Title	Random Forest-based Algorithm for Sleep Spindle Detection in Infant EEG
Authors(s)	Wei, Lan; Ventura, Soraia; Lowery, Madeleine M.; Mooney, Catherine; et al.
Publication date	2020-07-24
Publication information	Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS
Conference details	The Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2020), Virtual Conference, 20-24 July 2020
Publisher	IEEE
Item record/more information	http://hdl.handle.net/10197/12166
Publisher's statement	© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
Publisher's version (DOI)	10.1109/EMBC44109.2020.9176339

Random Forest-based Algorithm for Sleep Spindle Detection in Infant EEG

Lan Wei[1], Soraia Ventura[2,3], Madeleine Lowery[4], Mary Anne Ryan[2,3], Sean Mathieson[2,3], Geraldine B. Boylan[2,3] and Catherine Mooney[1]

***Abstract*— Sleep spindles are associated with normal brain development, memory consolidation and infant sleep-dependent brain plasticity and can be used by clinicians in the assessment of brain development in infants. Sleep spindles can be detected in EEG, however, identifying sleep spindles in EEG recordings manually is very time-consuming and typically requires highly trained experts. Research on the automatic detection of sleep spindles in infant EEGs has been limited to-date. In this study, we present a novel supervised machine learning-based algorithm to detect sleep spindles in infant EEG recordings. EEGs collected from 141 ex-term born infants and 6 ex-preterm born infants, recorded at 4 months of age (adjusted), were used to train and test the algorithm. Sleep spindles were annotated by experienced clinical physiologists as the gold standard. The dataset was split into training (81 ex-term), validation (30 ex-term), and testing (30 ex-term + 6 ex-preterm) set. 15 features were selected for input into a random forest algorithm. Sleep spindles were detected in the ex-term infant EEG test set with 92.1% sensitivity and 95.2% specificity. For ex-preterm born infants, the sensitivity and specificity were 80.3% and 91.8% respectively. The proposed algorithm has the potential to assist researchers and clinicians in the automated analysis of sleep spindles in infant EEG.**

I. INTRODUCTION

Loomis *et al.* [1] first described sleep spindles as rhythmic 12-14 Hz oscillations which last 0.5 to 3 second with a waxing and waning shape [2]. In full-term infants, sleep spindles typically can be observed at age 6-8 weeks. Several changes in spindle parameters occur with age [3], therefore, spindles can be used as a method to assess functional brain development [4]. It has been proposed that spindles are markers for the development and integrity of the central nervous system early in life [5] and are involved in infant sleep-dependent brain plasticity [6]. Sleep spindles are also important for memory consolidation [7] and the number of sleep spindles in the left frontocentral areas has been shown to highly correlated with overnight verbal memory retention [8]. Therefore, identification of abnormal sleep spindle architecture may indicate an early signature of poor neurodevelopment in infants. However, identifying sleep spindles manually is very time-consuming and research to date for the automatic detection of sleep spindles in infant EEGs has been limited [9], [10].

*This work was supported by INFANT Research Centre, Ireland

[1]School of Computer Science, University College Dublin, Dublin, Ireland lan.wei@ucdconnect.ie, catherine.mooney@ucd.ie

[2]Department of Paediatrics & Child Health, University College Cork, Cork, Ireland [3]INFANT Research Centre, Ireland

[4] School of Electrical & Electronic Engineering, University College Dublin, Dublin, Ireland

To the best of our knowledge, there have only been two previous publications on the automatic detection of sleep spindles in infant EEGs. In the first, Held *et al.* [9] presented an automated system trained on three and tested on two infants, which achieved a sensitivity of 87.7% and an 8.1% false-positive rate. This system combined two different approaches: detection criteria on the sigma-band filtered EEG signal, including fuzzy thresholds, and mimicking an expert's procedure for infants. The second approach, developed by Estevez *et al.* [10] used a Merge Neural Gas (MNG) algorithm and was trained on a single infant, and tested on another achieving 62.9% sensitivity.

Machine Learning based algorithms can review large volumes of data and discover specific trends and patterns that may not be apparent to humans. In this study, we developed a random forest based machine learning model to detect sleep spindles in the EEGs of 141 ex-term born and 6 ex-preterm born infants recorded at four month of age. This is the largest dataset of its kind that we are aware of.

II. MATERIALS AND METHODS

A. Subjects

Ethical approval was granted from the Clinical Research Ethics Committee of the Cork Teaching Hospitals, Cork, Ireland and consent from parents or guardians of the infants included in the study was obtained. A cohort of healthy full-term infants (n=141) were recruited soon after birth at Cork University Maternity Hospital (CUMH). EEG data were recorded from sleeping infants at four months. Six ex-preterm infants who were born at 32-36+6 weeks gestational age (GA) and had an EEG recorded at four months adjusted age were also included.

B. Data analysis

EEG (Lifelines, UK) was recorded with a sampling frequency of 500 Hz. Sleep spindles of ex-term infants were annotated by an experienced clinical physiologist on channel F4-C4 (R-Spindle), and sleep spindles of ex-preterm infants were annotated by an experienced research nurse on channel F4-C4, as the gold standard. Figure 1 shows the pattern of EEG signal in channel F4-C4 and Figure 2 shows the frequency spectrum of a sleep spindle and non-sleep spindle event. The number and duration of the sleep spindles in the EEGs used in this study are presented in Table I.

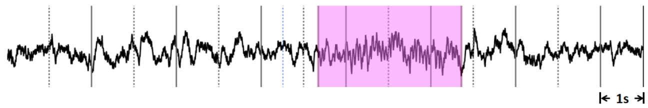
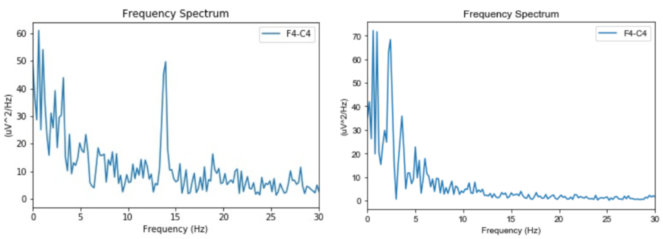


Fig. 1. EEG signal in channel F4-C4 (the signal in shaded block indicates the presence of of sleep spindle event).



(a) Sleep spindle

(b) Non-sleep spindle

Fig. 2. (a) Frequency spectrum of 3 seconds ex-term infant's EEG recording during a sleep spindle event (b) Regular EEG frequency spectrum from the same infant (3 seconds).

TABLE I

NUMBER AND DURATION OF SLEEP SPINDLES (R-SPINDLE) ANNOTATED BY EXPERIENCED CLINICAL PHYSIOLOGISTS

	Ex-term	Ex-premature
Number of Infants	141	6
Average number of sleep spindles per infant	159	137
Total number of sleep spindles	22,477	823
Total duration of sleep spindles (s)	67,997	2,901
Total duration of non-sleep spindles (s)	509,535	20,357

C. Data Pre-processing

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 channel F4-C4 EEG (R-Spindle) was segmented into epochs of 0.5s length with 0.25s overlap and used to estimate all features. The length of 0.5 seconds was chosen as this is the minimum required length of a sleep spindle [2].

D. Features Extraction

Time and frequency domain features were estimated. 6th order Butterworth filters (IIR) were used to filter the signals within the frequency bands of interest, which can reduce the interference of other waves: alpha (8–12 Hz), sigma (12.5-15 Hz) and spindle (10.5-16 Hz). 0.5 seconds epochs was used to develop all these 15 features, these features are as follows:

Standard features (4): The standard deviation of raw EEG absolute amplitudes (without removing DC offset) and mean and root mean squared of pre-processed absolute amplitude, were calculated in the time domain. The mean frequency of each epoch was calculated in the frequency domain. These features were selected to provide time- and frequency-based changes in EEG activity.

Features in sub-frequency bands (11): Sigma index, alpha band ratio and spindle band ratio were calculated based on prior work done by Patti *et al.* [11]. In addition, the following features were also selected to provide information in the sub-frequency bands: mean absolute amplitudes and Hilbert mean envelope amplitude in sigma band; mean absolute amplitudes and Hilbert mean envelope amplitude in spindle band; relative and absolute band power in sigma band; and relative and absolute band power in spindle band.

E. Dataset balancing

The duration of the spindle periods in the dataset was shorter than the duration of non-spindle periods resulting in a class imbalance problem that can make training a machine learning algorithm challenging. Therefore, Synthetic Minority Over-sampling Technique (SMOTE) [12] was used to balance the data. SMOTE is a method of oversampling, in which the minority class is oversampled by creating 'composite' examples [12]. Composite examples are generated in a less application-specific manner by operating in the feature space. The minority classes are over-sampled, and a comprehensive example is introduced by taking samples of each minority class and following the line of segments connecting the nearest neighbour of any k minority classes. The neighbour from the k nearest neighbour is randomly selected according to the required excessive sampling quantity. In this work, we use five nearest neighbours with the random state of 2.

F. Classification algorithms

Several classification algorithms were tested for spindle detection during this study, including neural networks and support vector machine techniques [11], [13], [14]. Although these methods have achieved success in many classification problems the random forest classifier had the best performance in this study. The random forest was implemented using the 'RandomForest' package of the sklearn library [15] within the Python 3 environment. 81 ex-term infants were used for training and 30 ex-term infants for validation, and the other 30 ex-term infants were used for independent testing of the model. Three parameters were optimized: the number of decision trees grown based on a bootstrap sample of the observations (n-estimators); the minimum number of samples required to split an internal node (min-samples-split); and the minimum number of samples required to be at a leaf node (min-samples-leaf). 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 algorithm for the detection of spindles in the EEG recordings. The n-estimators values were tested from 10 to 200, min-samples-split were tested from 2 to 150 and min-samples-leaf were tested from 2 to 20. The best performance on the validation set was achieved when n-estimators = 100, min-samples-split = 120, and min-samples-leaf = 20. In addition, the sleep spindles detected by the random forest algorithm with the duration greater than 0.5s were extended from the start time of the first component to the end time of the last component.

G. Performance evaluation

Automatic spindle detection is a binary classification task and we used the following metrics to evaluate the performance of our model: specificity (Spec), sensitivity (Sens), Matthews correlation coefficient (MCC) and accuracy (Acc).

$$
\begin{aligned}
Sens &= \frac{TP}{TP+FN} \\
Spec &= \frac{TN}{TN+FP} \\
TPR &= \frac{TP}{TP+FN} \\
FPR &= \frac{FP}{FP+TN} \\
MCC &= \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \\
Acc &= \frac{TP+TN}{TP+TN+FP+FN}
\end{aligned} \tag{1}
$$

where:

- True Positives (TP): the number of sleep spindles predicted as sleep spindles
- False Positives (FP): the number of non-sleep spindles predicted as sleep spindles
- True Negatives (TN): the number of non-sleep spindles predicted as non-sleep spindles
- False Negatives (FN): the number of sleep spindles predicted as non-sleep spindles.

The area under the receiver operating characteristics curve (AUROC) was calculated by plotting the TPR against the FPR while increasing the discrimination threshold from 0 to 1. The AUROC is equivalent to the probability that the model will rank a randomly chosen positive instance. MCC measures the correlation coefficient between the observed and predicted classifications. A value of 1 represents a perfect prediction, 0 a random prediction and -1 an inverse prediction and is a good indicator of the overall performance of the predictive methods in both sleep spindle and non-sleep spindle classes.

III. Results

A. AUROC Curve

The default threshold of the random forest algorithm is 0.50. However, the original data set contains a large number of non-spindle periods compared with the spindle events, so the algorithm tended to be biased towards predicting new data as "non-spindle". Therefore, the threshold was adjusted based on the AUROC curve in order to improve the sensitivity of the algorithm, as sensitivity is particularly essential for identifying all sleep spindle. In this study, setting the threshold of the random forest algorithm to 0.45 gave the best performance. Table II show the performance on the validation data set of the original random forest algorithm trained with a 0.50 threshold compared to the random forest algorithm trained with a 0.45 threshold.

TABLE II

Comparison between the original threshold (threshold = 0.50) and the adjusted threshold (threshold = 0.45) of the random forest algorithm on the validation set.

Threshold	Sens (%)	Spec (%)	Acc (%)	MCC
0.50	91.4	95.7	95.3	0.780
0.45	92.4	94.9	94.7	0.761

B. Performance on Infant EEG

Table III shows the performance of the random forest algorithm on the training set, validation set and test set on ex-term infant EEG recordings and the 6 ex-preterm infant test set EEGs.

TABLE III

Performance of the random forest algorithm on the train, validation, test set and ex-preterm born infants.

	Sens (%)	Spec (%)	Acc (%)	MCC
Train (N=81)	94.1	94.9	94.5	0.891
Val (N=30)	92.4	94.9	94.7	0.761
Test (N=30)	92.1	95.2	94.8	0.788
Ex-preterm (N=6)	80.3	91.8	90.4	0.632

C. Benchmarking

Research to-date for the automatic detection of sleep spindles on infants has been limited. However, several other studies have applied various algorithms to detect sleep spindles automatically in adults. Compared with previous work (Table IV), the models developed here used a larger number of infants and demonstrated higher sensitivity (92.1%).

TABLE IV

Compared with sleep spindles detection approaches done previously (S: training and testing set are the same)

Ref	**Subjects**	**Total No**	**Test No**	**Sens(%)**	**Spec(%)**
[10]	Infants	2	1	62.9	-
[9]	Infants	5	2	87.7	-
[16]	Children	56	19	88.2	89.7
[17]	Adults	6	2	79.0	-
[13]	Adults	-	-	95.4	-
[18]	Adult	1	S	83.4	92.9
[19]	Adults	6	-	76.9	90.0
[20]	Adults	9	S	81.2	81.2
[21]	Adults	12	S	70.0	98.6
[22]	Adults	12	-	92.9	-
[23]	Adults	10	S	98.96	88.49
[24]	Adults	6	S	70.2	98.62
[25]	Adults	19	S	84	90
[25]	Adults	8	S	76	92
[11]	Adults	15	12	71.2	96.73
[26]	Adults	20	S	65.1-74.1	-
[27]	Adults	110	-	68	-
This	Infants	141	30	92.1	95.2
work	Ex-preterm	6	6	80.3	91.8

IV. Discussion

We have developed a new random forest based method to detect sleep spindles in infant EEG. Fifteen EEG features from 81 EEG signals were used to train the model. As the quantity of spindle events is greater than the non-spindle events SMOTE was used to balance the training set. A validation set of 30 EEGs was used to adjust the parameters of the random forest algorithm. The results on the test set (N=30) show high specificity and sensitivity, 95.2% and 92.1%, respectively. The model was also tested on six ex-preterm born infants, which were not used in training (91.8% specificity and 80.3% sensitivity) showing that the model can generalise to ex-preterm born infants. MCC was used as an additional evaluation metric due to the imbalanced nature of the dataset. MCC takes into account true and false positives, and negatives are generally considered a balancing measure that can be used even if the classes are of very different sizes. The test set yielded an MCC of 0.788 on the ex-term and 0.632 on the ex-preterm infant EEGs demonstrating that the model performs well at identifying both the negative (non-spindle) and the positive (spindle) events.

A limitation of the current study is the range of ages of the infants. Our algorithm was trained on four-month old ex-term infant EEGs and tested on four-month old ex-term and ex-preterm infants. Therefore, we did not test the algorithm on infants of other ages, and it is not clear how accurate this algorithm is in detecting sleep spindles in younger or older infants. In future work, we would like to see if this algorithm works well on EEGs for other age groups of infants.

In this study, we describe a novel supervised machine learning-based algorithm to detect spindles in infant EEG recordings. The model can generalise well on infant EEG, including ex-preterm born infants not used in the development of the algorithm. In addition, the duration of spindles can be accurately detected. This will allow for faster, more reliable, and more reproducible detection of infant sleep spindles in long-duration, single-channel EEG recordings.

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. Part of this funding comes from our SFI SPOKES industry partner JOHNSON & JOHNSON CONSUMER SERVICES EAME LTD.

REFERENCES

[1] A. L. Loomis, E. N. Harvey, G. Hobart, Potential rhythms of the cerebral cortex during sleep., Science 81 (1935) 597–598.

[2] A. Rodenbeck, R. Binder, P. Geisler, H. Danker-Hopfe, R. Lund, F. Raschke, H.-G. Weeß, H. Schulz, A review of sleep eeg patterns. part i: a compilation of amended rules for their visual recognition according to rechtschaffen and kales, Somnologie 10 (4) (2006) 159–175.

[3] L. De Gennaro, M. Ferrara, Sleep spindles: an overview, Sleep medicine reviews 7 (5) (2003) 423–440.

[4] S. M. Fogel, C. T. Smith, The function of the sleep spindle: a physiological index of intelligence and a mechanism for sleep-dependent memory consolidation, Neuroscience & Biobehavioral Reviews 35 (5) (2011) 1154–1165.

[5] R. Wakai, W. Lutter, Slow rhythms and sleep spindles in early infancy, Neuroscience letters 630 (2016) 164–168.

[6] M. Friedrich, I. Wilhelm, J. Born, A. D. Friederici, Generalization of word meanings during infant sleep, Nature communications 6 (2015) 6004.

[7] M. Schabus, G. Gruber, S. Parapatics, C. Sauter, G. Klösch, P. Anderer, W. Klimesch, B. Saletu, J. Zeitlhofer, Sleep spindles and their significance for declarative memory consolidation, Sleep 27 (8) (2004) 1479–1485.

[8] Z. Clemens, D. Fabo, P. Halasz, Overnight verbal memory retention correlates with the number of sleep spindles, Neuroscience 132 (2) (2005) 529–535.

[9] C. M. Held, L. Causa, P. Estévez, C. Pérez, M. Garrido, C. Algarín, P. Peirano, 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.

[10] P. A. Estévez, R. Zilleruelo-Ramos, R. Hernández, L. Causa, C. M. Held, Sleep spindle detection by using merge neural gas, in: International Workshop on Self-Organizing Maps: Proceedings, 2007.

[11] C. R. Patti, S. S. Shahrbabaki, C. Dissanayaka, D. Cvetkovic, Application of random forest classifier for automatic sleep spindle detection, in: 2015 IEEE Biomedical Circuits and Systems Conference (BioCAS), IEEE, 2015, pp. 1–4.

[12] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, Smote: synthetic minority over-sampling technique, Journal of artificial intelligence research 16 (2002) 321–357.

[13] D. Gorur, U. Halici, H. Aydin, G. Ongun, F. Ozgen, K. Leblebicioglu, 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.

[14] D. Görür, U. Halici, H. Aydin, G. Ongun, F. Ozgen, K. Leblebicioglu, Sleep spindles detection using autoregressive modeling, in: Proc. of ICANN/ICONIP, Citeseer, 2003.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in python, Journal of machine learning research 12 (Oct) (2011) 2825–2830.

[16] L. Causa, C. M. Held, J. Causa, P. A. Estévez, C. A. Perez, R. Chamorro, M. Garrido, C. Algarín, P. Peirano, Automated sleep-spindle detection in healthy children polysomnograms, IEEE Transactions on Biomedical Engineering 57 (9) (2010) 2135–2146.

[17] E. Huupponen, A. Värri, S.-L. Himanen, J. Hasan, M. Lehtokangas, J. Saarinen, Optimization of sigma amplitude threshold in sleep spindle detection, Journal of sleep research 9 (4) (2000) 327–334.

[18] E. M. Ventouras, E. A. Monoyiou, P. Y. Ktonas, T. Paparrigopoulos, D. G. Dikeos, N. K. Uzunoglu, C. R. Soldatos, Sleep spindle detection using artificial neural networks trained with filtered time-domain eeg: a feasibility study, Computer methods and programs in biomedicine 78 (3) (2005) 191–207.

[19] S. Devuyst, T. Dutoit, J.-F. Didier, F. Meers, E. Stanus, P. Stenuit, M. Kerkhofs, 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.

[20] S. V. Schönwald, L. Emerson, R. Rossatto, M. L. Chaves, G. J. Gerhardt, Benchmarking matching pursuit to find sleep spindles, Journal of neuroscience methods 156 (1-2) (2006) 314–321.

[21] E. Huupponen, G. Gómez-Herrero, A. Saastamoinen, A. Värri, J. Hasan, S.-L. Himanen, Development and comparison of four sleep spindle detection methods, Artificial intelligence in medicine 40 (3) (2007) 157–170.

[22] R. Bódizs, J. Körmendi, P. Rigó, A. S. Lázár, The individual adjustment method of sleep spindle analysis: methodological improvements and roots in the fingerprint paradigm, Journal of neuroscience methods 178 (1) (2009) 205–213.

[23] L. B. Ray, S. M. Fogel, C. T. Smith, K. R. Peters, Validating an automated sleep spindle detection algorithm using an individualized approach, Journal of sleep research 19 (2) (2010) 374–378.

[24] S. Devuyst, T. Dutoit, P. Stenuit, M. Kerkhofs, 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.

[25] A. Tsanas, G. D. Clifford, Stage-independent, single lead eeg sleep spindle detection using the continuous wavelet transform and local weighted smoothing, Frontiers in human neuroscience 9 (2015) 181.

[26] C. R. Patti, T. Penzel, D. Cvetkovic, Sleep spindle detection using multivariate gaussian mixture models, Journal of sleep research 27 (4) (2018) e12614.

[27] K. Lacourse, J. Delfrate, J. Beaudry, P. Peppard, S. C. Warby, A sleep spindle detection algorithm that emulates human expert spindle scoring, Journal of neuroscience methods 316 (2019) 3–11.