Self-supervised representation learning from electroencephalography signals

Hubert Banville[1,2], Isabela Albuquerque[3], Aapo Hyvärinen[1,4],
Graeme Moffat[2], Denis-Alexander Engemann[1], Alexandre Gramfort[1]

[1] Inria, Université Paris-Saclay, Paris, France
[2] InteraXon Inc., Toronto, Canada
[3] INRS-EMT, Université du Québec, Montréal, Canada
[4] Dept. of CS and HIIT, University of Helsinki, Finland

Main result: Self-supervision can be used to learn physiologically relevant features from unlabelled EEG time series and improve classification performance on sleep staging.

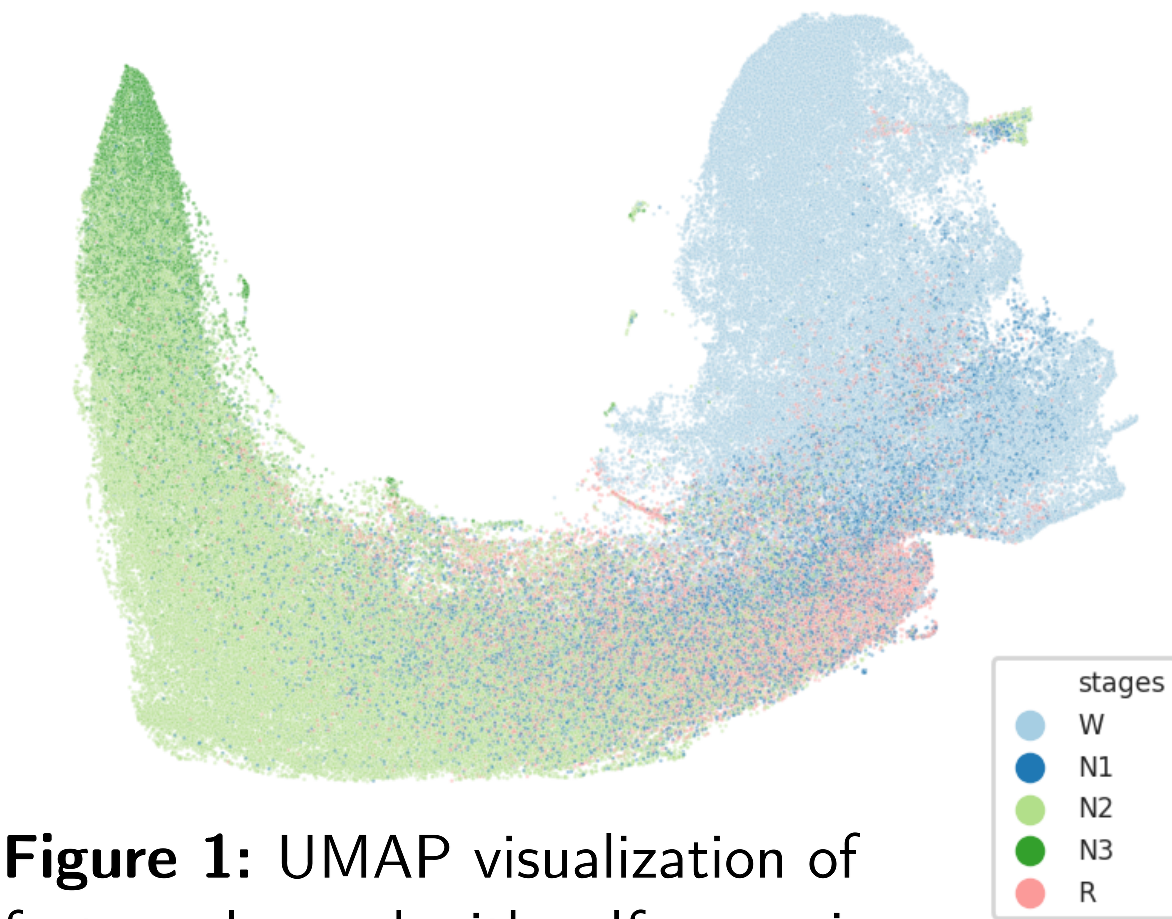


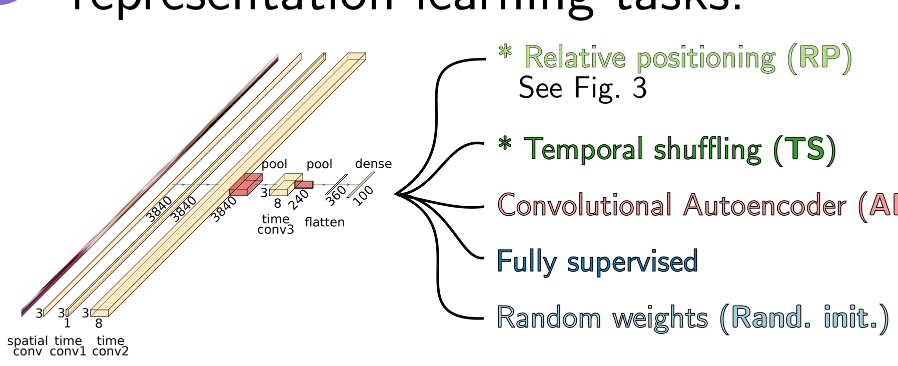
Figure 1: UMAP visualization of features learned with self-supervison (temporal shuffling) color-coded by sleep stage.

Introduction

- Electroencephalography (EEG) is a multivariate time series whose **labelling is expensive and time-consuming**.
- However, **large quantities of (unlabelled) EEG** are available, e.g., sleep recordings.
- How can we leverage unlabelled time series like EEG to **improve generalization** or **study physiological processes**?
- **Self-supervision** can be used to build a supervised learning task out of unlabelled data [Doersch et al., 2015] [Hyvärinen & Morioka, 2019].

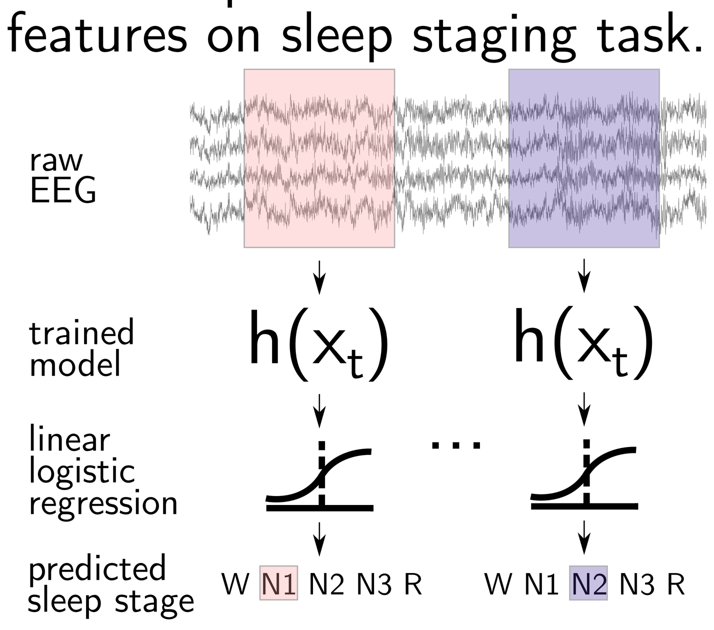
Methods

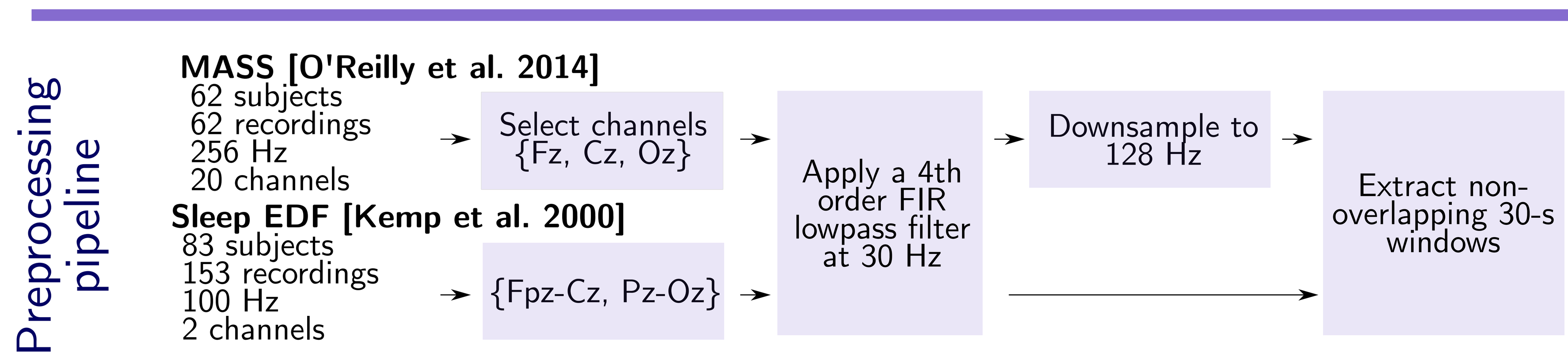
1 Pretrain $\mathbf{h(x_t)}$ on one of various representation learning tasks:



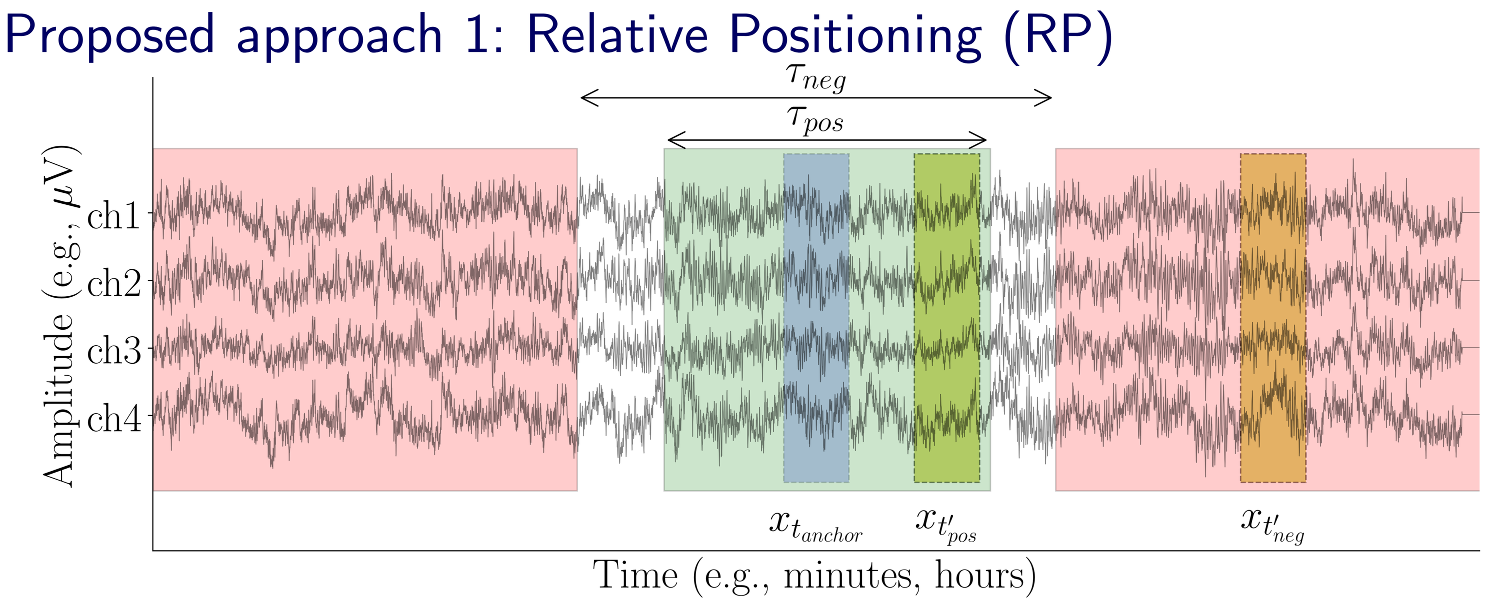
$\mathbf{h(x_t)}$ is a 4-layer CNN feature extractor [Chambon et al., 2018] with input size ~(3840 x 3) and output size (100 x 1). The model is separately pretrained on 5 different tasks, two of which (with *****) are self-supervised while the others are baselines.

2 Evaluate performance of learned features on sleep staging task.





Proposed approach 1: Relative Positioning (RP)



Given two windows x_t and $x_{t'}$, the model must predict whether the two windows are *close or far away in time*. The second approach - temporal shuffing (TS) - is similar, but introduces a third window (not shown).

$x_t \in \mathbb{R}^{3840\times 3}$: EEG window (samples x channels)
$x_{t_{anchor}}$: anchor window
$x_{t'_{pos/neg}}$: second sampled window
$\tau_{pos} \in \mathbb{R}$: **positive** context around $x_{t_{anchor}}$
$\tau_{neg} \in \mathbb{R}$: **negative** context around $x_{t_{anchor}}$

$$y_i = \begin{cases} 1, & \text{if } |t_i - t'_i| \leq \tau_{pos} \\ \text{-}1, & \text{if } |t_i - t'_i| > \tau_{neg} \end{cases}$$

$$\mathcal{L}(\Theta, w, w_0) = \sum_{(x_t, x_{t'}, y) \in \mathcal{Z}_N} \log(1 + \exp(-y[w^\top g(h(x_t), h(x_{t'})) + w_0]))$$

Results

Experiment 1:

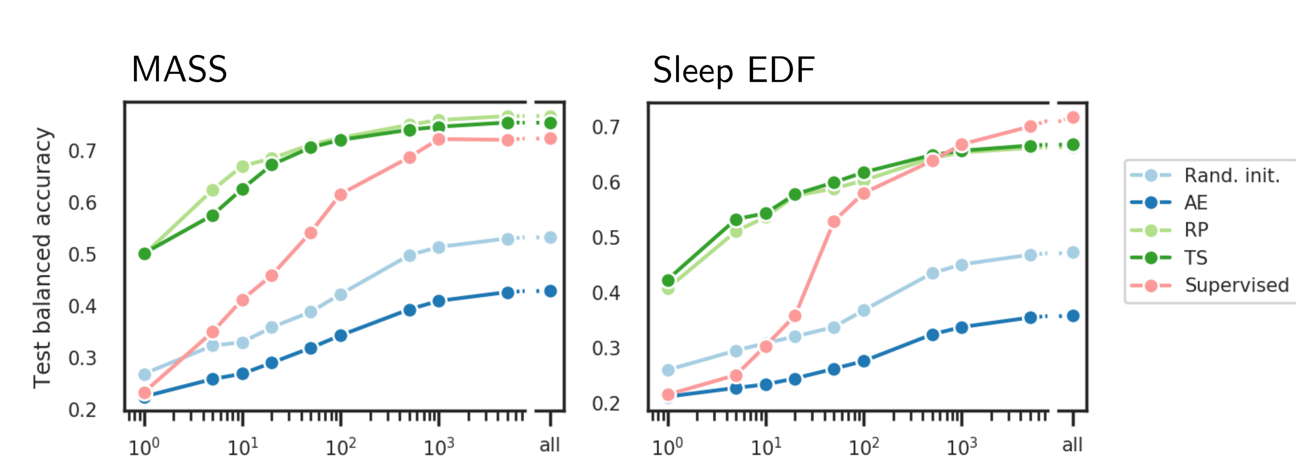
Our self-supervised models learn representations of EEG signals and **facilitate sleep staging**.

	τ_{pos}	τ_{neg}	bal acc$_{SSL}$	bal acc$_{staging}$
*RP	2	2	**79.49**	75.73
	4	15	78.60	**76.66**
	120	120	56.30	65.71
*TS	2	2	81.42	**75.90**
	4	15	**82.12**	75.37
	120	120	66.59	66.66
EEG features	-	-	-	79.43
Supervised	-	-	-	72.51

While hyperparameters τ_{pos} and τ_{neg} (in mins) impact performance in both the self-supervised and sleep staging tasks, sleep staging performance remains comparable to the one of baseline methods.

Experiment 2:

Self-supervised models achieve **much higher performance** than supervised ones **when few examples are available**.



The feature extractors trained with self-supervision outpeform other methods when fewer than ~1,000 examples are available per class.

Experiment 3:

Learned features are **physiologically relevant** and are related to **sleep macrostructure (Fig. 1)** and **age**.

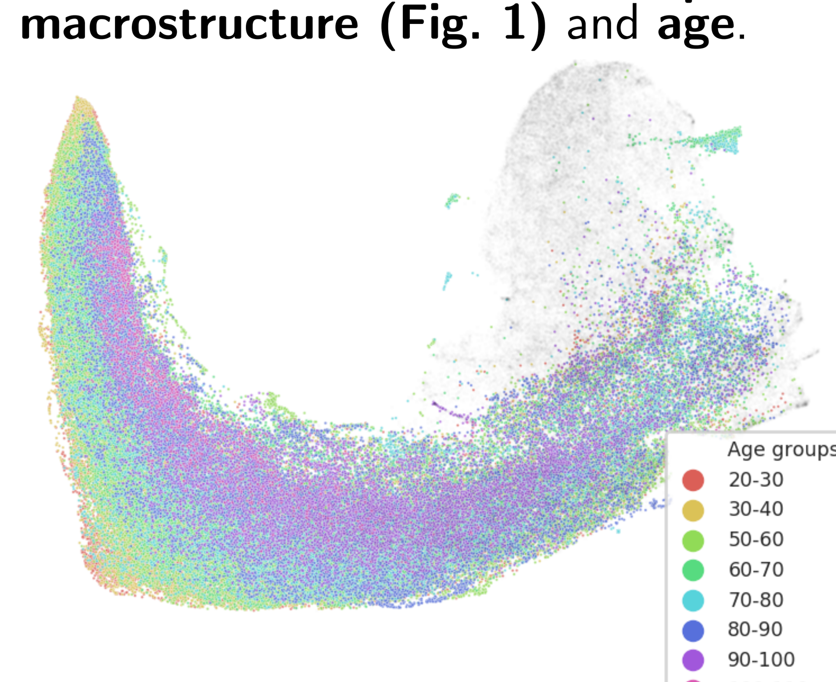


Figure 2: UMAP visualization of features learned with temporal shuffling on Sleep EDF, color-coded by age groups. See also Fig. 1 for sleep macrostructure color-coding.

Discussion

- We introduced **two self-supervised tasks** with a **deep learning architecture** for learning features from raw EEG.
- The learned features captured **sleep-** and **age-related** structure.
- **Classification performance improved** as compared to pure supervision.
- **Next steps**: What other pretext tasks would allow learning complementary information? Can we improve the algorithmic efficiency by using a smarter sampling methodology?

References

Chambon et al. (2018). "A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series." *IEEE Trans Neur Syst Rehab Eng.*

Doersch et al. (2015). "Unsupervised visual representation learning by context prediction." ICCV.

Hyvärinen et al. (2019). "Nonlinear ICA using auxiliary variables and generalized contrastive learning." *AISTATS.*

Kemp et al. (2000). "Analysis of a sleep-dependent neuronal feedback loop the slow-wave microcontinuity of the EEG." *IEEE Trans on Biomed Eng.*

O'Reilly et al. (2014). "Montreal Archive of Sleep Studies: an open-source resource for instrument benchmarking and exploratory research." *Journal of Sleep Research.*