Abstract

Sleepnet: automated sleep staging via deep learning

In this work, an efficient automated new approach for sleep stage identification based on the

new standard of the American academy of sleep medicine (AASM) is presented.

Sleep disorders, such as parasomnias, hypersomnia and sleep apnea, affect 50-70 million adults in the USA. (the importance of sleep disorders)

Overnight polysomnography (PSG), including brain monitoring using electroencephalography (EEG), is a central component of the diagnostic evaluation for sleep disorders. (PSG to detect sleep disorders)

While PSG is conventionally performed by trained technologists, the recent rise of powerful neural network learning algorithms combined with large physiological datasets offers the possibility of automation, potentially making expert-level sleep analysis more widely available.

Abstract

According to the feature of the sleep EEG, we developed a deep learning methodology of automatic sleep stage scoring.

We used class-balanced random sampling across sleep stages for each model in the ensemble to avoid skewed performance in favor of the most represented sleep stages, and addressed the problem of misclassification errors due to class imbalance while significantly improving worst-stage classification.

We used an openly available dataset from 39 healthy young adults for evaluation.

Our method has both high overall accuracy (86.2%), and high mean F1-score (86.5%) and mean accuracy across individual sleep stages (XX%) over all subjects.

2.3 Data Acquisition

The dataset that we used to evaluate our method is a publicly available sleep PSG dataset [14] from the PhysioNet repository [8] that can be downloaded from [XX]. The data was collected from electrodes Fpz-Cz and Pz-Oz. The sleep stages were scored according to the Rechtschaffen and Kales guidelines [22]. The epochs of each recording were scored by a single expert (6 experts in total). The sleep stages that are scored in this dataset are Wake (W), REM (R), non-R stages 1–4 (N1, N2, N3, N4), Movement and Not Scored. For our study, we removed the very small number of Movement and Not Scored epochs (Not Scored epochs were at the start or end of each recording), and also merged the N3 and N4 stages into a single N3 stage, as is currently the recommended by the American Academy of Sleep Medicine (AASM) [12, 25]. There were 61 movement epochs in our data in total, and only 17/39 recordings had movement artifacts. The maximum number of movement epochs per recording was 12. The rationale behind the decision of removing the movement epochs was based on two facts. First, these epochs had not been scored by the human expert as belonging to any of the 5 sleep stages, as it is recommended in the current AASM manual [12, p.31]. Second, their number was so small that they could not be used as a separate ‘movement class’ for learning. The public dataset includes 20 healthy subjects, 10 male and 10 female, aged 25–34 years. There are two approximately 20-hour recordings per subject, apart from a single subject for whom there is only a single recording. To evaluate our method we used the in-bed part of the recording. The sampling rate is 100 Hz and the epoch duration is 30 seconds.