Spindle-AI: Sleep Spindle Number and Duration Estimation in Infant EEG

Lan Wei, Soraia Ventura, Sean Mathieson, Geraldine B. Boylan, Madeleine Lowery and Catherine Mooney

***Abstract*— Objective: Sleep spindle features show developmental changes during infancy and have the potential to provide an early biomarker for abnormal brain maturation. Manual identification of sleep spindles in the electroencephalogram (EEG) is time-consuming and typically requires highly-trained experts. Automated detection of sleep spindles would greatly facilitate this analysis. Research on the automatic detection of sleep spindles in infant EEG has been limited to-date. Methods: We present a random forest-based sleep spindle detection method (Spindle-AI) to estimate the number and duration of sleep spindles in EEG collected from 141 ex-term born infants, recorded at 4 months of age. The signal on channel F4-C4 was split into a training set (81 ex-term) and a validation set (30 ex-term). An additional 30 ex-term infant EEGs (channel F4-C4 and channel F3-C3) were used as an independent test set. Fourteen features were selected for input into a random forest algorithm to estimate the number and duration of spindles and the results were compared against sleep spindles annotated by an experienced clinical physiologist. Results: The prediction of the number of sleep spindles in the independent test set demonstrated 93.3% to 93.9% sensitivity, 90.7% to 91.5% specificity, and 89.2% to 90.1% precision. The duration estimation of sleep spindle events in the independent test set showed a percent error of 5.7% to 7.4%. Conclusion and Significance: Spindle-AI has been implemented as a web server that has the potential to assist clinicians in the fast and accurate monitoring of sleep spindles in infant EEGs.**

***Index Terms*— EEG, infants, sleep spindles, Spindle-AI**

I. Introduction

Sleep spindles are an indicator of the development and integrity of the central nervous system in infants [1]. They were first described by Loomis *et al.* [2] as rhythmic 12-14 Hz oscillations which last 0.5 to 3 seconds with a waxing and waning shape [3]. They have been observed clearly in EEG during stages N2 and N3 [4] from the 4^{th} week post-term and are present in the EEG of all infants by nine weeks post-term [5]. Sleep spindles have been shown to change with aging [6], [7] possibly reflecting maturation changes such as synapses generation and elimination, and myelination [8]. For example, when sleep spindles first appear, they are of low amplitudes showing, however, an ascendant trajectory [4]. Spindle spectral power (measured as sigma power over central and occipital channels) increases in early childhood from 2 to 5 years [9]. Within participants ranging from 4 to 24 years old, a decrease of the spectral power over the frontal and centroparietal areas was seen [10]. It is also during adulthood that amplitudes stabilize, meaning that, for peak-to-peak analysis of the automatically detected sleep spindles, groups of 20-30 year old showed higher amplitudes than the 41-59 and the 60-73 year old participant groups, particularly in the frontal left region; but no statistically significant differences were seen between the latter groups in any of the areas spindles were studied: the frontopolar, frontal, central, parietal and occipital areas [11]. This evolution is also evident in the frequency domain, fast spindles predominate during infancy where their density peaks in groups from 4-12 months old, and decrease after that; while slow spindles tend to appear after the first year and keep increasing in density across the different age groups studied at least until the age of 25-48 months [12]. Mean frequency of sleep spindles peak at 4 to 12 months of age decreasing after that until at least 25 to 48 months [12]. Spindle duration decreases across infancy and early childhood and increases again after 3 years of age [9], [13], [14]. Sleep spindle density decreases from the second half of the 9^{th} month up to 1 year and 8.4 months, then shows a trend of increasing density up to 11 years [13]. Sleep spindle parameters might estimate neurodevelopment as their variations evolve with aging [13]. The rapid development of the infant brain, and the variability in sleep spindles makes infant sleep spindle detection challenging.

The visual detection of sleep spindles in EEG is a laborious and time-consuming task. Automated detection of sleep spindles would reduce the burden associated with the analysis of large datasets and facilitate more rapid identification of sleep abnormalities and an objective means to quantify spindle features. Previous studies have used various methods to measure the duration and number of sleep spindles automatically in adult sleep EEG [15]–[21]. However, research on the automatic detection of sleep spindles in infant EEGs to-date has been limited. To the best of our knowledge, there have only been two previously published studies on the automatic estimation of the number of sleep spindles in infant EEGs. In the first, Held *et al.* [22] presented an automated method to estimate the number of sleep spindles which was trained on three and tested on two infants. They achieved a sensitivity of 87.7% and an 8.1% false-positive rate. This method combined two different approaches: detection criteria on the sigma-band filtered EEG signal, including the application of fuzzy thresholds, and mimicking current procedures for manual identification of sleep spindles in infant EEG. However, only the number of sleep spindle events was estimated; sleep spindle duration was not considered. The second approach for estimating the number of sleep spindles in long EEG recordings, developed by Estevez *et al.* [23], used a Merge Neural Gas algorithm and was trained on a single infant, and tested on another achieving 62.9% sensitivity. In their work, the EEG recording was divided into 0.512s epochs, and a human expert labelled each epoch as containing sleep spindles or not. Epochs in which sleep spindles were contained in only part of the epoch were labelled as a sleep spindle event.

Manuscript received November 18, 2020; revised February 19, 2021, April 15, 2021 and June 22, 2021; accepted July 7, 2021. This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number 12/RC/2272 and 15/SP/3091. Part of this funding comes from our SFI SPOKES industry partner JOHNSON JOHNSON CONSUMER SERVICES EAME LTD.

Lan Wei is with the UCD School of Computer Science, University College Dublin, Dublin, Ireland (e-mail: lan.wei@ucdconnect.ie).

Soraia Ventura, Sean Mathieson and Geraldine B. Boylan are with the Department of Paediatrics & Child Health, University College Cork, Cork, Ireland and also with INFANT Research Centre, Cork University Hospital, Cork, Ireland. (e-mail: soraia.ventura@ucc.ie; sean.mathieson@ucc.ie; G.Boylan@ucc.ie).

Madeleine Lowery is with the UCD School of Electrical & Electronic Engineering, University College Dublin, Dublin, Ireland. (e-mail: madeleine.lowery@ucd.ie).

Catherine Mooney is with the UCD School of Computer Science, University College Dublin, Dublin, Ireland. (e-mail: catherine.mooney@ucd.ie).

However, if a sleep spindle was divided across different epochs, the sleep spindle was treated as two or more sleep spindles events. Moreover, in this way, the duration (start time and end time) of the sleep spindles can not be accurately estimated. There thus remains a need for an automated method of spindle detection in infant EEG that can accurately estimate the number and duration of spindle events.

In this study we develop a random forest-based sleep spindle detection method (Spindle-AI) to estimate the number and duration of sleep spindles in the EEGs of ex-term born infants recorded at four-month of age. The random forest algorithm is one of the most successful modern general-purpose classification algorithms [25] and has been widely used to identify sleep events in sleep EEG recordings [20], [26]–[28]. Moreover, it is generic enough to be applied to large-scale problems, easily adaptable to a variety of special learning tasks, and returns a measure of feature importance [25]. The approach combines several randomized decision trees and aggregates their predictions by averaging. The study builds on our previous preliminary work to develop a method to detect sleep spindles in infant EEGs [24]. Spindle-AI has been implemented as a web server freely available for academic use at http://lisda.ucd.ie/Spindle-AI/. Spindle-AI estimates the number of sleep spindles and identifies the specific time and duration of occurrence of the sleep spindle events, this has the potential to assist clinicians in the monitoring of sleep spindles in EEGs of infants faster than current methods.

II. Methods

The Spindle-AI method was developed and tested on EEG data recorded from 141 infants on EEG channels F4-C4 and F3-C3. Fourteen features were selected for input to a random forest algorithm. Synthetic Minority Over-sampling Technique (SMOTE) was used to balance the dataset. The model was trained using 81 infant EEG data as inputs to the random forest algorithm. After that, the signal was post-processed using sleep spindle detection criteria to estimate the number and duration of sleep spindles. An overview of Spindle-AI is presented in Fig. 1.

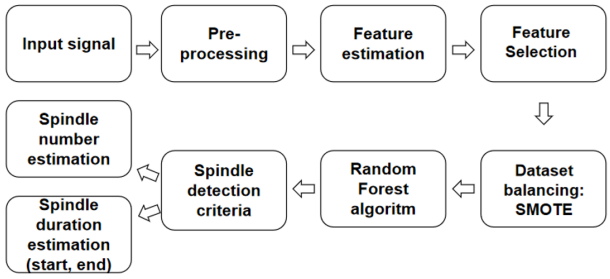

Fig. 1. Overview of Spindle-AI. Infants' sleep EEG data on channel F4-C4 were used to develop the random forest-based method. Infants' sleep EEG data on channel F4-C4 and channel F3-C3 were used as an independent test set. Features were estimated after removing artefacts, and feature selection was performed using a random forest-based wrapper method. SMOTE was applied to balance the dataset. The method was then trained using 81 infant EEG data as inputs to the random forest algorithm. The number, start time and end time of sleep spindles are estimated by post-processing the signals according to the sleep spindles detection criteria.

A. Participant details

Ethical approval was granted from the Clinical Research Ethics Committee of the Cork Teaching Hospitals, Cork, Ireland and written consent from parents or guardians of the infants included in the study was obtained. A cohort of healthy full-term infants (n=181) was recruited soon after birth at Cork University Maternity Hospital (CUMH). Inclusion criteria were gestational age higher than 37 weeks, being healthy and singleton. EEG data were recorded from sleeping infants at four months, with each infant's EEG recorded for around two hours. Three datasets were excluded due to abnormal EEG. A further 37 datasets were excluded due to software incompatibilities. Therefore, 141 ex-term infants (i.e. infants born after 37 weeks but before 42 weeks of gestation) EEG were included in this study.

B. Data collection

EEG data were recorded using a 31-channel polygraph system (Lifelines, UK) that included 21 EEG electrodes (FP2, FP1, F8, F7, F4, F3, FZ, A2, A1, T4, T3, C4, C3, CZ, T8, T7, P4, P3, PZ, O2, O1), ground and reference electrodes. Two electrooculogram (EOG) channels (below the outer canthus of the left eye and above the outer canthus of the right eye) monitored rapid and slow eye movements, chin tonicity was recorded using surface electromyography (EMG). Separate electrodes were applied for electrocardiogram (ECG), and a movement sensor was placed on the abdominal region for the recording of respiration. EEG, EMG, ECG, EOG data and movement sensor were recorded at a sampling rate of 500 Hz. Sleep spindles of ex-term infants were annotated on channel F4-C4 and channel F3-C3 by an experienced clinical physiologist as the gold standard. In the following text we refer to the sleep spindles in channel F4-C4 as "R-Spindle" and the sleep spindles in channel F3-C3 as "L-Spindle". The number and duration of the sleep spindles identified by the clinical physiologist are presented in Table I.

TABLE I
NUMBER AND DURATION OF SLEEP SPINDLES ANNOTATED BY AN EXPERIENCED CLINICAL PHYSIOLOGIST

	F4-C4	F3-C3
Infants number	141	141
Average number of sleep spindles per infant	167	155
Total number of sleep spindles	23,520	21,815
Total duration of sleep spindles (s)	67,997	66,320
Total duration of non-sleep spindles (s)	509,535	511,212

C. Data pre-processing and feature estimation

The EEG data were processed at the original sampling frequency of 500 Hz, a 50 Hz notch filter was applied to remove power line interference from the EEG recordings, and the DC offset was removed from each channel. The pre-processed EEG signals detected on channels F4-C4 and F3-C3 were segmented into 0.5s epochs with 0.25s overlap for feature estimation. The length of 0.5 seconds was chosen as this is the minimum required length of a sleep spindle [3]. EEG recordings from channel F4-C4 from 81 ex-term infants were used for training and 30 ex-term infants were used for validation. An additional 30 EEG recordings from channel F4-C4 and F3-C3 of ex-term infants were used for independent testing of the method. A selection of 43 time and frequency domain features previously used in EEG event identification, including sleep stage classification and seizure detection, were identified from the literature and estimated for each epoch. A list of the features estimated is provided in Table II.

In the time domain, the root mean square, the mean absolute amplitude, skewness, and kurtosis, of the pre-processed EEG signals were calculated for each epoch. In addition, max and min absolute amplitudes as well as max–min difference of amplitude (max-min difference) of the pre-processed EEG signals were estimated. The Teager-Kaiser energy operator (TKEO) is a nonlinear energy tracking

TABLE II
TIME AND FREQUENCY DOMAIN FEATURES ESTIMATED FROM EACH 0.5S EEG EPOCH OF PRE-PROCESSED EEG SIGNAL

Domain	Features	Reference
Time	SD, **RMS** and **mean absolute amplitudes**	[29], [30]
	Max and min absolute amplitudes	[29]–[31]
	Max–min difference of amplitude	[29]–[31]
	Skewness and Kurtosis	[32]
	Hjorth activity, complexity, mobility	[31], [32]
	Fractal Dimension	[33]
	Symmetry and Anti-symmetry	[22], [31]
	Number of peaks	[31]
	Mean and standard deviation TKEO values	[31]
Frequency	**Mean frequency**	[22], [29]
	Integral of the full band power (0-500 Hz)	[33]
	Mean value of the envelope in sigma band (12.5-15 Hz) and sleep spindle band (10.5-16 Hz)	[34]
	Mean absolute amplitudes in sigma and sleep spindle bands	[15], [35]
	Mean absolute amplitudes in delta (0-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz) bands	[36]
	Relative and absolute power in sigma and sleep spindle bands	[31]
	Relative and absolute power in delta (0-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz) bands	[37]
	Sigma Index, alpha band ratio and sleep spindle band ratio	[20], [35]

SD: standard deviation; RMS: root mean square; TKEO: Teager–Kaiser energy operator

operator capable of determining the instantaneous energy of a non-stationary signal, which has been applied to detect abrupt changes in biological signals [38]. The mean and standard deviation of the TKEO value in each epoch were also estimated.

$$\text{TKEO[n]} = x[n]^2 - x[n-1]x[n+1] \tag{1}$$

Where: x[n] is the n^{th} sample, x[n-1] is the $(n-1)^{th}$ sample and x[n+1] is the $(n+1)^{th}$ sample of the pre-processed EEG signal in the epoch.

The signal symmetry (Sym) and anti-symmetry (Antisym) were estimated [22], [31] along with the EEG signal complexity which was quantified using fractal dimension (FD) [33], [39] and Hjorth parameters (Mobility and Complexity)) [31], [32].

$$\text{Sym} = \frac{\sum_{i=0}^{N/2} f_+[i]}{N(max_{i=0}^{N/2}(f_+[i]))^2}$$

$$\text{where: } \mathrm{f}_+[i] = \frac{x[N/2+i]+x[N/2-i]}{2} \tag{2}$$

$$\text{Antisym} = \frac{\sum_{i=0}^{N/2} f_-[i]}{N(max_{i=0}^{N/2}(f_-[i]))^2}$$

$$\text{where: } \mathrm{f}_-[i] = \frac{x[N/2+i]-x[N/2-i]}{2} \tag{3}$$

$$FD = \frac{log_{10}^{N}}{log_{10}^{N} + log_{10}^{N/(N+0.4\delta))}} \tag{4}$$

$$\text{Mobility} = \sqrt{\frac{\text{Var}(\dot{x})}{\text{Var}(x)}} \tag{5}$$

$$\text{Complexity} = \frac{\text{Mobility}(\dot{x})}{\text{Mobility}(x)} \tag{6}$$

Where: x[N/2+i] is the $(N/2+i)^{th}$ sample and x[N/2-i] is the $(N/2-i)^{th}$ sample of the processed EEG signal in the epoch; N is the number of samples in each epoch; and δ is the number of sign changes in the signal derivative in that epoch; $\dot{x}$ is the time derivative of the pre-processed EEG signal x, and Var (x) is the variance of x estimated for that epoch.

The EEG data were filtered using a 4^{th} order Butterworth filters (IIR) within the frequency bands of interest: delta (0-4 Hz), theta (4–8 Hz), alpha (8–12 Hz), sigma (12.5-15 Hz), sleep spindle (10.5-16 Hz) [40] and beta (12–30 Hz). The mean absolute amplitude, signal envelope (estimated using the Hilbert transform), relative power in each band and absolute power in each band was estimated for each epoch. The mean frequency of each 0.5 s epoch was also estimated [15], [31], [34]–[37].

In addition to these classical metrics, we included the sigma index (Sigmaindex) [20], alpha band ratio and sleep spindle band ratio which have proven successful for adult sleep spindle detection in [20], [35].

$$\text{Sigmaindex} = \frac{mean(|F3(x)|)}{mean(|F1(x)|)) + mean(|F2(x)|)} \tag{7}$$

Where F1(x), F2(x) and F3(x) represent the pre-processed EEG signals filtered in the 4-10 Hz, 20-40 Hz and 12.5-15 Hz bands, respectively.

The alpha band ratio is estimated as the ratio of the root mean square (RMS) amplitude in alpha band (8-12 Hz) compared to total RMS amplitude of pre-processed signal in each epoch. The sleep spindle band ratio is estimated as the ratio of the RMS amplitude in sleep spindle band (10.5-16 Hz) compared to total RMS amplitude of pre-processed signal in each epoch.

D. Feature selection

The feature selection algorithm was built as a wrapper method around a random forest classifier. The mean precision loss and standard deviation were calculated, and a ranking of features was provided as output. In addition, Pearson's correlation coefficient [42] between all pairs of features was measured to identify highly correlated pairs of features.

E. Dataset balancing

The number and duration of sleep spindle events in the dataset were substantially less than the number and duration of non-sleep spindle events resulting in a class imbalance problem that can make training a machine learning algorithm challenging [44]. To address this, the Synthetic Minority Oversampling Technique (SMOTE) [45] was used to balance the data. SMOTE is a method of oversampling, in which the minority class is oversampled by creating 'synthetic' samples in the feature space. Synthetic data points are generated by following the line of segments connecting randomly chosen neighbours from k nearest neighbours according to the required excessive sampling quantity. In addition, the random state is used as a seed to the random number generator, which ensures that the splits generated are reproducible. In this work, we chose k = 5, and the random state = 2.

F. Classification algorithm of Spindle-AI

Before selecting the random forest algorithm, it was first benchmarked against three other algorithms (Multilayer Perceptron, Naive bayes and eXtreme Gradient Boosting). In preliminary testing the random forest algorithm had higher performance than the other algorithms (see APPENDIX Table VIII), it has a shorter processing time than the Multilayer Perceptron, is less likely to overfit a training dataset than the XGBoost algorithm [46] and has the advantage that

it can return the importance of the features after training the model. Random forest is an integrated technique that contains many decision trees and classifies by voting on the weakest unbiased classifiers [41]. The decision trees are executed on different bagging instances of the training set and the classification accuracy loss is caused by the feature importance measure being randomly arranged as the feature value.

The random forest classifier was implemented using the sklearn library [47] within the Python 3 environment. The structure of the random forest classifier is shown in Fig. 2. Three parameters were optimized; n-estimators, which is the number of trees in the forest; min-samples-split which is the minimum number of samples required to split an internal node; and min-samples-leaf which is the minimum number of samples required to be at a leaf node. These parameters (n-estimators, min-samples-split, and min-samples-leaf) were optimized based on the performance of the validation set, to improve the performance of the method for the estimation of sleep spindles in the EEG recordings. The validation set was found to achieve the best performance for n-estimators = 100, min-samples-split = 120, and min-samples-leaf = 20.

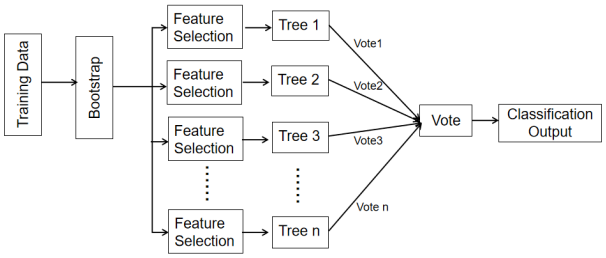

Fig. 2. The structure of a random forest classifier. Random Forest is an ensemble learning method that combines several randomized decision trees and aggregates their predictions by averaging [25].

G. Sleep spindle detection criteria: sleep spindle number and duration estimation

EEG typically contains artefacts which may interrupt and mask the sleep spindle trace [48]. To overcome this, consecutive sleep spindles estimated by the random forest algorithm with an interval less than 1 second were grouped together, and their duration was extended from the start time of the first component to the end time of the last component. Rodenbeck *et al.* [3] proposed that the length of 0.5 seconds is the minimum required length of a sleep spindle [3], therefore, if the duration of a sleep spindle identified by the random forest algorithm was less than 0.5 seconds it was relabelled as a non-sleep spindle event. After this initial post-processing step, the number of sleep spindles in the EEG recordings, and the start, end times and duration of each estimated sleep spindle event were estimated.

H. Performance evaluation

Sleep spindle number estimation: Sleep spindle number estimation: The sensitivity (Sens), specificity (Spec) and precision (Prec) of Spindle-AI in estimating the number of sleep spindle events was evaluated. Matthews correlation coefficient (MCC) was used as an additional evaluation metric due to the imbalanced nature of the dataset. The evaluation metrics were estimated as follows:

$$Sens = \frac{TP}{TP + FN} \times 100\% \tag{8}$$

$$Spec = \frac{TN}{TN + FP} \times 100\% \tag{9}$$

$$Prec = \frac{TP}{TP + FP} \times 100\% \tag{10}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{11}$$

where:

- True Positive (TP): Annotated by experts as a sleep spindle and predicted as a sleep spindle
- True Negative (TN): Not annotated by experts as a sleep spindle and not predict as a sleep spindle
- False Negative (FN): Annotated by experts as a sleep spindle and not predict as a sleep spindle
- False Positive (FP): Not annotated by experts as a sleep spindle and predicted as a sleep spindle

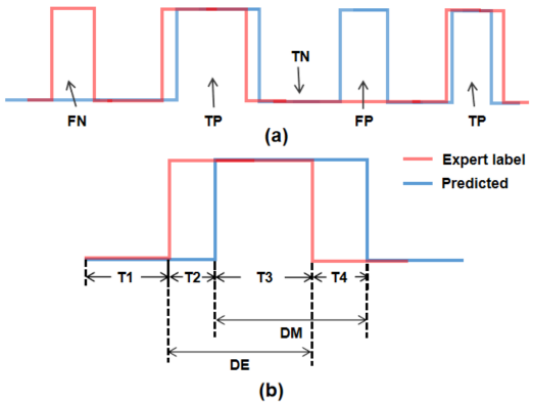

Fig. 3. Evaluation metrics used for: (a) sleep spindle number estimation; and (b) sleep spindle duration estimation.

Sleep spindle duration estimation: Sleep spindle duration estimation: The duration of sleep spindles estimated by Spindle-AI in the EEG recording was evaluated using the percent error on individual infant EEG. In addition, true negative rate (TNR), recall and Intersection over Union (IoU) were also used to evaluate the performance of duration estimation [18], [20], [21] as follows, Figure 3:

$$\text{Percent error} = \frac{|DE - DM|}{DE} \times 100\% \tag{12}$$

$$Recall = \frac{T3}{T2 + T3} \times 100\% \tag{13}$$

$$TNR = \frac{T1}{T1 + T4} \times 100\% \tag{14}$$

$$IoU = \frac{T3}{T2 + T3 + T4} \tag{15}$$

where DE is the duration of the sleep spindles identified by expert readers, DM is the duration of the sleep spindle estimated by Spindle-AI. T1 is the duration of the period between non-sleep spindle events labelled as non-sleep spindle event by both the expert reader and Spindle-AI, T2 is the duration of the period labelled as a sleep spindle event by the expert reader but labelled as a non-sleep spindle event by Spindle-AI, T3 is the duration of the period between sleep spindle events labelled as sleep spindle event by both the expert reader and Spindle-AI, T4 is the duration of the period labelled as a non-spindle event by the expert reader but labelled as a sleep spindle event by Spindle-AI.

III. Results

A. Feature Selection

In this study, the 25 top ranking features estimated by the random forest feature selection algorithm were chosen. In addition, Pearson's

correlation coefficient (corr) [42] between all pairs of features was measured to identify highly correlated pairs of features (absolute value of corr greater than 0.80). The pairs of features that were found to be highly correlated are presented in Table III. Therefore, signal envelope, max-min difference, symmetry, mobility, and number of peaks were removed. We further reduced the remaining 20 features by removing them individually, and reevaluating the performance of the random forest using MCC [43] on the validation set. An additional six features were removed (TKEO amplitude, variance, skewness, beta power, complexity, and fractal dimension) as removing them caused a drop in MCC < 0.0001, leaving a total of 14 features:

1) The mean absolute amplitude;
2) Root mean squared value;
3) Mean frequency;
4) Sigma index;
5) Alpha band ratio;
6) Sleep spindle band ratio;
7) Mean absolute amplitude in sigma band;
8) Mean value of the envelope in sigma band;
9) Relative power in sigma band;
10) Absolute power in sigma band;
11) Mean absolute amplitudes in sleep spindle band;
12) Mean value of the envelope in sleep spindle band;
13) Relative power in sleep spindle band;
14) Absolute power in sleep spindle band;

TABLE III
HIGHLY CORRELATED FEATURES.

Feature 1	Feature 2	Corr
Root mean square	Signal envelope	0.97
Variance	Max-min difference	0.85
Mean absolute amplitude	Symmetry	0.98
Mean frequency	Mobility	0.87
Fractal dimension	Number of peaks	-0.98

Corr: Pearson's correlation coefficient

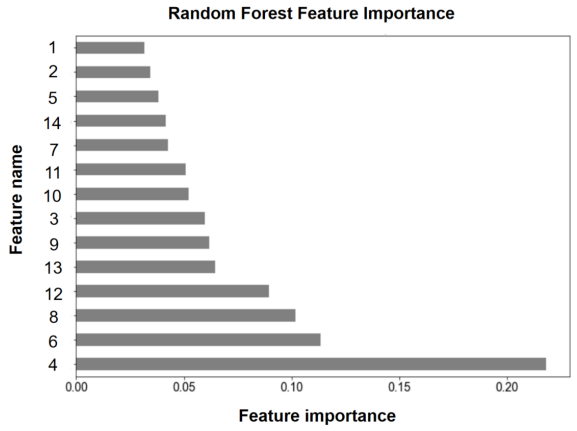

Fig. 4. A random forest feature importance plot of the selected features in the training dataset. Feature importance ranks the features by their contribution to the prediction of the sleep spindle events.

B. Estimation of the number of sleep spindles

Table IV summarizes the performance of the Spindle-AI method at estimating the number of R-Spindles and L-Spindles in the ex-term dataset. For the ex-term test set in channel F4-C4, 3,979 of the 4,263 R-spindle annotated by experts were correctly identified by Spindle-AI with 93.3% sensitivity, 91.5% specificity and precision of 90.1% (Table IV). We also evaluated the performance of the method on channel F3-C3 (L-Spindle) of the test set, which was not used for training. Experts labelled 4,268 L-spindles in the 30 ex-term infants. 4,491 L-spindles were predicted by the method, 4,006 of these were in agreement with the expert annotations, giving a 93.9% sensitivity, 90.7% specificity and 89.2% precision. Additionally, the test set yielded an MCC of 0.843 - 0.846 on the ex-term infant EEG demonstrating that the method performs well at identifying both the negative (non-sleep spindle) and the positive (sleep spindle) events. Table VI presents a comparison of the sensitivity, specificity and precision of Spindle-AI with recent studies on sleep spindle number estimation in EEG recordings. Figure 5 presents the sleep spindle detection criteria.

C. Estimation of sleep spindle duration

Table V summarizes the performance of Spindle-AI at estimating the duration of sleep spindles in the training set, validation set and test set for ex-term infant EEG recordings. The mean duration of each sleep spindle labelled by experts in the test set was 3.16s ($\pm$ 0.57s) on channel F4-C4 and 3.04s ($\pm$ 0.59s) on channel F3-C3. Compared with the expert annotation, Spindle-AI estimated the mean duration of sleep spindles as 3.28s ($\pm$ 0.61s) and 3.22s ($\pm$ 0.64s) on channel F4-C4 and F3-C3, respectively. Spindle-AI yielded a percent error of 5.7% (channel F4-C4) and 7.4% (channel F3-C3) in the test set. Table VII presents the results of the Spindle-AI method compared with those of recent studies on sleep spindle duration estimation in EEG recordings. The results for estimating the duration of sleep spindles by Spindle-AI on the test set show high recall (89.7%) and TNR (97.8%) in channel F4-C4 (R-spindle). Spindle-AI also performed well on channel F3-C3 (L-Spindle) of the ex-term infants in the test set, which were not used in training (90.5% recall and 97.3% TNR). The results show that Spindle-AI can generalize to both channels of F4-C4 and F3-C3 in ex-term born infants.

D. Implementation

Spindle-AI has been implemented as a web server and is freely available for academic use at http://lisda.ucd.ie/Spindle-AI/. The user can choose the sample frequency of their data and submit a CSV file that contains a single-channel EEG signal. Spindle-AI will predict if an event is a sleep spindle events and then return the start time and end time of each predicted sleep spindle event. Additionally, Spindle-AI returns the total number of sleep spindle events detected in the EEG.

IV. DISCUSSION

In this study, we present a random forest-based sleep spindle detection method, Spindle-AI, to estimate the number and duration of sleep spindles in infant EEG. The random forest algorithm can return a measure of feature importance, which is essential to understand the decision making of the algorithm. However, some artefacts of EEG mask the sleep spindle trace, which may increase the FP events (not annotated by experts as a sleep spindle, but predicted as a sleep spindle, see Figure 5C). Therefore, we combined the random forest algorithm with sleep spindle detection criteria to develop the final Spindle-AI method (Fig. 5D).

According to the performance on the validation set, we selected the 14 features that gave the best performance on the validation set to develop the Spindle-AI method. As can be seen from Figure 5E, sleep spindle events are characterized by an increase in signal power in the frequency range between 10 Hz and 15 Hz, and at frequencies around 30 Hz when compared with non-sleep spindle events. Consequently, the sigma index is important for distinguishing between the sleep

TABLE IV

PERFORMANCE OF SLEEP SPINDLE NUMBER ESTIMATION BY SPINDLE-AI ON THE TRAINING, VALIDATION, AND TEST SETS.

	Dataset	Actual	Automatic	TP	FN	FP	TN	Sens (%)	Spec(%)	Prec (%)	MCC
	Training set (n=81)	13,463	13,499	12,257	1,206	1,242	14,422	91.0	92.1	90.8	0.831
R-SS	Validation set (n=30)	4,304	4,640	4,050	254	590	4,982	94.1	89.4	87.3	0.830
	Test set (n=30)	4,263	4,418	3,979	284	439	4,736	93.3	91.5	90.1	0.846
L-SS	Test set (n=30)	4,268	4,491	4,006	262	485	4,726	93.9	90.7	89.2	0.843

R-SS: Channel F4-C4/R-Spindle; L-SS: Channel F3-C3/L-Spindle; Actual: The sleep spindles annotated by experts; Automatic: The sleep spindles predicted by Spindle-AI.

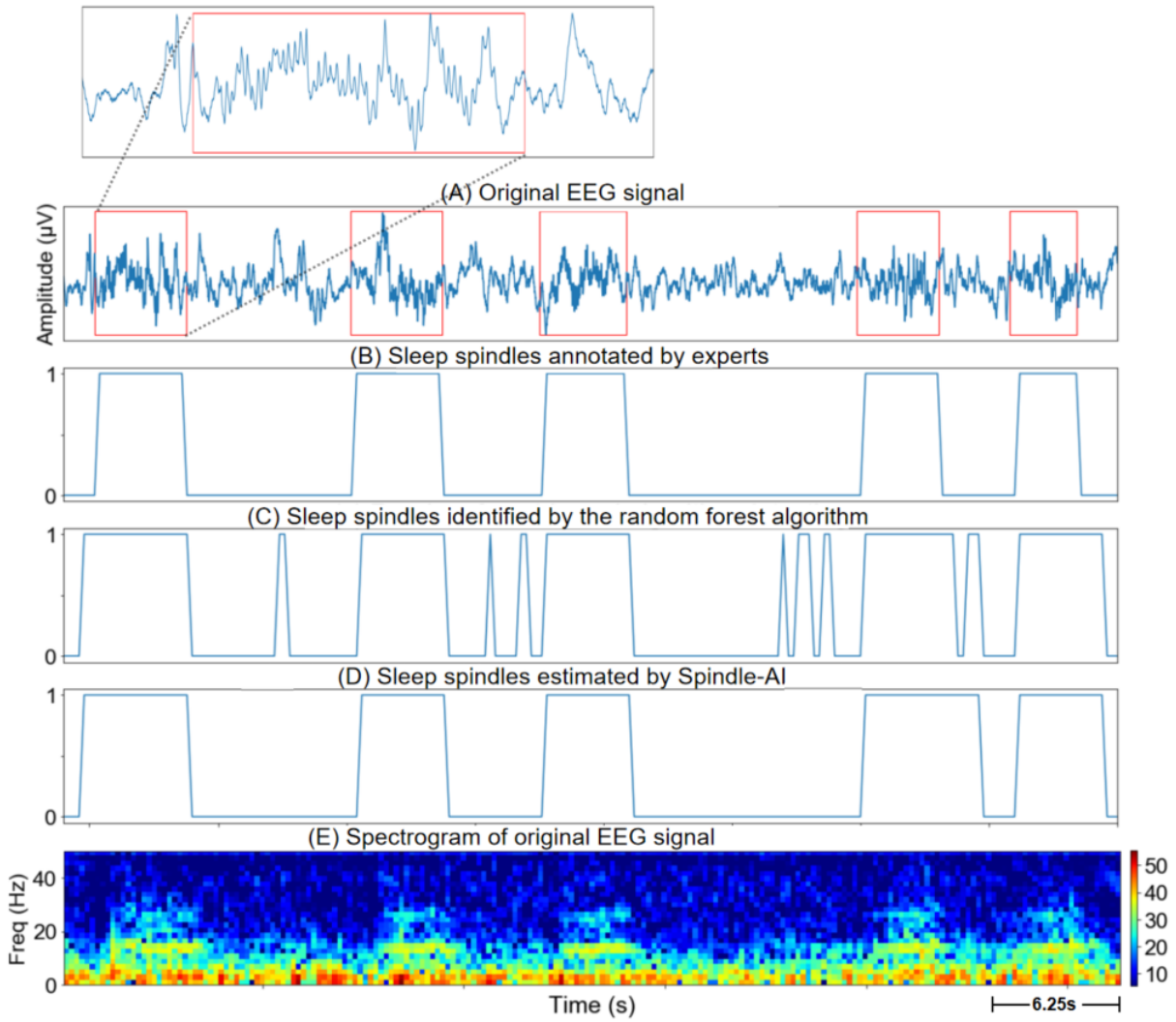

Fig. 5. Sleep spindle detection criteria: (A) original EEG signal (the signal in the red window indicates the presence of a sleep spindle event); (B) Sleep spindles annotated by experts; (C) Sleep spindles identified by the random forest algorithm (before application of the sleep spindle detection criteria); (D) Sleep spindles estimated by Spindle-AI; (E) Spectrogram of original EEG signal (the signal in yellow and red blocks around the frequency of 13 Hz indicate the presence of sleep spindle events).

TABLE V

PERFORMANCE OF SLEEP SPINDLE DURATION ESTIMATION BY SPINDLE-AI ON THE TRAINING, VALIDATION, AND TEST SETS.

	Dataset	Mean act dur (± std) (s)	Mean est dur (± std)	Mean of error (± std) (s)	Percent error (± std) (%)
	Traing set (N=81)	3.08 (± 0.60)	3.20 (± 0.60)	0.19 (± 0.13)	6.2 (± 4.3)
R-SS	Validation set (N=30)	3.07 (± 0.52)	3.21 (± 0.53)	0.22 (± 0.16)	7.4 (± 5.6)
	Test set (N=30)	3.16 (± 0.57)	3.28 (± 0.61)	0.18 (± 0.13)	5.7 (± 4.0)
L-SS	Ex-term (N=30)	3.04 (± 0.59)	3.22 (± 0.64)	0.22 (± 0.15)	7.4 (± 5.2)

R-SS: Channel F4-C4/R-Spindle; L-SS: Channel F3-C3/L-Spindle; Act dur: actual sleep spindle duration annotated by experts; Est dur: estimated sleep spindle duration by Spindle-AI.

TABLE VI
PREVIOUS WORK ON SLEEP SPINDLE NUMBER ESTIMATION.

Ref	Subjects	N	SS	Sens (%)	Spec (%)	Prec (%)
[16]	Adults	10	164	98.9	88.5	-
[49]	Adults	110	-	68.0	-	74.0
[19]	Adults	6	159	70.2	98.6	-
[15]	Adults	2	3,335	79.0	-	-
[17]	Adults	-	175	95.4	-	-
[50]	Adults	9	725	81.2	81.2	-
[51]	Adults	6	575	78.4	88.6	-
[52]	Adults	12	95	-	93.9	-
[35]	Adults	12	6,043	70.0	98.6	-
[53]	Adults	12	2,140	92.9	-	-
[54]	Adults	20	27,923	74.1	-	
[31]	Adults	8	355	53.0	96.0	37.0
[31]	Adults	19	11,207	77.0	96.0	46.0
[18]	Adults	1	-	87.5	97.3	-
[55]	Adults	2	1,089	96.5	98.1	-
[56]	Adults	8	3,875	96.0	92.9	-
[48]	Children	56	40,412	88.2	89.7	-
[23]	Infants	1	-	62.9	-	89.7
[22]	Infants	2	803	87.7	-	91.9
Tra	**Ex-term (R-SS)**	**81**	**13,463**	**91.0**	**92.1**	**90.8**
Val	**Ex-term (R-SS)**	**30**	**4,304**	**94.1**	**89.4**	**87.3**
Test	**Ex-term (R-SS)**	**30**	**4,263**	**93.3**	**91.5**	**90.1**
Test	**Ex-term (L-SS)**	**30**	**4,268**	**93.9**	**90.7**	**89.2**

N: Number of subjects; SS: Number of sleep spindles; R-SS: Channel F4-C4/R-Spindle and L-SS: Channel F3-C3/L-Spindle; Tra: Training set; Val: Validation set; Test: Test set;

TABLE VII
PREVIOUS WORK ON SLEEP SPINDLE DURATION ESTIMATION.

Ref	Subjects	N	Recall(%)	TNR(%)	IoU
[57]	Adults	19	90.1	96.2	-
[57]	Adults	8	77.9	94.2	-
[21]	Adults	19	84.0	90.0	-
[21]	Adults	8	76.0	92.0	-
[20]	Adults	15	71.2	96.7	-
[18]	Adult	1	-	-	0.37
Tra	**Ex-term (R-SS)**	**81**	**88.2**	**97.6**	**0.75**
Val	**Ex-term (R-SS)**	**30**	**90.2**	**97.6**	**0.74**
Test	**Ex-term (R-SS)**	**30**	**89.7**	**97.8**	**0.77**
Test	**Ex-term (L-SS)**	**30**	**90.5**	**97.3**	**0.75**

N: Number of subjects; R-SS: Channel F4-C4/R-Spindle and L-SS: Channel F3-C3/L-Spindle; Tra: Training set; Val: Validation set; Test: Test set;

spindle and non-sleep spindle waves and was identified as the highest ranking feature by the random forest algorithm (Figure 4).

The difference in datasets and means of evaluating performance make a direct comparison with other published sleep methods difficult. As discussed, sleep spindles change dramatically during maturation, and although there is no study directly comparing sleep spindles of infants at the 4 month mark with adults, there is vast literature exploring the maturation trajectories in several sleep spindle features (see for example [4], [9]–[13], [53], [58]). Due to the constant evolution of sleep spindles which may lead to significant differences with adult sleep spindles, we believe that there is a need for an infant-specific sleep spindle algorithm (See Appendix B).

Previous studies on infant data [22], [23] have tested methods on data from just 1 or 2 infants, with sensitivity of 62.9% to 87.7%. There are few published methods for estimating EEG sleep spindle number in infants, we therefore also compared the Spindle-AI method with sleep spindle number estimation methods reported for adult EEG [15]–[19], [31], [35], [49]–[56] (shown in Table VI). For the Spindle-AI method combining successive sleep spindle events separated by less than 1 second into a single spindle event avoided dividing spindle events across different epochs which may occur as with previously presented methods [16], [20], [22], [23], [49]. Moreover, these studies defined that if only part of the epoch contained sleep spindles, then this epoch will be labelled as a sleep spindle event. The work of [16], [20] and [55] used larger duration epochs (3s, 5s and 5s, respectively), if a sleep spindle and non-sleep spindle event present in the same epoch, these approaches may take both occurrences as one sleep spindle event. Therefore, they estimated the number of epochs which contain sleep spindles, rather than 'real' number of sleep spindles. Moreover, the duration of each sleep spindle and the specific time of occurrence of the sleep spindle events cannot be accurately expressed by these methods. In our study, we divided the EEG signal into small epochs, 0.5s with 0.25s overlap, reduced the problem of two sleep spindles occurring in one epoch.

A method based on the continuous wavelet transform [21] has previously been developed to identify the duration of sleep spindles in two adult EEG datasets: the MASS database [59] comprising 19 healthy controls and the DREAMS sleep spindle database [19] comprising 8 participants diagnosed with various sleep pathologies. They obtained recall of 84.0% and TNR of 90.0% on the MASS dataset and recall and TNR at 76.0% and 92.0%, respectively, on the DREAMS dataset. A deep learning-based method, SpindleNet, developed in [57] also estimated the duration of sleep spindle in the Mass and DREAMS datasets. SpindleNet yielded a recall of 90.1% and TNR of 96.2% on the MASS dataset. For the DREAMS dataset, SpindleNet obtained recall and TNR of 77.9% and 94.2%, respectively. Additionally, the random forest-based method described in [20], used an independent test data (N=12) to test the performance of the sleep spindle duration estimation. Their method has a TNR with 96.7% and with a recall of 71.2% for sleep spindle duration estimation in adult EEG. These methods were developed and tested on a smaller data (EEG data for 1-19 adult individuals, see Table VII). Moreover, these methods only estimated the duration of sleep spindles, and the number of the sleep spindles is not given. The Spindle-AI method estimate the number of sleep spindles, but also give the duration of sleep spindles on larger EEG data (N=141). The work of Ventouras *et al.* [18] estimates both the number and duration of sleep spindles by feed-forward networks in adult EEG. The number of sleep spindles was estimated with sensitivity from 79.2% to 87.5%. For the duration estimation, an inter-spindle interval (ISI) was defined as the time difference between the onset of two consecutive sleep spindles evaluated as ($\triangle ts$+ $\triangle te$)/tV. This is the same matrix as the IoU that we used to evaluate Spindle-AI. Ventouras *et al.* obtained a mean IoU of 0.37 ($\pm$0.31). However, this method was developed and tested using EEG data from only 1 adult. Individual differences may lead to differences in results, and whether this method is applicable to other EEG data is not clear.

The Spindle-AI method was developed on single-channel (F4-C4) EEG recordings of 111 infants (81 for training and 30 for validation), and tested on another 30 ex-term infants in channel F4-C4. As different sleep studies may analyse sleep spindle events on different channels [22], we explored if Spindle-AI can generalise to different channels, in this case from channel F4-C4 to channel F3-C3. Therefore, we include the signal in channel F3-C3 as an additional independent test set. There are a number of potential differences between R-Spindle (F4-C4) and L-Spindle (F3-C3) including that it is normal in this age group that R-Spindles and L-Spindles occur more independently or asynchronously [5]. Bódizs *et al.* [61] found that the duration, amplitude and density magnitudes differ between hemispheres (R-Spindle and L-Spindle) in a predominantly adult study. However, these differences were observed in an older age group compared to the cohort in this study. Our results show that in spite of these potential differences between R-Spindle and L-Spindle, Spindle-AI can generalize to EEG data from a channel that was not used in training (Tables IV and V).

Using the Spindle-AI method, large volumes of data can be quickly

reviewed potentially enabling trends and patterns to be identified in infant EEG recordings that may not be apparent otherwise. In a clinical setting, once a range of normative values for spindle features has been established across age groups, automated spindle analysis has potential as an assessment tool for brain maturation in 'at risk' individuals. As one element of sleep, spindles are considered part of the 'microstructure' of sleep. Sleep spindles are markers of maturation and may reflect neurologic pathologies/non-normative maturation that share the same neurobiological subtract [60]. Sleep spindle detection is also essential for the precise labeling of sleep stage N2 [4]. Therefore, the study of sleep spindles provides support for clinical interpretation and early identification of abnormal brain maturation in infants. An automated spindle detector might also constitute an important part of a broader automated algorithm to characterise and differentiate the various sub-stages of sleep, permitting analysis of the so called 'macrostructural' elements of sleep for the same purpose. As spindle features have also been shown to change in the short term following various learning tasks, accurate automated spindle detection would also progress research in cognitive function.

A limitation of the current study is the range of ages of the infants. Spindle-AI was trained on ex-term infant EEGs without testing on infants of other ages. In future work, we would like to apply Spindle-AI on other infant EEG data. Moreover, as machine learning is a 'black box' method, clinicians may have difficulty trusting the machine learning-based methods. In future work, we will combine Spindle-AI with explainable AI (XAI) [62] techniques to help users understand why certain events are predicted as a sleep spindle, which may help gain users' trust in the system and assist experts in analysing infant sleep spindles.

V. Conclusion

Spindle-AI has been developed for use on ex-term infant EEGs with the aim of assisting clinicians in the estimation of both the number and the duration of sleep spindles. We implemented a random forest-based sleep spindle detection method which incorporates novel post-processing and evaluation techniques. The dataset used to train and test Spindle-AI is substantially larger than datasets used in previous studies to identify sleep spindles in infant EEGs. In addition, Spindle-AI has been implemented as a web server and is freely available for academic use at http://lisda.ucd.ie/Spindle-AI/. The web server predicts the start time, end time and the total number of sleep spindles detected in long EEG recordings, allowing for fast and accurate analysis of infant sleep spindles in single-channel EEGs which may act an early biomarker for abnormal brain maturation.

APPENDIX

A. Preliminary analysis

TABLE VIII
PRELIMINARY ANALYSIS (WITHOUT POST-PROCESSING) OF THE PERFORMANCE OF THE RANDOM FOREST ALGORITHM FOR SLEEP SPINDLE DURATION COMPARED TO MULTILAYER PERCEPTRON, NAIVE BAYES AND EXTREME GRADIENT BOOSTING ON THE VALIDATION SET.

Method	Recall (%)	TPR (%)	MCC
NB	95.2	73.6	0.440
MLP	90.7	93.2	0.703
XGboost	91.2	95.5	0.770
RF	**91.3**	**95.7**	**0.779**

NB: Naive bayes; MLP: Multilayer Perceptron; XGboost: eXtreme Gradient Boosting; RF: Random forest;

B. Replication of adult sleep spindle detection method on infant EEGs

Due to the constant evolution of sleep spindles that may lead to significant differences between infant and adult sleep spindles, adult sleep spindle detection methods may not be suitable for infant sleep spindle detection. To explore this, we applied the adult sleep spindle detection method presented by Ventouras *et al.* [18] on infant EEG data (preliminary work, results not shown), as this work estimated both the number and duration of sleep spindles similar to the method presented here. Ventouras *et al.* used a feed-forward artificial neural network on adult EEG, signals with a frequency range of 10.5 - 16 Hz were used as input to the network. The number of sleep spindles was estimated with sensitivity from 79.2% to 87.5% in adult EEG data [18]. For the sleep spindle duration estimation on adult EEG, Ventouras *et al.* obtained a mean IoU of 0.37 ($\pm$0.31). We trained and tested this method on the infant EEG training and testing datasets used in the present study. When applying this method to the infant EEG data, we obtained sensitivity of 53.8% and specificity of 47.8% on the independent test set in channel F4-C4. For duration estimation, this method obtained IoU of 0.069. The poor performance of the adult sleep spindle detection method on the infant data likely reflects smaller differences between the amplitude of infant sleep spindle and non-sleep spindle activity in the 10.5 - 16 Hz sleep spindle band when compared with adult EEG. While a range of EEG features based upon the sleep spindle and sigma band activity ranked highly among those selected for the random forest algorithm, additional features including the mean frequency, RMS and mean absolute amplitude and alpha band activity were also included in the model (Figure 4). The results highlight the need for an infant-specific sleep spindle algorithm.

ACKNOWLEDGMENTS

This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number 12/RC/2272 and 15/SP/3091. The research was also funded in part by JOHNSON & JOHNSON CONSUMER SERVICES EAME LTD. We acknowledge the Research IT HPC Service at University College Dublin for providing computational facilities and support that contributed to the research results reported in this paper. We would also like to thank the families of all infants recruited for this study.

REFERENCES

[1] R. Wakai and W. Lutter, "Slow rhythms and sleep spindles in early infancy," *Neuroscience Letters*, vol. 630, pp. 164–168, 2016.
[2] A. L. Loomis, E. N. Harvey, and G. Hobart, "Potential rhythms of the cerebral cortex during sleep." *Science*, vol. 81, pp. 597–598, 1935.
[3] A. Rodenbeck *et al.*, "A review of sleep EEG patterns. part i: a compilation of amended rules for their visual recognition according to rechtschaffen and kales," *Somnologie*, vol. 10, no. 4, pp. 159–175, 2006.
[4] J. Louis *et al.*, "Ontogenesis of nocturnal organization of sleep spindles: a longitudinal study during the first 6 months of life," *Electroencephalography and Clinical Neurophysiology*, vol. 83, no. 5, pp. 289–296, 1992.
[5] R. J.Ellingson, "Development of sleep spindle bursts during the first year of life," *Sleep*, vol. 5, no. 1, pp. 39–46, 1982.
[6] H. Lenard, "The development of sleep spindles in the EEG during the first two years of life," *Neuropädiatrie*, vol. 1, no. 03, pp. 264–276, 1970.
[7] D. R. Metcalf, "EEG sleep spindle ontogenesis," *Neuropädiatrie*, vol. 1, no. 04, pp. 428–433, 1970.
[8] B. C. Clawson, J. Durkin, and S. J. Aton, "Form and function of sleep spindles across the lifespan." *Neural Plasticity*, 2016.
[9] I. J. McClain *et al.*, "Developmental changes in sleep spindle characteristics and sigma power across early childhood," *Neural Plasticity*, vol. 2016, 2016.
[10] S. Shinomiya *et al.*, "Development of sleep spindles in young children and adolescents," *Clinical Electroencephalography*, vol. 30, no. 2, pp. 39–43, 1999.

[11] N. Martin *et al.*, "Topography of age-related changes in sleep spindles," *Neurobiology of Aging*, vol. 34, no. 2, pp. 468–476, 2013.

[12] A. D'Atri *et al.*, "Different maturational changes of fast and slow sleep spindles in the first four years of life," *Sleep Medicine*, vol. 42, pp. 73–82, 2018.

[13] S. Scholle, G. Zwacka, and H. Scholle, "Sleep spindle evolution from infancy to adolescence," *Clinical Neurophysiology*, vol. 118, no. 7, pp. 1525–1531, 2007.

[14] D. R. Metcalf and K. Jordan, "EEG ontogenesis in normal children," *Drugs, Development, and Cerebral Function*, pp. 127–128, 1972.

[15] E. Huupponen *et al.*, "Optimization of sigma amplitude threshold in sleep spindle detection," *Journal of Sleep Research*, vol. 9, no. 4, pp. 327–334, 2000.

[16] L. B. Ray *et al.*, "Validating an automated sleep spindle detection algorithm using an individualized approach," *Journal of Sleep Research*, vol. 19, no. 2, pp. 374–378, 2010.

[17] D. Gorur *et al.*, "Sleep spindles detection using short time fourier transform and neural networks," in *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)*, vol. 2. IEEE, 2002, pp. 1631–1636.

[18] E. M. Ventouras *et al.*, "Sleep spindle detection using artificial neural networks trained with filtered time-domain EEG: a feasibility study," *Computer Methods and Programs in Biomedicine*, vol. 78, no. 3, pp. 191–207, 2005.

[19] S. Devuyst *et al.*, "Automatic sleep spindles detection—overview and development of a standard proposal assessment method," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 1713–1716.

[20] C. R. Patti *et al.*, "Application of random forest classifier for automatic sleep spindle detection," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 2015, pp. 1–4.

[21] A. Tsanas and G. D. Clifford, "Stage-independent, single lead EEG sleep spindle detection using the continuous wavelet transform and local weighted smoothing," *Frontiers in Human Neuroscience*, vol. 9, p. 181, 2015.

[22] C. M. Held *et al.*, "Dual approach for automated sleep spindles detection within EEG background activity in infant polysomnograms," in *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1. IEEE, 2004, pp. 566–569.

[23] P. A. Estévez *et al.*, "Sleep spindle detection by using merge neural gas," in *International Workshop on Self-Organizing Maps: Proceedings*, 2007.

[24] L. Wei *et al.*, "Random forest-based algorithm for sleep spindle detection in infant EEG," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 58–61.

[25] G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, no. 2, pp. 197–227, 2016.

[26] L. Fraiwan *et al.*, "Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 1, pp. 10–19, 2012.

[27] E. Hernandez-Pereira, I. Fernandez-Varela, and V. Moret-Bonillo, "A comparison of performance of sleep spindle classification methods using wavelets," in *International Conference on Innovation in Medicine and Healthcare*. Springer, 2016, pp. 61–70.

[28] A. R. Hassan and M. I. H. Bhuiyan, "A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features," *Journal of Neuroscience Methods*, vol. 271, pp. 107–118, 2016.

[29] V. Krajča *et al.*, "Neonatal EEG sleep stages modelling by temporal profiles," in *International Conference on Computer Aided Systems Theory*. Springer, 2007, pp. 195–201.

[30] A. Dereymaeker *et al.*, "An automated quiet sleep detection approach in preterm infants as a gateway to assess brain maturation," *International Journal of Neural Systems*, vol. 27, no. 06, p. 1750023, 2017.

[31] D. Lachner-Piza *et al.*, "A single channel sleep-spindle detector based on multivariate classification of EEG epochs: MUSSDET," *Journal of Neuroscience Methods*, vol. 297, pp. 31–43, 2018.

[32] B. D. Yetton *et al.*, "Automatic detection of rapid eye movements (REMs): A machine learning approach," *Journal of Neuroscience Methods*, vol. 259, pp. 72–82, 2016.

[33] M. Carrozzi, A. Accardo, and F. Bouquet, "Analysis of sleep-stage characteristics in full-term newborns by means of spectral and fractal parameters," *Sleep*, vol. 27, no. 7, pp. 1384–1393, 2004.

[34] M. Čić, J. Šoda, and M. Bonković, "Automatic classification of infant sleep based on instantaneous frequencies in a single-channel EEG signal," *Computers in Biology and Medicine*, vol. 43, no. 12, pp. 2110–2117, 2013.

[35] E. Huupponen *et al.*, "Development and comparison of four sleep spindle detection methods," *Artificial Intelligence in Medicine*, vol. 40, no. 3, pp. 157–170, 2007.

[36] E. Biagioni *et al.*, "Distribution of sleep and wakefulness EEG patterns in 24-h recordings of preterm and full-term newborns," *Early Human Development*, vol. 81, no. 4, pp. 333–339, 2005.

[37] A. Piryatinska *et al.*, "Automated detection of neonate EEG sleep stages," *Computer Methods and Programs in Biomedicine*, vol. 95, no. 1, pp. 31–46, 2009.

[38] S. Solnik *et al.* "Teager–kaiser energy operator signal conditioning improves EMG onset detection," *European Journal of Applied Physiology*, vol. 110, no. 3, pp. 489–498, 2010.

[39] F.S. Bao *et al.* "PyEEG: an open source python module for EEG/MEG feature extraction," *Computational Intelligence and Neuroscience*, vol. 2011, no. 406391, pp. 1–7, 2011.

[40] E. Huupponen *et al.*, "Determination of dominant simulated spindle frequency with different methods," *Journal of Neuroscience Methods*, vol. 156, no. 1-2, pp. 275–283, 2006.

[41] A. Mustaqeem *et al.*, "Wrapper method for feature selection to classify cardiac arrhythmia," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2017, pp. 3656–3659.

[42] J. L. Mattsson, P. J. Spencer, and R. R. Albee, "A performance standard for clinical and functional observational battery examinations of rats," *Journal of the American College of Toxicology*, vol. 15, no. 3, pp. 239–254, 1996.

[43] S. Boughorbel, F. Jarray, and M. El-Anbari, "Optimal classifier for imbalanced data using matthews correlation coefficient metric," *PloS One*, vol. 12, no. 6, p. e0177678, 2017.

[44] H. He *et al.*, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. IEEE, 2008, pp. 1322–1328.

[45] N. V. Chawla *et al.*, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[46] O. F. AlThuwaynee *et al.*, "Demystifying uncertainty in PM10 susceptibility mapping using variable drop-off in extreme-gradient boosting (XGB) and random forest (RF) algorithms," *Environmental Science and Pollution Research*, pp.1–23, 2021.

[47] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[48] L. Causa *et al.*, "Automated sleep-spindle detection in healthy children polysomnograms," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 9, pp. 2135–2146, 2010.

[49] K. Lacourse *et al.*, "A sleep spindle detection algorithm that emulates human expert spindle scoring," *Journal of Neuroscience Methods*, vol. 316, pp. 3–11, 2019.

[50] S. V. Schönwald *et al.*, "Benchmarking matching pursuit to find sleep spindles," *Journal of Neuroscience Methods*, vol. 156, no. 1-2, pp. 314–321, 2006.

[51] S. Devuyst *et al.*, "Automatic sleep spindle detection in patients with sleep disorders," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2006, pp. 3883–3886.

[52] B. Ahmed, A. Redissi, and R. Tafreshi, "An automatic sleep spindle detector based on wavelets and the teager energy operator," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2009, pp. 2596–2599.

[53] R. Bódizs *et al.*, "The individual adjustment method of sleep spindle analysis: methodological improvements and roots in the fingerprint paradigm," *Journal of Neuroscience Methods*, vol. 178, no. 1, pp. 205–213, 2009.

[54] C. R. Patti, T. Penzel, and D. Cvetkovic, "Sleep spindle detection using multivariate gaussian mixture models," *Journal of Sleep Research*, vol. 27, no. 4, p. e12614, 2018.

[55] M. M. Kabir *et al.*, "Enhanced automated sleep spindle detection algorithm based on synchrosqueezing," *Medical & Biological Engineering & Computing*, vol. 53, no. 7, pp. 635–644, 2015.

[56] B. Babadi *et al.*, "Diba: a data-driven bayesian algorithm for sleep spindle detection," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 483–493, 2011.

[57] P. M. Kulkarni *et al.*, "A deep learning approach for real-time detection of sleep spindles," *Journal of Neural Engineering*, vol. 16, no. 3, p. 036004, 2019.

[58] J. R. Hughes, "Development of sleep spindles in the first year of life," *Clinical Electroencephalography*, vol. 27, no. 3, pp. 107–115, 1996.
[59] C. O'Reilly *et al.*, "Montreal archive of sleep studies: an open-access resource for instrument benchmarking and exploratory research," *Journal of Sleep Research*, vol. 23, no. 6, pp. 628–635, 2014.
[60] L. M.J. Fernandez and A. Lüthi , "Sleep spindles: mechanisms and functions," *Physiological Reviews*, vol. 100, no. 2, 2020.
[61] R. Bódizs *et al.*, "The hemispheric lateralization of sleep spindles in humans," *Sleep Spindles & Cortical Up States*, vol. 1, no. 1, pp. 42–54, 2017.
[62] A. M. Antoniad *et al.*, "Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: A systematic review," *Applied Sciences*, vol. 11, no. 11, 2021.