Journal of Research High School

Vol. 2024 (5) 207 – 211
ISSN 2688-3651 [online]

Development of a Classifier to Identify Sleep Stages from EEG Data

Anika Kumar[1] *

[1]Woodinville High School, Woodinville, WA, USA
*Corresponding Author: me@anikakumar.net

Advisor: Alejandro Carnicer-Lombarte, ac723@cam.ac.uk

Received October 3, 2023; Revised February 12, 2024; Accepted, March 20, 2024

Abstract

Understanding the various stages of sleep is crucial in diagnosing and treating sleep disorders like insomnia and sleep apnea. The current methods used to determine sleep stages involve using a polysomnogram (PSG), a sleep diagnostic tool in healthcare, to examine a patient's activity throughout their sleep. In a polysomnogram study, an electroencephalogram (EEG) records the brain's electrical activity to determine signal amplitudes at various frequency bands (e.g., alpha, beta, theta, delta). Sleep technicians then examine electroencephalogram data to determine each sleep state during the patient's sleep. However, this process is labor-intensive. This paper investigated spectral and temporal features of a group of patients' electroencephalogram signals, such as the amount of frequency present and its changes and behaviors over time. It is revealed through the paper's results that analyzing these features with a classifier would provide experts with a more efficient way to score sleep. Distinguishing corresponding sleep stages from wake/sleep characteristics would give physicians a more accurate and accessible scoring system than existing human performance sleep scoring methods. This research anticipates a transformation in healthcare by increasing productivity in sleep scoring, allowing patients with severe sleep disorders to be diagnosed and treated faster. By creating a classification system that categorizes each sleep stage, this paper provided a way for physicians to move from tedious human sleep scoring to automatic sleep scoring, improving patients' sleep-based health issues.

Keywords: Electroencephalography, EEG, Sleep, Sleep stages, Neuroscience

1. Introduction

Sleep is a natural biological phenomenon all humans perform (Santaji and Desai, 2020). Humans dedicate one-third of their life to sleep. This sleep is not just resting- it is essential for a human's performance, thinking, and participation in daily activities. However, in the US, a significant problem exists. Over 50 million individuals suffer from severe or ongoing sleep disorders (National Heart, Lung, and Blood Institute (NHLBI), 2022). Sleep disorders are connected to many long-term health consequences, including an increased risk of hypertension, diabetes, obesity, depression, heart attack, and stroke (Colten and Altevogt, 2006). Additionally, these disorders significantly impact aspects of society. For instance, higher rates of automobile crashes can be commonly linked to sleep disorders. Sleep disorders also correlate to decreased work productivity, leading to daytime fatigue and impaired cognitive and motor functions. Therefore, recognizing the importance of sleep staging is essential in medicine to help physicians clinically analyze sleep to diagnose patients with sleep disorders.

Sleep scoring generally follows the standardized scoring rules presented by the American Association of Sleep Medicine (AASM) (Malafeev et al., 2018). In the ASSM guideline, sleep is divided into stage W (wake), stages N1-N3 (non-rapid eye movement or non-REM), and stage R (rapid eye movement or REM). Stage W is defined with high beta (associated with alertness) and alpha waves (associated with relaxation), which become more present as the patients are awake yet close their eyes. Stage N1, the transition from waking to sleeping, is characterized by high alpha and theta waves (associated with drowsiness) (Gonzalez et al., 2019). This is when human sleep begins, though a patient can easily be awakened as it is light sleep (Malafeev et al., 2018). Stage N2 is marked by the continued

Vol. 2024 (5) 207 – 211
ISSN 2688-3651 [online]

presence of theta waves and sleep spindles, along with the addition of some delta waves (associated with sleep) K complexes. Sleep spindles are high-frequency bursts believed to mediate many sleep-related functions like memory consolidation (Andrillon et al., 2011). K complexes are long, sharp downward waves that last less than half a second and can occur as a response to a stimulus in the environment, like a door closing (Gonzalez et al., 2019). Stage N3, the final stage of NREM sleep, is identified by high delta waves when a patient experiences sleep inertia (a transient phase of mental fogginess) and physical recovery. Compared to sleep stage N1, sleep stage N3 will be more challenging to wake. Finally, Stage R is distinguished by its resemblance to Stage W; high beta waves mainly define it. In this stage, a patient experiences dreams reminiscent of the thinking performed when awake, causing similar brainwave patterns in EEG data.

Conventionally, a patient's sleep stage will be identified by a trained sleep scorer who will manually sleep score (Sharma et al., 2021). Although this method works, it is time-consuming and tedious (Choo et al., 2023). Therefore, I believe that an automated classifier would be a better way to complete this process. An automatic classification system allows a patient's entire sleep quality to be determined instantly than human scoring methods. Giving physicians the data, they can use to treat extreme sleep-related health conditions. In this paper, EEG data from a database is preprocessed and separated into specific EEG frequencies to recognize patients' sleep and wake characteristics corresponding to the various sleep stages. By identifying the patterns that indicate periods of sleep and wakefulness for each patient, a classifier can be applied to sort and categorize the patient's sleep into different segments to evaluate their overall sleep quality and see if the results are more accurate than current human sleep scoring methods.

2. Materials and Methods

2.1 Data Set

The EEG dataset is obtained from the Haaglanden Medisch Centrum (HMC) sleep center published online for public access (Alvarez-Estevez et al., 2022). The dataset contains EEG signals from a random group of 151 patients (85 males, 66 females) with varying sleep disorders. Selectors made no additional requirements to assess patients' PSG recordings. Recordings were held during patients' times at the hospital and when they were not admitted. The PSG data consisted of placing four electrodes (F4/M1, C4/M1, O2/M1, and C3/M2) on a patient's scalp to assess EEG (Figure 1). EEG signals were all sampled at 256 hertz. All PSG data was stored in EDF file format without any additional filtering to then process in the programming platform, Matlab, to extract the data into figures and tables for analysis. Sleep stage scores were done manually by trained sleep technicians following ASSM guidelines.

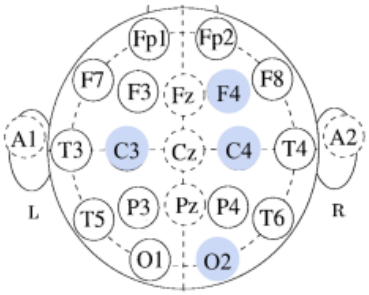

Figure 1. The 10–20 International System for EEG Electrode Placement. An internationally standardized method that ensures consistent spacing and placement of electrodes. Letters describe brain regions (Frontal, Temporal, Parietal, and Occipital) and sides (Left and Right). Odd numbers indicate the left side, even numbers indicate the right side, and "z" indicates the midline (Ward et al., 1999).

2.2 Data Processing and Feature Extraction

Preprocessing started with importing the patient's data from the EDF files. The overall sampling frequency was set to 256 Hz by the HMC sleep center. Additionally, the frequency of each sleep wave sampling was defined (Table 1). Data was then extracted from each electrode for filtering (Figure 2). A Butterworth filter allowed low frequencies and removed high-frequency noise, like movement artifacts (unnecessary electrical signals from a person moving). The root mean square (RMS), which grabs a signal, squares it, and then takes root to

Table 1. The table indicates the EEG signal frequency bands universally used in sleep stage classification

Bands	Frequencies (Hz)
Delta (δ)	0-4
Theta (θ)	4-8
Alpha (α)	8-13
Beta (β)	13-22

Vol. 2024 (5) 207 – 211
ISSN 2688-3651 [online]

produce positive-only values, was taken from all the EEG bands to understand the magnitude of the signals by ignoring their negative signal direction (Hindarto et al., 2014). A window rolling means was then run determined from the data, which took the average of each interval to reduce and smooth the data to detect deviations and patterns more easily.

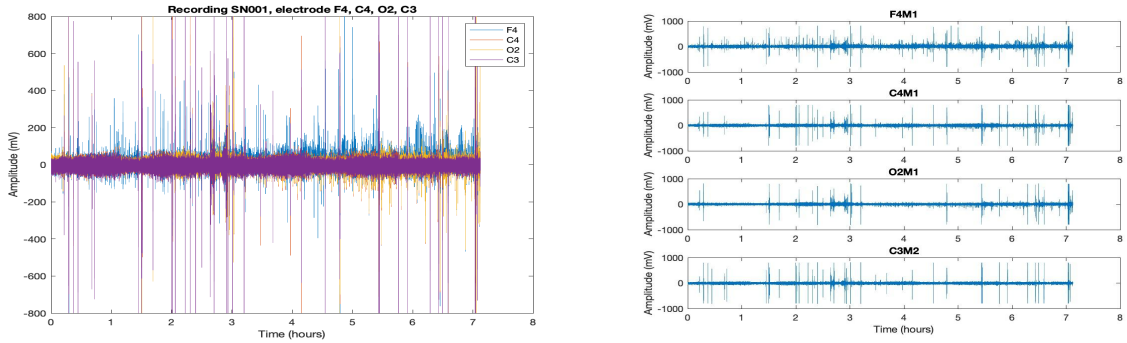

Figure 2. The first graph represents the unfiltered EEG data from a patient's electrodes. The second graph represents the extraction of each electrode from the patient's EEG data for easier analysis. The x-axis of the graphs is the amount of time the patient slept, and the y-axis is the amplitude to measure the energy height of the sleep waves.

2.3 Classifying Data

The classifier uses thresholds determined by looking at the mean graph of each band (Figure 3). Based on the graphs, the data can be analyzed to determine what frequency would make the band low or high to categorize into respective sleep stages. Using thresholds, conditions were made for each sleep stage; for example, Stage W was defined by high beta and alpha waves. Additionally, if the data showed all characteristic waves or none, it was labeled as no state.

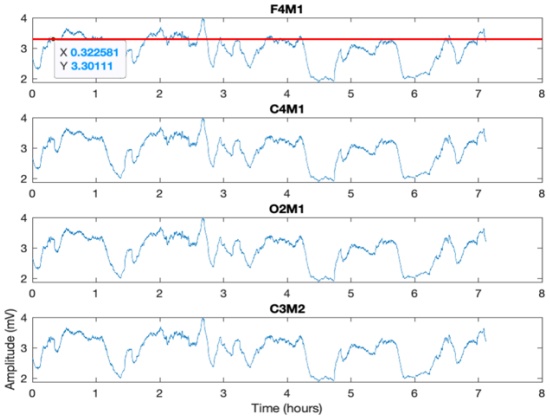

Figure 3. Periods of high alpha wave activity were determined by the threshold of 3.3 mV.

3. Results

The data from the dataset and the classifier interpretation were graphed for comparison. The classifier's data matches the database's scores for a patient for most of the 30-second intervals seen in Figure 4. Overall, the classifier does not align perfectly with the database, with some intervals being categorized as no state due to not fitting into any of the categories; this may be because not all humans have a specific band or may have much lower activity in frequencies than the normal person in each sleep stage, indicating they have different thresholds that tell technicians when these patients move sleep stages. The classification from Alvarez-Estevez and Rijsman and the classifier presented here disagreed on certain phases. The classifier's result appears more detailed and accurate in specific areas. As seen in Figure 4, the patient's third hour of sleep was categorized as a combination of Stages W and R than the database's classifier, which listed this interval as just wake-sleep. Additionally, while the developed classifier tended to categorize recordings with sleep spindles and K complexes as N1, Alvarez-Estevez, and Rijsman often classified these as N2. Classifier data was also compared with other patient's EEG recordings, supporting the above results. A general trend of N3 sleep not having a broader impact was found; this may have to do with the delta bands that barely showed any significant increases overall.

Vol. 2024 (5) 207 – 211
ISSN 2688-3651 [online]

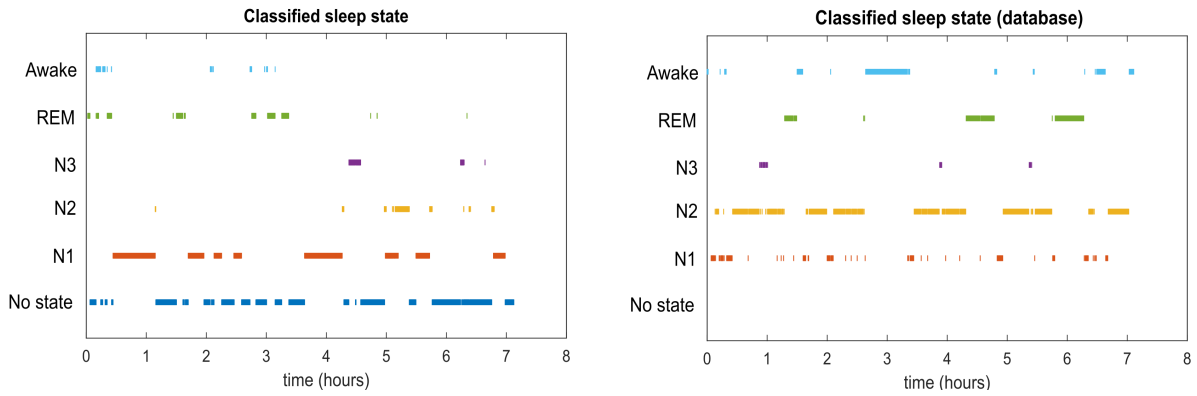

Figure 4. Comparison of sleep stages between classifier data (left graph) and database data (right graph).

4. Discussion

When comparing the classifier's data to the human-scored data, the data does not align perfectly with the data from the database. This may have to do with the errors made by the database's original scoring by sleep technicians. The database confirms that sleep scoring was done manually and accurately; however, through further analysis, it seems that may not be the case. For example, in the database, the data in the 30-second intervals occasionally switches from Stage R to Stage N1, which is abnormal as REM sleep is the deepest stage of sleep compared to N1. Therefore, Stage R should not switch directly to N1 but gradually go from N3 to N2 and then to N1 sleep as it is rare to exit deep sleep immediately.

Based on the disagreed results, the classifier gives more data in certain areas than the classification from Alvarez-Estevez Rijsman. In certain long intervals, the patient's sleep is determined by one sleep stage, while the classifier can pick up each change during each second, offering a more detailed view of how the patient's sleep changes. This may have to do with the fact that all the patients tend to sleep for around 7 hours, and when scored by hand, it is not ideal for analyzing each interval; therefore, inferences may be used, undermining the data's accuracy. This is in contrast to the developed classifier that uses a mechanism that will be able to check each interval to what it should be defined as to more accurately distinguish every interval. The results also indicate that in the classifier, Stage N2 sleep is occasionally distinguished as Stage N1 sleep compared to the database. The researchers may have classified parts of the data as N2 sleep instead of N1. This may have to do with the fact that specific waves, like the other stages of sleep, do not define N2 sleep. Instead, it is identified through the two distinct features of the waveform, sleep spindles and K complexes, which are sleep behaviors often seen in EEG results through shapes and patterns that sometimes are not always seen. When observing the delta activity of a patient who had this misclassification, it can be observed that few long delta waves were seen, which represent the K complexes in N2 sleep. This may have to do with the fact that specific waves, like the other stages of sleep, do not define N2 sleep. Instead, it is identified through the two distinct features of the waveform that are not always seen.

The common pattern of low Stage N3 sleep was also discovered from the findings. With N3 sleep being determined by activity from delta waves, no big effect was made by N3 as delta activity remained around the same throughout the interval. When observing the delta activity of each patient, it showed barely any significant increases overall. Since the data examined is from patients hospitalized for sleep disorders and with N3 sleep being the deepest NREM sleep, it may be possible that delta activity is rare. For example, patients facing insomnia cannot reach deep sleep and will not have any significant increases in delta activity and, therefore, be unable to experience N3 sleep or any further deeper sleep stage.

5. Conclusion

In the broader context, the sleep classifier performed accurately in capturing the overall data trends to clearly see patients' wake/sleep characteristics to identify the sleep quality or even when the patient is awake at night. Automatic

Vol. 2024 (5) 207 – 211
ISSN 2688-3651 [online]

sleep stage classifiers have already been published in research, making this field competitive. Nonetheless, all of them use machine learning methods and algorithms that consist of mechanisms that are more complex and harder to utilize. This classifier is developed using simpler code to ultimately accomplish the same goal that will be more accessible and convenient for physicians while giving the same accurate results. This paper showed that experts can improve sleep-scoring efficiency by analyzing spectral and temporal features. Using a classifier to identify sleep stages based on these features offers a more reliable and convenient choice than human-based methods. This research will transform sleep scoring from a time-consuming to a productive way for physicians to analyze patients' PSG data to determine sleep stages to more effectively diagnose and treat patients with serious sleep disorders.

References

Andrillon, T., et al. (2011). Sleep spindles in humans: insights from intracranial EEG and unit recordings. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 31(49), 17821–17834. https://doi.org/10.1523/JNEUROSCI.2604-11.2011.

Alvarez-Estevez, D., & Rijsman, R. (2022). Haaglanden Medisch Centrum sleep staging database (version 1.1). *PhysioNet*. https://doi.org/10.13026/t79q-fr32.

Choo B. P., et al. (2023). Benchmarking performance of an automatic polysomnography scoring system in a population with suspected sleep disorders. *Frontiers in Neurology* 14:1123935. doi: 10.3389/fneur.2023.1123935.

Colten, H. R., et al. (2006). Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. N*ational Academies Press (US)*.

Gonzalez, A., et al., (2019). Sleep Stage Scoring. *Medscape.* https://emedicine.medscape.com/article/1188142-overview#a2.

Hindarto, H., et al. (2014). EEG signal identification based on root mean square and average power spectrum by using BackPropagation. *Journal of Theoretical and Applied Information Technology*. 66. 782-787.

Malafeev A., et al. (2018). Automatic Human Sleep Stage Scoring Using Deep Neural Networks. *Frontiers in Neuroscience* 12:781. doi: 10.3389/fnins.2018.0078.

National Heart, Lung, and Blood Institute (NHLBI). (2022). Sleep deprivation and deficiency. https://www.nhlbi.nih.gov/health-topics/sleep-deprivation-and-deficiency.

Patel A. K., et al. (2023). Physiology, Sleep Stages. In *StatPearls*. *StatPearls* Publishing.https://www.ncbi.nlm.nih.gov/books/NBK526132/.

Santaji, S., et al. (2020). Analysis of EEG Signal to Classify Sleep Stages Using Machine Learning. *Sleep Vigilance* 4, 145–152. https://doi.org/10.1007/s41782-020-00101-9.

Sharma, M., et al. (2021). Automatic Sleep-Stage Scoring in Healthy and Sleep Disorder Patients Using Optimal Wavelet Filter Bank Technique with EEG Signals. *International journal of environmental research and public health*, 18(6), 3087. https://doi.org/10.3390/ijerph18063087

Ward, D., et al. (1999). Enhancement of deep epileptiform activity in the EEG via 3-D adaptive spatial filtering. *IEEE* Transactions on bio-medical Engineering. 46. 707-16. 10.1109/10.764947.