Bioradiolocation-Based Multi-Class Sleep Stage Classification Using Time and Frequency Features with Random Forest Classifier

Conference Paper · December 2022

DOI: 10.1109/ICECE57408.2022.10089093

4 authors:

Shakhrul Iman Siam
The Ohio State University
6 PUBLICATIONS **33** CITATIONS
SEE PROFILE

Md. Saiful Bari Siddiqui
BRAC University
4 PUBLICATIONS **6** CITATIONS
SEE PROFILE

Mushfiqul Abedin Khan
The University of Texas at Dallas
9 PUBLICATIONS **65** CITATIONS
SEE PROFILE

M. I. H. Bhuiyan
Bangladesh University of Engineering and Technology
121 PUBLICATIONS **3,696** CITATIONS
SEE PROFILE

Bioradiolocation-Based Multi-Class Sleep Stage Classification Using Time and Frequency Features with Random Forest Classifier

Md Shakhrul Iman Siam[1], Md Saiful Bari Siddiqui, Mushfiqul Abedin, Mohammed Imamul Hassan Bhuiyan[2]
Department of Electrical and Electronic Engineering,
Bangladesh University of Engineering and Technology, Dhaka 1205, Bangladesh
Email: [1]shakhrul.buet.eee@gmail.com, [2]imamul@eee.buet.ac.bd

2022 12th International Conference on Electrical and Computer Engineering (ICECE) | 979-8-3503-9879-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICECE57408.2022.10089093

***Abstract*—Sleep disorders are a common problem that disrupts our regular sleeping patterns. To diagnose sleep disorders, Long-term monitoring of sleep could be useful. In this paper an automated scheme of sleep staging is presented based on Bioradiolocation signals using time and frequency domain feature extraction and Random Forest Classifier. This experiment is validated using data of 32 subjects without sleep-related breathing disorders. A Random Forest based algorithm is used for two, three, four and five-stage classification. We achieved the best performance so far (89.35% accuracy and 0.65 Cohens kappa) on 2-stage, 75.3% accuracy on 3-stage, 56.18% on 4-stage, and 54.2% accuracy on 5-stage classification with BRL Signals. These results show high potential in real-life sleep stage monitoring systems.**

***Index Terms*—Sleep stage, Bioradiolocation (BRL), Polysomnography, EEG, Random Forest.**

I. Introduction

Sleep disorder is a well-known health problem. People can be affected by insomnia, hypersomnia, narcolepsy, sleep apnea categorized as dyssomnias, sleepwalking and rem sleep behavior disorder which is classed as parasomnias as well as bruxism and circadian rhythm sleep disorders [1]. Consequently, sleep disorder diagnosis has been of noteworthy significance in recent times. One of the primary stages in sleep disorder diagnosis involves the recognition of the various stages of sleep.

Sleep stages can be classified differently following different rules. Rapid Eye Movement (REM) and non-REM (NREM) sleep are often considered to be the two main phases of normal sleep. They alternate cyclically throughout the night [2]. Broadly speaking, sleep can be classified into three phases considering the aforementioned and wakefulness (W). Following the R & K and AASM rules, NREM can be broken down into multiple stages. In general, sleep scoring is done clinically using polysomnographic records which include electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG), and electrocardiogram (ECG) signals. NREM is divided into three stages following the American Academy of Sleep Medicine Scoring Rules (AASMSR). S1 and S2 are known as Light Sleep (LS) and S3 is called Deep Sleep (DS) for which the muscle hypotonia increases. Events such as dreams, eye movements, muscle atony, etc. occur during paradoxical or REM sleep.

Most of the existing automated sleep scoring algorithms rely on either on one of the multichannel or multiple physiological signals or both [3]. However, subject mobility generally has an impact on sleep scoring using multichannel signals. In the case of scoring techniques using multiple physiological signals, the drawbacks are raised by setting up a large number of electrodes in the subjects which can result in interference of the signals. Apart from these, Contactless sleep monitoring might be an effective way in order to develop a sleep monitoring system without causing any discomfort, the primary reason is the fact that these methods are non-invasive. Bioradiolocation (BRL) and Ballistocardiography (BCG) are two of the most used techniques for the non-contact monitoring of Sleep stages.

The objective of this paper is to present sleep stage classification using BRL data which is scarce in the literature [4] [5] [6]. Time and frequency domain features are extracted from the BRL sleep data and employed by a random forest classifier to identify the sleep stages. The performance of the Proposed method is compared with those of the state-of-the-art BRL-based techniques on a publicly available BRL dataset [7], and quite promising accuracy, F1-score, and kappa coefficient are reported.

II. Materials and Methods

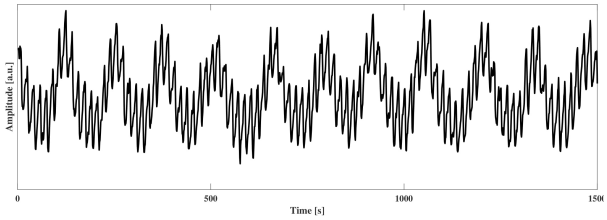
A. Dataset

The database contains records of non-contact sleep monitoring by a Bioradar from 32 healthy subjects without sleep-related breathing disorders [7]. The records are accompanied by results of sleep scoring, based on polysomnography according to the rules of the American Academy of Sleep Medicine. All patients have undergone a PSG study with a sampling frequency of 10 Hz which has 8 operating frequencies ranging between 3.6 GHz to 4.0 GHz.

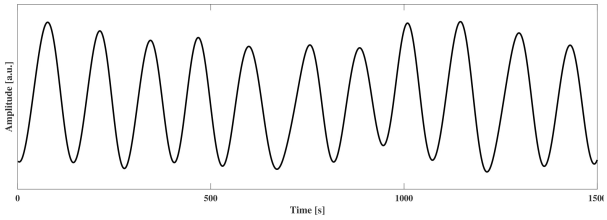
B. Signal Pre-processing

Each BRL signal consists of 16 signals which were recorded simultaneously. Instead of using a particular signal from those 16 signals, The best parts of all signals were combined into one signal for performance improvement. This was done by-

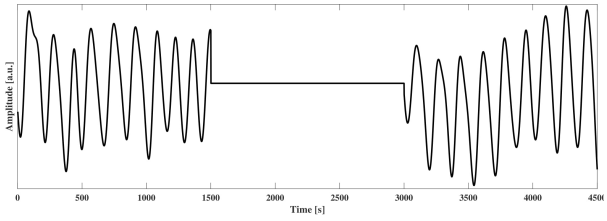
- We removed high-frequency noise from Each of the BRL signals (Fig.1a) by applying a Butterworth Low pass filter of 10th order with a cutoff frequency of 0.6 Hz.
- The median peak to through amplitude over the whole period of recording was subtracted from each signal to remove the baseline.
- From the 16 available signals, we selected a signal that has a maximum mean amplitude as Signal-W_1



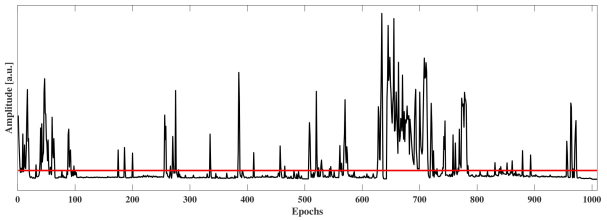
(a) BRL signal before preprocessing



(b) BRL signal after filtering (Signal-W_1)



(c) Signal after artifact removal (Signal-W_2). Artifact portion is replaced by zero



(d) calibrating body movement by comparing DTW measure with a threshold

Fig. 1: Signal preprocessing

(Fig.1b) and Dynamic Time Warping measurement was performed on it.

- As BRL data are sensitive to body movement, we need to calibrate them and remove body motion artifacts before extracting features. We used Dynamic time Warping features to make a decision if an epoch contains body movement or not. It is seen that epochs containing body movements have irregular waveshape and cause a larger value of DTW. We set a threshold, if the DTW measure of an epoch is larger than the threshold (Fig.1d), it is considered to have an artifact and replaced all data of that epoch with zeros. This signal is denoted as Signal-W_2 (Fig.1c).

C. Feature Extraction

We used both the signal-W_1 (Fig.1b) and Signal-W_2 (Fig.1c) to extract various Time and Frequency domain features. Some of the features were extracted from Signal-W_1 while other features were extracted from Signal-W_2. A total of 25 features were extracted for each epoch.

- Dynamic Time Warping (DTW) and Dynamic Frequency Warping (DFW) were calculated using signal-W_1 to measure the similarity between the target epoch and its nearby epochs. [8] DTW and DFW calculate the maximum time-domain similarity and maximum Frequency domain similarity respectively between the target epoch and its nearby epochs.
- Mean, Median, Standard deviation, and range (difference between the maximum and minimum amplitude) were calculated using signal-W_1 for each epoch.
- Frequency band power-based features were extracted for each epoch using signal W_1. They are- Power in the Low-frequency band between 0.05 Hz to 0.15 Hz, High-frequency band between 0.15 Hz to 0.5 Hz, and Very high-frequency band above 0.5 Hz.
- Sample entropy [9] measures of the peak and trough sequences which measure the regularity of the peak and through sequences.
- Depth based features [8] were extracted from signal-W_2. Peaks and Throughs were obtained by finding signal turning points. (Fig.3a). The median and IQR of peaks and through sequence, the median of peak-to-through difference, median area of inhalation, and exhalation period are calculated over 15 epochs.
- Inhalation and Exhalation portion of the signal-W_2 was identified as shown in Fig.3b. The Median and IQR area during the Inhalation and Exhalation period was calculated.
- Normalized epoch index, obtained by dividing by the total number of epochs for each subject.
- Singular value decomposition [10] and Katz fractal dimension [11].

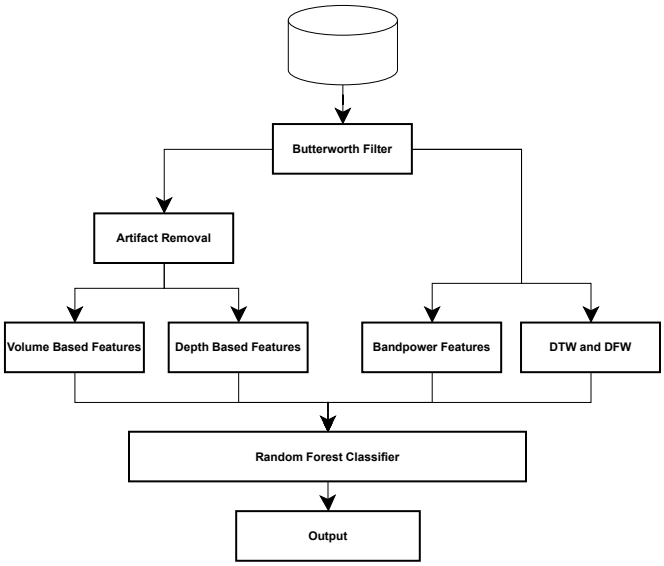


Fig. 2: Workflow of the proposed method.

To eliminate inter-subject variation, we performed Z-normalization by subtracting the mean from each feature and then dividing by the standard deviation. A workflow of our proposed method is shown in Fig.2.

D. Classifier

In our study, we use the Random Forest (RF) classifier [14], which is widely used in the classification task with high dimensional data. Also, we can visualize the relative

TABLE I: Comparison of 2 stage Classification (wake-sleep)

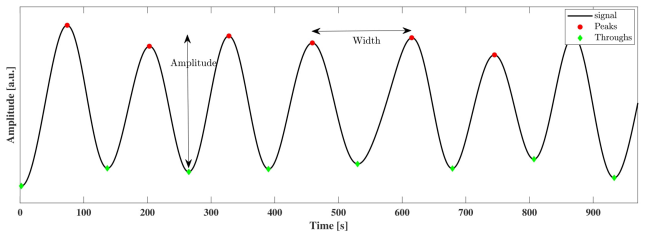
First author	Signals	N	Method	Accuracy	kappa	F1 Score
Pallin [5]	BRL	18	-	82.1	-	-
Hashizaki [12]	BRL	49	-	90.3±5.1	0.37±0.19	-
Zafforoni [6]	BRL	40	Train/Test	88.8	0.53	-
Kagawa [13]	BRL	13(2 in test)	-	79.4	-	-
Tataraidze [4]	BRL	32	LOSOCV	86.3±8.5	0.57±0.16	65.4 ± 14.0
This paper	**BRL**	**32**	**Train/Test**	**92.5**	**0.773**	**0.88**
			LOSOCV	**89.35±0.098**	**0.65±0.18**	**0.82 ± 0.1**

TABLE II: Comparison of 3 stage Classification (Wake-REM-NREM)

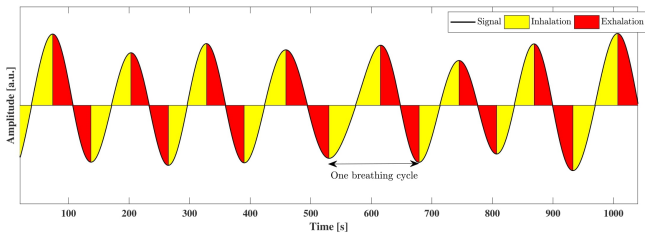
First author	Signals	N	Method	Accuracy	kappa	F1 Score
Zafforoni [6]	BRL	40	Train/Test	78.3	0.53	-
Kagawa [13]	BRL	13(2 in Test)	-	68.1	-	-
Tataraidze [4]	BRL	32	LOSOCV	75.9±9.6	0.55±0.14	-
This paper	**BRL**	**32**	**Train/Test**	**85.15**	**0.718**	**0.81**
			LOSOCV	**75.3±0.08**	**0.48±0.14**	**0.63±0.11**

TABLE III: Comparison of 4 stage Classification

First author	Signals	N	Method	Accuracy	kappa	F1 Score
Zafforoni [6]	BRL	40	Train/Test	63.3	0.47	-
Kagawa [13]	BRL	13(2 in test)	-	34.3	-	-
Tataraidze [4]	BRL	32	LOSOCV	63.5±0.08	0.49±0.12	-
This paper	**BRL**	**32**	**Train/Test**	**81.1**	**0.72**	**0.8**
			LOSOCV	**56.18±0.071**	**0.32±0.1**	**0.47±0.09**



(a) Peaks and Through of the signal



(b) Inhalation and Exhalation area

Fig. 3: Feature extraction from the signals

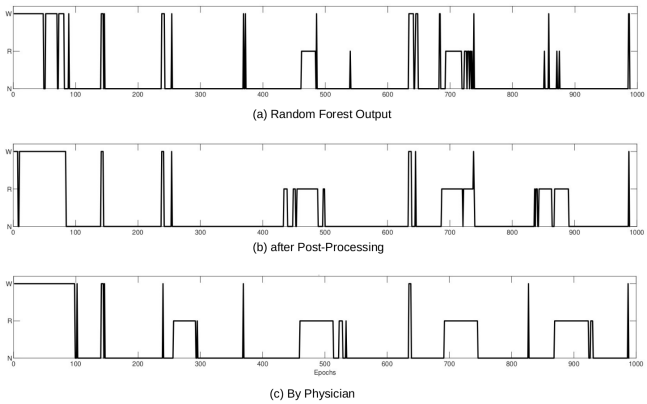


Fig. 4: classifier output for WRN classification

TABLE IV: Results of the Proposed Method for 5 stage classification

Method	Accuracy	kappa	F1 Score
Train/Test	**77.88**	**0.7**	**0.66**
LOSOCV	**54.2±0.07**	**0.31±0.1**	**0.4±0.06**

importance of the features using the feature importance function of Random Forest(Fig.5). The output of the classifier is post-processed to improve classification performance as :

- If the predicted class probability of an epoch is larger than a certain threshold, the modified output is the same as the classifier output.
- If the class probability is less than the threshold, a window of 15 epochs (7 on each side including the target epoch) is used for averaging the class probabilities, and then the class with the highest average probability is set for that epoch.

In Fig.4(a) output of Random Forest classifier for 3 stage classification is shown. We can see the abrupt changes in stages in several regions which is post-processed in Fig.4(b) as described. Fig.4(c) shows the corresponding sleep stages marked by a physician which indicates that the post-processing technique has improved the classification performance.

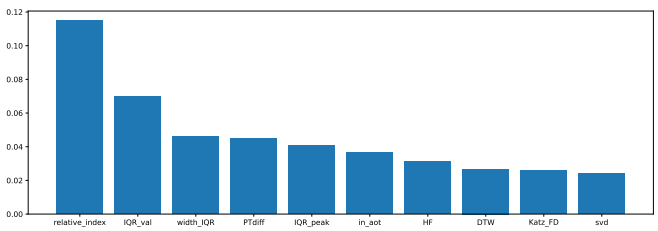


Fig. 5: Top 10 Features for WS classification

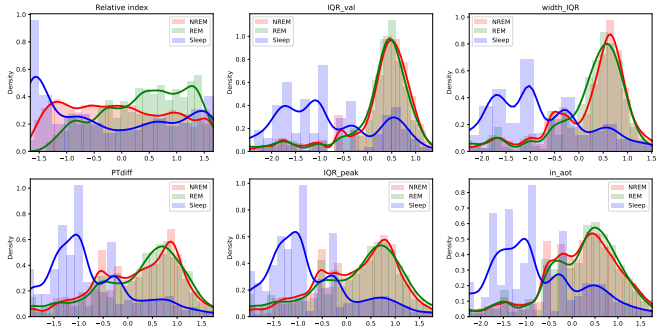


Fig. 6: Feature Distribution for different sleep stages

III. Results and Discussion

In this paper 2, 3, 4, and 5-stage classifications were performed using BRL signals. Table-I and Table-II show results and comparisons of 2 stage classification (Wake-Sleep) and Three stage classification (Wake-REM-NREM) respectively. The result of four and five-stage classification (Wake-REM-N1-N2-N3) is presented in Table-III and Table-IV.

Both Train-Test Split and A Leave-One-Subject-Out-Cross-Validation (LOSOCV) procedures were used for testing classification performance. In Train-Test Split a training set with the size of 80% was used for training and the rest of 20% is used for testing. For LOSOCV, Using 32 subjects, a training set was formed using features extracted from 31 subjects, and the rest remaining subject was used for testing. This procedure was repeated 32 times ensuring each subject was tested once. Classification accuracy, Cohen's kappa value, and F-1 Score were computed for performance analysis. Mean and standard deviation was computed for all the metrics and shown as (mean $\pm$ std) in tables. We got 89.35% accuracy and 0.65 cohen kappa for Wake-Sleep Classification (Table-I) which is better than other papers [4] [12] [6] on BRL dataset. Hashizaki et al. [12] achieved 90.3% accuracy on a different BRL dataset, but their kappa measure was too low. Pallin et al. [5] and tataraidze et al. [4] achieved 82.1% and 86.3% mean accuracy on Wake-Sleep classification. In the train test split method we get 92.5% accuracy and 0.773 kappa measure which is better than zafforoni et al. [6]. The relative importance of the features used is illustrated in Fig.5 for the top 10 features obtained using the feature importance function of Random Forest classifier. Features such as relative index, IQR of peaks, valleys and widths work well to differentiate between the sleep and wake stages. Fig.6 shows some of the important feature distributions for the Wake-REM-NREM stages. From the feature distribution plot, we can understand why these features work so well to differentiate Wake (blue line) from REM(green line) and NREM(red line) sleep stages. For the three-stage classification, we get 75.3% accuracy and 0.48 kappa score using LOSOCV method which is almost the same as tataraidze et al. [4]. However, for the train-test split method, both our accuracy and kappa measure is significantly better than zafforoni et al. [6](accuracy 85.15% vs 78.3% and kappa 0.718 vs 0.53). For four stage classification, in LOSOCV method our accuracy measure is 56.18% which is lower than tataraidze et al. [4]. However, zafforoni et al. [6] reported results using the train-test split method and achieved 63.3% accuracy and 0.47 kappa score. We get better results in both accuracy(81.1%) and kappa(0.72) scores than other papers that use the train-test split method. For Five-Stage Classification, we achieve 54.2% accuracy and 0.31 Cohen Kappa measure for LOSOCV method and 77.88% accuracy and 0.7 Cohen Kappa measure for the train-test split. This is the first paper to the best of our knowledge to classify 5 sleep stages and achieve decent results.

IV. Conclusions

In this paper, we tried to improve the accuracy and performance of sleep staging based on the BRL signal. The proposed method allows for classifying 2,3,4 and 5 different stages. We also succeeded to achieve more stable and reliable accuracies compared to the other state of the arts. BRL is a non-invasive method, which is why this study can go a long way to monitor sleep disorders without much hassle. Overall, it can be said without a doubt that this research aids in the development of sleep monitoring systems.

References

[1] L. Zhu and P. C. Zee, "Circadian rhythm sleep disorders," *Neurol. Clin.*, vol. 30, pp. 1167–1191, Nov. 2012.

[2] H. Sawai, M. Matsumoto, and E. Koyama, "The relationship between each length of REM - NREM sleep cycle and sleep stage," in *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*, pp. 171–172, IEEE, Mar. 2021.

[3] R. S. Arslan, H. Ulutaş, A. S. Köksal, M. Bakır, and B. Çiftçi, "Automated sleep scoring system using multi-channel data and machine learning," *Comput. Biol. Med.*, vol. 146, p. 105653., July 2022.

[4] A. Tataraidze, L. Anishchenko, L. Korostovtseva, B. Kooij, M. Bochkarev, and Y. Sviryaev, "Sleep stage classification based on bioradiolocation signals," vol. 2015, 08 2015.

[5] M. Pallin, E. O'Hare, A. Zaffaroni, P. Boyle, C. Fagan, B. Kent, C. Heneghan, P. de Chazal, and W. T. McNicholas, "Comparison of a novel non-contact biomotion sensor with wrist actigraphy in estimating sleep quality in patients with obstructive sleep apnoea," *J. Sleep Res.*, vol. 23, pp. 475–484, Aug. 2014.

[6] A. Zaffaroni, L. Gahan, C. Heneghan, L. Collins, E. O'Hare, I. Fietze, C. Garcia, and T. Penzel, "Automated sleep staging classification using a non-contact biomotion sensor," *ResearchGate*, Sept. 2014.

[7] "Sleep Bioradiolocation Database v1.0.0," Sept. 2022. [Online; accessed 19. Sep. 2022].

[8] X. Long, J. Yang, T. Weysen, R. Haakma, J. Rolink, P. Fonseca, and R. Aarts, "Measuring dissimilarity between respiratory effort signals based on uniform scaling for sleep staging," *Physiological Measurement*, vol. 35, pp. 2529–2542, 11 2014.

[9] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, June 2000.

[10] V. Klema and A. Laub, "The singular value decomposition: Its computation and some applications," *IEEE Transactions on Automatic Control*, vol. 25, no. 2, pp. 164–176, 1980.

[11] I. Wijayanto, R. Hartanto, and H. A. Nugroho, "Higuchi and Katz Fractal Dimension for Detecting Interictal and Ictal State in Electroencephalogram Signal," in *2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE)*, pp. 1–6, IEEE, Oct. 2019.

[12] M. Hashizaki, H. Nakajima, M. Tsutsumi, T. Shiga, S. Chiba, T. Yagi, Y. Ojima, A. Ikegami, M. Kawabata, and K. Kume, "Accuracy validation of sleep measurements by a contactless biomotion sensor on subjects with suspected sleep apnea," *Sleep Biol. Rhythms*, vol. 12, pp. 106–115, Apr. 2014.

[13] M. Kagawa, N. Sasaki, K. Suzumura, and T. Matsui, "Sleep stage classification by body movement index and respiratory interval indices using multiple radar sensors," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 7606–7609, 2015.

[14] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, pp. 5–32, Oct. 2001.