IEEE BioCAS 2024

VowelNet: Enhancing Communication with Wearable EEG-Based Vowel Imagery

Integrated Systems Laboratory (ETH Zürich)

Thorir Mar Ingolfsson, Victor Javier Kartsch Morinigo, Xiaying Wang, Andrea Cossettini, Luca Benini

thoriri@iis.ee.ethz.ch

PULP Platform

Open Source Hardware, the way it should be!

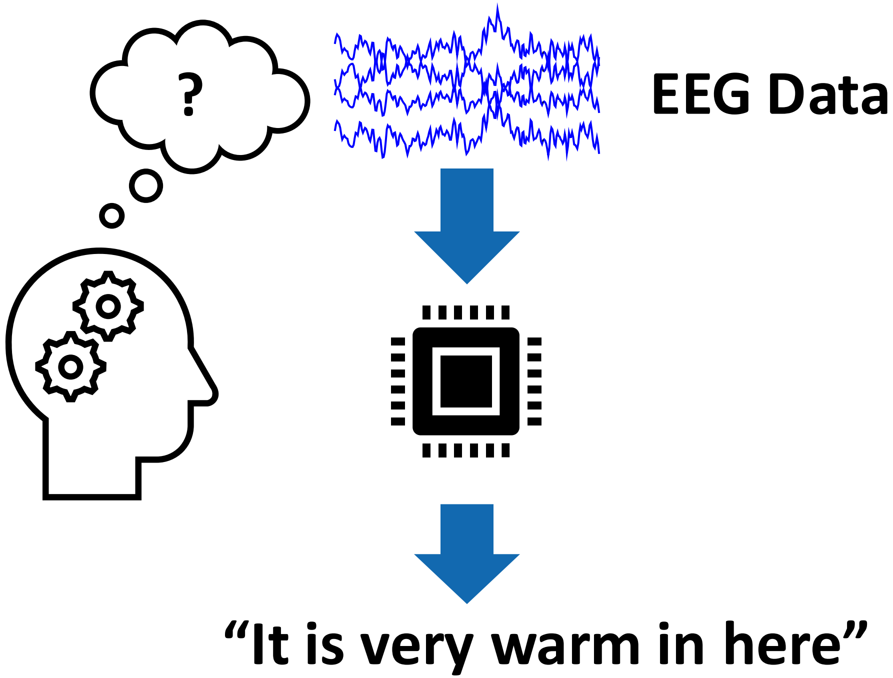
@pulp_platform

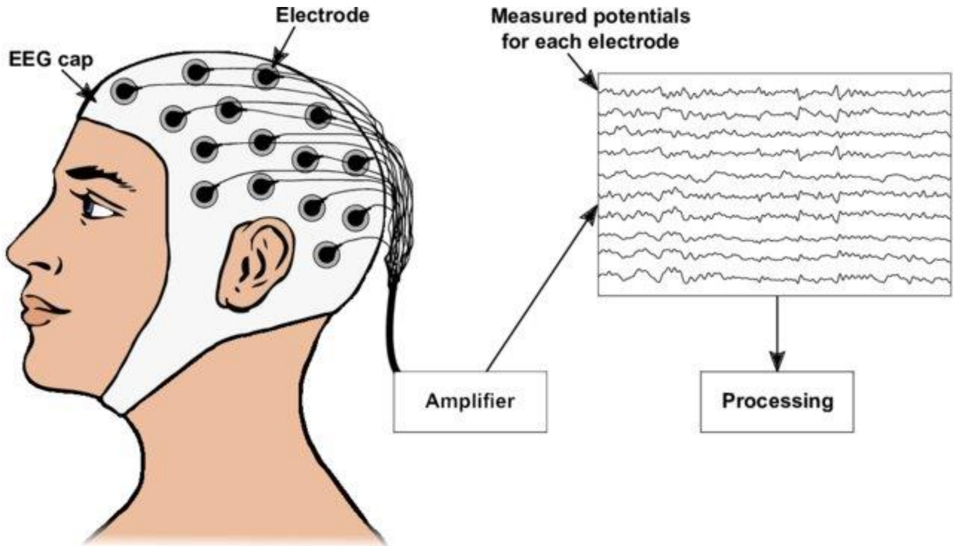
pulp-platform.org

youtube.com/pulp_platform

We are controlled by electrical activity

- **Everything we do/think is electrical activity**
  - Brain signals decode complex thoughts





Source: Nagel, Sebastian. Towards a home-use BCI: fast asynchronous control and robust non-control state detection.

- **Electroencephalography (EEG) method of monitoring the electrical activity**
  - Signals typically range from 10 µV to 100 µV ← low signal to noise ratio

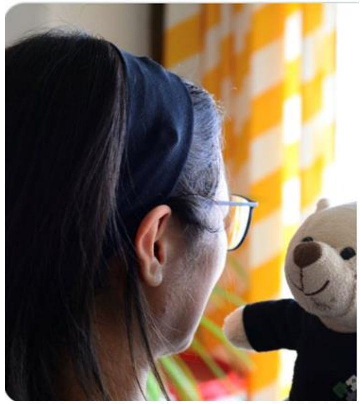
Wearable EEG Devices come at a price


Source: https://www.flickr com/photos/tim_ uk/8135749317

Conventional EEG Caps

- Stigmatizing
- Power hungry
- Not intended for normal day use

Wearable solutions



Dry 8 channel EEG System

- Non-stigmatizing
- Low-power
- Data processed locally
  - Privacy
  - Latency

Data acquisition more **susceptible** to **artifacts**

Models need to be **robust** and **accurate** to compensate

Not full head coverage

Requirements for a Successful Wearable Device

What we target!

Safe

Privacy preserving

Accurate

Long battery life

Real-time response

Two goals:

Goal 1. How do we **maximize accuracy** of speech classification for wearable devices.

Goal 2. Lightweight **end-to-end** approach with operating on **raw data for real-time response!**

Speech Decoding from Wearables → Vowel Decoding

First step to **full speech decoding** from scalp EEG is to do →

Vowel Classification (/a/, /e/, /i/, /o/, and /u/)

How we think of vowels is **not universally the same**

/a/ → [a]   /i/ → [ɪ]   /e/ → [ɛ]

/a/ → [æ]/[a]   /i/ → [iː]/[ɪ]

/e/ → [ɤ]/[ə]

How do **you think** of the different vowels?

(Is it always the same?)

Fully-Dry 8 Channel Wearable EEG Vowel Dataset

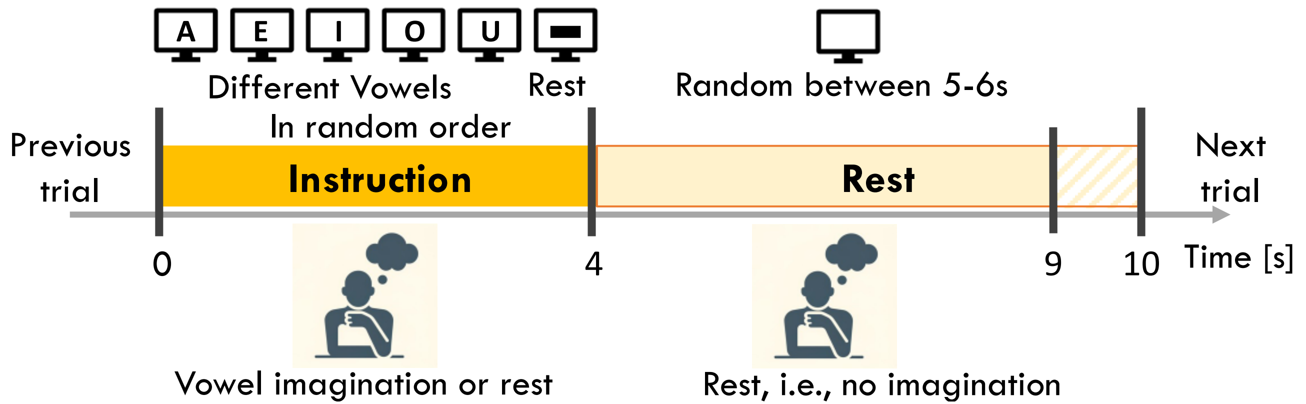
5 Subjects (Different origins)

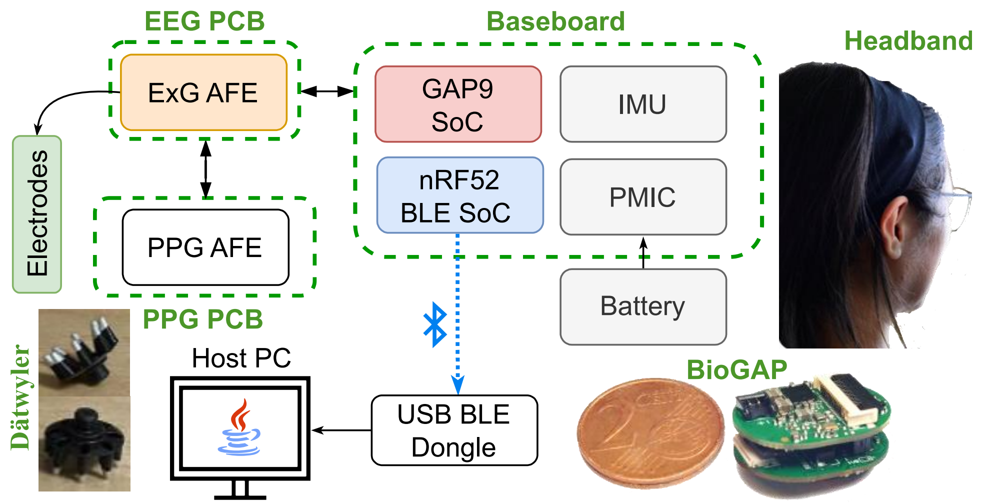
Five sessions – 30 trials each

Think about saying the vowel shown on screen

Data gathered at 500 Hz

BioGAP with a **tailored headband** and **8 channels** of dry electrodes.





Three main different ways of evaluating the vowel classification task

We consider three different training methodology and evaluation:

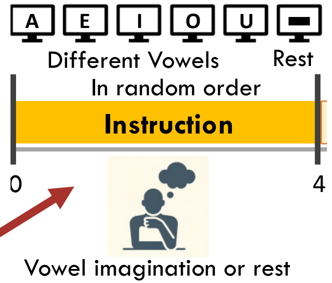
1. **Vowel vs Rest** (/a/ - rest, /e/-rest, /i/-rest, /o/-rest, /u/-rest)
And all vowels grouped together vs rest.

2. **Inter-vowels** (/a/-/e/, /a/-/i/ etc.) (All pairs)

3. **Six class classification** (/a/, /e/, /i/, /o/, /u/, rest)

Goal 1

Trimming vs **No Trimming**



EEG trimmed from front or back

VowelNet: A Highly Parallelizable and MCU-Deployable Neural Network Architecture

Focus on temporal filters

We present the novel *VowelNet*:

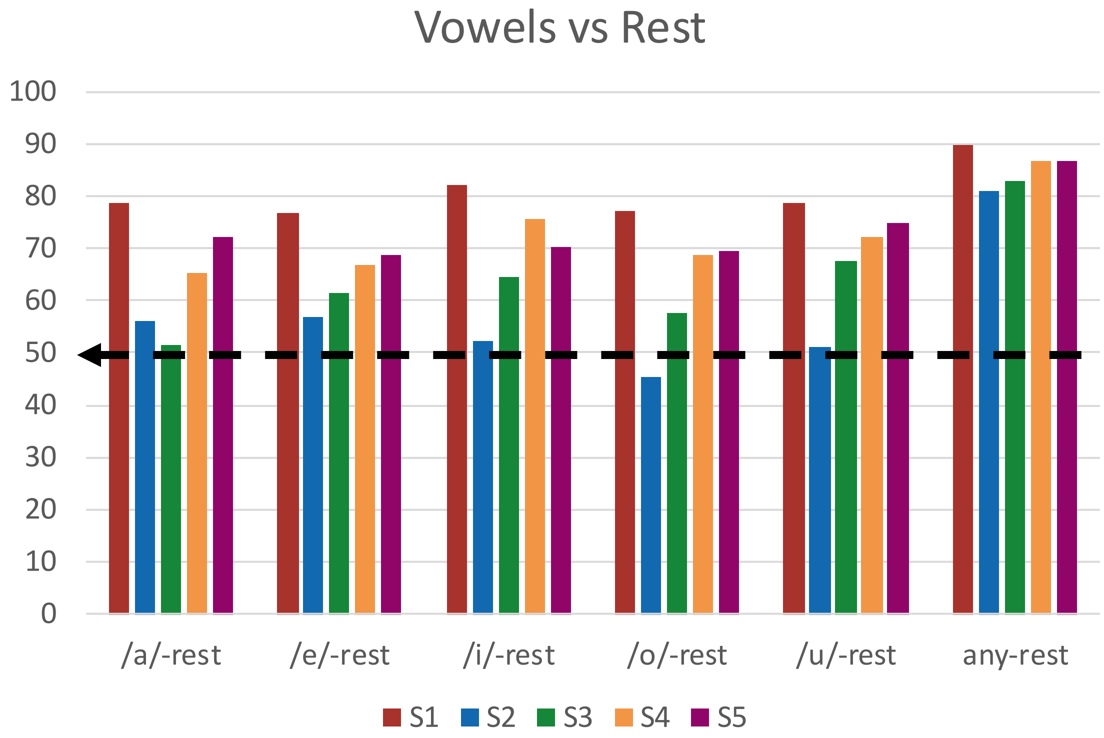
Block	Filters	Kernel	Output
Temporal Conv + BN	32	(1,64)	(32,C,T)
Depthwise Conv + Act	64	(C, 1)	(64,1,T)
Pooling & Dropout	-	(1,4)	(64,1,T//4)
Separable Conv	64	(1,16)	(64,1,T//4)
Conv + BN + Act	64	(1,1)	(64,1,T//4)
Pooling & Dropout	-	(1,8)	(64,1,T//32)
Dense	-	-	N

Goal 2

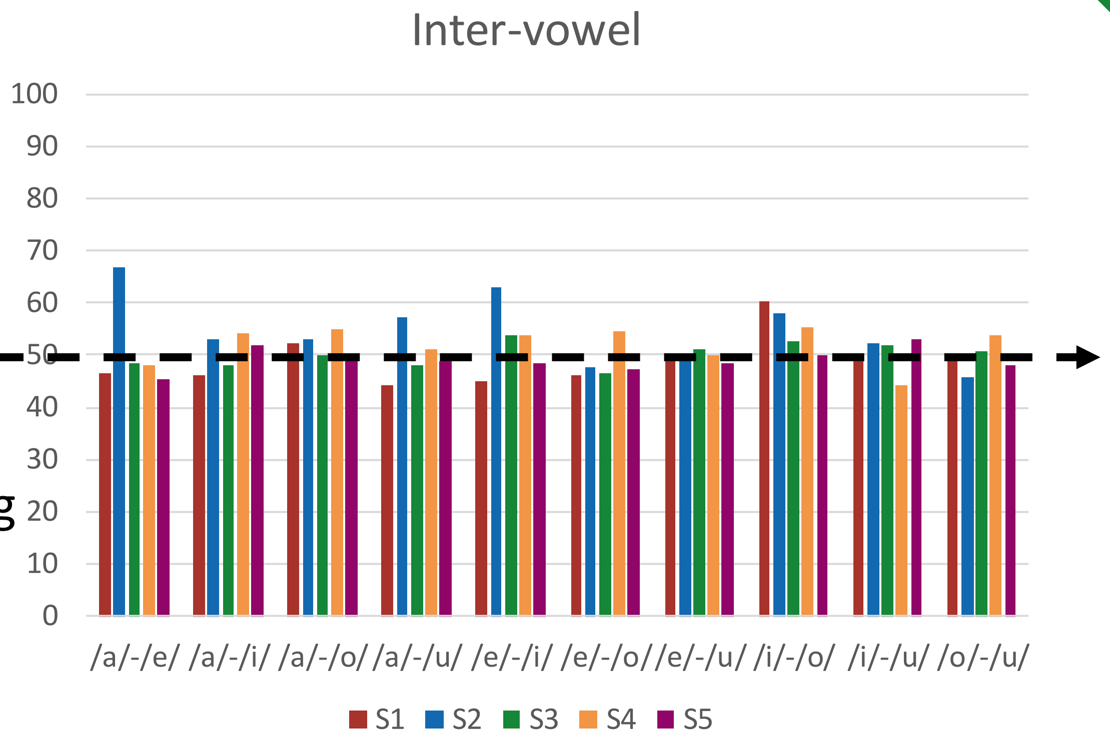
***VowelNet* very efficient due to usage of parallelizable operations (Conv + Pool)**

***VowelNet* only utilizes 18k weights → fits on resource constrained devices**

Rest seems to be more easily distinguishable



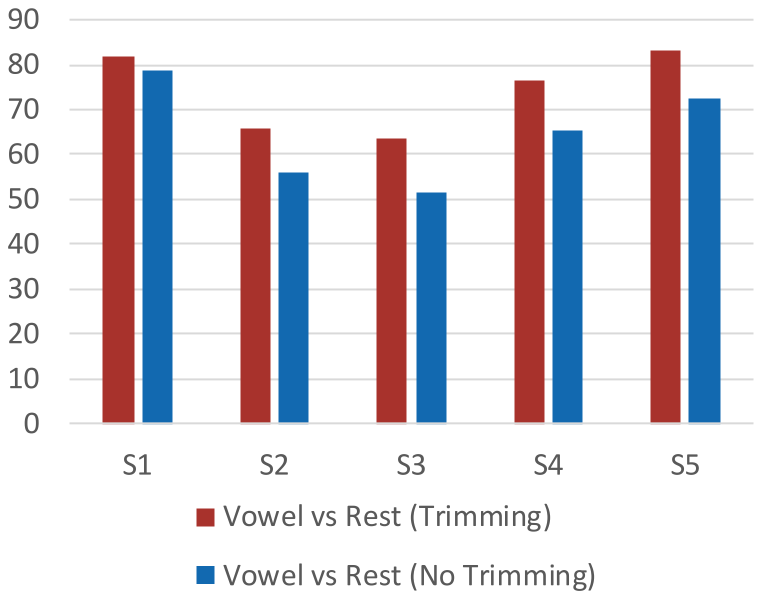
Random Guessing

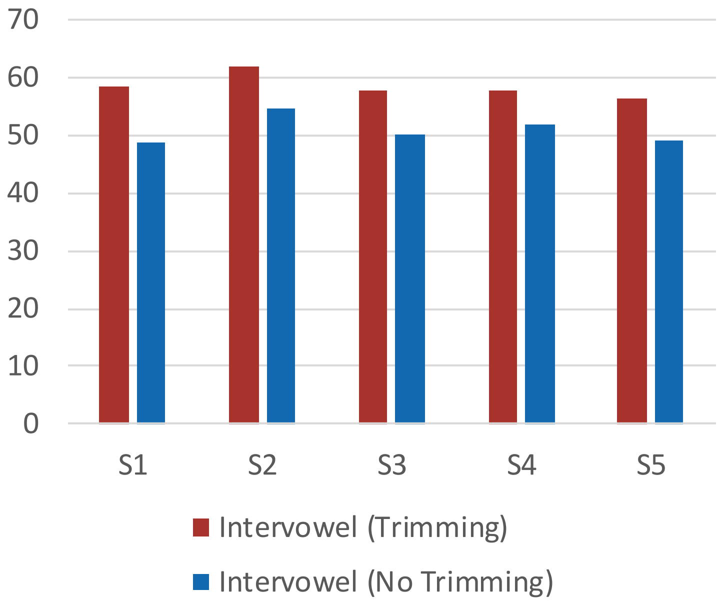


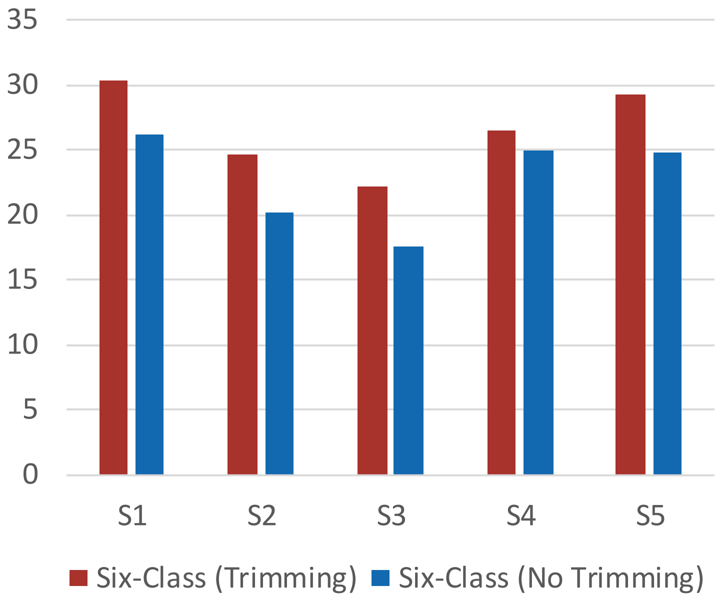
	S1	S2	S3	S4	S5
Six-class	**26.1%**	**20.1%**	**17.5%**	**24.9%**	**24.8%**

Random Guessing : 16.67%

To trim or not to trim



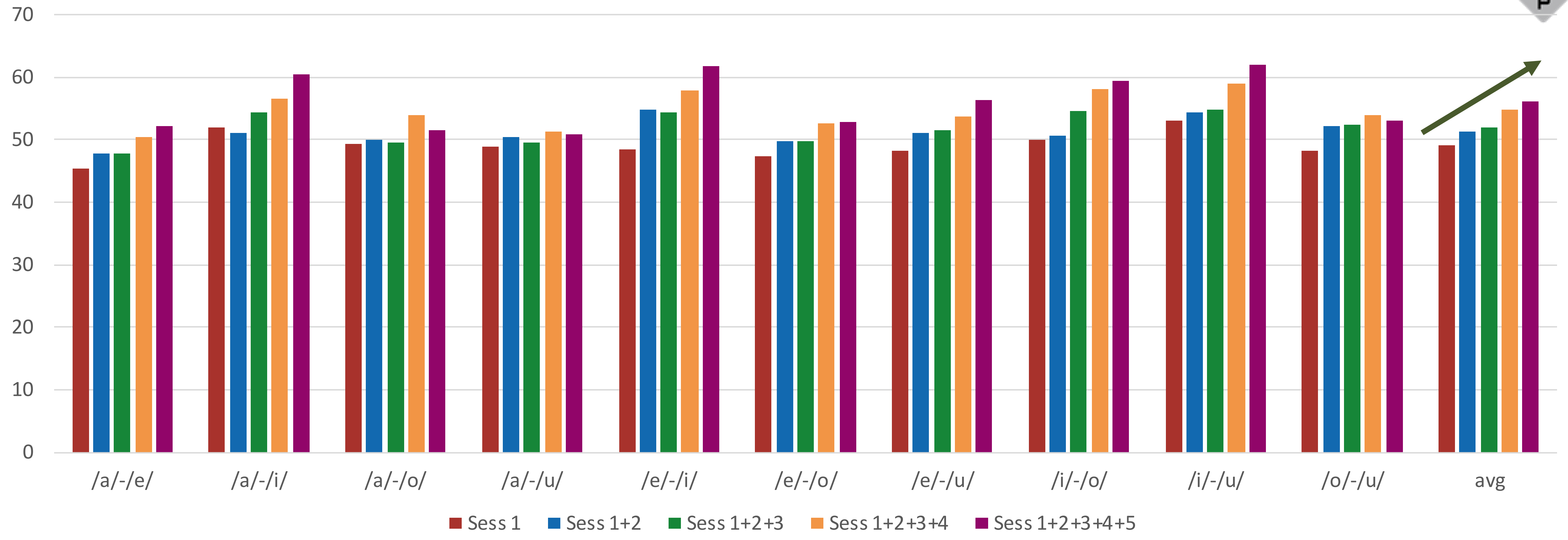




- **Vowel vs Rest**: 3.30% improvement
- **Inter-vowel**:19.22% improvement
- **Six-Class**: 16.09% improvement

Cutting around 3s from end of signal (back of signal)

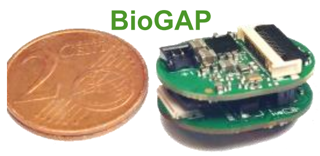
More data better results



Adding sessions on average showcases **14.2% improvement**

Embedded Deployment of VowelNet

GAP9 ← MCU on **BioGAP**



Quantlab → **DORY**

Quantized to 8bit **Deployment**

- 9 Core RISC-V Cluster
- 240 MHz
- Memory:
  - L1: 128kB
  - RAM: 1.5 MB
  - Non Volatile: 2 MB

Energy efficiency	**45.39 GMAC/s/W**
Inference:	**40.9 ms**
Energy/Inference:	**0.71 mJ**
Average Power:	**17.43 mW**
Throughput:	**791.38 MMAC/s**

Conclusions

- **91% accuracy in vowel vs rest classifcation**
- **Up to 68% accuracy in inter-vowel classification**

Energy efficiency	**45.39 GMAC/s/W**
Inference:	**40.9 ms**
Energy/Inference:	**0.71 mJ**
Average Power:	**17.43 mW**
Throughput:	**791.38 MMAC/s**

150mAh battery

BioGAP (GAP9) Implementation → 25.6 hours

Contact info