Train-On-Request: an On-Device Continual Learning Workflow for Adaptive Real-World Brain-Machine Interfaces

Lan Mei[1], Cristian Cioflan[1], Thorir Mar Ingolfsson[1], Victor Kartsch[1], Andrea Cossettini[1], **Xiaying Wang**[1], Luca Benini[1,2]

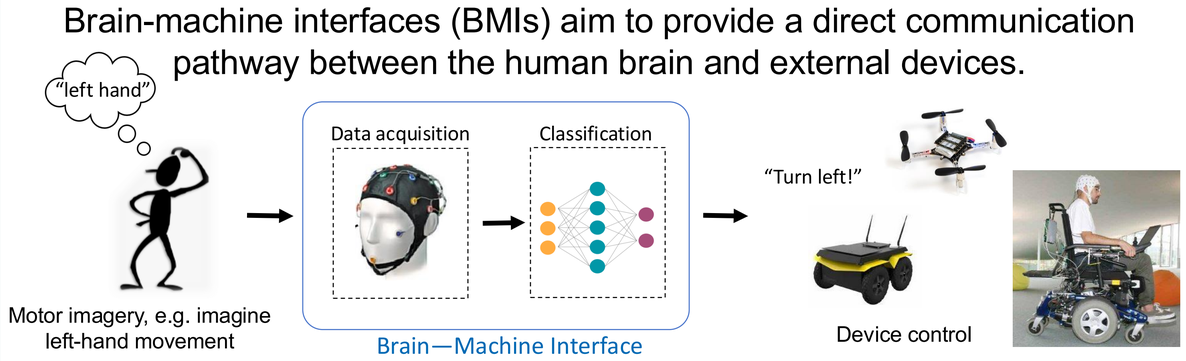
[1]Dept. ITET, ETH Zürich, Switzerland [2]DEI, University of Bologna, Italy

IEEE BioCAS 2024 Xi'an
Biomedical Circuits and Systems Conference
October 24-26, 2024 | Xi'an China

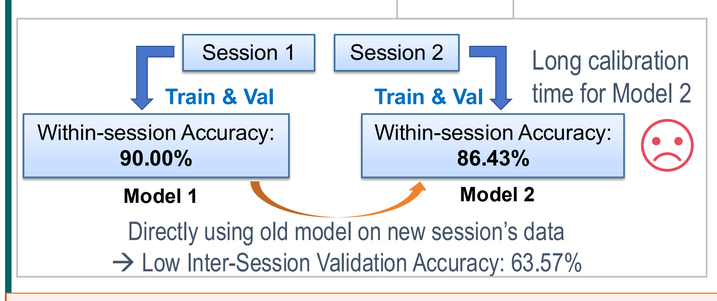
Poster No. B2P-C01

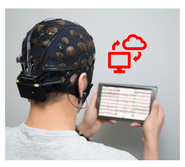
Motivation

Brain-machine interfaces (BMIs) aim to provide a direct communication pathway between the human brain and external devices.



Requirements for Wearable BMIs: Safe | Accurate | Comfortable | Privacy preserving | Long battery life | Real-time response





Stigmatizing | Bandwidth | Power hungry | Privacy | Short lifetime | Latency

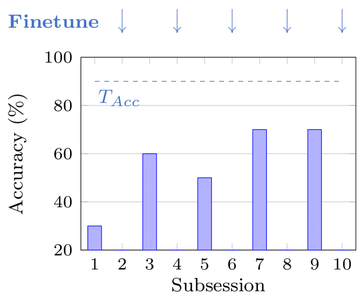
New Trend: BMI with Edge Computing!

Continual On-Device Learning (ODL) with short calibration time is needed!

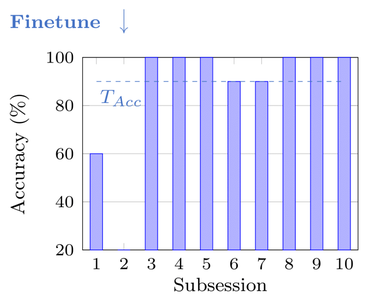
Results

Unsuccessful / Successful TOR Session

More successful sessions observed in last 3 sessions



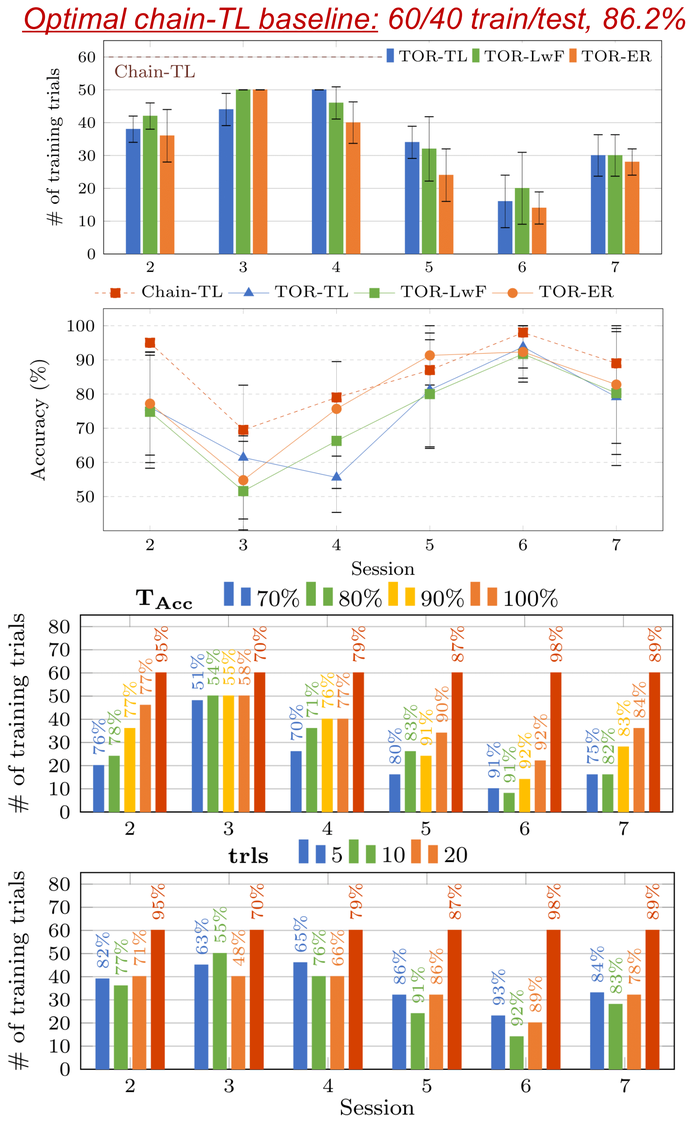
Acc.: 56%
Acq. time: 8.3min
Adapt. time: 16.2s



Acc.: 93.3%
Acq. time: 1.6min
Adapt. time: 3.2s

TOR Performance over Multiple Sessions

Optimal chain-TL baseline: 60/40 train/test, 86.2% acc., ~10min acq. time per session



#Train. Trials & Acq. Time:

- **TOR methods** -> fewer #training trials, less acquisition time
- **TOR-ER** -> **best among all**, #trials **46%** fewer than chain-TL, acq. time **5.3min** per session

Accuracy:

- **TOR-ER** -> **best among TOR**
- **TOR** accuracy **increases over time**, **TOR-ER comparable with chain-TL**, especially for the last 3 sessions (**88.8%**)
- 7.6 bit/min ITR for **TOR-ER**

Accuracy threshold T_{Acc}:

- T_{Acc} increases -> Required **#training data** increases, **accuracy** increases in most sessions

#Trials per subsession **trls**:

- **trls = 10** -> an **optimal** accuracy-training time trade-off, better than **trls = 5** and **trls = 20**

On-device Measurements

800mV, 370MHz on GAP9

~1mJ, 21.6ms per training step

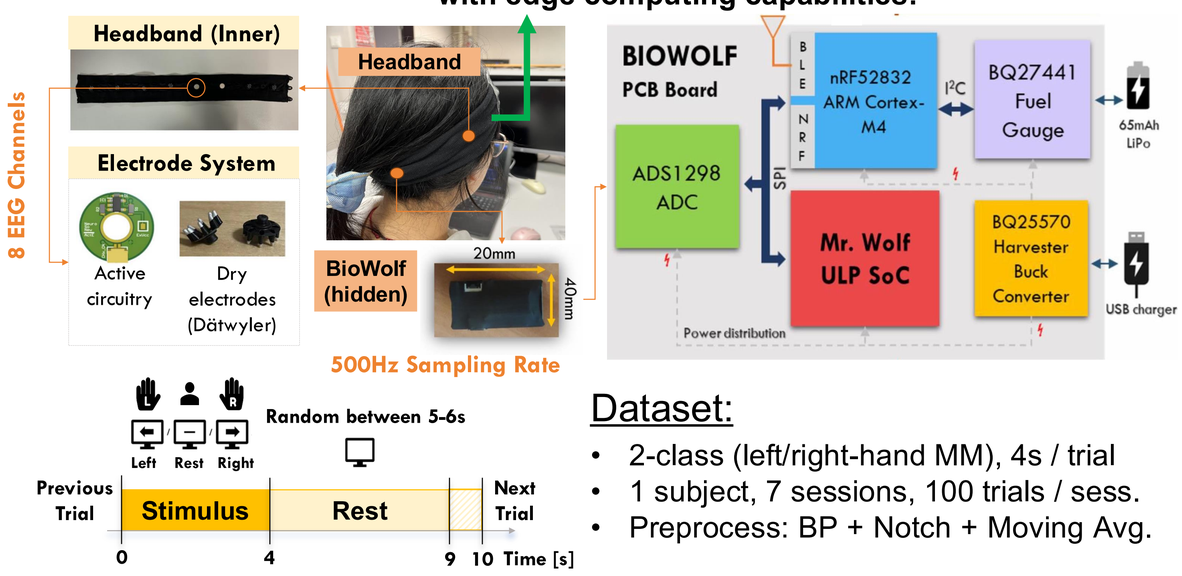
Adaptation time: 3.2s for 15 epochs << 1.6min acquisition time per subsession

On-site adaptation for an improved user experience!

Methods

Dataset Acquisition

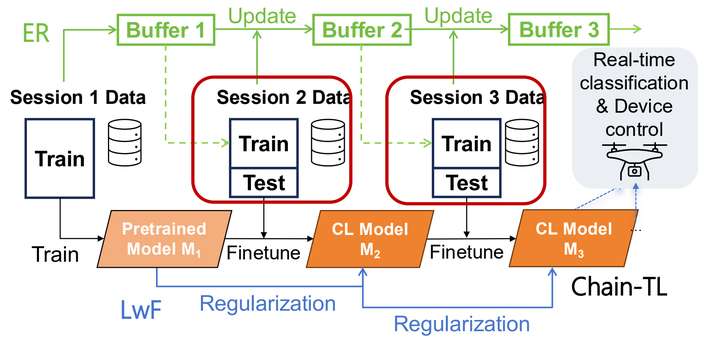
Non-stigmatizing and comfortable BMI with edge computing capabilities!



Dataset:

- 2-class (left/right-hand MM), 4s / trial
- 1 subject, 7 sessions, 100 trials / sess.
- Preprocess: BP + Notch + Moving Avg.

Learning Framework and TOR Workflow

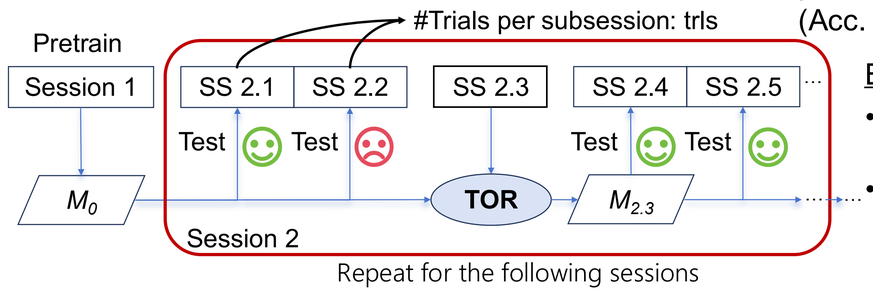


- Baseline: chain-TL
- CL mitigates catastrophic forgetting: Experience Replay (ER), Learning w/o Forgetting (LwF)

Train-On-Request (TOR)

Optimization: Only train when needed

- Finetune the model when performance is unsatisfactory (Acc. ≤ T_{Acc})

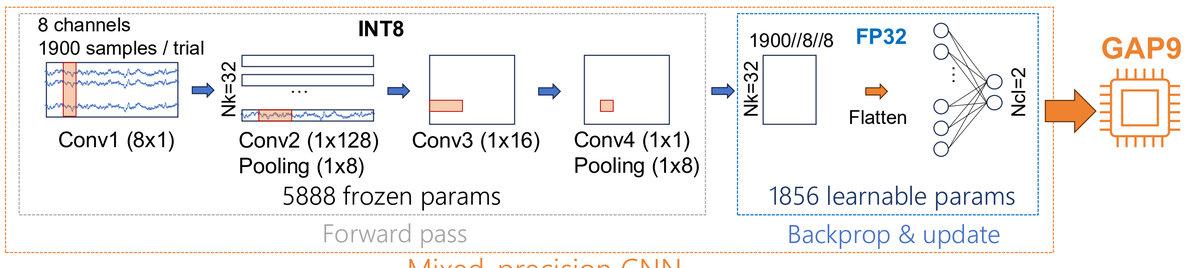


Benefits:

- Near real-time, online adaptation
- Combined with TL and CL (TOR-TL, TOR-ER, TOR-LwF)

Classification Model & On-device Learning

MI-BMINet, SoA lightweight CNN



Conclusion

- **Novel TOR ODL workflow** introduced for EEG-based BMI systems.
- Tackling **inter-session variability** through **near real-time data acquisition** and **online finetuning** for out-of-distribution sessions.
- Achieved a **46% reduction in calibration time**, lowering it to **1.6 minutes**, with **up to 92% accuracy.**
- Deployed on GAP9 MCU with **50 mW** power, **1 mJ** per learning step, and **21.6 ms** runtime.
- A robust, user-friendly, and energy-efficient solution to increase user acceptance of BMIs in non-clinical settings.

Acknowledgment

This work was supported by the Swiss National Science Foundation under grant n. 193813 (PEDESITE) and grant n. 207913 (Tiny-Trainer), by the ETH-Domain Joint Initiative program (project UrbanTwin), and by the ETH Future Computing Laboratory (EFCL).