Adapting a General-Purpose EEG Foundation Model for Pediatric Sleep Staging: Development of a Channel-Agnostic Multimodal Framework

Colin Flipo, Hugo Launay, Guillaume Jubien, Aurore Guyon, Pr. Patricia Franco, Ulysse Gimenez

Abstract accepted as Poster presentation at the International Conference on Signal, Image Processing and Pattern Recognition, February 23-24th, Paris

Introduction

Sleep stage scoring, the process of classifying sleep into its different stages, is essential for assessing sleep quality and diagnosing sleep disorders. Manual sleep staging is time-consuming and requires expert annotators, making automation highly desirable. In pediatrics, this task is particularly challenging due to strong age- and development-related variability in EEG patterns.

Although deep learning approaches have achieved substantial performance gains in adult sleep staging, pediatric sleep staging remains more difficult and typically underperforms adult counterparts because of pronounced age-dependent neurophysiological variability.

Recently, EEG foundation models pretrained on diverse large-scale neurophysiological datasets have demonstrated strong transferability across tasks, even when target-domain data are limited.

Bioserenity E1 is a foundation model trained on heterogeneous EEG signals, where sleep data represent less than 20% of the pretraining corpus. While this model has shown strong performance on several downstream neurophysiological applications, its applicability to sleep staging has not yet been systematically explored.

In this work, we investigate whether such a foundation model can be effectively adapted to pediatric sleep staging.

Methods

The dataset comprises 262 pediatric sleep records, totaling 2537 hours of PSG, with subjects aged from 1 to 18 years. Input signals include 6 EEG, 2 EOG, and 1 EMG channels, segmented into standard 30-second epochs and manually labeled according to AASM criteria.

An external validation dataset consists of a first part of 52 healthy pediatric records and a second part of 100 healthy pediatric PSGs and 100 narcolepsies pediatric PSGs collected through a collaboration with HFME (Hôpital Femme-Mère-Enfant, Lyon).

To address the diversity of sleep montages and the multimodal nature of polysomnography (PSG), we developed a channel-agnostic version of the foundation model specifically designed to handle heterogeneous signal configurations, without relying on sleep-specialized pretraining.

We constructed a pediatric sleep staging framework by attaching a downstream model to the pretrained Bioserenity E1 model.

The foundation model weights were frozen, and downstream layers included convolutional neural networks, a final attention mechanism, and bidirectional GRU layers to model temporal dependencies.

For comparison, an internally developed task-specific baseline based on Robust SleepNet (RSN) first trained on adults PSG (n=689) and then finetuned on previously described pediatric database using transfer learning.

Results

The foundation-based model achieved a macro-F1 score of 0.746 on the internal test set, compared to 0.79 for the RSN baseline. On the first 52 healthy records of the HFME external dataset, the RSN baseline reached 0.767, while the foundation-based model achieved 0.65.

Despite being a general-purpose model not specifically pretrained for sleep staging, the channel-agnostic foundation-based model achieved promising performance on pediatric PSGs.

Discussion

This study demonstrates the feasibility of adapting a general-purpose brain foundation model to pediatric sleep staging, despite minimal sleep-specific pretraining. While performance remains below a finetuning based on sleep specific model, the framework establishes a flexible EEG foundation relevant for pediatric sleep analysis.

As more advanced, modality-aware, and sleep-specialized foundation model variants are developed, performance and generalization, particularly for rare disorders, are expected to improve. Future work will evaluate the model on the full HFME dataset, focusing on narcolepsy detection to assess generalization to clinically relevant, data-scarce conditions.