CRoNE 2019
Conference Proceedings

Online sleep spindles detection with short and long time average ratio

Felipe A. Torres[1*], **Patricio Orio**[2,3], **María José Escobar**[1]

***For correspondence:**
felipe.torrese@sansano.usm.cl
(FAT)

[1]Department of Electronic Engineering, Universidad Técnica Federico Santa María; [2]Centro Interdisciplnario de Neurociencia de Valparaíso; [3]Instituto de Neurociencia, Universidad de Valparaíso

Abstract Sleep spindles occurrence correlates with the consolidation of recently acquired information. The memory consolidation literature supports that there are more sleep spindles after a learning task. Thus, the detection of them does not only allow the classification of the N2 sleep stage, further provides a quantification value of memory replay and memory consolidation during sleep. Event detection is an important processing step performed in the analysis of diverse kinds of waveforms. The short and long term average ratio is the most widely used event detection approach to analyze passive seismic data and trigger the storing or discarding of data. Its popularity comes from its simplicity and the usage of a fixed threshold determined by the intention of the data usage and not based on the signal dynamics. This work explores the usage of this event detection approach on the online detection of sleep spindles. The advantages of the detection performance with this feature over using the same binary classification method using other fast calculation features come from its statistical properties. The classification features compared are the root mean square amplitude, relative spindle power, and the Teager-Kayser energy operator.

Introduction

Sleep spindles are short-duration events (0.5-2 s) in the specific frequency range of 11-16 Hz that occur during the NREM sleep and they are detectable in EEG registers. The principal characteristic of the N2 sleep stage is the high presence of these events ***Devuyst et al.*** (***2011***). The ratio of occurrence and other characteristics of the spindles could indicate some health or sleep disorders (***De Gennaro and Ferrara, 2003***), also is known that burst on the hippocampus, slow oscillations and spindles has a time order in occurrence and are related to memory consolidation (***Sara, 2017***) making the spindles a marker of the capacity of learning and memory processing (***Cairney et al., 2015***). Then, find spindles with accuracy and precision could help to evaluate differences in the sleep after learning or memory tasks and to detect pathologies. The human visual sleep scoring usually employs the power in the spindle frequency range as a helper for the experts (***Purcell et al., 2017***). Another manual method is the use of crowd-sourced annotations from non-experts (***Zhao et al., 2017***). Event detection is an important processing step performed in waveform analysis. In particular, in seismology from the appearance of digital signal acquisition systems, there was a need to reduce the length of recordings to improve the storage and data transmission capabilities of the electronic devices ***Akram et al.*** (***2019***).

Classification algorithms have the assumption that there is some space of features where the patterns are separated. Online or real-time applications require fast computation of these features. Thus, it is a great achievement if this separation is notorious in a small dimensional space or even in just one dimension.

This work takes this last approach. The proposed algorithm does a binary classification (spindle

or not spindle) using only an extracted feature from the input signal. Then it also annotates the time of occurrence of each spindle to make a detector.

Related works

This work uses a single feature from the EEG signal as many previous methods of spindles detection (***Devuyst et al., 2011***) (***O'Reilly and Nielsen, 2015***). This work is also similar to works that perform statistical analysis of spindles to obtain a Bayesian detection algorithm (***Babadi et al., 2012***) or an HMM-SVM detection scheme (***Mporas et al., 2013***). Other works used features of more computational demand as wavelets (***Al-Salman et al., 2019***), a combination of features (***Liu et al., 2017***) or the non-negative matrix factorization (NMF) periodogram of the correntropy function (***Ulloa et al., 2016***).

The use of a public database is important to compare the results of different methods. In that sense, this work use the same dataset used by (***Devuyst et al., 2011***), ***O'Reilly and Nielsen*** (***2015***), (***Liu et al., 2017***) and (***Al-Salman et al., 2019***).

Methods and Materials

Data

The event detection methods were evaluated over a synthetic EEG signal and also in a real EEG dataset. The baseline of the synthetic signal has a frequency spectrum slope of $\sim f^3$ typical of slow-wave sleep. It is generated with a sum of sinusoidal signals:

$$x_{EEG}(t) = \frac{1}{498} \sum_{n=2}^{n=499} \frac{1}{(0.1n)^{3/(f_s)}} \{cos((2\pi(1 + (\sigma_f)r_f)(0.1n)t + (1 + \pi r_p(t))), \quad (1)$$

where $f_s = 100Hz$ is the sampling frequency, $\sigma_f = 0.01$ is the standard deviation given to the central frequencies of the expression 1, r_f is a random number, and $r_p(t)$ is a random time series generated from a normal Gaussian distribution. The baseline signal adds with a spindles signal:

$$\begin{aligned} x_{spindle}(t) &= \frac{1}{6} \left(\sum_k f_{spindle}(t) * p(t-k) \right) x_{sigma}(t), \\ x_{sigma}(t) &= \sum_{m=\{-1,0,1\}} \left((1 - |m|d_a)cos(2\pi(12 + 0.5m)(1 + \sigma_{sf1})t) + (\frac{1}{\sqrt{(2)}} - |m|d_a)cos(2\pi(24 + m)(1 + \sigma_{sf2})t) \right), \end{aligned} \quad (2)$$

where d_a is the amplitude difference of the main lobe with the lateral lobes, σ_{sf1} and σ_{sf2} are random deviations for the main frequency of 12Hz and its first harmonic for spindles. $f_{spindle}(k)$ is a triangular envelope defined by the expression (3) with a duration $t_{spindle} = 0.75s$. Expression (4) defines the probability of occurrence of spindles where $U(t)$ is a random time series uniformly distributed between [0,1].

$$f_{spindle}(t) = \begin{cases} \frac{t_{spindle} a_{spindle}}{2} + a_{spindle} t & t < t_{spindle}/2 \\ t_{spindle} a_{spindle} - a_{spindle} t & t \geq t_{spindle}/2 \end{cases}, \quad (3)$$

$$p(t) = \begin{cases} 1 & U(t) > 0.998 \\ 0.01 & 0.99 < U(t) < 0.998 \\ 0 & U(t) \leq 0.99 \end{cases}, \quad (4)$$

The tested synthetic EEG signal has 8 segments of 300s with different values for $a_{spindle}$: 0.125, 0.250, 0.375, 0.5, 0.625, 0.750, 0.875, and 1.0.

The real EEG signals come from the DREAMS dataset (***Devuyst and Dutoit, 2011***). This dataset consists of eight registers of 30-minutes. The data were sampled at frequencies of 50, 100, and 200 Hz. The dataset includes the visual scoring of spindles from two experts. They marked the start time and the duration of spindles. A duration of 1-second is annotated for many of the spindles but they have another time length. This work uses annotations without modification from just one expert.

Spindles detection

General Procedure

The event detection methods in this work use a general framework of a single dimension signal and a single threshold to classify a sample as forming part of an event occurrence. Consecutive samples classified as spindles must exceed the minimum duration to finally designate a section of the input signal as an event. A time for possibles gaps is also included to actuate as replace of an hysteresis mechanism.

A common procedure before any step is to obtain the z-score value of the signal. Given that there is not any assumption about the signal properties, the mean value and the standard deviation are calculated at each sample, then there is the need of advance in various samples to achieve consistency in these statistical estimators. The z-scored signal $x(t)$ is band-pass filtered in the sigma band (11-16Hz) with a Chebyshev type I fourth order filter to obtain $x_\sigma(t)$. The calculation of features at each sample uses both signals. The detection follows the procedure show in Algorithm 1.

Algorithm 1 Spindle Detection Algorithm

Inputs: $\mathbf{feature(t)}$, $\mathbf{t_{min}}$, $\mathbf{t_{gap}}$
Outputs: $threshold_{feature}(t = t_{final})$, $is\_spindle(t)$, $count\_spindles$

```
if feature(t)>threshold_feature(t) then                ▷ The threshold could be time dependent
    if hold==False then
        start=t;
    end if
    hold=True;
else if hold==True then
    gap=gap + 1;
    if gap>t_gap then
        gap=0;
        hold=False;
        if t − start>t_min then
            count_spindles=count_spindles + 1;
            is_spindle(start : t)=1;
        end if
    end if
end if
```

Features calculation

The short and long term ratio was compared with other features. The short-time and the ratio were predefined for calculation of each feature. ST is the length in samples of selected short-time, $LT = \frac{ST}{ratio}$ is the length in samples of selected long-time. $STA0(n)$ is the ratio feature using as long-time all past samples, $STA1(n)$ uses as short-time a single sample. $STA0(n)$, $STA1(n)$, and $Teager(n)$ pass through a moving average filter of length ST.

Root mean square (RMS) value

$$RMS(n) = \sqrt{\frac{\sum_{k=n-ST}^{n} x_\sigma(k)^2}{ST}}, \tag{5}$$

Short and long-time average ratio (STA/LTA)

$$STA0(n) = \frac{|x_\sigma(k)|}{\frac{1}{n}\sum_{k=0}^{n}|x_\sigma(k)|} \quad \Big| \quad STA1(n) = \frac{|x_\sigma(k)|}{\frac{1}{LT}\sum_{n=n-LT}^{n}|x_\sigma(k)|} \quad \Big| \quad STA/LTA(n) = \frac{\frac{1}{ST}\sum_{k=n-ST}^{n}|x_\sigma(k)|}{\frac{1}{LT}\sum_{k=n-LT}^{n}|x_\sigma(k)|}, \tag{6}$$

Relative spindle power (RSP)

This version of RSP is different of the based in Discrete Fourier Transform bins used by (***O'Reilly and Nielsen, 2015***).

$$RSP(n) = \frac{\sum_{k=n-ST}^{n} x_{\sigma}(k)^2}{\sum_{k=n-ST}^{n} x(k)^2}, \tag{7}$$

Teager-Kaiser energy operator

$$Teager(n-1) = (x(n-1)^2 - (x(n-2)x(n))), \tag{8}$$

Detection metrics

The detection metrics employed here are the same as used by (***O'Reilly and Nielsen, 2015***). An additional metric of distribution separation, d_{θ}, uses the probability distributions (pdf) of the feature signals in the search for a threshold value.

$$d_{\theta} = \frac{P\{\bar{y} \leq \theta\} - P\{y \leq \theta\} + P\{y > \theta\} - P\{\bar{y} > \theta\}}{P\{\bar{y} \leq \theta\} + P\{y > \theta\}} = \frac{2(P\{\bar{y} \leq \theta\} - P\{y \leq \theta\})}{1 + (P\{\bar{y} \leq \theta\} - P\{y \leq \theta\})}, \tag{9}$$

where θ is the threshold value, y is the distribution of the detection feature when there is a spindle and $\bar{y}$ the distribution of feature values when there is not a spindle. The best threshold, known the probability distributions f_y and $f_{\bar{y}}$, is which maximizes this metric. The maximum value is 1 and it is achieved only if the pdf's have disjointed domain and the threshold value is in the gap between them.

Results

Statistical analysis

The probability density function (pdf) of samples of an EEG signal is assumed to be *Gaussian* and reinforced by the histograms in Figure 1 C. Thus, the features calculations perform those operations to a normal aleatory variable. Then, the root mean square (RMS) should have a *Chi pdf*, short and long-time average (STA/LTA) and relative spindle power (RSP), as calculated here, should have a *Fractional Gamma pdf*, and the Teager-Kaiser energy operator should have a *Chi-squared pdf*.

The probability density function of samples of an EEG signal is Figure 1 D shows the histogram of the values of the features calculated from spindles and not spindles samples. The first 3 seconds of the synthetic signal were removed because the z-score is badly estimated for the first samples. The mean and standard deviation are still not consistent. The first 30 seconds, the first epoch, are removed from real signals for the same reason. The distance between the means of each distribution and the length of the tails say something about the classification difficulty using a single threshold. The metric d_{θ} allows the selection of a better threshold. However, it needs caution because the features depend on the length of the time window and the metrics are threshold dependent (Figure 2). Figure 1 E shows the classification performance with different thresholds. The most classical performance metrics have their best value at threshold values above the preferable threshold selected with d_{θ}.

Detection results

Figure 2 A presents ROC curves that are another perspective of the results of Figure 1 C. An additional short-time value and an additional ratio were included to show the different behavior of the detection performance. Figure 2 B shows the maximum value achieved for d_{θ} occurring with STA/LTA at the short-time window of 0.05 seconds and a ratio equal to 0.001.

Discussion

The proposed spindle detection method considers a single EEG channel and a single extracted feature. This consideration could be a weakness in typical EEG studies where multiple channels are recorded and more sophisticated analyses could be performed (***Mucarquer et al., 2019***). On

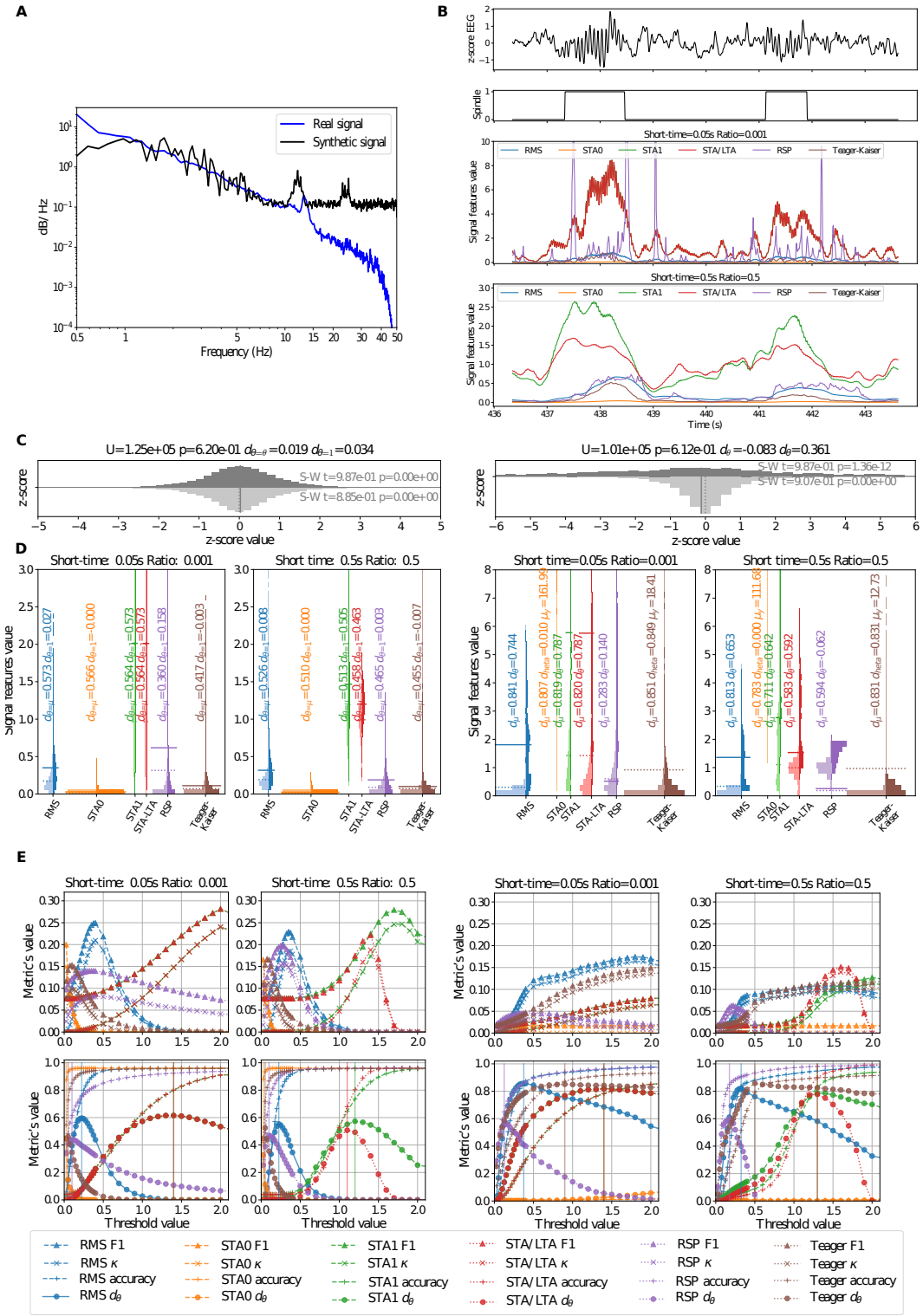

Figure 1. A. The spectrums of a real signal and a synthetic signal (Welch method, Hamming 10s windows with 80% of overlap). **B**. A portion of a real signal with two sleep spindles and the calculated features. **C**. Histogram of z-scored real signals in the left and synthetic signals in the right. Samples marked as spindles in a darker color and as not spindle in a lighter color. Mann-Whitney U test was performed with 500 randomly selected samples of each class to test if without any feature extraction there is possible that a sample of the not-spindles signal is lesser than a spindle sample. The t and p values of the Shapiro-Wilk test for normality are also annotated. **D**. Histograms of detection features from real signals in the left and from synthetic signals in the right. The filled line is the mean of features for spindle samples. The dotted line is the mean of not spindle samples (μ). The metric of distribution separation d_θ is annotated for θ=μ and θ=1. **E**. Detection metrics overall real signals in the left and overall synthetic signals in the right. STA/LTA and $STA1$ have similar behavior with small ratios. STA0 has very long tails for the synthetic signals. *In **C** y-axis for STA0 has a range of [0,80].

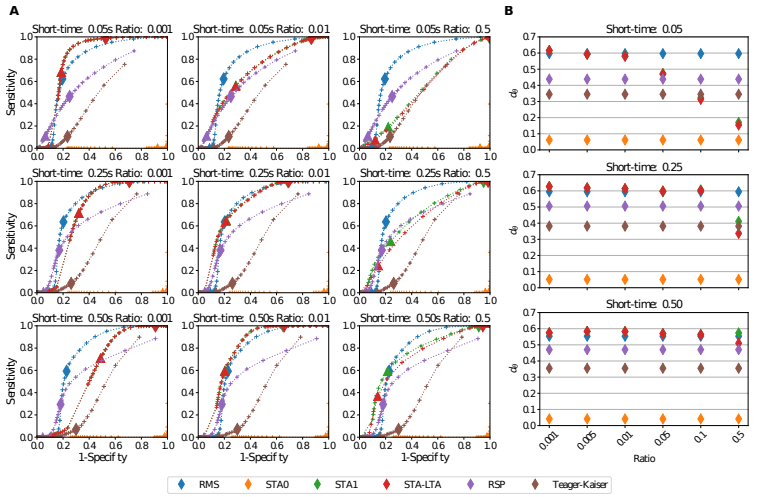

Figure 2. A. ROC curves overall real signals. Each row has different short-time value and every column a distinct ratio value. The triangle markers show the points in the ROC curves when the threshold ($\theta = 1.4$) is selected for the maximum value of d_θ for the STA/LTA feature. The diamonds markers are the points when the threshold ($\theta = 0.2$) is selected for the maximum value of d_θ for the RMS feature. Both thresholds selected at the smallest short time and the smallest ratio. **B**. Value of max$\{d_\theta\}$ along different thresholds for different short-times and ratios. Note that RMS, RSP and, Teager-Kaiser are ratio independent features but all plotted in all cases to compare with STA/LTA.

the other hand, it is a necessary alternative in studies where recording devices of few channels are available like the OpenBCI, Emotiv EPOC (***Xu and Zhong, 2018***) or single channel as the Neurosky MindWave device (***Torres et al., 2014***; ***Avendaño et al., 2018***). In the framework of educational research, these devices known as portable EEG technology (PEEGT) are not appropriate for their single-use. These are the present alternatives to include electrophysiological data (***Xu and Zhong, 2018***) in addition to other data and tools. Sleep research uses polysomnography for electrophysiological data acquisition, but PEEGT could be an advantage to include more subject samples (***Debellemaniere et al., 2018***).

The STA/LTA feature performs better with the smallest ratios for the real signals dataset. That does not occur in the synthetic signals, where the Teager-Kaiser energy operator features performs the better. This could indicate that there is a need for more pre-processing of the input signal to remove noise and other physiological artifacts not considered in the construction of the synthetic signals like ECG, EMG and, ocular movements. Without any other pre-processing another good performance feature is RMS.

The accuracy metric goes near to 1 with higher thresholds due to the best classification of True Negatives samples that are much higher in quantity than True Positive samples. Interestingly, Cohen-κ and F1 metrics have better values for higher thresholds than for the d_θ metric (Figure 1 E). Furthermore, all metrics are consistent in performance between cases. The ROC curves allow having another perspective than a simple value, although they represent the same information, and clarifies why the common choice of the RMS feature over another feature in the task of spindle detection. The maximum of the d_θ metric occurs at a similar threshold value for the RMS feature in any combination of short-times and ratios. The threshold value that gives the best d_θ metric for STA/LTA is more variable across cases (we do not show but the behavior of probability distributions in Figure 1 D explain it).

Funding

This work was supported by CONICYT grant BASAL FB0008 and CONICYT grant Doctorado Nacional Folio No 2118640.

References

Akram J, Peter D, Eaton D. A k-mean characteristic function for optimizing STA/LTA based detection of microseismic events. . 2019; .

Al-Salman W, Li Y, Wen P. Detecting sleep spindles in EEGs using wavelet fourier analysis and statistical features. Biomedical Signal Processing and Control. 2019; 48(February):80–92. https://doi.org/10.1016/j.bspc.2018.10.004, doi: 10.1016/j.bspc.2018.10.004.

Avendaño GO, Ballado AH, Cruz JCD, Bentir SAP, Camposanto JCB, Carreos AL, Domingo LAB, M Garcia KD. Sleep Onset Period Detection Using Slow Eyelid Movement (SEM) Through Eye Aspect Ratio with Electroencephalogram (EEG). In: *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM)*; 2018. p. 1–6. doi: 10.1109/HNICEM.2018.8666429.

Babadi B, Babadi B, McKinney SM, Tarokh V, Ellenbogen JM, Ellenbogen JM. DiBa: A Data-Driven Bayesian Algorithm for Sleep Spindle Detection. IEEE Transactions on Biomedical Engineering. 2012; 59(2):483–493. doi: 10.1109/TBME.2011.2175225.

Cairney SA, Ashton JE, Roshchupkina AA, Sobczak JM. A dual role for sleep spindles in sleep-dependent memory consolidation? Journal of Neuroscience. 2015; 35(36):12328–12330.

De Gennaro L, Ferrara M, Sleep spindles: an overview. Elsevier; 2003.

Debellemaniere E, Chambon S, Pinaud C, Thorey V, Dehaene D, Léger D, Chennaoui M, Arnal PJ, Galtier MN. Performance of an Ambulatory Dry-EEG Device for Auditory Closed-Loop Stimulation of Sleep Slow Oscillations in the Home Environment. Frontiers in Human Neuroscience. 2018; https://doi.org/10.3389/fnhum.2018.00088, doi: 10.3389/fnhum.2018.00088.

Devuyst S, Dutoit T, Stenuit P, Kerkhofs M. Automatic sleep spindles detection - Overview and development of a standard proposal assessment method. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*; 2011. p. 1713–1716. doi: 10.1109/IEMBS.2011.6090491.

Devuyst S, Dutoit T, The DREAM sLeep spindle database; 2011. www.tcts.fpms.ac.be/~devuyst/Databases/DatabaseSpindles/.

Liu MY, Huang A, Huang NE. Evaluating and Improving Automatic Sleep Spindle Detection by Using Multi-Objective Evolutionary Algorithms. Frontiers in human neuroscience. 2017; 11.

Mporas I, Korvesis P, Zacharaki EI, Megalooikonomou V. Sleep spindle detection in EEG signals combining HMMs and SVMs. Communications in Computer and Information Science. 2013; 384:138–145. doi: 10.1007/978-3-642-41016-1_15.

Mucarquer JA, Prado P, Escobar MJ, El-Deredy W, Zañartu M. Improving EEG Muscle Artifact Removal With an EMG Array. IEEE Transactions on Instrumentation and Measurement. 2019; (1):1–10. doi: 10.1109/tim.2019.2906967.

O'Reilly C, Nielsen T. Automatic sleep spindle detection: benchmarking with fine temporal resolution using open science tools. Frontiers in human neuroscience. 2015; 9.

Purcell S, Manoach D, Demanuele C, Cade B, Mariani S, Cox R, Panagiotaropoulou G, Saxena R, Pan J, Smoller J, et al. Characterizing sleep spindles in 11,630 individuals from the National Sleep Research Resource. Nature Communications. 2017; 8.

Sara SJ. Sleep to remember. Journal of Neuroscience. 2017; 37(3):457–463.

Torres FA, Sánchez C, Baus KP. Adquisición y análisis de señales cerebrales utilizando el dispositivo MindWave. Maskana. 2014; 0(0):83–93.

Ulloa S, Estevez PA, Huijse P, Held CM, Perez CA, Chamorro R, Garrido M, Algarín C, Peirano P. Sleep-Spindle Identification on EEG Signals from Polysomnographic Recordings using Correntropy. Conf Proc IEEE Eng Med Biol Soc 2016. 2016; p. 3736–3739.

Xu J, Zhong B. Review on portable EEG technology in educational research. Computers in Human Behavior. 2018; 81:340–349. https://doi.org/10.1016/j.chb.2017.12.037, doi: 10.1016/j.chb.2017.12.037.

Zhao R, Sun J, Zhang X, Wu H, Liu P, Yang X, Qin W. Sleep spindle detection based on non-experts: A validation study. PloS one. 2017; 12(5):e0177437.