Assessment of submentalis muscle activity for sleep-wake classification of healthy individuals and patients with sleep disorders

Ritika Jain [1] and Ramakrishnan Angarai Ganesan [1]

***Abstract*— This work proposes a method utilizing only the submentalis EMG channel for the classification of sleep and wake states among the healthy individuals and patients with various sleep disorders such as sleep apnea hypopnea syndrome, dyssomnia, etc. We extracted autoregressive model parameters, discrete wavelet transform coefficients, Hjorth's complexity and mobility, relative bandpowers, Poincaré plot descriptors and statistical features from the EMG signal. We also used the energy of each epoch as a feature to distinguish between the sleep and wake states. Mutual information based feature selection approach was considered to obtain the top 25 features which provided maximum accuracy. For classification, we employed an ensemble of decision trees with random undersampling and boosting technique to deal with the class-imbalance problem in the sleep data. We achieved an overall accuracy of about 85% for the healthy population and about 70% on an average across different pathological groups. This work shows the potential of EMG chin activity for sleep analysis.**

***Clinical relevance*— Automatic and reliable sleep-wake classification can reduce the burden of sleep experts in analyzing overnight sleep data ($\sim$ 8 hours) and also assist them to diagnose various neurological disorders at an early stage. Utilizing EMG channel provides an easier and convenient long-term recording of data without causing much disturbance in sleep; unlike EEG, which is inconvenient and hampers the natural sleep.**

I. INTRODUCTION

There is no doubt that sleep plays a very important role in our physical and mental well-being; and any sort of disturbance in sleeping pattern can lead to numerous diseases. The connections between sleep disruption and both disease and mortality have become more firmly established [1]. Therefore, an accurate diagnosis of sleep disorders is extremely crucial and can help in the early detection of neurological disorders such as dementia, schizophrenia, depression or Parkinson's disease. The most common sleep disorder by-far is insomnia, which gets often complicated by the presence of another sleep disorder, such as sleep apnea or restless leg movement syndrome [1]. Currently, the gold standard for sleep evaluation is overnight polysomnography (PSG) which records multiple physiological signals such as electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG), electrocardiogram (ECG) and respiration. The sleep experts visually inspect the whole night PSG data and annotate it based on the scoring guidelines provided by Rechtschaffen and Kales (R&K) or American Academy of Sleep Medicine (AASM) [2]. However, this manual scoring is tedious, subjective, expensive as well as time-consuming. An automated sleep scoring can help in reducing the time and cost involved. Numerous studies have been conducted to provide a reliable and automatic sleep scoring method [3–6]. These studies have utilized various features such as spectral band powers, complexity, multi-scale entropy, and discrete wavelet transform coefficients, mainly derived from the EEG channel. Recently, many studies are employing neural networks and deep learning techniques for the classification of sleep stages. However, majority of these studies have considered only EEG channels or combination of EEG and EOG channels. Only one study [7] has explored the usability of EMG (anterior tibialis) signal for sleep analysis. EEG is considered to be the most informative signal for decoding the brain activity into different stages of sleep. But for the estimation of sleep and wake states, we can utilize EOG or EMG signal rather than EEG; which is convenient and easy to set-up and does not require the assistance of any specialist.

In general for healthy individuals, the amplitude of chin EMG signal reduces in sleep, with minimum in REM (rapid eye movement) sleep stage; and increases in the wake state. This can also be seen in Fig. 1, which shows the amplitude of EMG signal in sleep and wake states for a healthy subject. However, it is reported that there is an increase of chin EMG activity for RBD (rapid eye movement behavior disorder) patients in REM stage of sleep unlike normal individuals. RBD can be an early warning for the emergence of Parkinsonism [8]. Thus, the chin EMG can provide valuable information for detecting early forms of neurodegenerative conditions such as Parkinson's.

In this work, we have utilized only the submentalis muscle activity (chin EMG) for discriminating sleep from wake states among healthy subjects and patients with various sleep disorders.

II. MATERIALS AND METHODS

A. Experimental Data

We utilized two publicly available datasets, namely ISRUC [9] and the DREAMS Patients database [10]. The ISRUC dataset comprises three subgroups out of which we used subgroup-3, which includes PSG data of 10 healthy subjects. The recordings have been scored on the basis of 30s epochs by two different experts as per the AASM standard [2]. We have considered only those epochs which are scored same by both the experts as our ground truth. The DREAMS Patient dataset includes data from 27 patients with different pathologies; mixed apnea: 2, periodic limb movement syndrome (PLMS): 9, sleep apnea hypopnea syndrome (SAHS): 9 and

*This work was not supported by any organization

[1]Ritika Jain (email: ritikajain@iisc.ac.in) and Ramakrishnan A G (email: agr@iisc.ac.in) are with the Department of Electrical Engineering, Indian Institute of Science Bangalore 560012, India

Dyssomnia: 6, and pathology of one patient is not specified in the dataset. This dataset, however, has been scored by only one expert using AASM standard, which is used as the ground truth. Out of 27 patients, two (patients 12 and 14) had missing or bad EMG data and hence removed from our analysis. Since this work focuses only on sleep and wake states, we merged all the multiple stages of sleep (N1, N2, N3 and REM) as sleep state. Table I shows the epoch distributions for sleep and wake states in the two datasets.

TABLE I

SLEEP AND WAKE EPOCH DISTRIBUTION OF 10 SUBJECTS FROM ISRUC DATASET AND AND 25 PATIENTS FROM DREAMS PATIENTS DATASET

Dataset	Wake Epochs	Sleep Epochs	Total Epochs
DREAMS Patients	6510	19436	25946
ISRUC	1670	6185	7855

B. Methodology

In this study, we used only submentalis EMG channel (chin EMG) data to classify the epochs into sleep or wake state. The workflow of the proposed method is as follows: The EMG data is first filtered using a 4^{th} order IIR butterworth bandpass filter with passband frequencies ranging from 10 to 99.5 Hz. Also, notch filter with cutoff frequency of 50 Hz is applied to remove the line noise. Then, we segmented the signal into 30s epochs and removed epochs marked as no score. Different features are extracted from this segmented EMG signal, followed by the mutual information based feature selection. The selected features are finally fed to the RUSBoost (random undersampling with boosting technique) classifier for performing the classification of epochs into sleep or wake state. The subsections below present a detailed description of the proposed method.

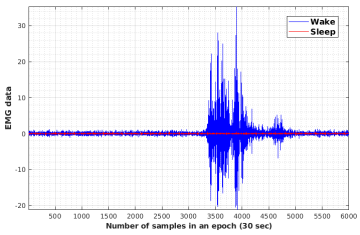

Fig. 1. Submentalis EMG signal in a sleep and wake epoch (30 sec) for a healthy subject.

1) Feature extraction: After filtering and segmentation of the EMG signal into 30 sec epochs, multiple features are derived to capture temporal, spectral as well as non-linear characteristics of the signal. A list of all the features utilized in this work are presented in Table II. We used autoregressive (AR) modelling to characterize the degree of predictability of signal from its past samples. This parametric model essentially represents the signal as a linear combination of the past observations weighted by coefficients $a_1, a_2, ...a_p$ where p is the model order. In this work, we used 8^{th} order AR model based on our previous study [3]. We also employed Hjorth's parameters namely Hjorth's mobility (M) and complexity (C) to assess the statistical properties of the EMG signal. These are calculated as shown in (1) and (2).

$$M = \sqrt{var(y')/var(y)} \tag{1}$$

$$C = M(y')/M(y) \tag{2}$$

where y and y' are the signal and first difference respectively.

We also used $4th$ level decomposition of the EMG signal using Daubechies-2 orthogonal wavelet, which provides four detail and one approximation coefficients. For each of these coefficients, we calculated minimum, maximum, average and standard deviations, giving a total of 20 features. The choice of wavelet is based on our previous study [5].To extract the spectral information of the EMG signal, we calculated bandpowers in two frequency bands; i.e. 10-40 Hz and 40-80 Hz, with respect to the total power in the 10-99.5 Hz range. Also, we calculated features such as RMS value (RMSval), sum of absolute differences (SAD), and log root of squared differences (LRSD), to extract the statistical properties of EMG signal. These features are calculated using (3) to (5).

$$LRSD = \log_{10}\left(\sqrt{\textstyle\sum_{n=1}^{N}(x(n) - x(n-1))^2}\right) \tag{3}$$

$$SAD = \textstyle\sum_{n=1}^{N} \mid x(n) - x(n-1) \mid \tag{4}$$

$$RMSval = \sqrt{\tfrac{1}{N}\textstyle\sum_{n=1}^{N} x(n)^2} \tag{5}$$

where $x(n)$ is the EMG signal and N is number of samples in an epoch.

It is evident from Fig. 1 that the energy of the EMG signal is higher in the wake state as compared to sleep. Therefore, we computed the energy of each epoch and used the log transformed epoch-energy ($EpEn$) as a feature to distinguish between the sleep and wake states. It is calculated for each epoch (of 30 sec) as shown in (6)

$$EpEn_\tau = \log\left(\textstyle\sum_{n=30f_s(\tau-1)+1}^{30f_s\tau} x(n)^2\right) \tag{6}$$

where $x(n)$ is the n^{th} EMG sample, f_s is the sampling frequency (here 200 Hz), and the squares of the samples are summed for the entire epoch length of 30 sec and log transformed to get the feature for the corresponding epoch τ. Here, $EpEn_\tau$ is the value of epoch energy of the τ-th epoch. Finally, we also incorporated Poincaré plot based descriptors to consider the long and short-term variabilities in EMG signal. A Poincaré plot (PP) is generated by plotting a signal $x(k)$ against its lagged version, $x(k+\tau)$. We extracted various descriptors to quantify PP) in terms of its width, length, area, asymmetry, etc. The most commonly used standard descriptors of PP are SD1 and SD2, which measure the standard deviations of distribution of points around the line perpendicular to the identity (i.e. width of PP) and around the identity line (referred to as length of PP). Also, the ratio of SD1 and SD2 (i.e. SD1/SD2) and the area of the ellipse formed by the scatter plots (with semi-major axis

TABLE II

DETAILS OF THE FEATURES USED IN THIS STUDY

Feature	Dimension
Epoch-energy (EpEn)	1
Autoregressive model coefficients	8 x 1 = 8
Wavelet approx. and detail coefficients	4 x 5 = 20
Hjorth's mobility (M)	1
Hjorth's complexity (C)	1
Relative band power in 10-40 Hz	1
Relative band power in 40-80 Hz	1
Sum of absolute differences (SAD)	1
Log root of squared differences (LRSD)	1
Poincaré plot descriptors	1 x 7 = 7
Root mean square value (RMSval)	1

as SD2 and semi-minor axis as SD1) are used as features. In addition, we used energy of the slopes and intercepts of linear fit across the delayed Poincaré plots (delay τ varied from 1 to 100). To quantify the skewness of the distribution of the PP, we utilized Ehler's index proposed in [11]. We used the novel feature of grid distribution entropy, proposed in the work by Yan et. al. [12]. This feature is the Shanon entropy of the points in each grid after gridding the original Poincaré plot.

2) Feature selection: Table II shows that the dimension of feature-set is 43. In order to remove redundant and irrelevant features from the set, we used mutual information (MI) based feature selection. This method calculates mutual information between the features and their labels, with a goal to maximise information between the feature and label. The features are ranked according to the MI score and top k features are used for classification. After experimenting with various values of k, we selected k as 25, thereby reducing the dimension from 43 to 25.

3) Classification: For classifying sleep from wake states, we used an ensemble of decision trees. Further in order to deal with the problem of class imbalance, we adopted random undersampling with boosting technique (RUSBoost) for the ensemble classifier. We used RUSBoost classifier because it performed better than the SVM classifier in our earlier work [5]. The different hyperparameters of RUSBoost are the number of weak learners (trees), learning rate, and maximum number of splits for each tree. The values are chosen to be 2000 weak learners, with learning rate of 0.1 and maximum splits equal to the number of training samples. These values are fine-tuned by using a grid search across a range of values for each of these hyperparameters.

III. EXPERIMENTAL RESULTS

In this work, we experimented with one dataset of healthy subjects and another of patients with different sleep disorders for the identification of sleep and wake states using only chin EMG signal. In order to evaluate the performance of the proposed method, we used leave-one-subject-out (LOSO) validation approach. In this approach, the model is trained on all but one (left out) while tested on the left out subject. The final performance of model is reported as the average accuracy across all the test subjects. Figure 2 shows the classification accuracies for each of the 25 patients from the DREAMS Patients dataset as well as 10 healthy subjects from ISRUC dataset using LOSO method. For the healthy population, an average accuracy of $84.5\% \pm 4.4\%$ is obtained. In case of patients, the classification accuracy of $\geq 70\%$ is achieved for 18 out of 25 patients, with maximum of 91% for a patient with PLMS and minimum of 26% for a dyssomnia patient.

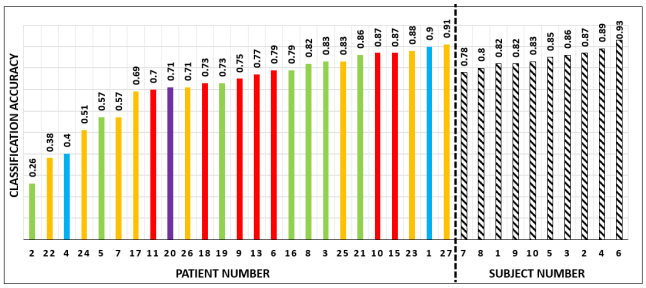

Fig. 2. Classification accuracies for each of the 25 patients with different pathologies (blue: mixed apnea; red: SAHS, yellow: PLMS; green: dyssomnia; purple: not specified) from DREAMS Patients dataset and 10 healthy subjects from ISRUC dataset using LOSO method with only chin EMG signal. *(Coloured bars represent patients with different pathologies and shaded black bars represent healthy subjects).*

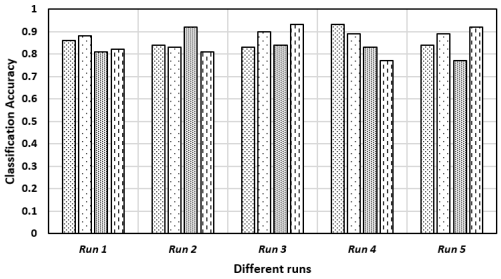

Fig. 3. Classification accuracies of 4 unseen test subjects from ISRUC dataset using hold-out approach in five different runs with only chin EMG signal.

The classification accuracy of the model is consistently high across the SAHS patient group. Also, it performs well on dyssomnia patients except for patient number 2. For patient No. 2 (dyssomnia) and 22 (PLMS), it resulted in poor accuracy ($< 40\%$). The mean accuracies for patient groups with mixed apnea, dyssomnia, SAHS, and PLMS are $65 \pm 35.4\%, 69 \pm 23.4\%, 78 \pm 6.6\%$, and $69 \pm 18.8\%$, respectively. There is a lot of variability across the individual patients in each group. The variability is maximum for mixed apnea group, followed by dyssomnia, and then PLMS and least for SAHS. However, the large variability in mixed apnea group is mainly due to the limited size (only two patients with mixed apnea); and for dyssomnia, because of an outlier patient (patient 2), on whom the model performs the worst. Due to the limited number of patients and high variability in the classification performance in different pathological groups, it is difficult to conclude if any specific pathology is difficult to classify. However, it is evident that the model is able to provide good results for patients, especially SAHS (with mean, minimum and maximum classification accuracy of about 78%, 70%, and 86% respectively.) Apart from

the LOSO technique, we also evaluated the performance of the model with hold-out validation approach. For ISRUC dataset, 4 subjects (out of 10) are randomly selected for testing; while the model is trained on the rest 6 subjects; and this is repeated for 5 different runs. Thus, in each run the performance of model is evaluated on four unseen subjects. Figure 3 shows the classification accuracies of unseen test subjects in different runs. The mean accuracies in the five runs are $84.4\%, 84.9\%, 87.4\%, 85.4\%$, and 85.7%. Since we did not find any existing study using submentalis EMG for sleep-wake classification, we compared our performance with the literature irrespective of modality. The comparison results are presented in Table III. It is evident that the maximum accuracy is achieved by using EEG; however among other modalities, EMG provides quite good results for both normal controls and patients. Only one study [7] has used EMG for sleep-wake classification, but instead of submentalis (chin) they used anterior-tibialis (leg muscle) for recording EMG activity. They obtained an accuracy of around 80.5% for healthy subjects $(N = 7)$ and around 77.6% for patients $(N = 5$; all apnea patients). Comparatively, our method achieved higher accuracy (85%) with more number of subjects $(N = 10)$ for controls; while the performance on patient population is lesser in our work. However, they have considered only 5 patients (all apnea), while our study included 25 patients (with diverse sleep disorders: PLMS, apnea, dyssomnia). Further, submentalis EMG recording maybe a better option than anterior tibialis for patients with PLMS because of the presence of leg movements in REM sleep.

It is difficult to directly compare the results since many studies have used their private datasets and not reported performance on the standard publicly available datasets. Also, most of the studies have not included patient population. Furthermore, the validation approaches used in these studies also vary; some have used LOSO approach (which gives more robust measure of model's generalizability), while others used epochwise train-test splitting or k-fold crossvalidation (which is biased estimation of model's performance since train and test set share data from same subject).

TABLE III

COMPARISON OF THE PERFORMANCE OF PROPOSED METHOD WITH EXISTING LITERATURE

Authors	Dataset	Validation	Modality	Healthy/Patients	Accuracy
Zoubek et. al. [13]	Private	90-10 epochwise split	EEG+EMG+EOG	47/Nil	80%
Brignol et. al. [14]	Private	90-10 epochwise split	EEG	7/Nil	94%
Adnane et. al. [15]	MIT-BIH	80-20 split epochwise	ECG	16/Nil	79.3%
Hwang et. al. [7]	private	LOSO	EMG (Anterior Tibialis)	7/5	80.5%/77.6%
Aktaruzzaman et. al. [16]	private	LOSO	HRV	Nil/20	71.3%
Chakar et. al. [17]	private	80-20 split epochwise	wrist-actigraphy	38/Nil	41.1%
Chakar et. al. [17]	private	80-20 split epochwise	jaw movements	38/Nil	58.9%
Ganesan et. al. [4]	Sleep-EDF	10-fold CV epochwise	EEG	8/Nil	93.9%
Chen et. al. [18]	Private	LOSO	HRV+Acceleration	11/Nil	86.6%
Ganesan et. al. [3]	Sleep-EDF	10-fold CV epochwise	EEG+EOG	8/Nil	98.1%
Proposed work	**ISRUC & DRMS-PAT**	**LOSO**	**EMG (Submentalis)**	**10/25**	**85%/70%**

IV. CONCLUSION

In this work, we analysed the suitability of submentalis muscle activity for the detection of sleep and wake states among healthy people and patients. To the best of our knowledge, none of the studies in the literature have explored the capability of submentalis muscle for sleep-wake classification. The proposed method achieves an average accuracy of about 85% for the healthy controls by utilizing only the chin EMG signal. Further, it provides satisfactory performance across the different patient groups, especially for the SAHS patients. The results indicate that the chin EMG activity has a good potential to be used for the analysis of sleep-wake patterns. In our future work, we plan to include more patients' data to validate the usability of the proposed method for clinical population.

REFERENCES

[1] Susan L Worley. "The extraordinary importance of sleep: the detrimental effects of inadequate sleep on health and public safety drive an explosion of sleep research". In: *Pharmacy and Therapeutics* 43.12 (2018), p. 758.

[2] Conrad Iber et al. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. Vol. 1. American academy of sleep medicine Westchester, IL, 2007.

[3] Ramakrishnan Angarai Ganesan and Ritika Jain. "Binary State Prediction of Sleep or Wakefulness Using EEG and EOG Features". In: *2020 IEEE 17th India Council International Conference (INDICON)*. IEEE. 2020, pp. 1–7.

[4] Ritika Jain and Ramakrishnan Angarai Ganesan. "Sleep-awake Classification using EEG Band-power-ratios and Complexity Measures". In: *2020 IEEE 17th India Council International Conference (INDICON)*. IEEE. 2020, pp. 1–6.

[5] Ritika Jain and Ramakrishnan A G. "Reliable sleep staging of unseen subjects with fusion of multiple EEG features and RUSBoost". In: *Biomedical Signal Processing and Control* 70 (2021), p. 103061.

[6] Ritika Jain and Ramakrishnan Angarai Ganesan. "An Efficient Sleep Scoring Method using Visibility Graph and Temporal Features of Single-Channel EEG". In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE. 2021, pp. 6306–6309.

[7] SuHwan Hwang et al. "Sleep/wake estimation using only anterior tibialis electromyography data". In: *Biomedical engineering online* 11.1 (2012), pp. 1–15.

[8] Bradley F Boeve, Michael H Silber, and Tanis J Ferman. "REM sleep behavior disorder in Parkinson's disease and dementia with Lewy bodies". In: *Journal of geriatric psychiatry and neurology* 17.3 (2004), pp. 146–157.

[9] Sirvan Khalighi et al. "ISRUC-Sleep: A comprehensive public dataset for sleep researchers". In: *Computer methods and programs in biomedicine* 124 (2016), pp. 180–192.

[10] S Devuyst. *The DREAMS Databases and Assessment Algorithm [Data set]*. 2005.

[11] Karolina Kramarić et al. "Heart rate asymmetry as a new marker for neonatal stress". In: *Biomedical signal processing and control* 47 (2019), pp. 219–223.

[12] Chang Yan et al. "Novel gridded descriptors of poincaré plot for analyzing heartbeat interval time-series". In: *Computers in biology and medicine* 109 (2019), pp. 280–289.

[13] Lukáš Zoubek et al. "Feature selection for sleep/wake stages classification using data driven methods". In: *Biomedical Signal Processing and Control* 2.3 (2007), pp. 171–179.

[14] Arnaud Brignol, Tarik Al-Ani, and Xavier Drouot. "EEG-based automatic sleep-wake classification in humans using short and standard epoch lengths". In: *12th International Conference on Bioinformatics & Bioengineering (BIBE)*. IEEE. 2012, pp. 276–281.

[15] Mourad Adnane, Zhongwei Jiang, and Zhonghong Yan. "Sleep–wake stages classification and sleep efficiency estimation using single-lead electrocardiogram". In: *Expert Systems with Applications* 39.1 (2012), pp. 1401–1413.

[16] Md Aktaruzzaman et al. "The addition of entropy-based regularity parameters improves sleep stage classification based on heart rate variability". In: *Medical & biological engineering & computing* 53.5 (2015), pp. 415–425.

[17] Bassam Chakar et al. "Validation of midsagittal jaw movements to measure sleep in healthy adults by comparison with actigraphy and polysomnography". In: *Sleep Science* 10.3 (2017), p. 122.

[18] Zhenghua Chen et al. "A novel ensemble deep learning approach for sleep-wake detection using heart rate variability and acceleration". In: *IEEE Transactions on Emerging Topics in Computational Intelligence* 5.5 (2020), pp. 803–812.