Hybrid Features for Cross-Institutional Cognitive Impairment Prediction from Polysomnography

Goeun Yeh*, Muhi Kim, Jiyoung Park

University of Ulsan College of Medicine, Ulsan, KR

Aims : Predicting future cognitive impairment from polysomnography (PSG) is challenging due to heterogeneous multi-site recording conditions. Our team (Momochi-SleepAI) aimed to develop a simple, generalizable algorithm that explicitly addresses cross-institutional distribution shift without relying on complex deep learning architectures.

Methods : Training data comprised 622 PSG recordings from three institutions (S0001, I0006, I0002). Features were extracted from CAISR (Complete AI Sleep Report) automated sleep annotations and retained only if they showed consistent directional correlation with cognitive impairment labels across all three sites. This yielded 32 features combining absolute sleep measures (arousal index, sleep efficiency, AHI), subject-internal relative measures (REM arousal ratio, sleep structure asymmetry, arousal clustering ratio), and demographic covariates (age, BMI, sex). Absolute features captured overall sleep severity; relative features were designed to resist inter-site distribution shift. To address annotation heterogeneity, two models were trained: a full model (32 features) when CAISR annotations were available, and a fallback model (11 features) using SaO2-derived surrogates when absent. Both used a soft-voting ensemble of RandomForest (n_estimators=200, max_features=sqrt, min_samples_leaf=12) and GradientBoosting (n_estimators=150, learning_rate=0.05). Performance at unseen institutions was estimated via institution-held-out cross-validation, training on S0001 and evaluating on I0006 and I0002 as independent external test sites.

Results : The hybrid feature set improved worst-case external AUROC from 0.50 (absolute features only) to 0.66, achieving AUROC 0.67 on I0006 and 0.66 on I0002. The official leaderboard score was 0.51, with the gap attributed to a site-specific EDF filename error, since corrected. Permutation importance identified REM arousal ratio, arousal clustering ratio, and stage entropy variability as the most consistent predictors across both external sites.

Conclusion : External validation results suggest that hybrid sleep fragmentation features provide site-invariant signal for cognitive impairment prediction, with robust generalization achievable through simple, interpretable PSG features alone.

Table 1. Cross-institutional validation AUROC by model (Training site: S0001)

Model	I0006	I0002
Absolute feature only	0.67	0.50
Hybrid (ours)	0.67	0.66