

**Radar-based sleep stage classification in children undergoing polysomnography
a pilot-study**

de Goederen, R.; Pu, S.; Silos Viu, M.; Doan, D.; Overeem, S.; Serdijn, W. A.; Joosten, K. F.M.; Long, X.; Dudink, J.

DOI

[10.1016/j.sleep.2021.03.022](https://doi.org/10.1016/j.sleep.2021.03.022)

Publication date

2021

Document Version

Final published version

Published in

Sleep Medicine

Citation (APA)

de Goederen, R., Pu, S., Silos Viu, M., Doan, D., Overeem, S., Serdijn, W. A., Joosten, K. F. M., Long, X., & Dudink, J. (2021). Radar-based sleep stage classification in children undergoing polysomnography: a pilot-study. *Sleep Medicine*, 82, 1-8. <https://doi.org/10.1016/j.sleep.2021.03.022>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Original Article

Radar-based sleep stage classification in children undergoing polysomnography: a pilot-study



R. de Goederen ^{a,e}, S. Pu ^b, M. Silos Viu ^d, D. Doan ^e, S. Overeem ^{b,c}, W.A. Serdijn ^d, K.F.M. Joosten ^a, X. Long ^{b,1}, J. Dudink ^{e,*}

^a Pediatric Intensive Care Unit, Erasmus MC, Sophia Children's Hospital, Rotterdam, the Netherlands

^b Department of Electrical Engineering, Eindhoven University of Technology, the Netherlands

^c Sleep Medicine Center Kempenhaeghe, Heeze, the Netherlands

^d Section Bioelectronics, Department of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, the Netherlands

^e Department of Neonatology, Wilhelmina Children's Hospital, University Medical Center Utrecht Utrecht, the Netherlands

ARTICLE INFO

Article history:

Received 11 January 2021

Received in revised form

18 March 2021

Accepted 20 March 2021

Available online 29 March 2021

Keywords:

Sleep stages

Radar

Respiration

Breathing rate

Obstructive sleep apnea

Central sleep apnea

ABSTRACT

Study objectives: Unobtrusive monitoring of sleep and sleep disorders in children presents challenges. We investigated the possibility of using Ultra-Wide band (UWB) radar to measure sleep in children.

Methods: Thirty-two children scheduled to undergo a clinical polysomnography participated; their ages ranged from 2 months to 14 years. During the polysomnography, the children's body movements and breathing rate were measured by an UWB-radar. A total of 38 features were calculated from the motion signals and breathing rate obtained from the raw radar signals. Adaptive boosting was used as machine learning classifier to estimate sleep stages, with polysomnography as gold standard method for comparison.

Results: Data of all participants combined, this study achieved a Cohen's Kappa coefficient of 0.67 and an overall accuracy of 89.8% for wake and sleep classification, a Kappa of 0.47 and an accuracy of 72.9% for wake, rapid-eye-movement (REM) sleep, and non-REM sleep classification, and a Kappa of 0.43 and an accuracy of 58.0% for wake, REM sleep, light sleep and deep sleep classification.

Conclusion: Although the current performance is not sufficient for clinical use yet, UWB radar is a promising method for non-contact sleep analysis in children.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sleep is thought to play a crucial role in infants' and children's brain development [1,2]. The human sleep cycle has two main stages: rapid-eye-movement (REM) sleep and Non-REM (NREM) sleep [3]. The latter can be subdivided in three stages: N1 and N2 (light sleep) and N3 (deep sleep, slow wave sleep). REM sleep and NREM sleep each seem to contribute to different aspects of brain maturation [4]. During early brain development REM sleep is thought to provide important endogenous neural stimulation, laying the groundwork for early neural circuitry [4]. NREM sleep seems to be involved in regulating synaptic homeostasis [4].

* Corresponding author. Wilhelmina Children's Hospital, University Medical Center Utrecht, Lundlaan 6, 3584, EA, Utrecht, the Netherlands.

E-mail address: j.dudink@umcutrecht.nl (J. Dudink).

¹ Shared last authorship.

While acute or transient changes in sleep may impact a person's cognitive and behavioral performance [5,6], long-term sleep deprivation or sustained changes in sleep patterns carry the risk factor for several diseases and suboptimal development (eg, diabetes, obesity, mood swings, suboptimal growth and school performance) [7,8]. Analyzing the sleep duration and the distribution of sleep stages may well offer useful clinical insights and deeper understanding of sleep physiology and pathology in young children.

To date, a level 1 polysomnography (PSG) with clinical evaluation remains the gold standard method to study sleep. PSG has been widely used in clinical and laboratory settings through multi-channel bio-signal recordings manually assessed by sleep technologists. Performing a PSG can be a very challenging procedure for children, and a level 1 full PSG cannot be performed easily outside a laboratory setting [9–11].

These limitations promote the demand for an unobtrusive, or even non-contact sleep monitoring system for home monitoring. Assessing sleep in the child's safe environment and without sensors

attached to its body for several days can provide valuable information and could serve as screening method for the necessity of an in-hospital PSG. Several techniques have been suggested for this purpose, such as actigraphy, radar, ballistocardiography, or Doppler laser [12]. These techniques make use of non-EEG vital signs and behavioral measurements that are easier to obtain unobtrusively, such as heart rate, respiration and body movements.

Radar is a contactless method for the monitoring of human vital signs and relies on the modulation effect, due to the chest-wall displacement of a radio signal sent by a transmitter towards the patient. Depending on the type of signal it transmits, four types of radar can be distinguished: pulsed wave, continuous wave (CW), frequency-modulated continuous-wave (FMCW) or stepped-frequency continuous-wave (SFCW) [13].

CW radar systems, such as Doppler radar, have frequently been used for vital sign monitoring because of their relatively low power consumption and simple radio architecture [14]. Several studies have shown that CW radar is a feasible method for vital sign estimation [15–19]. However, complex signal processing techniques are required to accurately detect and measure these vital signs, and these techniques increase the power consumption significantly. These technical difficulties hinder the use of CW radar for widespread public usage [13]. Kagawa et al. investigated sleep stage classification using two Doppler radars, but showed lower accuracy compared to other non-contact methods [20].

Moreover, frequency modulation of the radar is needed for the detection of multiple subjects and human presence [21]. SFCW is not commonly used for human vital sign detection. FMCW on the other hand already demonstrated its potential for vital sign monitoring [22–25]. Turppa et al. [24] demonstrated that FMCW can provide accurate vital sign monitoring during sleep. However, unlike the radars discussed above, pulse-based radars are able to detect vital signs of humans behind solid objects and have high power efficiency [13]. Furthermore, a comparison by Wang et al. showed that Impulse radio ultra-wide band radar (IR-UWB) had better accuracy ratios than FMCW [26].

UWB radar is the most commonly used pulse-based radar and is considered a promising and reliable technique that can capture patients' vital signs without contacting the body [27]. By measuring reflected RF signals, the distance and movement of an object can be determined [14]. This way, movements of the chest and abdomen can be quantified. The UWB radar technique has a high range resolution and a strong penetration with good capability to distinguish between multiple targets [28].

Several studies have used radar technology to measure respiratory information [29–34]. One of these, by Immoreev and Tao [33] measured the agreement between UWB radar and impedance pneumography in four infants under the age of two months. The authors reported that >95% of sample pair differences lay within the $\pm 95\%$ confidence interval. Kang et al. have used IR-UWB radar to measure sleep and sleep apnea and compared it with a PSG with good results in 21 adults [35]. At the moment, however, studies validating UWB radar sleep data against PSG as the gold standard are still very limited, especially in children.

The aim of our study was to make a pediatric sleep stage classification algorithm based on UWB radar data with PSG as gold standard method; a secondary aim was to assess the UWB radar accuracy in automatically determining sleep stages.

2. Methods

2.1. Study population, inclusion and exclusion criteria

The study was conducted at the Erasmus MC - Sophia Children's Hospital Rotterdam. The institutional medical ethics review board

approved this study (Rotterdam, the Netherlands. File No. MEC-2017-159). Children scheduled to undergo a clinical PSG were eligible for inclusion in this study. Parents or guardians of the children provided informed consent was given by all, and children from the age of 12 years provided consent themselves. A subject was excluded in case of technical failure of the sensor or the PSG.

2.2. Study set-up

Connected to a stand-alone laptop that stored the radar data, the UWB-sensor was placed on a standard at the head of the bed at approximately 1.5 m distance from the child, as shown in Fig. 1. While the radar was functioning, the clinically indicated level 1 PSG was carried out as usual.

2.3. Polysomnography

The level 1 PSG included the following: ECG, 14-channel EEG, nasal airflow (thermistor), video recording, chest and abdominal wall motion (plethysmography), a capillary blood gas test, arterial blood oxygen-hemoglobin saturation using pulse oximetry (SpO_2), and transcutaneous partial pressure of carbon dioxide ($tcpCO_2$). The monitoring was a single-night recording. Sleep stages were scored on each 30-s non-overlapping epoch according to the rules of the American Association of Sleep Medicine (AASM) [3]. Respiratory events were also scored according to the rules of the latest AASM criteria and the apnea-hypopnea index (AHI) was calculated [3]. The child was classified as having obstructive sleep apnea (OSA) if the obstructive AHI was higher than 1 event per hour [37], and central sleep apnea (CSA) if the central AHI was higher than 1 [38]. The limited restricted comparison to only OSA vs. non-OSA, and CSA vs. non-CSA. These groups were compared to the total group. To explore the effect of age, the population was split in a group younger than 1 year old and a group older than 1 year old.

2.4. Radar system

The XeThru X2M200 and X4M200 radar modules (Novelda AS, Oslo, Norway) [39] were used for this research. They rely on observing periodic movements when a person is resting and breathing. These modules make use of pulse-Doppler processing, in which coherent pulses produced by a local oscillator are transmitted through an antenna. These pulses propagate through space until they meet reflectors. Some of the transmitted energy will be reflected back to the receiver along with phase modulation caused by motion. The received RF signals are then down-converted by the local oscillator to a baseband signal. The baseband signal is split into two quadrature signals [40] (see Fig. 2). The XeThru modules we used to convert the RF signals into baseband frequencies of 20 Hz in X2M200 and 17 Hz in X4M200. Both radar modules provide In-Phase and Quadrature (IQ) or amplitude/phase (AP) data.

The amplitude baseband data was used in this study because magnitudes carry the most comprehensible information. Body motion and respiration rate were obtained by processing these baseband data. The motion signal was created by integrating the differences between two subsequent time frames across the amplitude baseband signals [36]. It is expressed by

$$MVM_{win} = \sum_{t=t_i-win+1}^{t_i} |[A(t) - A(t-1)]|$$

where MVM_{win} represents the movement quantity within a time window win at a specific sample point t . A is the amplitude of the baseband data, and t_i is the sample point in time.

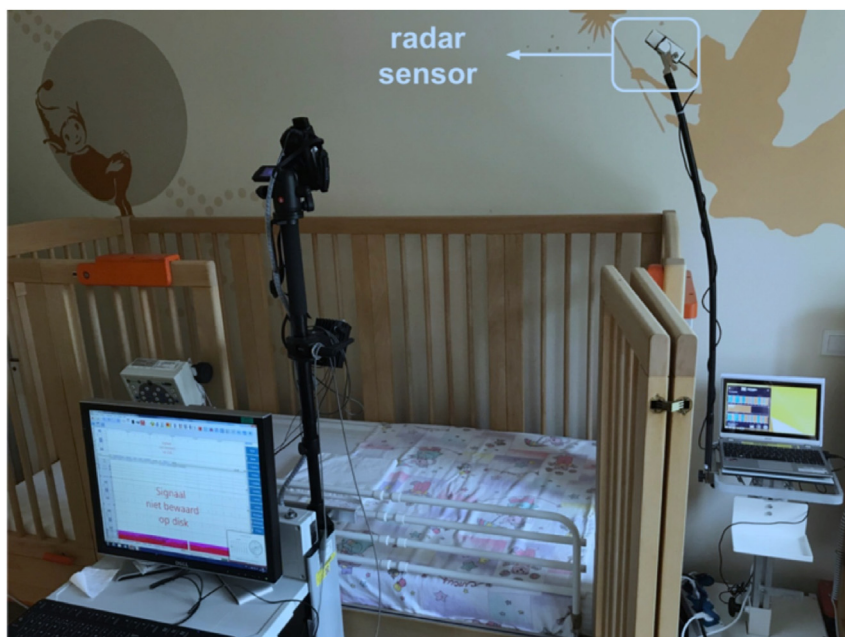


Fig. 1. Radar sensor setup at the Erasmus MC - Sophia Children's Hospital, the Netherlands [36].

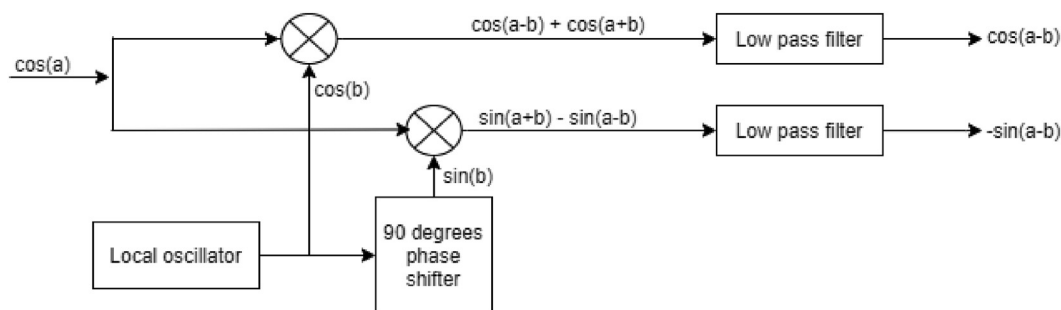


Fig. 2. Block diagram of I/Q mixer analog implementation in a radar system [41].

A Fast Fourier Transform (FFT) analysis on amplitude baseband data was then performed to create a 'range-frequency' matrix. Static objects were removed, and small movements of breathing were measured. By means of localizing the peak of the range-frequency matrix, the breathing rate (BR, in respirations per minute) was detected (see Fig. 3).

2.5. Feature extraction

Thirty-eight motion and respiratory features were extracted in this study, summarized in Table 1. Nine features came from the motion signals 'twenty-nine from the BR.

1) *Respiratory features*: The variance and average of the BR from a 6 min epoch were computed. Also, the variance and mean value of BR were computed over an epoch. Besides, the sample entropy, approximate entropy, and multiscale sample entropy scaled from 1 to 10 with different time delays were computed over an epoch [42–44]. These features were obtained by various methods to measure the complexity and similarity of time-series signals. The mean values of Teager energy in both the time domain and frequency domain within an epoch were calculated [45,46]. Lastly, a robust algorithm, Katz's fractal dimension, was applied to

estimate the fractal dimension of the signals [47]. All these features were subjected to a Z-score normalization by subtracting the mean of the feature values and dividing by their standard deviation for each recording. The normalization step aims to decrease physiological variance from subject to subject by reducing data redundancy and improving data integrity.

2) *Motion feature*: On the analysis of the baseband data, two motion signals were obtained by the motion detection algorithm in different settings. One was the signal with a window of 1 s, which represented the amount of the movements within 1 s, *MVM1*, also called fast movement signal. The other one was the movement quantity for a time window of 20 s, *MVM20*. These motion signals were normalized by their maximum values. The average, area, variance, and entropies of motion signals were computed. Besides, we computed the fast movement ratio, ie, the relative time within 10 min in which *MVM1* exceeds a threshold of 10% of the 20-s window signal.

2.6. Classifier

An adaptive boosting (AdaBoost) algorithm based on decision trees was adopted for the classification tasks due to its multiple advantages in good classification performance, low susceptibility to

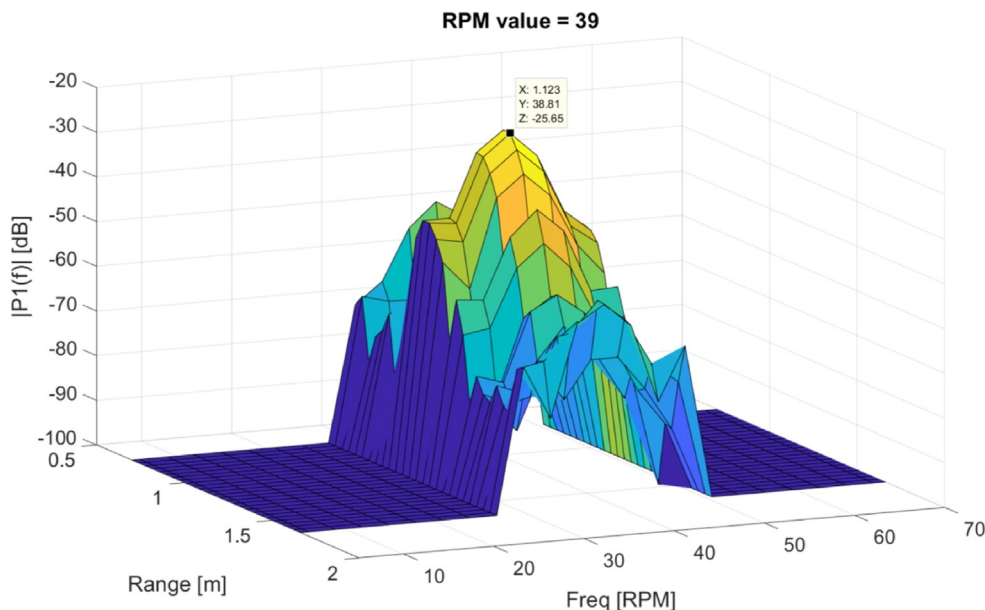


Fig. 3. An example of a frequency analysis, in which the breathing rate was 39 respirations per minute [36].

Table 1
Description of all features used for the machine learning algorithm.

Feature Index (#)	Description
1	BR Variance of BR 6 min centered at the epoch
2	BR Average of BR 6 min centered at the epoch
3	BR Variance of BR
4	BR Average of BR
5	BR Sample entropy of BR
6	BR Approximate entropy of BR
7–16	BR Multiscale sample entropy scaled from 1 to 10 with unity delay
17–26	BR Multiscale sample entropy scaled from 1 to 10 with a delay of 2 units
27	BR Teager Energy in the frequency domain
28	BR Teager Energy in the time domain
29	BR Katz's fractal dimension
30	Motion Average of MVM20 signal
31	Motion Average of MVM1 signal
32	Motion Area of MVM20 signal
33	Motion Area of MVM1 signal
34	Motion Variance of MVM20 signal
35	Motion Variance of MVM1 signal
36	Motion Information entropy of MVM1 signal
37	Motion Information entropy of MVM20 signal
38	Motion Fast movement ratio within the last 10 min

overfitting, and relatively high computational efficiency. For comparison, we used two other popular and widely used classification algorithms, the k-nearest neighbors (KNN) algorithm and support vector machine (SVM).

2.7. Evaluation

Leave-one-out cross validation was conducted, which implies that each patient in turn was used as a test set while the remaining ones form a training set during each iteration of the Leave-one-out cross validation. Classification results and performance metrics from each patient during each test iteration were obtained. Three sleep stage classification tasks were performed and evaluated: 1) Wake-sleep (WS) classification, 2) Wake-REM-NREM (WRN) classification, and 3) Wake-REM-Light-Deep (WRLD) classification. To

evaluate the performance of the classifiers, the overall accuracy was computed. As we were dealing with class imbalance, Cohen's Kappa coefficient of agreement was also computed. In addition, to show the discriminative power of a single feature towards the output, we calculated the feature importance for each iteration. Furthermore, an unpaired Wilcoxon rank-sum test served to test for significant differences in performance between different groups.

3. Results

3.1. Subjects

Initially, 40 patients were included. Six had to be excluded because of software issues. These issues were resolved for the subsequent measurements. Two with severe mental retardation did not have a single episode of normal REM-sleep during the recording, and were excluded as well. Thus, data of 32 patients were analyzed. Characteristics of these patients are presented in Table 2.

3.2. Data

Fig. 4 displays the performance of the three different classifiers (SVM with a linear kernel, AdaBoost, and KNN) when using all 38 features. KNN presented a relatively low Kappa coefficient for all the classification tasks, in particular for WRLD classification with a significantly lower Kappa ($p < 0.01$) than the other two classifiers. AdaBoost and SVM showed similar performance, but because the computational cost of AdaBoost was much lower than that of SVM, we used AdaBoost in the remainder of this study.

Fig. 5 shows the importance of each feature within the model for the three classification tasks. The feature numbers in Fig. 5 correspond to the features listed in Table 1. This quality indicates that the motion features and respiratory features are complimentary. In general, motion features occupy the leading positions in the classifications due to their capability of distinguishing between wake and sleep states. For example, the motion feature average of MVM20 signal within 30 s (#30), the variance of MVM1 signal within 30 s (#35), and the fast movement ratios within the last

Table 2
Patient characteristics.

Characteristics	
N	32
Age (range)	3.25 y (2 mo–14 y)
<1 y	15
1 y–12 y	15
>12 y	2
Syndromes/Diagnosis	
Bronchopulmonary dysplasia	6
Pierre Robin Sequence	4
Spinal Muscular Atrophy type II	4
Down syndrome	3
Crouzon syndrome	2
Apert syndrome	1
Myasthenia gravis	1
Neuroendocrine hyperplasia of infancy (NEHI)	1
Prader–Willi syndrome	1
Saethre–Chotzen syndrome	1
Tracheomalacia	1
other (Seizures, ex-preterm, MECP2-duplication)	3
No syndrome	2
Unknown	2
Sleep stages, % (mean in minutes)	
Wake	10.6 (20.0, range 3.7–44.1)
REM	6.3 (19.4, range 7.0–29.0)
NREM	9.2 (59.8, range 43.0–78.2)
Light Sleep	11.1 (30.4, range 2.1–59.6)
Deep Sleep	9.4 (29.4, range 14.8–52.1)
TST (min)	501.8 (88.5)
Arousal index	8.7 (6.6)
Parameters	
oAHI	mean 3.7 (range 0.0–43.5)
oAHI ≤ 1	20
1 < oAHI ≤ 5	8
5 < oAHI ≤ 10	0
oAHI > 10	4
cAHI	mean 2.3 (range 0.0–11.3)
cAHI > 1	21

10 min (#38) were of high importance. Interestingly, the BR feature variance of BR 6 min centered at the epoch (#1), and Katz's fractal dimension within 30 s (#29), were given higher scores in WRN and WRLD classifications than in WS classification.

The summary of the classification performance per patient group is presented in Table 3. For all patients, we achieved a mean Kappa of 0.67 ± 0.14 with an overall accuracy of $89.82 \pm 5.5\%$ for the WS classification, a mean Kappa of 0.47 ± 0.12 with an accuracy of $72.93 \pm 6.8\%$ for the WRN classification, and a mean Kappa of 0.43 ± 0.11 with an accuracy of $57.99 \pm 7.7\%$ for the WRLD classification. The summary of the classification performance presented as positive predictive value (PPV) and sensitivity is shown in Table 4.

Fig. 6 illustrates the classification results (Kappa values) and the sleep stage distribution for each patient. Accuracy for Non-OSA was higher than that OSA for all classifiers: 90.71% vs. 86.13% for WS; 71.67% vs. 68.80% for WRN, and 57.98% vs. 53.66% for WRLD. Accuracy for Non-CSA was lower than that for CSA, with 88.49% vs. 89.12% for WS; 69.64% vs. 71.48% for WRN; and 54.29% vs. 57.72% for WRLD classification.

4. Discussion

This study shows that sleep stage classification in children can be accurately assessed using UWB-radar technology and shows that it can be a reliable technique to contactlessly assess children's sleep. From the machine learning classifiers that we tested in this study, AdaBoost proved to be the most accurate classifier for the machine learning algorithm. This study effectively identified the most useful

features for the classification tasks. While movement features were more important for the wake-sleep differentiation, breathing rate features were more important to differentiate between sleep stages. Differentiating between wake and sleep was the most accurate. Differentiating between more sleep stages became less accurate, with the WRLD classification being the least accurate.

The accuracy of our machine learning algorithm reached was similar to that reported in other studies [20,48–53]. Some of these studies have used UWB radar, similar to our study; others have used ECG signals, video-based actigraphy, or combined different types of sensors and microphones to get the results. However, most studies show similar accuracies for the three types of sleep stage classification: approximately 90% for the WS classification, 60–70% for the WRN classification, and 55–75% for WRLD classification [49].

In the present study, we have made use of UWB radar technology to classify sleep stages. At the moment, several other modalities may be used for this purpose. Probably the most popular method at the moment is actigraphy with the use of a watch. ECG and methods with other kinds of bands, such as ballistocardiography, are also used to estimate sleep [49,54–56]. In the end, it comes down to different methods of quantifying physiologic signals and using a machine learning algorithm to estimate sleep variables. The advantage of UWB radar is that it is a truly unobtrusive device to measure a person's breathing and movement during sleep. This non-obtrusive tool has the potential of increasing patient compliance in sleep diagnostics. The device can be placed on a fixed location next to the bed, and does not have to be attached to the person's body. Furthermore, close to or within the device, several other factors can be measured, such as room temperature, sound, and light exposure. Note that UWB radar is unable to measure oxygen saturation. To measure oxygen saturation, a single extra sensor could be added to the set-up.

For sleep stage classification, many studies have focused on extracting informative features from respiratory signals. For instance, features such as breathing rate, respiratory self-similarity and regularity and inhalation/exhalation rate and volume have been applied [50,57]. Very few sleep staging studies using radar technology have been performed in infants and children. Some of these studies applied other contactless monitoring approaches (eg, capacitive ECG and video camera) to acquire vital signs for sleep stage classification. Werth et al. [53] developed an algorithm to automatically detect sleep stages from capacitive ECG data from 8 preterm infants (gestational age 30 ± 2.5 weeks). In that study, a performance of Kappa 0.44 was achieved for classifying active sleep, quiet sleep, and caretaking and wakefulness. Recently, Long et al. [51] used video-based actigraphy to identify wakefulness and sleep states in 10 healthy term infants (<18 months), whereby they obtained a Kappa of 0.73 and an overall accuracy of 92% [51].

This study has several limitations. For one thing, the wide age range of subjects, from newborn infants to adolescents. For the sleep stage classification, age could not be included in our machine learning algorithm. In our exploratory analysis, however, there was no difference in accuracy between children younger than 1 year of age and children older than 1 year of age.

Second, there was also a wide range of diseases among the subjects. As this pilot study was performed in a tertiary referral university medical center, the study population does not reflect the average pediatric population. Further research has to be done to also validate the algorithm in a healthy population.

Thirdly, some of the subjects in this study had OSA. This means that their breathing pattern and movement pattern were disturbed by apneas and arousals during the measurements. Due to the relatively small sample size, we could not correct for apneas and arousals in this particular study. OSA seemed to have some negative effect on the accuracy of the algorithms, but the patient numbers in

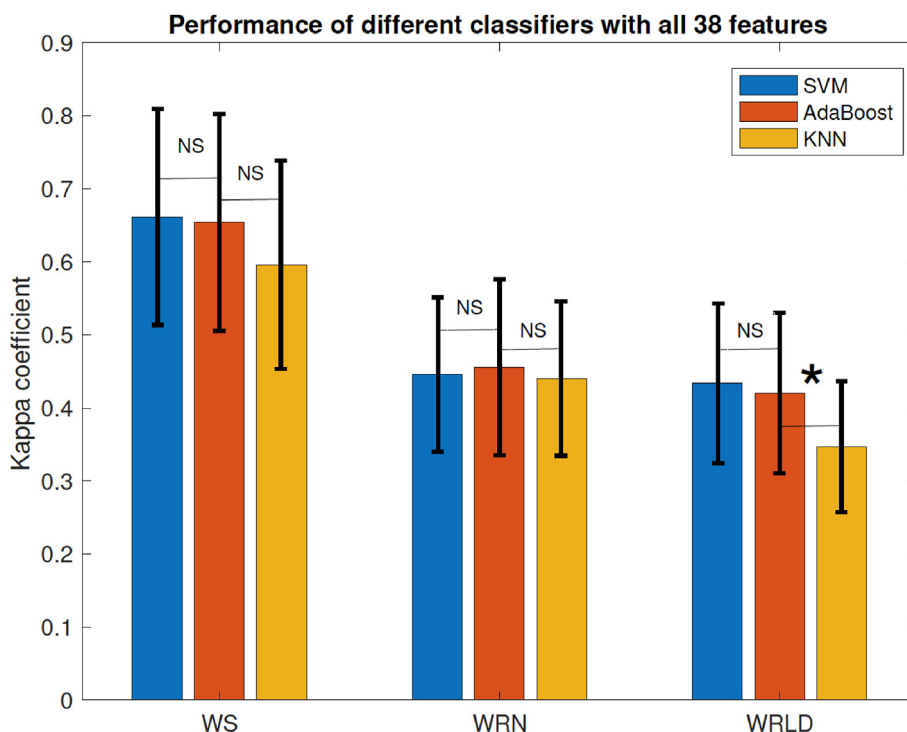


Fig. 4. Performance of sleep stage classification using different classifiers. Abbreviation: NS = not significant at $p = 0.05$, $*p < 0.01$.

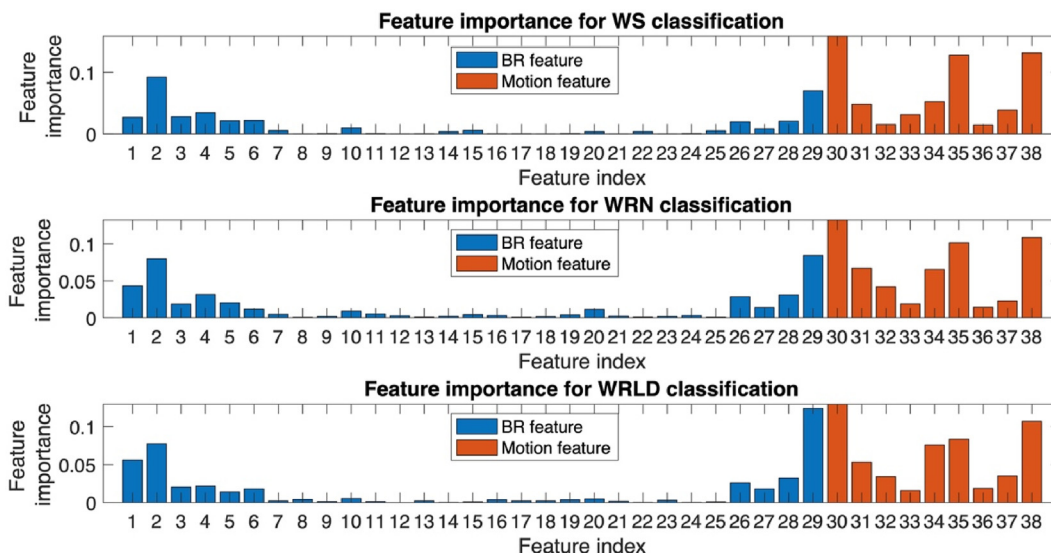


Fig. 5. The feature importance (feature weight in AdaBoost classifier) for WS, WRN, and WRLD classifications derived from data of all patients.

Table 3
Performance of the Adaboost algorithms in different groups, presented as Kappa and accuracy.

	WS		WRN		WRLD	
	Kappa	Acc. [%]	Kappa	Acc. [%]	Kappa	Acc. [%]
All	0.67 (0.14)	89.82 (5.5)	0.47 (0.12)	72.93 (6.8)	0.43 (0.11)	57.99 (7.7)
Non-OSA	0.70 (0.14)*	90.71 (5.0)	0.48 (0.11)*	71.67 (6.8)	0.43 (0.10)*	57.98 (7.6)
OSA	0.54 (0.16)	86.13 (7.7)	0.39 (0.12)	68.80 (6.4)	0.36 (0.11)	53.66 (7.1)
Non-CSA	0.60 (0.13)	88.49 (6.3)	0.38 (0.11)	69.64 (8.1)	0.35 (0.10)	54.29 (12.6)
CSA	0.67 (0.16)	89.12 (6.2)	0.48 (0.11)*	71.48 (6.6)	0.43 (0.09)	57.72 (6.7)
Young (<1 yr)	0.69 (0.12)	89.28 (5.3)	0.48 (0.10)	70.74 (6.5)	0.44 (0.07)	58.56 (5.5)
Old (>1 yr)	0.66 (0.16)	90.53 (5.3)	0.45 (0.11)	73.37 (6.4)	0.38 (0.10)	55.00 (6.7)

* $p < .05$.

Table 4
Performance of the Adaboost algorithms presented as positive predictive value (PPV) and sensitivity. The results are specified per classification outcome.

Adaboost	WS		WRD			WLRD			
	Wake	Sleep	Wake	REM	Deep	Wake	REM	Light	Deep
PPV (mean, SD)	0.72, 0.15	0.94, 0.04	0.73, 0.17	0.57, 0.25	0.73, 0.08	0.69, 0.18	0.54, 0.21	0.43, 0.14	0.63, 0.15
Sensitivity (mean, SD)	0.77, 0.17	0.92, 0.07	0.77, 0.15	0.22, 0.14	0.88, 0.06	0.82, 0.14	0.46, 0.19	0.34, 0.13	0.76, 0.14

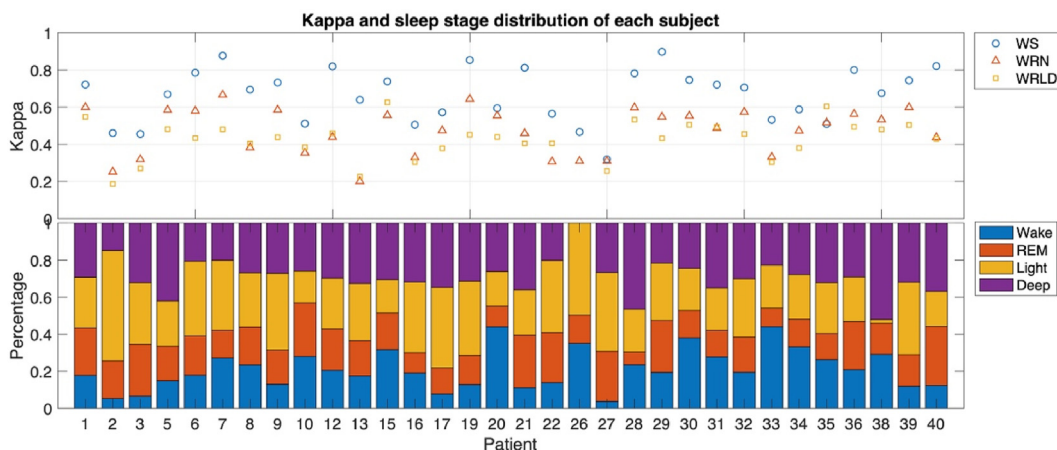


Fig. 6. Sleep stage classification performance and sleep stage distribution of each patient.

this study are too small to correct for having OSA. In future studies, investigators should try to only include subjects that do not suffer from OSA or any other disease.

Lastly, the hypnograms of the polysomnography are scored by three different observers. Although they were blinded for the UWB-radar results, this may have caused some bias in the gold standard method of determining sleep stages and might have lowered the accuracy of the sleep stage classification by the UWB-radar sensor. In an ideal situation, every hypnogram should have been scored by two observers independently, and their results combined through consensus.

Even given these limitations, the yielded results tally well with those of other studies and we therefore believe that our results could have been even better in a larger and more homogenous study group.

Our current study has only focused on sleep stage classification, even though many other aspects in a night's sleep can be evaluated. Future studies could focus on adding new parameters to the algorithm, such as the total sleep time, sleep onset latency, and wake time after sleep onset. Future studies could also focus on detecting obstructive and central sleep apneas, and might be able to detect sleep apnea. Some studies in adults have already proven the feasibility of detecting apneas using an UWB radar [58,59].

In conclusion, this study shows that sleep classification using an UWB radar could be a feasible non-contact method in children. Since this was a pilot study with a heterogeneous study population, the accuracy of the sleep stage classification is not yet sufficient for clinical use. However, the results are promising for the future of this radar technique in sleep diagnostics. This study paves the way for more in-depth and more detailed studies on sleep quantification using radar-technologies.

Author credit statement

R. de Goederen: Conceptualization, Methodology, Writing – original draft. S. Pu: Methodology, Software, Formal analysis, Writing – original draft. M. Silos Viu: Conceptualization, Software. D. Doan: Writing – review & editing. S. Overeem: Methodology,

Formal analysis, Supervision. W.A. Serdijn: Conceptualization, Methodology, Supervision. K.F.M. Joosten: Conceptualization, Project administration. X. Long: Methodology, Formal analysis, Supervision. J. Dudink: Conceptualization, Supervision, Project administration.

Financial\Non-financial disclosure

None.

Acknowledgements

This work was done in part in the Eindhoven MedTech Innovation Center (e/MTIC).

Conflict of interest

No competing interests to declare.

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: <https://doi.org/10.1016/j.sleep.2021.03.022>

References

- [1] Matricciani L, Blunden S, Rigney G, et al. Children's sleep needs: is there sufficient evidence to recommend optimal sleep for children? *Sleep* 2013;36(4):527–34.
- [2] Watson NF, Badr MS, Belenky G, et al. Joint consensus statement of the American academy of sleep medicine and sleep research society on the recommended amount of sleep for a healthy adult: methodology and discussion. *Sleep* 2015;38(8):1161–83.
- [3] Berry RB, Budhiraja R, Gottlieb DJ, et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events. Deliberations of the sleep apnea definitions task force of the American academy of sleep medicine. *J Clin Sleep Med* 2012;8(5):597–619.
- [4] Knoop MS, de Groot ER, Dudink J. Current ideas about the roles of rapid eye movement and non-rapid eye movement sleep in brain development. *Acta Paediatr* 2021 Jan;110(1):36–44. <https://doi.org/10.1111/apa.15485>. Epub 2020 Aug 8. PMID: 32673435; PMCID: PMC7818400.
- [5] Alhola P, Polo-Kantola P. Sleep deprivation: impact on cognitive performance. *Neuropsychiatric Dis Treat* 2007;3(5):553–67.

- [6] Dawson D, Reid K. Fatigue, alcohol and performance impairment. *Nature* 1997;388(6639):235.
- [7] Bonnet MH, Arand DL. We are chronically sleep deprived. *Sleep* 1995;18(10):908–11.
- [8] Medic G, Wille M, Hemels ME. Short- and long-term health consequences of sleep disruption. *Nat Sci Sleep* 2017;9:151–61.
- [9] Tamaki M, Nittono H, Hayashi M, et al. Examination of the first-night effect during the sleep-onset period. *Sleep* 2005;28(2):195–202.
- [10] Agnew Jr HW, Webb WB, Williams RL. The first night effect: an EEG study of sleep. *Psychophysiology* 1966;2(3):263–6.
- [11] Scholle S, Scholle HC, Kemper A, et al. First night effect in children and adolescents undergoing polysomnography for sleep-disordered breathing. *Clin Neurophysiol* 2003;114(11):2138–45.
- [12] Werth J, Atallah L, Andriessen P, et al. Unobtrusive sleep state measurements in preterm infants - a review. *Sleep Med Rev* 2017;32:109–22.
- [13] Kebe M, Gadhafi R, Mohammad B, et al. Human vital signs detection methods and potential using radars: a review. *Sensors* 2020;20(5).
- [14] Li C, Lubecke VM, Boric-Lubecke O, et al. A review on recent advances in Doppler radar sensors for noncontact healthcare monitoring. *IEEE Trans Microw Theor Tech* 2013;61(5):2047–60.
- [15] Droitcour AD, Seto TB, Park BK, et al. Non-contact respiratory rate measurement validation for hospitalized patients. *Annu Int Conf IEEE Eng Med Biol Soc* 2009;2009:4812–5.
- [16] Iwata Y, Ishibashi K, Sun G, et al. Contactless heartbeat detection from CW-Doppler radar using windowed-singular spectrum analysis(.). *Annu Int Conf IEEE Eng Med Biol Soc* 2020;2020:477–80.
- [17] Kim JY, Park JH, Jang SY, et al. Peak detection algorithm for vital sign detection using Doppler radar sensors. *Sensors* 2019;19(7).
- [18] Nejadgholi I, Sadreazami H, Baird Z, et al. Estimation of breathing rate with confidence interval using single-channel CW radar. *J Healthc Eng* 2019;2019:2658675.
- [19] Li M, Lin J. Wavelet-transform-based data-length-variation technique for fast heart rate detection using 5.8-GHz CW Doppler radar. *IEEE Trans Microw Theor Tech* 2018;66(1):568–76.
- [20] Kagawa M, Suzumura K, Matsui T. Sleep stage classification by non-contact vital signs indices using Doppler radar sensors. In: *Conf Proc : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference*. 2016; 2016. p. 4913–6.
- [21] Lee H, Kim BH, Yook JG. Path loss compensation method for multiple target vital sign detection with 24-GHz FMCW radar. In: *Proceedings of the 2018 IEEE asia-pacific conference on antennas and propagation (APCAP)*; 2018. p. 100–1.
- [22] Wang S, Pohl A, Jaeschke T, et al. A novel ultra-wideband 80 GHz FMCW radar system for contactless monitoring of vital signs. *Annu Int Conf IEEE Eng Med Biol Soc* 2015;2015:4978–81.
- [23] Sacco G, Piuze E, Pittella E, et al. An FMCW radar for localization and vital signs measurement for different chest orientations. *Sensors* 2020;20:3489.
- [24] Turppa E, Kortelainen JM, Antropov O, et al. Vital sign monitoring using FMCW radar in various sleeping scenarios. *Sensors* 2020;20:6505.
- [25] Wang Y, Wang W, Zhou M, et al. Remote monitoring of human vital signs based on 77-GHz mm-wave FMCW radar. *Sensors* 2020;20.
- [26] Wang D, Yoo S, Cho SH. Experimental comparison of IR-UWB radar and FMCW radar for vital signs. *Sensors* 2020;20:6695.
- [27] Schleicher B, Nasr I, Trasser A, et al. IR-UWB radar demonstrator for ultra-fine movement detection and vital-sign monitoring. *IEEE Trans Microw Theor Tech* 2013;61:2076–85.
- [28] Li J, Zeng Z, Sun J, et al. Through-wall detection of human being's movement by UWB radar. *Geosci Rem Sens Lett IEEE* 2019;9:1079–83.
- [29] Hafner N, Mostafanezhad I, Lubecke VM, et al. Non-contact cardiopulmonary sensing with a baby monitor. In: *Conf Proc : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference*. 2007; 2007. p. 2300–2.
- [30] Huang X, Sun L, Tian T, et al. Real-time non-contact infant respiratory monitoring using UWB radar. In: *International Conference on Communication Technology Proceedings, ICCT*. 2016-Febru; 2016. p. 493–6.
- [31] Nijssure Y, Tay WP, Gunawan E, et al. An impulse radio ultrawideband system for contactless noninvasive respiratory monitoring. *IEEE Trans Biomed Eng* 2013;60(6):1509–17.
- [32] Ziganshin EG, Numerov MA, Vygolov SA. UWB baby monitor. 2010 5th international conference on ultrawideband and ultrashort Impulse signals. 2010. p. 6–10.
- [33] Immoreev I, Tao T-H. UWB radar for patient monitoring. *Aerospace and Electronic Systems Magazine, IEEE* 2008;23(11):11–8.
- [34] Kim JD, Lee WH, Lee Y, et al. Non-contact respiration monitoring using impulse radio ultrawideband radar in neonates. *R Soc Open Sci* 2019;6(6):190149.
- [35] Kang S, Lee Y, Lim YH, et al. Validation of noncontact cardiorespiratory monitoring using impulse-radio ultra-wideband radar against nocturnal polysomnography. *Sleep Breath* 2020;24(3):841–8.
- [36] Silos Viu M. SOMNUS: an ultra-wideband radar-based approach for neonatal sleep state classification. Delft: Delft University of Technology; 2017.
- [37] Dehlink E, Tan H-L. Update on paediatric obstructive sleep apnoea. *J Thorac Dis* 2016;8(2):224–35.
- [38] McLaren AT, Bin-Hasan S, Narang I. Diagnosis, management and pathophysiology of central sleep apnea in children. *Paediatr Respir Rev* 2019;30:49–57.
- [39] Sjöman H, Soares N, Suijkerbuijk M, et al. The breathing room: breathing interval and heart rate capturing through ultra low power radar. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal QC, Canada: Association for Computing Machinery; 2018. D207.
- [40] Ragaini E, Woodman RF. Demodulation of complex baseband radar signals for the analysis of multiple narrow spectral lines. *Radio Sci* 1997;32(2):783–9.
- [41] Neji S, Kooji RPC. Radar based human vital sign detection in cars: a system Analysis. Delft: Delft University of Technology; 2020.
- [42] Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiological signals. *Circulation* 2000;101(23):E215–20.
- [43] Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of biological signals. *Phys Rev* 2005;71(2):021906.
- [44] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 2000;278(6):H2039–49.
- [45] Kamath C. ECG beat classification using features extracted from teager energy functions in time and frequency domains. *IET Signal Process* 2011;5(6):575–81.
- [46] Some useful properties of Teager's energy operators. In: Kaiser JF, editor. *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*; 1993. 27–30 April 1993.
- [47] David MG, Naiara Maria de S, Luiz Carlos MV. Heart rate variability analysis: Higuchi and Katz's fractal dimensions in subjects with type 1 diabetes mellitus. *Rom J Diabetes Nutr Metab Dis* 2018;25(3):289–95.
- [48] Chung KY, Song K, Shin K, et al. Noncontact sleep study by multi-modal sensor fusion. *Sensors* 2017;17(7).
- [49] Fonseca P, van Gilst MM, Radha M, et al. Automatic sleep staging using heart rate variability, body movements, and recurrent neural networks in a sleep disordered population. *Sleep* 2020;43(9):zsaa048. <https://doi.org/10.1093/sleep/zsaa048>.
- [50] Tataraidze A, Korostovtseva L, Anishchenko L, et al. Bioradiolocation-based sleep stage classification. In: *Conf Proc IEEE Eng Med Biol Soc*. 2016; 2016. p. 2839–42.
- [51] Long X, Otte R, Sanden EV, et al. Video-based actigraphy for monitoring wake and sleep in healthy infants: a laboratory study. *Sensors* 2019;19(5).
- [52] Ranta J. Sleep classification in infants using a bed mattress sensor and analysis of ECG. Espoo: Aalto University; 2018.
- [53] Werth J, Serteyn A, Andriessen P, et al. Automated preterm infant sleep staging using capacitive electrocardiography. *Physiol Meas* 2019;40(5).
- [54] Haghayegh S, Khoshnevis S, Smolensky MH, et al. Accuracy of wristband fitbit models in assessing sleep: systematic review and meta-analysis. *J Med Internet Res* 2019;21(11):e16273.
- [55] Miller DJ, Lastella M, Scanlan AT, et al. A validation study of the WHOOP strap against polysomnography to assess sleep. *Sci Sports Sci* 2020 Nov;38(22):2631–6. <https://doi.org/10.1080/02640414.2020.1797448>. Epub 2020 Jul 26. PMID: 32713257.
- [56] Park KS, Choi SH. Smart technologies toward sleep monitoring at home. *Biomed Eng Lett* 2019;9(1):73–85.
- [57] Long X, Yang J, Weysen T, et al. Measuring dissimilarity between respiratory effort signals based on uniform scaling for sleep staging. *Physiol Meas* 2014;35(12).
- [58] Kang S, Kim DK, Lee Y, et al. Non-contact diagnosis of obstructive sleep apnea using impulse-radio ultra-wideband radar. *Sci Rep* 2020;10(1):5261.
- [59] Zhou Y, Shu D, Xu H, et al. Validation of novel automatic ultra-wideband radar for sleep apnea detection. *J Thorac Dis* 2020;12(4):1286–95.