Single-Channel EEG Detection of REM Sleep Behaviour Disorder: The Influence of REM and Slow Wave Sleep

Single-channel EEG Detection of REM Sleep Behaviour Disorder: the Influence of REM and Slow Wave Sleep

Irene Rechichi[1 [0000−0002−9420−4922]], Federica Amato[1 [0000−0003−4409−8794]], Alessandro Cicolin[2 [0000−0001−6196−3925]], and Gabriella Olmo[1 [0000−0002−3670−9412]]

[1] Dept. Control and Computer Engineering, Politecnico di Torino, Italy
firstname.lastname@polito.it
[2] Sleep Disorders Centre, Dept. Neuroscience, University of Turin, Italy
firstname.lastname@unito.it

Abstract. Sleep Disorders have received much attention in recent years, as they are related to the risk and pathogenesis of neurodegenerative diseases. Notably, REM Sleep Behaviour Disorder (RBD) is considered an early symptom of alpha-synucleinopathies, with a conversion rate to Parkinson's Disease (PD) up to 90%. Recent studies also highlighted the role of disturbed Non-REM Slow Wave Sleep (SWS) in neurodegenerative diseases pathogenesis and its link to cognitive outcomes in PD and Dementia. However, the diagnosis of sleep disorders is a long and cumbersome process. This study proposes a method for automatically detecting RBD from single-channel EEG data, by analysing segments recorded during both REM sleep and SWS. This paper inspects the underlying microstructure of the two stages and includes a comparison of their performance to discuss their potential as markers for RBD. Machine Learning models were employed in the binary classification between healthy and RBD subjects, with an 86% averaged accuracy on a 5-fold cross-validation when considering both stages. Besides, SWS features alone proved promising in detecting RBD, scoring a 91% sensitivity (RBD class). These findings suggest the applicability of an EEG-based, low-cost, automatic detection of RBD, leading to potential use in the early diagnosis of neurodegeneration, thus allowing for disease-modifying interventions.

Keywords: EEG · Machine Learning · Sleep Disorders · RBD · REM Sleep Behaviour Disorder · Automatic Classification

1 Introduction

Sleep is a transient state of altered consciousness opposed to wake, and provides a restorative function to the human organism. Given its complex nature, it is a reservoir of significant clinical data. Human sleep cyclically alternates between two different states – which also entail two different kinds of brain activity:

rapid-eye movement sleep (REM) and non-REM sleep (NREM). The latter, according to the American Academy of Sleep Medicine (AASM) guidelines, is further divided into three stages, from light to deep: N1, N2 and Slow Wave Sleep (SWS) [1]. REM Sleep, on the contrary, exhibits mixed features and is often referred to as *paradoxical sleep*; it supports memory consolidation and is characterised by skeletal muscle atonia, as well as electrical brain activity resembling wake [21].

It has been demonstrated that the quality of both REM and SWS contributes to neurological outcomes in older adults [20], and their disruption not only leads to sleep disorders, but plays a role in the pathogenesis of neurodegenerative diseases.

Recent studies highlighted the role of sleep in clearing toxic metabolites – such as Amyloid-β – from the brain [14]. This occurs primarily during SWS through the glymphatic system, which is instead inhibited throughout wakefulness [28]. Consequently, poor sleep is associated to a variety of Sleep Disorders, which, in turn, entail a lower Quality of Life (QoL) [15], as well as an increased risk for neurodegeneration.

Sleep disorders may manifest in different ways. Among these, REM Sleep Behaviour Disorder (RBD), a REM parasomnia characterised by lack of physiological muscle atonia during REM Sleep [18], features a prevalence of approximately 2% in the elderly population worldwide [9]. RBD in its isolated phenotype is considered an early prodrome of α-synucleinopathy – e.g., Parkinson's Disease (PD), Dementia with Lewy Bodies (DLB) and multiple system atrophy [10, 23]; indeed, the phenoconversion rate to PD is around 90% [7].

Polysomnography (PSG) is the gold standard to diagnose sleep disorders. It consists in recording various biosignals during sleep; including the electroencephalogram (EEG), pivotal to assess sleep stages, the electromyogram (EMG), and the electrooculogram (EOG). However, a PSG exam is costly and cumbersome; besides, the recordings are manually scored by a sleep technologist, thus significantly hindering the diagnostic process. Different studies undertook the automatic classification of sleep disorders through Machine Learning (ML) algorithms, in an attempt to accelerate and support the diagnosis [12]. They either employed PSG features [27] – i.e., clinical variables descriptive of the quality of sleep – or PSG biosignals [26]. Automatic RBD detection has also been addressed, primarily through EMG features [4, 5], but also relying on EEG data only [3]. This study aims at performing an automatic classification between healthy and RBD subjects, based on single-channel EEG. Features extracted from the REM and SWS segments are input to ML models, in order to explore their applicability and determine which of the two stages is more relevant to the application at hand. To this end, in this work the feature sets derived from the two stages are analysed both separately and collectively.

2 Materials and Methods

2.1 Subjects and Data

For the purpose of this work, both a public and a private datasets were employed. The public dataset is the CAP Sleep Database [25], available on PhysioNet [8]. It includes PSG recordings of 22 subjects affected by RBD (19 males, aged 70 $\pm$ 6 years) and 16 healthy subjects (9 males, aged 32 $\pm$ 5 years). The recordings were studied along with the provided manual annotations of sleep stages, scored according to the AASM standards [1]. The additional data for this study were taken from a private dataset – the Turin Sleep Disorders Database (TuSDi Database) – encompassing PSG recordings of 10 healthy subjects (6 males, aged 37 $\pm$ 16 years) and 10 RBD subjects (8 males, aged 62 $\pm$ 6 years). Data were collected at the Center for Sleep Disorders at Molinette Hospital (Turin, Italy); the procedure has been conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of A.O.U. Città della Salute e della Scienza di Torino (approval No. 00384/2020). Informed consent for observational study was obtained from the participants. Inclusion criteria were suspected or diagnosed RBD, with polysomnographic evidence of REM Sleep Without Atonia; exclusion criteria included dementia or other psychiatric conditions that could affect the correct execution of the PSG exam. All participants received detailed information on the study purpose and execution, and informed consent was obtained.

PSG recordings in the TuSDi Database were manually scored and annotated by a sleep technologist, according to the AASM standards. To sum up, this work considered for the analysis 26 healthy and 32 RBD subjects.

Being a single-channel EEG classification, only the recordings from the central EEG channel have been employed. Therefore, the C3-A2 channel (or C4-A1, if the former was not available) was selected for the subsequent analysis. Given the aim of the study, only sleep segments related to the REM stage and SWS were taken into account for the feature extraction step, detailed in Section 2.3.

2.2 Data pre-processing

First of all, the sleep segments' duration was inspected; all subjects presented with at least 5 minutes of REM episodes and SWS and were therefore included in the study. The majority of EEG signals were collected at a sampling frequency of 512 Hz; signals presenting with a different sampling frequency were resampled at 512 Hz. This work relies on raw EEG data, meaning that the analysis did not require additional processing such as artefact removal or spatial filtering. However, in order to decrease high-frequency noise, all recordings have been pre-processed through a low-pass, zero-phase Chebyschev Type 1 Filter (cut-off frequency: 40 Hz). To prevent the acquisition conditions from affecting the analysis, the signals amplitude was converted to μV.

2.3 Feature Extraction

In this work, a total of 367 features were extracted. The features comprised both polysomnographic features – i.e., clinical parameters that describe the overall sleep structure – and features extracted from the EEG during REM and SWS segments. The feature extraction procedure is detailed in the following paragraphs.

Polysomnographic Features As previously introduced, a first set of polysomnographic variables was computationally extracted from the hypnogram – i.e., the sleep scoring array. This set consists of variables that are commonly employed by sleep experts to assess overall sleep structure and quality – e.g., Sleep Onset Latency (SOL), Sleep Efficiency (SE); other parameters that describe the architecture of sleep stages – e.g., the proportion of each stage per sleep time, Minutes of REM Sleep (MREM), Minutes of SWS (MSWS) – and, lastly, variables that describe the variability within sleep stages and the sleep fragmentation – e.g., Arousal Index (ARI), Sleep Transition Index (STI). The complete set of employed polysomnographic features is presented in Table 1, and a detailed definition is provided for each parameter.

Electroencephalographic Features Electroencephalographic features were extracted from the available central EEG channel (C3-A2 or C4-A1) from both REM and SWS segments, respectively. The features were extracted in three domains: Time, Frequency and Non-Linear. Given the fact that the AASM criteria score sleep by inspecting the EEG signal in 30 s epochs, the data in this study were analysed accordingly. Therefore, each sleep segment (i.e., REM or SWS) was divided into 30 s epochs, and time-domain features were extracted from each epoch. Spectral (i.e., frequency-domain) and Non-Linear features were extracted on 2 s sub-epochs and then averaged across the corresponding 30 s macro-epoch, thus ensuring stationarity for the EEG signal as well as a reasonable spectral resolution. At the end of the feature extraction process, for each variable, the values corresponding to the 30 s epochs were assembled into one feature array, and three statistics were computed – i.e., mean value, standard deviation (STD) and 75^{th} percentile. A list of the employed EEG features is provided in Table 2, grouped according to their domain – i.e., Time, Frequency or Non-Linear.

Temporal features mainly account for the information provided by the amplitude of the signal or its waveform. For the purpose of this work, the latter is evaluated through the Form (FF), Crest (CF) and Impact factors (IF), defined in Equations (1), (2), (3) – where x is the considered signal.

$$FF = \frac{x_{RMS}}{|x|_{mean}} \tag{1}$$

$$CF = \frac{x_{peak}}{x_{RMS}} \tag{2}$$

Table 1. Polysomnographic features employed in the study, along with their acronym and description.

Feature (Acronym)	**Description**
Sleep Onset Latency (SOL)	The amount of time required to fall asleep (minutes)
Wake After Sleep Onset (WASO)	The amount of time the subject is awake during the night (minutes)
Total Sleep Time (TST)	Total hours of sleep
Time in bed (TIB)	Lights-off to lights-on interval (hours)
Sleep Efficiency (SE)	The percentage of time spent asleep while in bed (%)
Arousal Index (ARI)	Frequency of occurrence of arousals
Minutes of REM Sleep (MREM)	Total duration of REM Sleep (minutes)
Minutes of SWS Sleep (MSWS)	Total duration of SWS (minutes)
Proportion of N1 Sleep (PN1)	N1 Sleep per TST (%)
Proportion of N2 Sleep (PN2)	N2 Sleep per TST (%)
Proportion of SWS Sleep (PN3)	SWS Sleep per TST (%)
Proportion of REM Sleep (PNR)	REM Sleep per TST (%)
NREM Fragmentation Index (NFI)	A measure of the number of transitions from NREM to any other NREM stage per hour of NREM sleep
REM Fragmentation Index (RFI)	A measure of the number of transitions from REM to any other sleep stage per hour of REM
Wake Proportion (WP)	Awake time during the night (%)
Sleep Transition Index (STI)	A measure of the number of transitions from REM to NREM (and vice versa) per hours of sleep
Average Length N1 (ALN1)	Average length of N1 segments (minutes)
Average Length N2 (ALN2)	Average length of N2 segments (minutes)
Average Length SWS (ALN3)	Average length of SWS segments (minutes)
Average Length REM (ALREM)	Average length of REM segments (minutes)

$$IF = \frac{x_{peak}}{|x|_{mean}} \tag{3}$$

The Hjorth Parameters – i.e., Activity, Mobility, Complexity – were also computed on the EEG signal and its first- and second-order derivatives. They are defined in Equations (4), (5), (6).

$$Activity = var(x) \tag{4}$$

$$Mobility = \sqrt{\frac{var(\frac{dx}{dt})}{var(x)}} \tag{5}$$

$$Complexity = \frac{Mobility(\frac{dx}{dt})}{Mobility(x)} \tag{6}$$

As regards the spectral features, the Power Spectral Density (PSD) was estimated on each mini-epoch through the Welch Periodogram (50% overlap, 1 s Hamming Window). Then, statistics on the power spectrum density (PSD) were computed, the mean and median frequencies were retrieved, as well as the absolute and relative powers for each clinically relevant EEG band and entropy measures. The Teager-Kaiser Energy Operator was also computed on the mini-epochs, for the whole available spectrum (0–40 Hz), as previously introduced in [17].

Sleep Substructure Features Though previously introduced as homogeneous states, both the REM Sleep and SWS stages are characterised by an underlying substructure. In particular, the REM stage features two micro-states: the tonic stage (TREM) and the phasic stage (FREM) [22]. In further detail, aside from the diversity regarding the morphology of the signal, the two aforesaid micro-states significantly differ from a spectral point of view. In fact, as also presented in [17], the FREM and TREM micro-states lie in the frequency ranges 2–8 Hz and 7–16 Hz, respectively; these two sub-bands are noteworthy for feature extraction purposes. Likewise, deep sleep – i.e., SWS – presents with an underlying substructure, with the 1 Hz threshold being particularly significant. While commonly and clinically located in the δ-band (0.5–4 Hz), SWS is indeed further divided into two sub-bands: slow oscillations (SOs) and slow-wave activity (SWA). From a spectral point of view, SOs are the slowest waves of deep sleep ($<$ 1 Hz), whereas SWA lies in the 1–4 Hz range [2]. Recent works highlighted the significance of this dychotomy in relation to Amyloid-β aggregation [14]. In particular, the latter shows positive correlation with the SWA frequency range; thus proving its significance in the neurodegeneration process. In accordance with these statements, additional features regarding the spectral substructure of the REM and SWS stages were extracted, and are listed in Table 3.

Table 2. Employed features, along with the domain and proper reference. ⋄: adapted from the cited study, †: first proposed in this study.

Category	Feature (Name and description)	Reference
Time	Amplitude metrics: mean, standard deviation, skewness, kurtosis, range, maximum and minimum value	various
	Zero Crossing Rate	[24]
	Hjorth Parameters	[13]
	Percentiles (25^{th}, 75^{th}, 95^{th})	various
	Form, Crest and Impact Factors	various
	Coastline	[29]
Frequency	Fast Fourier Transform: numerical and statistical measures (mean and median frequencies, total power, ...)	various
	Spectral Edge Frequencies (SEF25, SEF75, SEF95)	⋄[17]
	Spectral Edge Frequencies differentials ($75-25$, $95-25$, $95-50$)	†
	Absolute Power for each clinically relevant band ($\delta, \theta, \alpha, \beta, \gamma$)	various
	Relative Power for each clinically relevant band ($\delta, \theta, \alpha, \beta, \gamma$)	various
	Entropy measures	⋄[17]
Non-Linear	Teager-Kaiser Energy Operator: numerical and statistical measures	⋄[11]

Table 3. Additional spectral features, extracted according to sleep sub-structure. ⋄: adapted from the cited study, †: first proposed in this study.

Sleep Stage	Feature	Reference
REM stage	Absolute and Relative Power in TREM, FREM	⋄[17]
	Mean, Median Frequencies and Spectral Percentiles (SEF_x) in TREM, FREM	⋄[17]
	Total Power Ratio TREM/FREM	⋄[17]
Slow Wave Sleep	Absolute and Relative Power in SO, SWA	⋄[19]
	Mean, Median Frequencies, Spectral percentiles (SEF_x), statistical measures in SO, SWA	†

2.4 Data post-processing and Feature Sets

Section 2.3 described the steps employed in the feature extraction process. Owing to the fact that the extracted features belong to diverse domains – e.g., clinical scores and computational parameters – Z-score normalization was applied to the whole feature set (367 features), thus transforming the features to the continuous range [0, 1]. This procedure allowed the feature arrays to follow a normal distribution, easing the subsequent implementation of distance-based measures and classifiers.

At the end of the post-processing steps, three feature sets (FeatSet_i) are obtained:

- **FeatSet**$_1$: Polysomnographic + REM Sleep features;
- **FeatSet**$_2$: Polysomnographic + Slow Wave Sleep features;
- **FeatSet**$_3$: Polysomnographic + REM Sleep + Slow Wave Sleep features.

The three feature sets are then used for binary classification purposes, in an attempt to discriminate between healthy and RBD subjects.

2.5 Feature Selection

As regards Feature Selection, a minimal-optimal approach was adopted, which aims at selecting the set of features that, once grouped, have the highest predictive power. The employed feature selection method is the Minimum Redundancy Maximum Relevance (mRMR) [16], based on Pearson Correlation and F-test, which maximises the mutual information provided by the features in spite of their redundancy. This method ranks the features according to their relevance to the target variable. The resulting top-5 features for each subset are shown in Table 4. As appreciable, when considering FeatSet_3 (i.e., REM, SWS + PSG), three features out of five belong to SWS.

2.6 Classification

The aim of this work was to perform an automatic binary classification between healthy and RBD subjects. To this end, supervised Machine Learning methods were applied to the three feature sets presented in Section 2.4. Four different models were tested, namely: Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Naive-Bayes (NB), Decision Tree (DT), along with an ensemble method, Bootstrap Aggregating (BAG). The latter is a meta-algorithm specifically designed to reduce variance within the dataset, thus preventing overfitting. The model randomly samples the original dataset, creating from such different sub-sets; it then parallel-trains the sub-sets separately and finally yields prediction scores based on majority voting. As regards the other models, a brief description follows:

- **SVM**: it aims at finding the hyperplane which best separates data of the two classes, while maximising the margin – i.e., the distance between the data points (support vectors) and the hyperplane.

Table 4. Top-5 ranked features with the mRMR feature selection method, in all three employed feature sets. For FeatSet_3: ⋆ REM features, ∘ SWS features

FeatSet	**Top-5 Features**
FeatSet_1 (PSG + REM)	Relative Power (α) Minutes in REM Sleep WASO Mobility (2^{nd} order) SEF75
FeatSet_2 (PSG + SWS)	Relative Power SWA (75^{th} pctl) Relative Power (θ), STD Relative Power (α), STD Minimum Amplitude Median (75^{th} pctl)
FeatSet_3 (PSG + SWS + REM)	Relative Power SWA (75^{th} pctl), ∘ Relative Power (θ), STD, ∘ Mobility (2^{nd} order), ⋆ Relative Power (α), ⋆ Minimum Amplitude, ∘

- **KNN**: it is a non-parametric method that classifies observations based on their similarity to its closest data-points in the datasets (i.e., *neighbours*).
- **NB**: it is a probabilistic classifier based on the Naive-Bayes theorem.
- **DT**: this model is based on decision rules; it starts from a root-node and classifies observations by testing them at each decision node, eventually leading to the terminal node – i.e., the label.

Hyperparameters optimisation (Bayesian approach) was applied to each model. Finally, to prevent overfitting k-fold cross-validation (CV) ($k = 5$) was applied. This technique samples the dataset into k different sub-sets; it performs training on $k - 1$ subsets and tests the model on the remaining one. Considering the random sampling adopted by the CV, and to prevent the classification performance from being affected by weak generalization capability, this procedure was iterated 10 times.

3 Results

As introduced earlier, the models presented in Section 2.6 were employed in a binary classification task, with the aim of discriminating healthy from RBD subjects, (CAP Sleep Database and TuSDi Database). This Section presents the classification performance, for each considered Feature Set (*cf.* Section 2.4).

3.1 Classification Performance: REM Features

The first Feature Set considered for the analysis (FeatSet_1) comprised PSG features and features extracted from the REM segments. The classification perfor-

mance of the five models tested is displayed in Table 5; the values refer to a 5-fold CV, averaged over 10 iterations. The macro-averaged accuracy (across all tested classifiers) is 80.29% ± 0.03. The best overall performance was attained through the KNN classifier (Accuracy: 83.91% ± 0.81, Sensitivity: 86.46% ± 2.95). As regards the KNN model optimisation, the searched parameters were: K (number of neighbours, 1 to 29) and distance metrics, resulting in the optimised parameters K=3 and Chebyshev Distance.

Table 5. Classification performance (%) of the employed classifiers as regards FeatSet_1 (PSG + REM features).

	SVM	**KNN**	**NB**	**DT**	**BAG**
Accuracy	81.03 ± 0.5	83.91 ± 0.81	74.14 ± 1.72	82.18 ± 0.81	80.17 ± 0.86
Sensitivity	84.38 ± 3.13	86.46 ± 2.95	73.44 ± 1.56	85.42 ± 5.31	82.81 ± 1.56
Specificity	76.92 ± 3.85	80.77 ± 3.14	75 ± 1.92	78.21 ± 7.2	76.92 ± 0.1
Precision	81.94 ± 1.94	84.78 ± 1.75	78.33 ± 1.67	83.28 ± 4.25	82.16 ± 0.91
F1	83.06 ± 0.52	85.55 ± 0.9	75.81 ± 1.61	84.07 ± 0.87	82.16 ± 0.91
AUC	0.87 ± 0.05	0.87 ± 0.01	0.76 ± 0.02	0.83 ± 0.03	0.89 ± 0.02

3.2 Classification Performance: REM + SWS Features

Secondly, the classifiers were tested on FeatSet_3, which adds to the previous one the features extracted from the SWS stage. The classification performance consistently increases for all classifiers (Table 6). This feature set, indeed, presents with an overall macro-averaged accuracy of 85.70% ± 0.04 (+6% with respect to FeatSet_1), and a 8.3% increase for the AUC metric. The best classification performance is achieved through the DT (Accuracy: 90.80% ± 0.8, Sensitivity: 95.83% ± 2.95). As for best model optimisation, the searched hyperparameters were Split Criterion (method) and N (maximum number of splits, range: 1–57); the resulting optimised parameters were cross entropy as Split Criterion and N=4.

Table 6. Classification performance (%) of the employed classifiers as regards FeatSet_3 (PSG + REM + SWS features).

	SVM	**KNN**	**NB**	**DT**	**BAG**
Accuracy	89.08 ± 1.63	85.34 ± 2.59	79.31 ± 1.72	90.80 ± 0.8	83.62 ± 0.86
Sensitivity	92.71 ± 2.95	93.75 ± 0.4	82.81 ± 4.69	95.83 ± 2.95	85.94 ± 1.56
Specificity	86.42	75 ± 5.77	75 ± 1.92	84.62 ± 3.14	80.77
Precision	88.11 ± 0.33	82.33 ± 3.38	80.32 ± 0.32	88.54 ± 1.82	84.61 ± 0.24
F1	90.33 ± 1.56	87.63 ± 1.92	81.47 ± 2.11	91.99 ± 0.79	85.26 ± 0.89
AUC	0.98 ± 0.01	0.92 ± 0.03	0.82 ± 0.03	0.92 ± 0.02	0.93 ± 0.05

3.3 Classification Performance: SWS Features

Finally, given the final aim of the study – i.e, determine which stage (REM vs SWS) is more informative in the automatic detection of RBD – the classifiers were tested also on the remaining feature set (FeatSet_2), encompassing PSG features and variables extracted from the SWS stage alone. The performance metrics are displayed in Table 7. As appreciable, the classifiers achieved reasonably good results, with an overall macro-averaged accuracy of 81.10% ± 0.03. Though the average accuracy (across classifiers) only shows a slight overall increase with respect to the REM Subset, the SVM classifier clearly outperforms the performances of the classifiers on FeatSet_1. In fact, it achieved an Accuracy of 86.21% ± 2.11 and Sensitivity of 91.23% ± 5.24, as well as an AUC value of 0.94 ± 0.02. As regards SVM, the searched parameters were Kernel Function and Maximum Penalty; the optimised model featured a cubic Kernel Function and a Maximum Penalty of 2.56 (search range: 0.001–1000).

Table 7. Classification performance (%) of the employed classifiers as regards FeatSet_2 (PSG + SWS features only).

	SVM	**KNN**	**NB**	**DT**	**BAG**
Accuracy	86.21 ± 2.11	80.46 ± 4.94	78.74 ± 4.30	79.52 ± 7.21	80.60 ± 3.31
Sensitivity	91.23 ± 5.24	83.71 ± 12.26	76.67 ± 6.73	80.83 ± 9.78	79.13 ± 5.58
Specificity	83.36 ± 1.57	80.83 ± 2.94	81.16 ± 3.15	79.06 ± 6.25	82.19 ± 1.92
Precision	76.92 ± 2.72	74.36 ± 7.90	76.92 ± 5.44	74.04 ± 6.72	77.88 ± 3.19
F1	83.36 ± 2.38	77.58 ± 3.58	76.50 ± 4.14	77.04 ± 7.31	78.34 ± 2.98
AUC	0.94 ± 0.02	0.85 ± 0.02	0.81 ± 0.02	0.77 ± 0.09	0.9 ± 0.02

4 Discussion

As described above (*cf.* Sections 3.1–3.3), the employed classifiers generally achieved a good global performance on all tested feature sets. Indeed, the best classifier for the REM subset scored an overall Accuracy of 84%, with 86.5% sensitivity to the RBD class. As appreciable from the data presented in Table 6, the classification performance significantly increases when including the SWS features into the analysis, yielding a 91% Accuracy and 96% Sensitivity (best score). Furthermore, the SWS features alone proved efficient in addressing the binary classification task, achieving an overall 86% Accuracy and 91% Sensitivity (SVM classifier). As displayed in Fig. 1, that compares the best models on each tested feature set, FeatSet_2 (SWS features) outperforms the REM subset, suggesting its applicability and relevance in the automatic detection of RBD.

This consideration is in line with the results presented in [3], which performed a stage-agnostic feature extraction, and highlighted the δ-range – and in particular the 1.5–2 Hz sub-band – as the most important in a healthy vs RBD

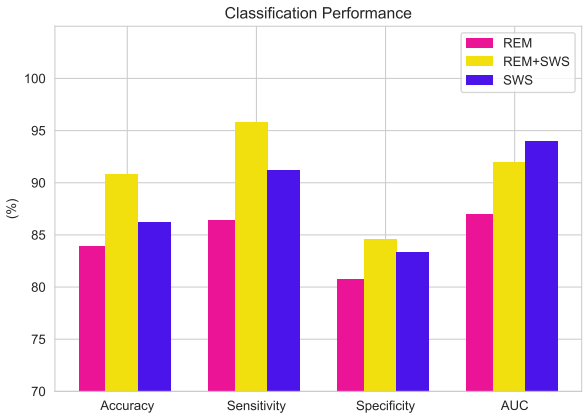

Fig. 1. Performance comparison of the best model for each analysed Feature Set.

classification approach. In addition, the classification performance was compared to the cited study, which addressed the task employing the same publicly available data (CAP Sleep Database) [3]. Therefore, the cited study analysed the EEG recordings across the whole spectrum (0.5–50 Hz), with no stage distinction; our results outperform their metrics when considering features from both SWS and EEG (FeatSet_3), with a balanced accuracy (averaged across all models) of 86% versus 83%, as well as a 26% improvement on Specificity, and +2% on Sensitivity. When considering the metrics on the SWS feature set only, the SVM classifier yielded comparable performances as regards Accuracy and Sensitivity, and achieved a 22% increase on Specificity with respect to [3].

5 Conclusion and Future Work

The clinical diagnosis of RBD involves a overnight full-PSG exam and additional anamnestic interviews. Though being the gold standard for the detection and monitoring of sleep disorders, PSG is invasive and impractical. Indeed, not only it entails a high number of recording electrodes, but the manual and visual scoring process is labourious, thus significantly hindering the diagnosis. Given that RBD itself foreruns the onset of α-synucleinopathy up to 14 years, there is the need to accelerate the detection process, providing accurate and lightweight alternatives to manual scoring.

This paper proposed a method for the automatic detection of RBD from single-channel EEG data. Data were extracted from the recordings during REM Sleep and SWS. The implemented ML methods achieved high performance, suggesting the applicability of single-channel electroencephalography in detecting RBD subjects.

Moreover, the performance attained by the SWS features alone is quite promising, therefore implying its applicability to the study of RBD and the related neurodegenerative process. These findings indicate the feasibility of a lightweight screening tool, thus easing the scoring and diagnostic process. Furthermore, automatic classification would reasonably facilitate follow-up procedures, and allow for early detection and early disease-modifying interventions which would have beneficial impact on the QoL of patients [6].

Future work will address a larger dataset to further investigate the SWS dychotomy and its role in the development of RBD, and will combine the investigation of the breathing frequency patterns during both REM Sleep and SWS.

References

1. Berry, R.B., Brooks, R., Gamaldo, C.E., Harding, S.M., Marcus, C., Vaughn, B.V., et al.: The aasm manual for the scoring of sleep and associated events. Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine **176**, 2012 (2012)
2. Björn, R., Jan, B.: About sleep's role in memory. Physiol. Rev **93**, 681–766 (2013)
3. Buettner, R., Grimmeisen, A., Gotschlich, A.: High-performance diagnosis of sleep disorders: a novel, accurate and fast machine learning approach using electroencephalographic data. In: Proceedings of the 53rd Hawaii International Conference on System Sciences (2020)
4. Cooray, N., Andreotti, F., Lo, C., Symmonds, M., Hu, M.T., De Vos, M.: Detection of rem sleep behaviour disorder by automated polysomnography analysis. Clinical Neurophysiology **130**(4), 505–514 (2019)
5. Cooray, N., Andreotti, F., Lo, C., Symmonds, M., Hu, M.T., De Vos, M.: Proof of concept: Screening for rem sleep behaviour disorder with a minimal set of sensors. Clinical Neurophysiology **132**(4), 904–913 (2021)
6. Fayyaz, M., Jaffery, S.S., Anwer, F., Zil-E-Ali, A., Anjum, I.: The effect of physical activity in parkinson's disease: a mini-review. Cureus **10**(7) (2018)
7. Galbiati, A., Verga, L., Giora, E., Zucconi, M., Ferini-Strambi, L.: The risk of neurodegeneration in rem sleep behavior disorder: a systematic review and meta-analysis of longitudinal studies. Sleep medicine reviews **43**, 37–46 (2019)
8. Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. circulation **101**(23), e215–e220 (2000)
9. Haba-Rubio, J., Frauscher, B., Marques-Vidal, P., et al.: Prevalence and determinants of rapid eye movement sleep behavior disorder in the general population. Sleep **41**(2) (2018)
10. Hu, M.T.: Rem sleep behavior disorder (rbd). Neurobiology of Disease **143**, 104996 (2020)
11. Lajnef, T., Chaibi, S., Ruby, P., Aguera, P.E., Eichenlaub, J.B., Samet, M., Kachouri, A., Jerbi, K.: Learning machines and sleeping brains: automatic sleep stage classification using decision-tree multi-class support vector machines. Journal of neuroscience methods **250**, 94–105 (2015)

12. López-Garcıa, D., Ruz, M., Ramırez, J., Górriz, J.: Automatic detection of sleep disorders: Multi-class automatic classification algorithms based on support vector machines. In: International Conference on Time Series and Forecasting (ITISE 2018). vol. 3, pp. 1270–1280 (2018)
13. Motamedi-Fakhr, S., Moshrefi-Torbati, M., Hill, M., Hill, C.M., White, P.R.: Signal processing techniques applied to human sleep eeg signals—a review. Biomedical Signal Processing and Control **10**, 21–33 (2014)
14. Ngo, H.V.V., Claassen, J., Dresler, M.: Sleep: slow wave activity predicts amyloid-β accumulation. Current Biology **30**(22), R1371–R1373 (2020)
15. Pavlova, M.K., Latreille, V.: Sleep disorders. The American Journal of Medicine **132**(3), 292–299 (2019)
16. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence **27**(8), 1226–1238 (2005)
17. Rechichi, I., Zibetti, M., Borzì, L., Olmo, G., Lopiano, L.: Single-channel eeg classification of sleep stages based on rem microstructure. Healthcare Technology Letters **8**(3), 58 (2021)
18. Schenck, C.H., Bundlie, S.R., Ettinger, M.G., Mahowald, M.W.: Chronic behavioral disorders of human rem sleep: a new category of parasomnia. Sleep **9**(2), 293–308 (1986)
19. Schreiner, S.J., Imbach, L.L., Valko, P.O., Maric, A., Maqkaj, R., Werth, E., Baumann, C.R., Baumann-Vogel, H.: Reduced regional nrem sleep slow-wave activity is associated with cognitive impairment in parkinson disease. Frontiers in Neurology **12**, 156 (2021)
20. Scullin, M.K., Gao, C.: Dynamic contributions of slow wave sleep and rem sleep to cognitive longevity. Current sleep medicine reports **4**(4), 284–293 (2018)
21. Siegel, J.M.: Rem sleep: a biological and psychological paradox. Sleep medicine reviews **15**(3), 139 (2011)
22. Simor, P., van der Wijk, G., Nobili, L., Peigneux, P.: The microstructure of rem sleep: Why phasic and tonic? Sleep medicine reviews **52**, 101305 (2020)
23. Stefani, A., Högl, B.: Sleep in parkinson's disease. Neuropsychopharmacology **45**(1), 121–128 (2020)
24. Šušmáková, K., Krakovská, A.: Discrimination ability of individual measures used in sleep stages classification. Artificial intelligence in medicine **44**(3), 261–277 (2008)
25. Terzano, M.G., Parrino, L., Sherieri, A., Chervin, R., Chokroverty, S., Guilleminault, C., Hirshkowitz, M., Mahowald, M., Moldofsky, H., Rosa, A., et al.: Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (cap) in human sleep. Sleep medicine **2**(6), 537–553 (2001)
26. Urtnasan, E., Joo, E.Y., Lee, K.H.: Ai-enabled algorithm for automatic classification of sleep disorders based on single-lead electrocardiogram. Diagnostics **11**(11), 2054 (2021)
27. Widasari, E.R., Tanno, K., Tamura, H.: Automatic sleep disorders classification using ensemble of bagged tree based on sleep quality features. Electronics **9**(3), 512 (2020)
28. Xie, L., Kang, H., Xu, Q., Chen, M.J., Liao, Y., Thiyagarajan, M., O'Donnell, J., Christensen, D.J., Nicholson, C., Iliff, J.J., et al.: Sleep drives metabolite clearance from the adult brain. science **342**(6156), 373–377 (2013)
29. Yetton, B.D., Niknazar, M., Duggan, K.A., McDevitt, E.A., Whitehurst, L.N., Sattari, N., Mednick, S.C.: Automatic detection of rapid eye movements (rems): A machine learning approach. Journal of neuroscience methods **259**, 72–82 (2016)