UNIVERSITY OF COPENHAGEN

U-Sleep: resilient high-frequency sleep staging

Mathias Perslev[1], Sune Darkner[1], Lykke Kempfner[2], Miki Nikolic[2], Poul Jennum[2] and Christian Igel[1]

[1]Department of Computer Science, University of Copenhagen

[2]Danish Center for Sleep Medicine, Rigshospitalet

Contact Information
Mathias Perslev
☎ +45 40450066
✉ map@di.ku.dk

Introduction

Sleep staging segments a period of sleep into a sequence of phases providing the basis for most clinical decisions in sleep medicine. Our main objectives were to:

1. Develop an accurate sleep staging model which is resilient to high clinical variability.
2. Evaluate the obtained model against human expert consensus scores.
3. Investigate if high-frequency sleep staging scores carry diagnostic information.

Materials and Methods

Model: U-Sleep [1] is a fully convolutional neural network based on predecessor *U-Time* [2], a U-Net [3] for time-series segmentation. Typically, each sleep segment is 30s. U-Sleep can provide segmentations with higher frequencies than used for training.

Dataset: Fit to annotated sleep staging dataset containing recordings from 15,660 participants of 16 independent clinical studies. Patients vary in age, sex, BMI, disease history and ethnicity. Recordings vary in sample rate, electrode placement, acquisition equipment, pre- and post-processing, etc.

Table 1: Datasets overview.

Type	Studies	Datasets	Subjects	Records	Length (days)
Internal - Train/Test	12	13	15,322	19,578	8,129.3
Hold-Out	4	8	338	346	115.3

Optimization: Fit to batches selected from all training datasets. A random 1 EEG and 1 EOG channel was picked for each sample in each batch.

Evaluation: 13 datasets for training and evaluation; 8 additional datasets for evaluation only, where 2 were consensus scored.

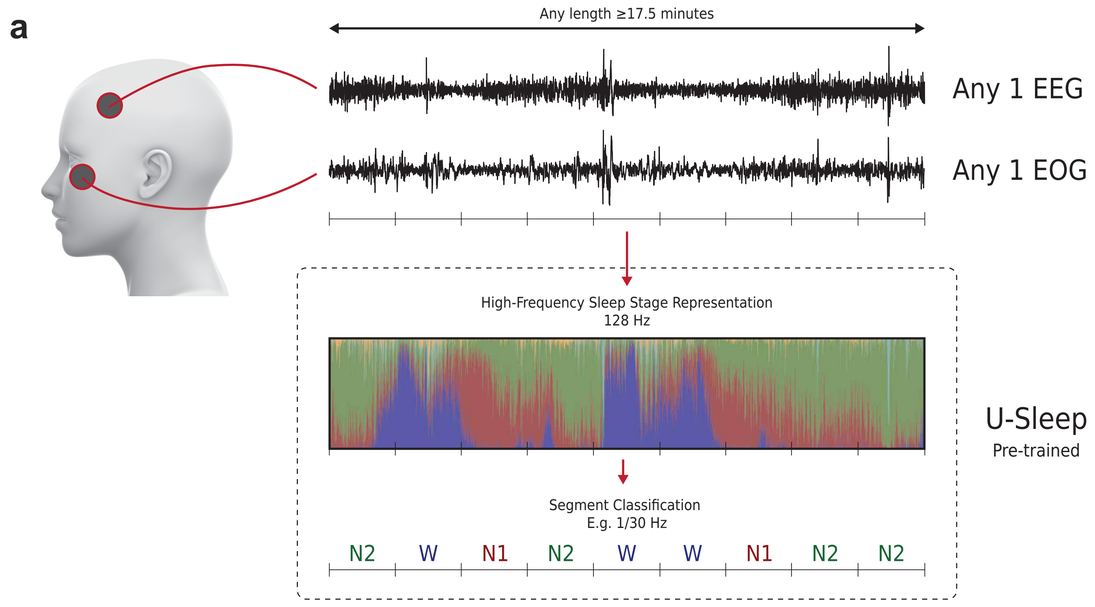

Figure 1: (a) Overview of prediction pipeline. (b) Model architecture schematic.

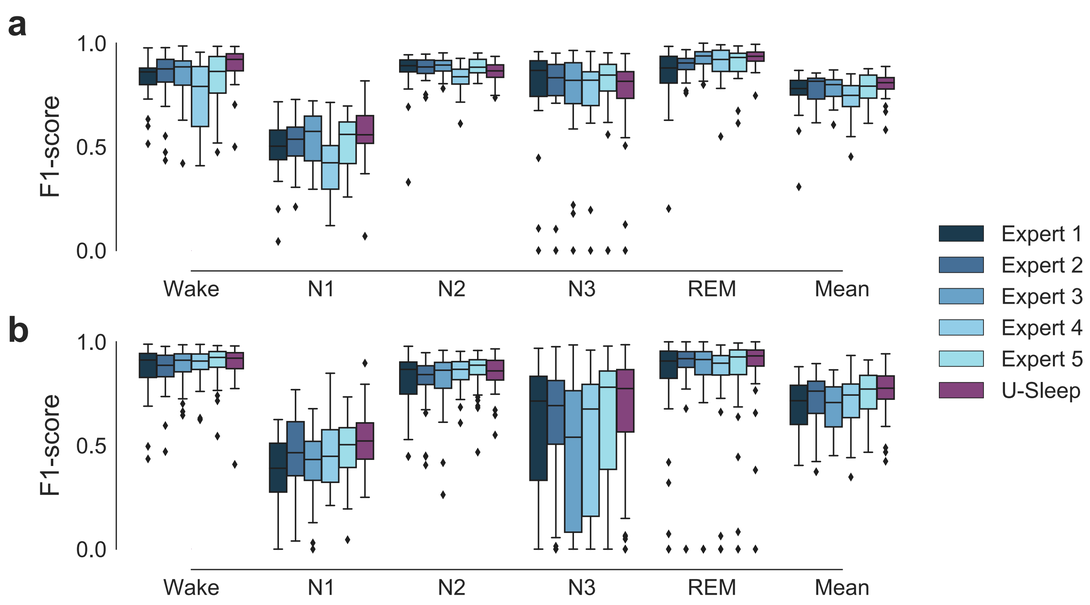

Figure 2: (a) Results from dataset `DOD-H` on 25 healthy subjects. (b) Results from dataset `DOD-O` on 55 patients suffering from obstructive sleep apnea. Sleep stages produced by U-Sleep and the five individual experts were compared to consensus-scored hypnograms.

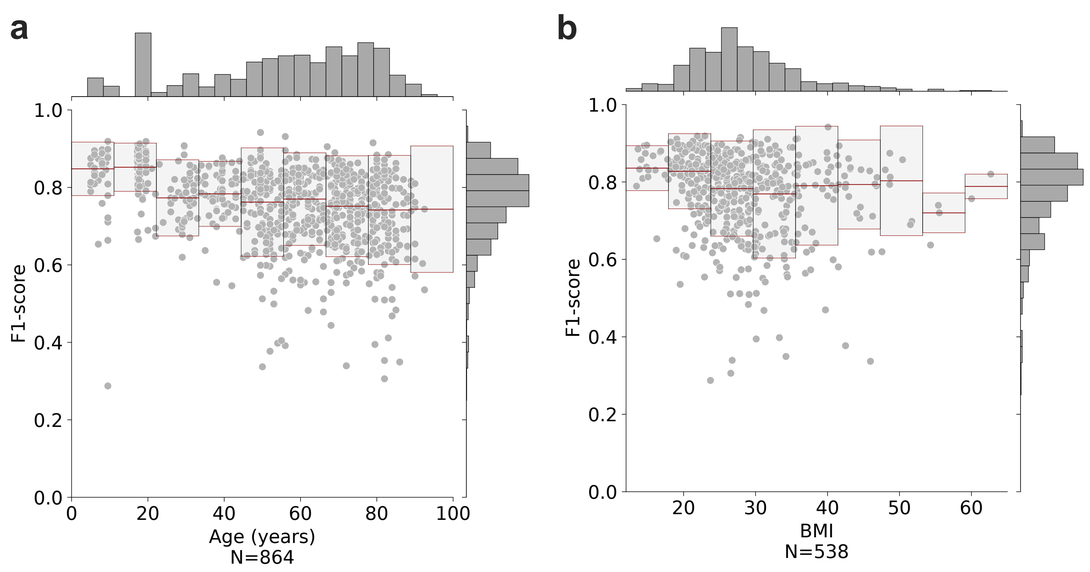

Figure 3: Macro F1 as a function of test subject age (a) and BMI (b).

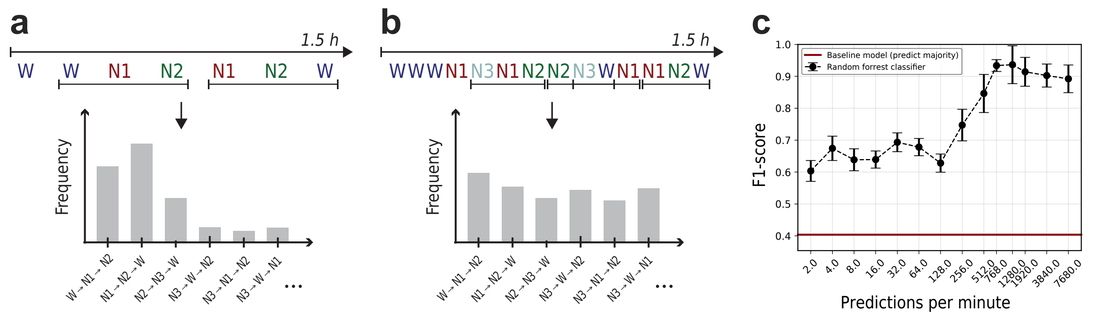

Figure 4: Classification of healthy and sleep disordered patients based on high-frequency sleep scores. Illustration of the extraction of sleep stage triplet transition frequencies from low (a) and high (b) frequency outputs from U-Sleep. (c) Random Forrest classification performance as a function of sleep staging frequency.

Results

- Mean ± STD F1 across 21 test sets: 0.90 ± 0.04 (Wake), 0.53 ± 0.07 (N1), 0.85 ± 0.04 (N2), 0.76 ± 0.07 (N3), and 0.90 ± 0.02 (REM).
- Similar performance to best of 5 human experts on healthy (`DOD-H`) and sleep disordered (`DOD-O`) patients from a held-out clinic. Performs at least as well as other models trained on the consensus scored data.
- Largely invariant to electrode positioning (STD over F1 scores across electrodes was at most 0.03 for all datasets).
- Separation of healthy & sleep disordered patients using Random Forrest classifier was easier using features computed from high-frequency scores.
- U-Sleep is freely available at `https://sleep.ai.ku.dk`.

Future Work

- Needs validation on more patient groups (e.g., diabetes, Parkinson's, and narcolepsy).
- Investigation of performance on patients with atypical EEG patterns, e.g., due to stroke.
- Does U-Sleep capture spatial sleep staging variations?
- U-Sleep was not hyperparameter-tuned due to long optimization time. Can the model be improved further?

References

1. Perslev, M. *et al.* U-Sleep: resilient high-frequency sleep staging. *npj Digital Medicine* **4** (2021).
2. Perslev, M., Jensen, M., Darkner, S., Jennum, P. J. & Igel, C. *U-Time: A Fully Convolutional Network for Time Series Segmentation Applied to Sleep Staging* in *NeurIPS* **32** (2019), 4415–4426.
3. Ronneberger, O., Fischer, P. & Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation* in *MICCAI* **9351** (Springer, 2015), 234–241.