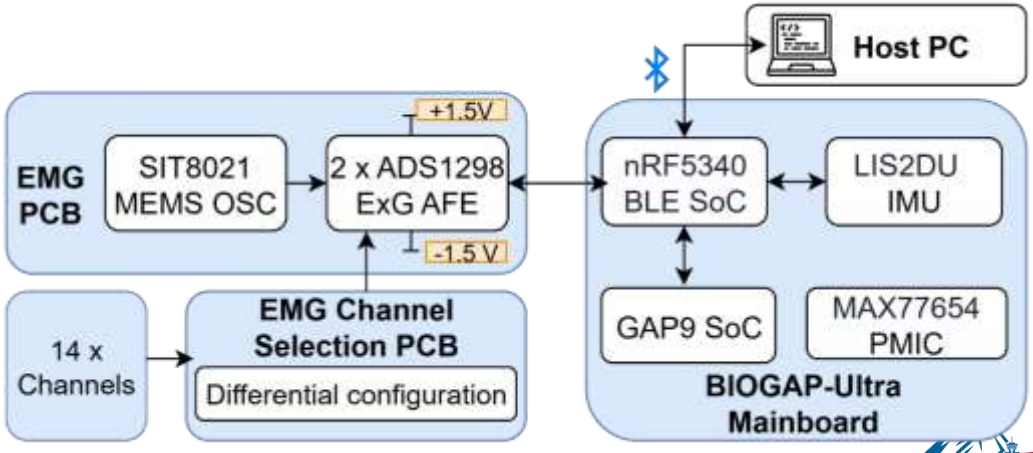
- Acquisition Platform: **BIOGAP Ultra** [6]
  - Main Board: NRF53 for BLE streaming



Interface Design

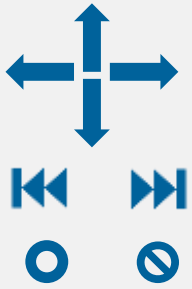
- Wearable **NeckBand**
  - 14 differential EMG channels
  - 4 bias electrodes
  - SoftPulse Dry Electrodes - Datwyler

- Acquisition Platform: **BIOGAP Ultra** [6]
  - Main Board: NRF53 for BLE streaming
  - EMG-PCB: Two ADS1298 for EMG recording
  - Ultra-low power consumption **(<22mW)**



Data Collection Protocol

8 HMI – related Commands

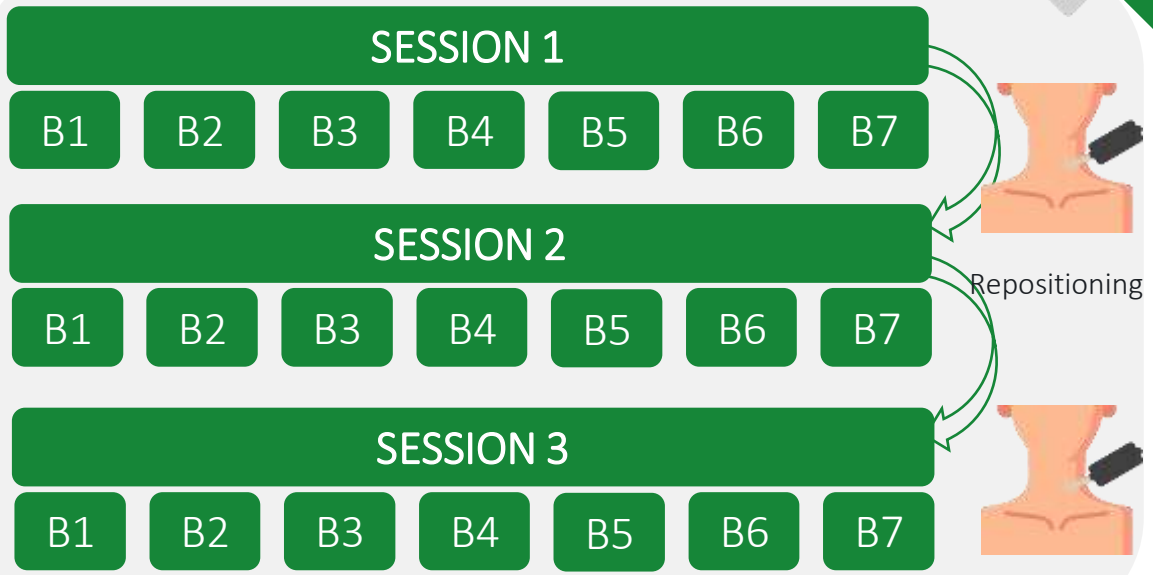


Up, down, left, right

Forward, backward

Start, stop

Acquisition Batches (B) 20 randomized repetitions per word → 160 utterances



Two Experiments

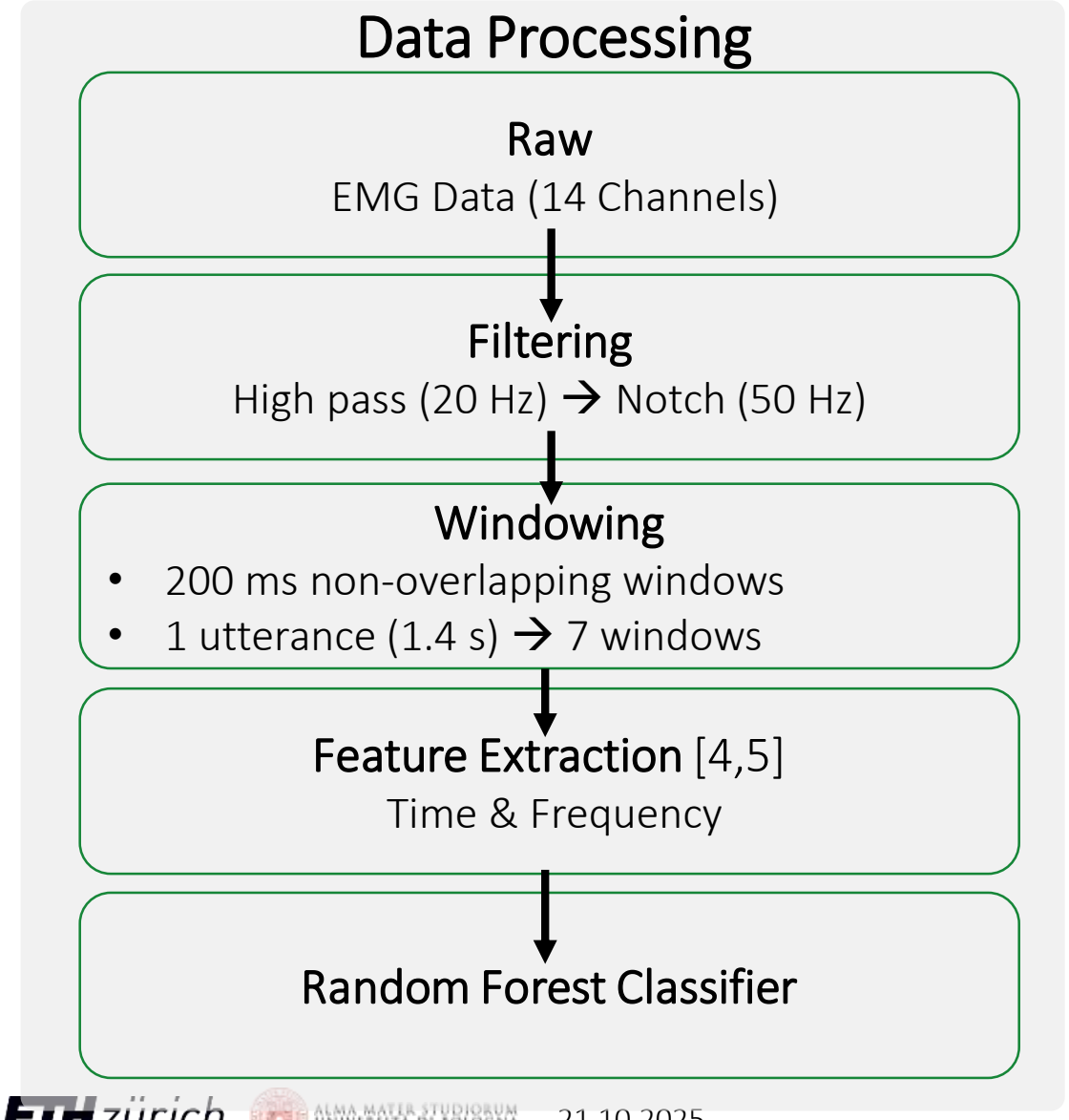
Vocalized

Silent

Dataset size

- ≈ 290 minutes – silent or vocalized
- 3360 utterances per experiment

Data Processing and Evaluation Method



Data Processing and Evaluation Method

Data Processing

Raw
EMG Data (14 Channels)

↓

Filtering
High pass (20 Hz) → Notch (50 Hz)

↓

Windowing
- 200 ms non-overlapping windows
- 1 utterance (1.4 s) → 7 windows

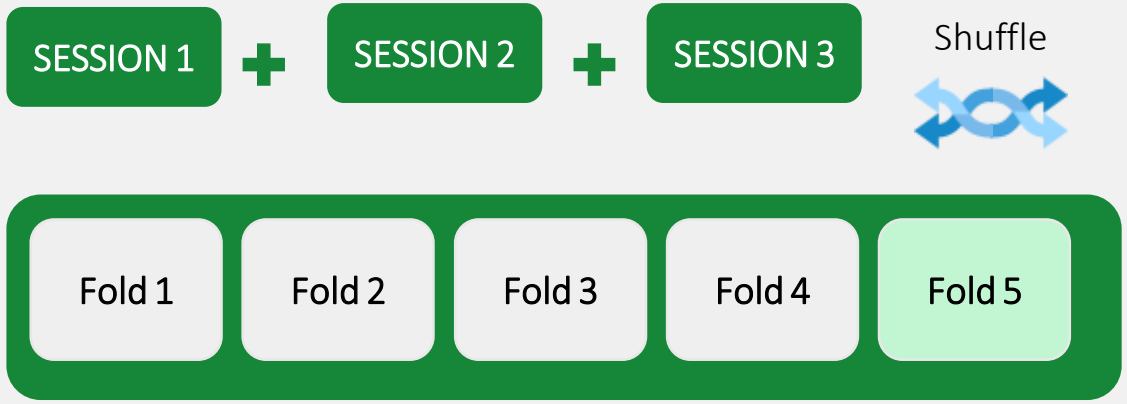
↓

Feature Extraction [4,5]
Time & Frequency

↓

Random Forest Classifier

Aggregated Analysis: 5-fold CV



Data Processing and Evaluation Method

Data Processing

Raw
EMG Data (14 Channels)

↓

Filtering
High pass (20 Hz) → Notch (50 Hz)

↓

Windowing
- 200 ms non-overlapping windows
- 1 utterance (1.4 s) → 7 windows

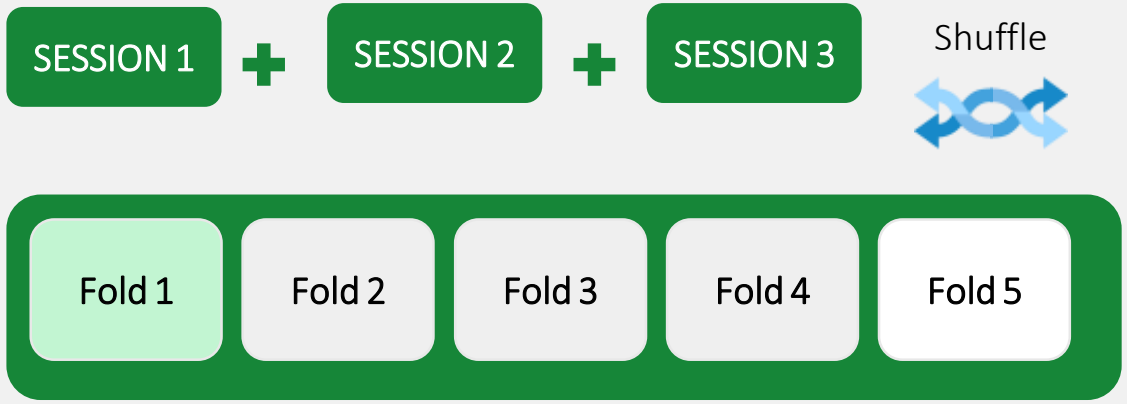
↓

Feature Extraction [4,5]
Time & Frequency

↓

Random Forest Classifier

Aggregated Analysis: 5-fold CV



Inter-Session Analysis

Data Processing and Evaluation Method

Data Processing

Raw
EMG Data (14 Channels)

↓

Filtering
High pass (20 Hz) → Notch (50 Hz)

↓

Windowing
- 200 ms non-overlapping windows
- 1 utterance (1.4 s) → 7 windows

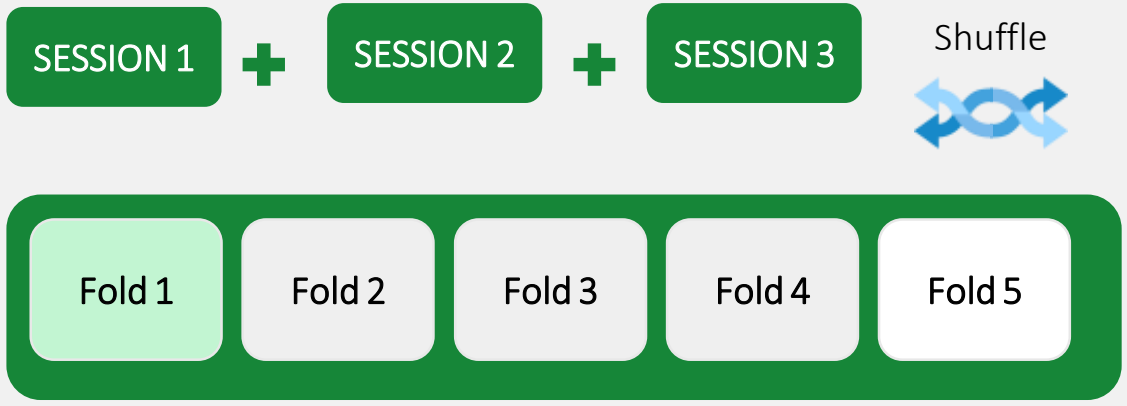
↓

Feature Extraction [4,5]
Time & Frequency

↓

Random Forest Classifier

Aggregated Analysis: 5-fold CV



Inter-Session Analysis

Data Processing and Evaluation Method

Data Processing

Raw
EMG Data (14 Channels)

↓

Filtering
High pass (20 Hz) → Notch (50 Hz)

↓

Windowing
- 200 ms non-overlapping windows
- 1 utterance (1.4 s) → 7 windows

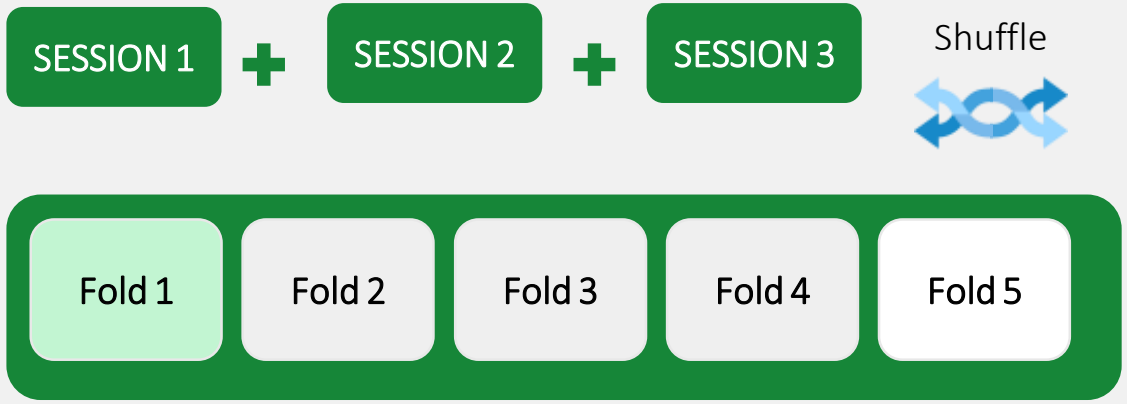
↓

Feature Extraction [4,5]
Time & Frequency

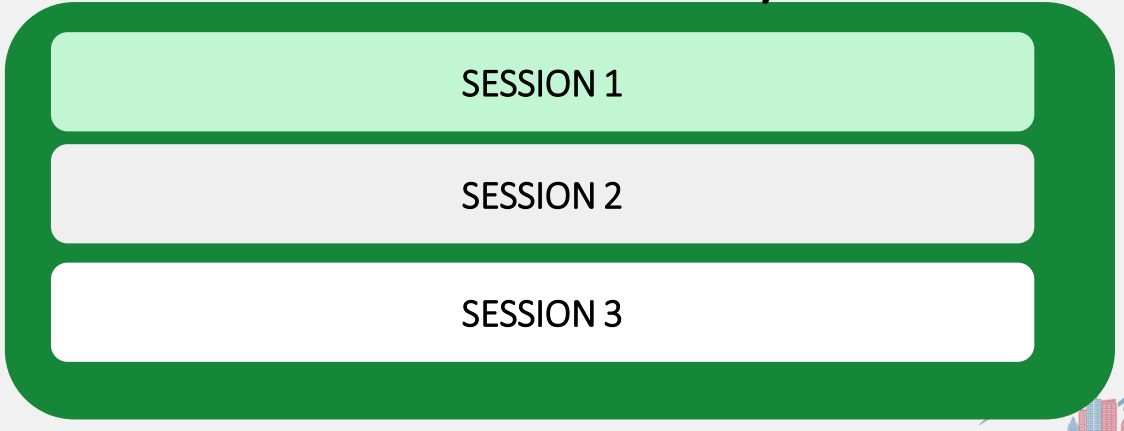
↓

Random Forest Classifier

Aggregated Analysis: 5-fold CV



Inter-Session Analysis



Data Processing and Evaluation Method

Data Processing

Raw
EMG Data (14 Channels)

↓

Filtering
High pass (20 Hz) → Notch (50 Hz)

↓

Windowing
- 200 ms non-overlapping windows
- 1 utterance (1.4 s) → 7 windows

↓

Feature Extraction [4,5]
Time & Frequency

↓

Random Forest Classifier

Aggregated Analysis: 5-fold CV

SESSION 1 + SESSION 2 + SESSION 3 — Shuffle

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5

Inter-Session Analysis

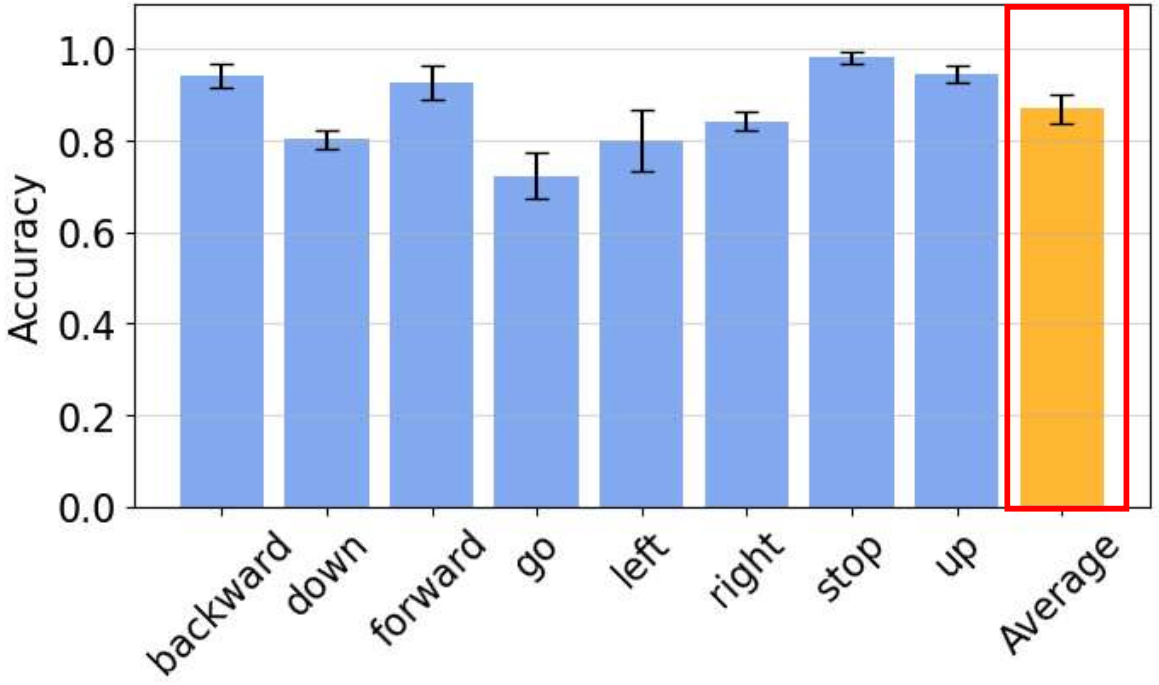
SESSION 1

SESSION 2

SESSION 3

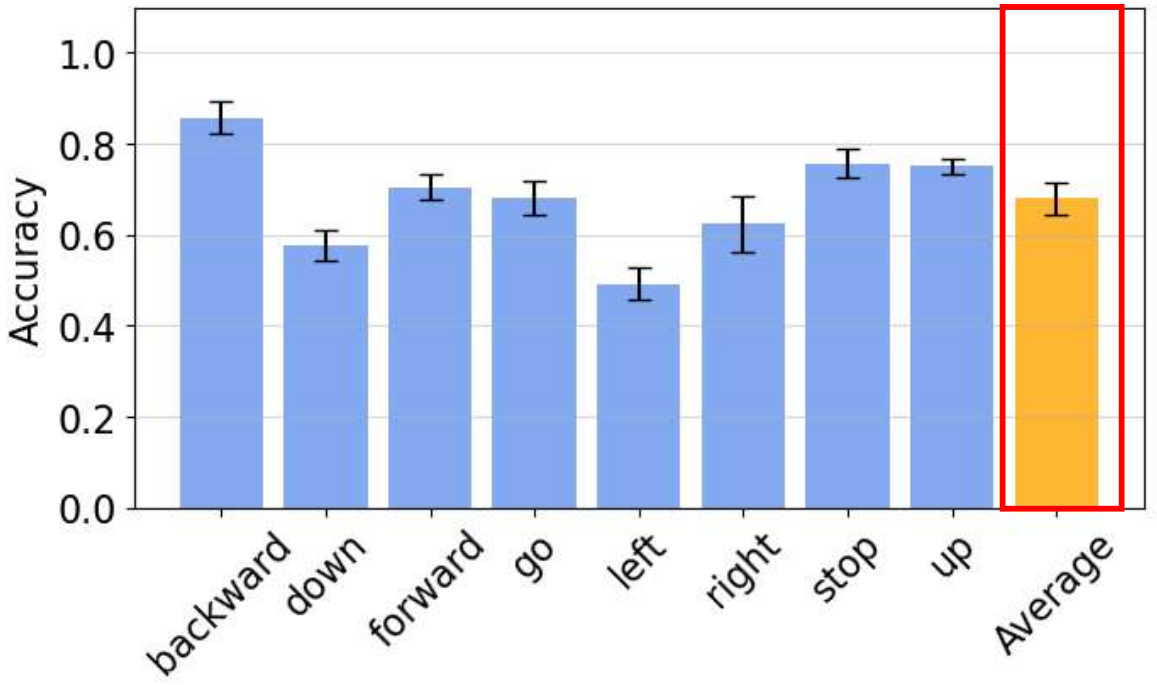
Results: Aggregated Analysis

Vocalized EMG Predictions



- Average Accuracy: **87 ± 3 %**
- Liu et al. [4]: 89.6%, non-wearable system
- Wu et al .[5]: 93%, not fully dry

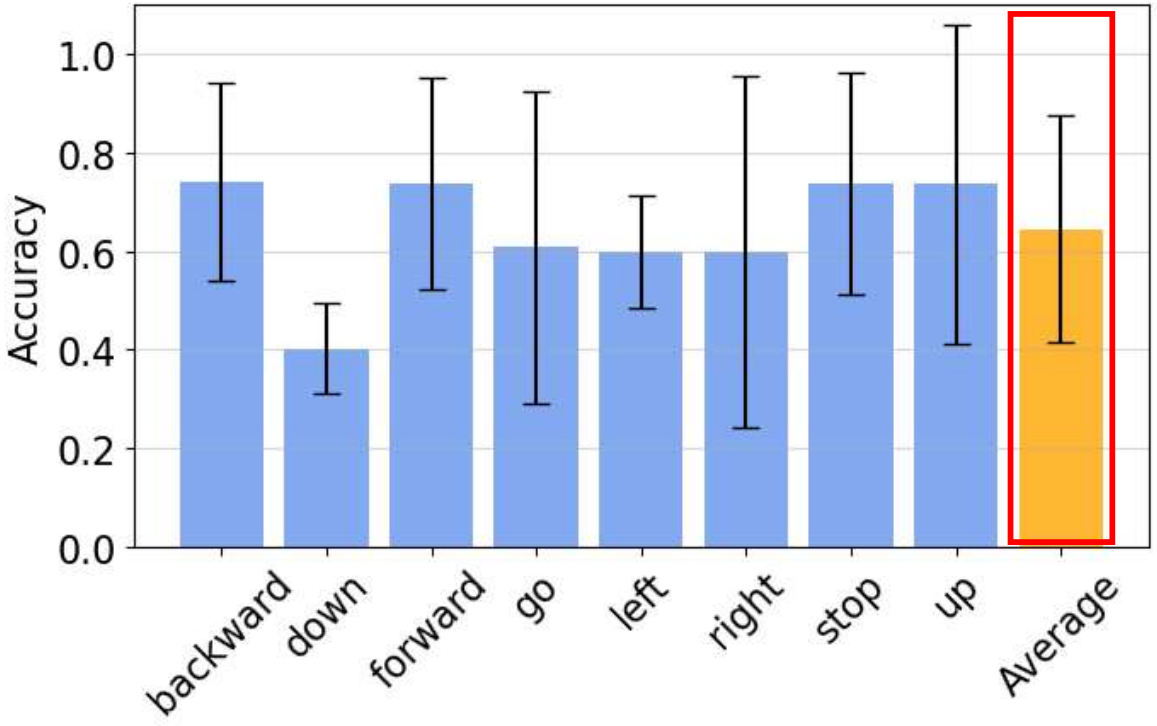
Silent EMG Predictions



- Average Accuracy: **68 ± 3 %**
- Weaker signals
- Practice Required

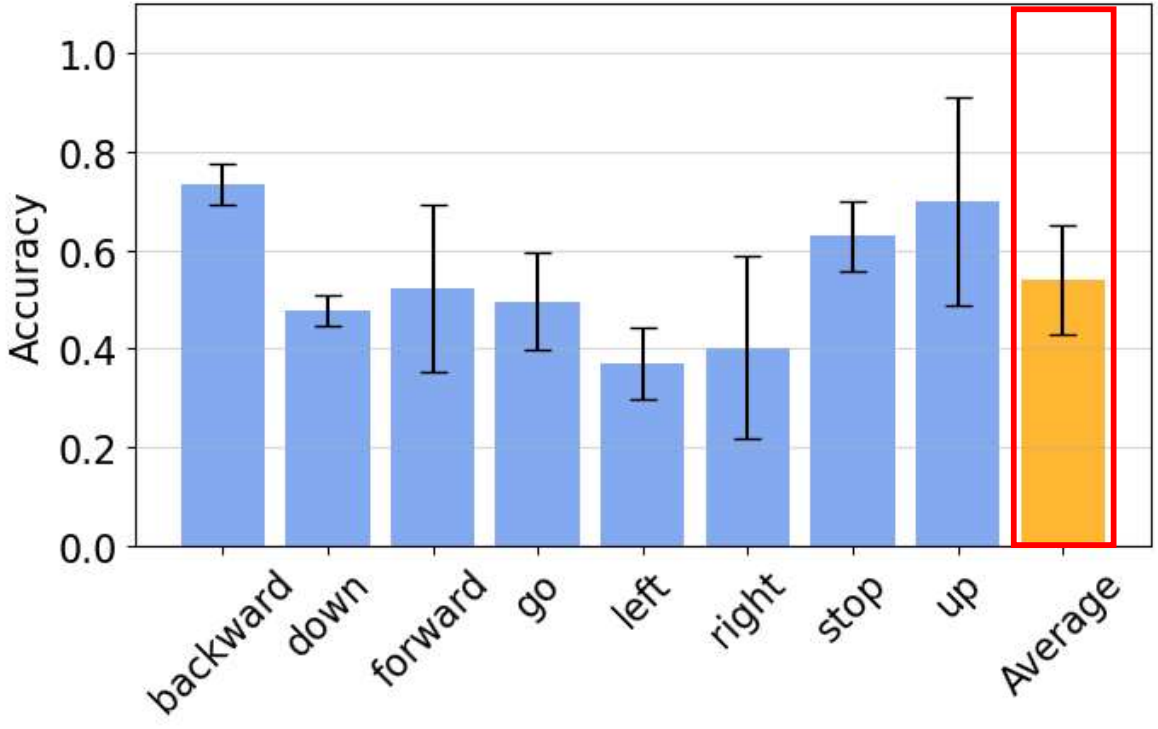
Results: Inter-Session Analysis

Vocalized EMG Predictions



- Average Accuracy: **64 ± 18 %**
- High Standard Deviation highlights challenges introduced by repositioning

Silent EMG Predictions



- Average Accuracy: **54 ± 7 %**
- Lower Standard Deviation compared to Vocalized

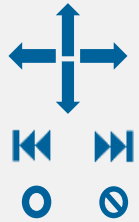
Conclusion and Future Work

Wearable Neckband for EMG-Speech Detection

- BIOGAP Ultra
- 14 dry EMG channels
- **22 mW** power consumption (data collection & BLE streaming)

Validation on Vocalized & Silent Speech

- **83%** on **8 classes Vocalized EMG**
  - Comparable to SoA [4,5]
  - Fully wearable!



- Further exploration with **sensors repositioning**

Future Work

- Increasing channel count
- Miniaturized electrodes

- Complementary sensing modalities:
  - Ultrasound
  - IMU

- Edge processing:
  - RISCV PULP architecture
  - ML accelerators

Extensive Data Collection

ETH zürich

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

IEEE SENSORS 2025
Vancouver, Canada || October 19 - 22, 2025

Thanks for Your Attention!

Integrated Systems Laboratory (ETH Zürich)

F. Meier, **G.Spacone**, S.Frey, L.Benini, A. Cossettini

gspacone@iis.ee.ethz.ch

PULP Platform

Open Source Hardware, the way it should be!

BioGAP Ultra

Open-Source

repo

@pulp_platform

pulp-platform.org

youtube.com/pulp_platform

References

[1] . A. Gonzalez-Lopez, A. Gomez-Alanis, J. M. M. Donas, J. L. Perez-Cordoba, and A. M. Gomez, “Silent speech interfaces for speech restoration: A review,” IEEE access, vol. 8, pp. 177 995–178 021, 2020.

[2] R. Netsell and B. Daniel, “Neural and mechanical response time for speech production,” Journal of Speech and Hearing Research, vol. 17,no. 4, pp. 608–618, 1974.

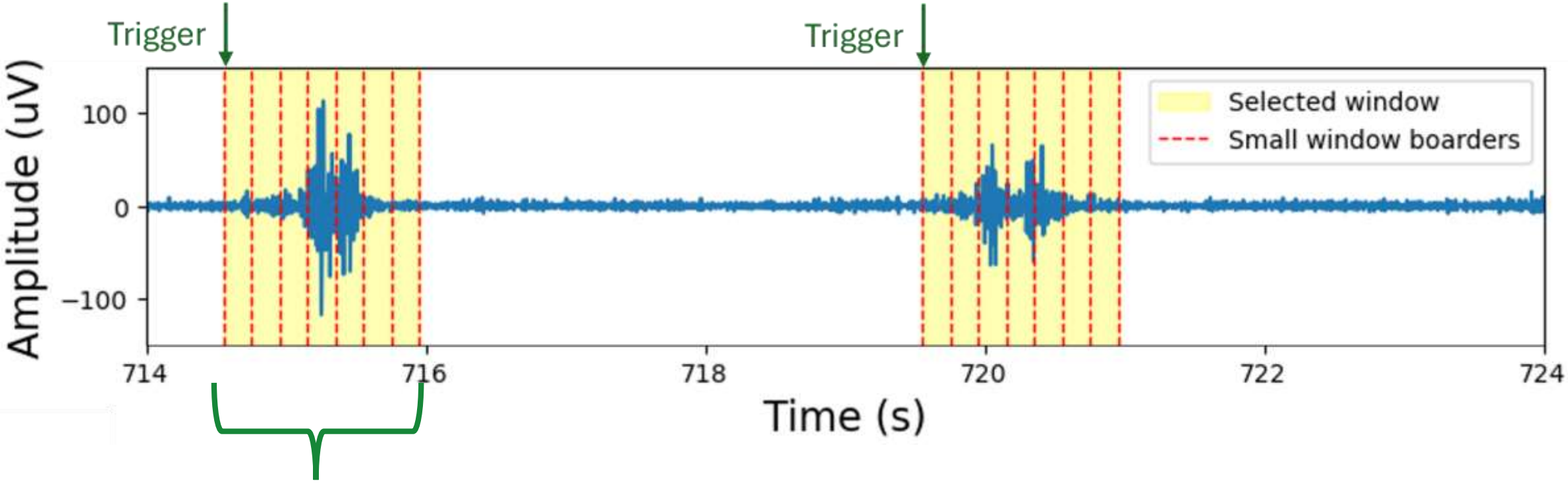
[3] A. Kapur, S. Kapur, and P. Maes, “Alterego: A personalized wearable silent speech interface,” in Proceedings of the 23rd International Conference on Intelligent User Interfaces, 2018, pp. 43–53

[4] H. Liu, W. Dong, Y. Li, F. Li, J. Geng, M. Zhu, T. Chen, H. Zhang,L. Sun, and C. Lee, “An epidermal semg tattoo-like patch as a new human–machine interface for patients with loss of voice,” Microsystems& nanoengineering, vol. 6, no. 1, p. 16, 2020.

[5] C. Tang, J. Mallah, D. Kazieczko, W. Yi, T. R. Kandukuri, E. Occhipinti,B. Mishra, S. Mehta, and L. G. Occhipinti, “Wireless silent speech interface using multi-channel textile emg sensors integrated into headphones,” 2025. [Online]. Available: https://arxiv.org/abs/2504.13921

[6] S. Frey, G. Spacone, A. Cossettini, M. Guermandi, P. Schilk, L. Benini, and V. Kartsch, “Biogap-ultra: A modular edge-ai platform for wearable multimodal biosignal acquisition and processing,” 2025.[Online]. Available: https://arxiv.org/abs/2508.13728

Labelling and Windowing



1.4 s → divided into 7 windows of 200 ms