Sequence Modeling of CAISR Epoch Features for Predicting Cognitive Impairment from Polysomnography

Jingsu Kang, Hao Wen

Tianjin Medical University
Tianjin, China

Aim: This paper introduces a sequence modeling approach for predicting future cognitive impairment from polysomnography (PSG), serving as the "Revenger" team's solution to the George B. Moody PhysioNet Challenge 2026.

Methods: A key challenge of this task is the pronounced heterogeneity of PSG signals across the five contributing institutions. To obtain a site-agnostic representation, we summarized the pre-computed Complete AI Sleep Report (CAISR) algorithmic annotations into a 21-dimensional feature vector for each 30-second epoch, capturing sleep-stage, arousal, respiratory, limb-movement, and temporal-position information. The full-night epoch sequence was modeled with two types of deep sequence encoders: a Convolutional Recurrent Neural Network (CRNN) that combines a ResNet-like convolutional backbone with a bidirectional LSTM, and a Transformer encoder with stacked self-attention layers. Both models incorporated demographic features (age, sex, BMI) via feature-wise linear modulation (FiLM). We evaluated multiple model scales (small/medium/large) and selected the final entry based on validation AUROC. During training, sequences were cropped to at most 768 epochs to improve batch efficiency. Models were trained with the binary cross-entropy loss, label smoothing of 0.05, AdamW optimizer with weight decay of 0.01, batch size 16, and the OneCycle learning rate scheduler. A fixed stratified 80/20 train-validation split, stratified jointly by site and label, was used for model development.

Results: The best submission entry of our team "Revenger" used a medium-sized CRNN model with 0.4 million trainable parameters and achieved an AUROC of 0.555 on the hidden validation set (ranking: 102/244). It outperformed the larger Transformer and heavier CRNN variants evaluated in the unofficial phase.

Conclusion: CAISR-derived epoch features provide a practical, site-invariant representation for PSG-based cognitive impairment prediction, and the results suggest that moderate-capacity sequence models currently offer the most reliable trade-off between representation power and generalization under severe cross-site heterogeneity.