POLITECNICO DI TORINO

Department of Automatics and Informatics

Master's Degree Course in Biomedical Engineering

Master' Thesis

Automatic classification of REM Sleep Behaviour Disorder exploiting Heart Rate Variability and Electromyography

Supervisor
Prof. Gabriella Olmo

Co-supervisor
Eng. Irene Rechichi

Candidate
Umberto Mosca

March 2023

INDEX

Contents

INDEX **2**

ABSTRACT **3**

INTRODUCTION **4**

I. SLEEP IN GENERAL 4

II. SLEEP ROLE 6

III. AMYLOID-B 6

IV. RBD 7

IV.1 ETIOLOGY 7

IV.2 EPIDEMIOLOGY 8

IV.3 PATHOPHYSIOLOGY 8

IV.4 CLINICAL FEATURES AND DIAGNOSIS 9

MATERIALS AND METHODS **11**

I. SUBJECTS AND DATA 11

II. DATA PRE-PROCESSING 12

III. FEATURE EXTRACTION 12

III.1 ECG FEATURES 12

III.2 EMG FEATURES 15

IV. FEATURE SELECTION 16

V. CLASSIFICATION 17

RESULTS **19**

I. CLASSIFICATION PERFORMANCE 19

DISCUSSION **20**

REFERENCES **24**

ABSTRACT

Recent works highlighted that one of the fundamental roles of sleep is the removal of toxic metabolites produced during wakefulness. Consequently, a correlation was found between the quality of both rapid-eye movement (REM) sleep and deep sleep (i.e., slow wave sleep, SWS) and the development of neurodegenerative diseases.

In this regard, the study of RBD has assumed great importance as it is considered a symptom that precedes, even by decades, the development of α-synucleinopathies, and a conversion rate to Parkinson's disease of around 90% has been found.

The diagnosis of REM sleep behaviour disorder (RBD) is a lengthy and cumbersome process involving the analysis of polysomnography (approximately 8 hours of recording on more than 10 tracings) by a sleep expert. Consequently, the aim of this study is to develop a model capable of easing the physician's workload by performing an automatic binary classification between healthy and RBD patients.

To date, some steps have already been taken in this direction often by exploiting electroencephalography (EEG) and electromyography (EMG) signals. However, following the indications of the latest research highlighting a relationship between heart rate variability (HRV) and RBD, this work aims to assess the predictive properties of the ECG signal, either singly or conjugated with the EMG signal, in detecting RBD.

Different combinations of feature selection algorithms and machine learning models (preferred over deep learning due to the explainability of the results provided) were evaluated, obtaining accuracy values over 85%. This supports and confirms the research conducted on HRV and could further reduce the costs associated with the diagnosis of RBD, leading to the possibility of implementing screening procedures on the over-60 population. This would result in a very early diagnosis and, consequently, more effective treatment.

INTRODUCTION

I. Sleep in general

Sleep is a cyclically recurring condition in most living beings and is essential in many physiological functions necessary for proper growth and development of the organism.

The study of sleep is a relatively recent discipline, born in the 1930s with Berger.

The definition of sleep used to be expressed on an exclusively behavioural basis (considering motility, response to external stimuli, eye closure, posture, and reversibility of the state). Recently, it has become essential to supplement the behavioural definition model with physiological parameters that define certain peculiarities of the different sleep stages. The instrumental tests used for this purpose are:

EEG (electroencephalogram): analyses cortical dynamic activity and, by extension, subcortical activity.

EOG (electrooculogram): records eye movements, used to differentiate certain sleep phases that would otherwise be very similar and to identify REM sleep. Registration is based on the potential difference between the cornea (positive) and retina (negative).

EMG (electromyogram): analyses basic muscle tone.

Sleep is roughly divided into REM sleep (25% - with Rapid Eyes Movements) and NREM sleep (75% - in which the eyes do not stand still but do not make rapid movements). NREM sleep is further subdivided into N1 sleep (5% - wakefulness transition sleep), N2 sleep (50% - more stable sleep), N3 sleep (20% - Slow Wave Sleep). The percentages of the different phases change throughout life and according to the duration of wakefulness: the child spends more time in the N3 phase than the adult, the elderly show a reduction in slow wave sleep and an increase in transitions to the waking state - up to 6-7).

The most important phases in homeostatic terms are N3 sleep and REM sleep, the

others are expendable: in case of a sleepless night, during the following night, N3 sleep is recovered with a doubling of its normal duration and a large part of REM sleep, to the detriment of N1-N2 sleep (transition phases). These different phases are organised in 70-90 minutes cycles that follow each other until awakening.

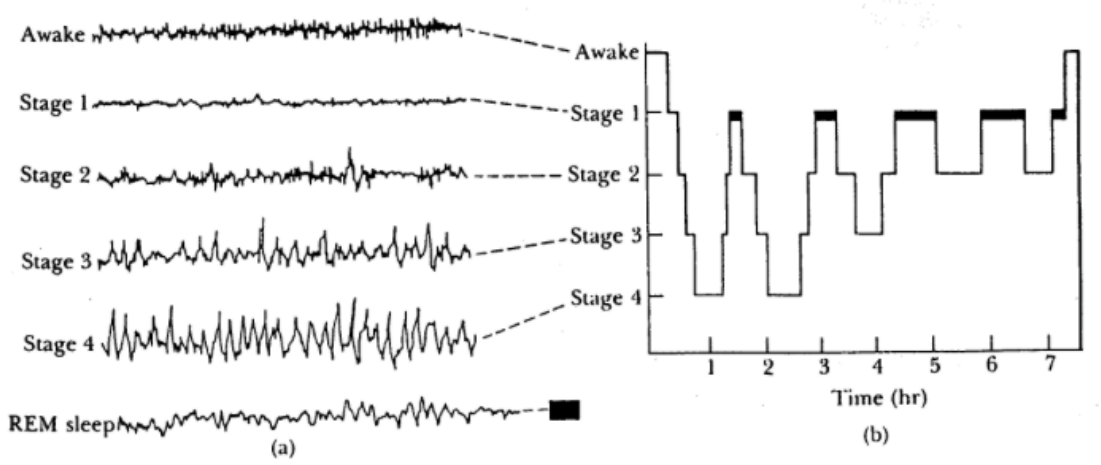

Figure 1 *Example of physiological hypnogram*

Observing the graph in Figure 1, which shows a physiological hypnogram, it can be seen that initially one moves from wakefulness to deep sleep N3 (N4 sleep has recently been merged with N3 sleep), then follows the reverse path back to the N1 stage from which one moves on to REM sleep which closes the cycle. In the first part of sleep, N3 sleep is more represented, while REM sleep prevails in the second part. This is due to physiological necessity and has great pathological impacts as REM sleep parasomnias frequently appear in this part of sleep.

It now remains to be determined how the different sleep stages can be differentiated. For this purpose, the most decisive electrophysiological signal is certainly the EEG, which is characterised by different frequency bands. The analysis of these provides critical information about the regulation of wakefulness by the several brain areas and the succession of sleep stages.

- **Awake stage**: rapid, low-amplitude brain activity à α band (9-15 Hz) and β band (15-30Hz).
- **N1 stage**: decrease in amplitude (less than 75 to 50 µV) and frequency à θ band (4-9 Hz).

- **N2 stage**: It is characterized by predominant θ activity and occasional quick bursts of faster activity. K complexes and sleep spindles occur for the first time and are typically episodic. K complexes are sharp, monophasic or polyphasic slow waves, with a sharply negative (upward) deflection followed by a slower positive (downward) deflection. They characteristically stand out from the rest of the background.
- **N3 stage**: even slower and more extensive activity δ band (0.5-4 Hz). K-complexes may be detected. When the sleep becomes even deeper, cortical activity decreases further becoming particularly wide and slow (what used to be called N4 sleep).
- **REM stage**: desynchronisation of brain activity that appears similar to wakefulness associated with Rapid Eyes Movements and generalised atonia.

II. Sleep role

Over the years, numerous research studies have given sleep a wide variety of functions, including energy conservation, immunological processing, memory consolidation, emotion regulation, and threat simulation.

More recently, it was hypothesised, and subsequently demonstrated [2], that another fundamental role of sleep was the elimination of toxic metabolites that accumulate during wakefulness.

One of the most studied neurodegenerative effects is the formation of senile plaques due to the accumulation, and subsequent aggregation, of the protein Amyloid-β(Aβ).

III. Amyloid-β

The amyloid precursor protein (APP) is a long transmembrane protein that, when processed incorrectly, gives rise to β-peptide (most frequently in the form Aβ42), which causes the accumulation of Aβ.

The processing of APP follows the so-called amyloid cascade and is carried out by different isoforms of the secretase enzyme, of which α-secretase causes the formation of a non-amyloidogenic fragment, while the conjugated β- and γ-secretase

produces β-peptide.

Aggregation into fibrils/platelets of β-peptide usually occurs from fragments (seeds) of 8-12 subunits, which are first deposited within the neuron and later, with neuron death, escape into the extracellular space.

The result of β-peptide aggregation is neuronal damage to synaptic function and nerve transmission, with parallel activation of microglia in attempt to clear the protein aggregates. The progressive deposition of Aβ evokes a neuroinflammatory response from glial cells, with production of neurotoxic cytokines (IL-1, IL-6, TNF-α). Neuronal death and degeneration then appear. Indeed, a depletion of acetylcholine has been noted in the basal ganglia and in the nucleus basalis of Meynert, responsible at least in part of the cognitive and mnestic decline.

Elements called extracellular vesicles have also been identified that can contain and transport β-peptide protein aggregates, transmitting the pathology to contiguous areas of the CNS.

Early studies looked for and found a relationship between Aβ accumulation and macroscopic sleep measurements. Going into greater detail, analysing the EEG trace in the 0.5-4Hz band, Slow Wave Activity (SWA), an inverse proportionality was discovered between spectral power for frequencies below 1Hz and Aβ aggregation and, at the same time, a direct proportionality for frequencies above 1Hz. This phenomenon confirms the results of early studies and can be translated into a correlation between sleep quality and the onset of several neurodegenerative diseases [3].

IV. RBD

IV.1 Etiology

In this context REM Sleep Behaviour Disorder (RBD) plays an important role in increasing the risk of developing neurodegenerative syndrome, with a conversion rate of 6.3% per year and a total of 74% converting after 12-year follow-up. It is defined, by the International Classification of Sleep Disorders 3rd Edition (ICSD-3) [1], as a parasomnia manifested by vivid, often frightening dreams accompanied by simple or

complex motor behaviours during REM sleep.

RBD can be divided into three categories:

1. Idiopathic RBD: it is most suggestive in neurodegenerative synucleinopathies, including dementia with Lewy bodies, Parkinson's disease, olivopontocerebellar degeneration, multiple-system atrophy, and Shy-Drager syndrome. [5]
2. Drug-induced RBD: it is common in individuals who are taking antidepressants as serotonin reuptake inhibitors (fluoxetine), tricyclic antidepressants (mirtazapine, protriptyline, amitriptyline, nortriptyline, desipramine, imipramine), and monoamine oxidase inhibitors (phenelzine and selegiline). [6,7]
3. Secondary RBD due to medical condition: it may be considered a distinct phenotype of RBD. It is characterized by less violent or complex behaviour during REM sleep, earlier age of onset, equal sex distribution, and hypocretin (orexin) deficiency (a lab diagnosis specific for narcolepsy type 1). [8]

IV.2 Epidemiology

The prevalence of this disease is around 0.5-1.25% in the entire world population and increases to 2% if only the population over 60 is considered. In fact, the average age of diagnosis is around 52-65 years and a higher incidence, around 78%, can be noted in the male sex [4], but it is worth mentioning the fact that female cases are likely underreported and underdiagnosed [9].

Since isolated RBD is a prodromal syndrome of alpha-synuclein neuropathology, it is widespread among patients with Parkinson’s disease (PD) (33-50%), multiple system atrophy (80-95%), and dementia with Lewy bodies (80%) [10].

IV.3 Pathophysiology

Physiologically, REM sleep is associated with skeletal muscle atonia, producing paralysis. This confers a protective measure, preventing people from acting what they dream in the setting of dream mentation. Quiescence of motor activity during REM sleep may also facilitate sleep-related memory consolidation.

The physiologic suppression of motor activity during REM sleep is the cumulative result of multiple neuronal circuits that predominantly originate in the pons and ultimately terminate on spinal cord motor neurons. While the exact lesion that causes RBD is unknown, evidence suggests that in both isolated and medication-induced RBD, the loss of REM sleep atonia is related to dysfunction of the subcoeruleus complex in the rostral pons [11,12].

The neurodegenerative α-synucleinopathies consist of glial cytoplasmic inclusions aggregates of insoluble alpha-synuclein protein [11]. However, it is unclear if the link between this neurodegenerative classification and RBD results from these accumulated aggregates or through another pathology. Animal studies showed that progressive alpha-synuclein aggregation and neuronal apoptosis occurred in the ventricular reticular nucleus in the brainstem and resulted in RBD symptoms in a chronic rat model of Parkinson disease [13].

While the relationship between RBD and synucleinopathies is more frequently described, other non-synucleinopathies such as progressive supranuclear palsy (PSP) [14], familial amyotrophic lateral sclerosis [15], frontotemporal dementia [16], myotonic dystrophy [17] were reported less commonly.

IV.4 Clinical features and diagnosis

Patients with RBD exhibit abnormal motor behaviours during REM sleep, predominantly while in bed. These behaviours include talking, screaming, swearing, gesturing, arm flailing, punching, kicking, and leaping or falling off the bed. Although violent behaviours are most common, nonviolent activities such as laughing, whistling, singing, and masturbation may occur occasionally (about in 1 on 5 patients). An early warning bell for suspecting the presence of this pathology may be a positive history of injuries reported by both the patient and his or her bed partner, who, in 6 out of 10 cases, reports being a victim of nocturnal assaults. The aggressive attitudes are a direct consequence of the dreams content that the study [18] describes to be of a more violent nature than those reported by the control group. This difference does not appear to be correlated with greater aggression, as measured by a questionnaire, shown throughout the day. In fact, it appears to be in range and even lower than the average of the healthy group.

Screening questionnaires or a history of abnormal sleep behaviours (particularly when obtained from the bedpartner) can be highly suggestive of RBD in the appropriate context, in particular, the question "Have you ever been told, or suspected yourself of acting out the content of your dreams while asleep?" is considered in itself sensitive to suspect RBD [11].

Anyway, the guidelines [1] require polysomnography (PSG) to provide a formal diagnosis, and the following criteria are necessary:

- Repeated episodes of sleep-related vocalization and/or complex motor behaviours.
- Behaviours are documented by polysomnography to occur during REM sleep or, based on clinical history of dream enactment, are presumed to occur during REM sleep.
- Presence of REM Sleep Without Atonia (RSWA) on PSG
- Absence of epileptiform activity during REM sleep, unless RBD can be clearly distinguished from any concurrent REM sleep-related seizure disorder.
- The sleep disturbance is not better explained by another sleep disorder, medical or neurologic disorder, mental disorder, or substance use disorder.

This, to date, mandatorily implies the analysis of PSG by a physician or expert in the field and, as it involves the all-night recording of about 8 EEG-derived traces, at least one EMG, EOG and ECG trace and often a video, the amount of work is burdensome and time-consuming.

To try to provide a solution to this challenge, several approaches have recently been tried out that exploit artificial intelligence algorithms to process PSG data. Since the ultimate goal is not only a correct health-RBD classification, but also a good results explicability in order to meet the physician's needs, the use of Machine Learning (ML) models has often been preferred over Deep Learning ones.

MATERIALS AND METHODS

I. Subjects and Data

Polysomnographic data from two different databases, one public and the other private, were used to conduct the study.

The public one is the CAP Sleep Database, available on PhysioNet [19], and from it the PSG traces of 15 healthy (7 males, aged 31 ± 5 years) and 22 RBD (19 males, aged 70 ± 6 years) patients were used, for a total of 37 subjects. For the sake of completeness, it should be specified that this database contains data from 16 healthy patients, one of whom was excluded from the study due to the lack of EMG and ECG traces.

The private one, called Turin Sleep Disorders Database (TuSDi Database), is composed by the polysomnographic recordings collected at the Centre for Sleep Disorders at Molinette Hospital (Turin, Italy) from which were extracted the data of 10 healthy and 19 pathological patients, out of whom 10 affected by RBD and 9 by RSWA.

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.

Therefore, in summary, the database exploited has 66 subjects, 25 healthy and 41 pathological.

The data most indicative of the presence of RBD are found in slow-wave sleep (SWS) and REM sleep. Bearing in mind that the primary purpose of this study is to evaluate the effectiveness of features extracted from EMG and ECG polysomnographic records in correctly identifying a healthy patient from one with RBD, sleep staging was not performed automatically in order to avoid introducing an error variable.

The manual staging and classification, according to the AASM standard, by a sleep expert from both databases was used for this purpose.

II. Data pre-processing

In addition to the presence of the ECG and EMG traces, another parameter for inclusion within the study is the presence, throughout the recording, of at least 5 minutes of SWS sleep and REM sleep. Therefore, the signals were first examined and subsequently resampled, if necessary, at 512 Hz to unify them.

Once the channels corresponding to the signals in question were found, they were filtered in order to reduce the high and low frequency noise while leaving the specific band of the signal under investigation unchanged.

For both signals, a zero-phase Chebyschev Type 1 bandpass filter was applied. For the ECG signal, the frequency band between 0.5 and 30 Hz was retained, while in the case of the EMG signal, the frequency band between 10 and 100 Hz was preserved. Since the filter applied on the EMG signal does not attenuate the 50 Hz line noise, it was decided to apply a Notch filter if it had not been already done previously.

As a final step to prepare the data for the subsequent feature extraction algorithms, all signals employed were converted in µV, thus preventing the analysis from being affected by the acquisition conditions.

III. Feature Extraction

A total of 257 features were extracted in this work, of which 270 from the ECG signal and 87 from the EMG signal. A detailed description of the extracted features follows in the subsequent paragraphs.

The algorithm was developed in the MATLAB R2022b (MathWorks, Natick, MA) software environment.

III.1 ECG features

In recent scientific literature, an association between RBD and autonomic nervous system dysfunction has emerged [19]. Going into detail, patients with idiopathic RBD predominantly showed symptoms in the gastrointestinal, genito-urinary and

cardiovascular regions, all of which are in some way connected to or governed by the autonomic nervous system [20, 21].

An important indicator of its functioning is the Heart Rate Variability (HRV) and recent studies have reported a lowering of some of its parameters, especially in those linked with the sympathetic system, in RBD patients.

As mentioned above, the sleep phases with the greatest relevance in the case of RBD are those belonging to SWS sleep (equivalent to the N3 stage) and REM sleep. For this reason, the ECG signal was not analysed in its entirety, but features were calculated on ECG traces resulting from the concatenation of the sleep stages of interest.

Furthermore, in previous studies performed on EEG, a partial correlation with RBD was also found on features calculated within the N2 stage of sleep.

For both databases, the specialist, relying on all available traces, assigned each 30-second PSG recording epoch to the corresponding sleep stage. In accordance with this, five ECG traces were created as a result of the concatenation of different sleep stages:

- concatenation of N2 stages
- concatenation of N3 stages
- concatenation of REM stages
- concatenation of N3 + REM stages
- concatenation of N2 + N3 + REM stages

First, to reach the calculation of HRV-related features, it is necessary to determine the R-R intervals (i.e. the time lapse between one peak of the R-wave and the next), and, consequently, it is essential to find the time instants of each R-peak.

To do this as accurately as possible, the wavelet transform was exploited to enhance the R peaks in the ECG waveform.

Subsequently, thanks to the R-R intervals just found, three HRV features related to the time domain (meanHR, SDNN, RMSSD) and six related to the frequency domain (VLF, LF, HF, normalizedLF, normalizedHF, LF-HF ratio) were calculated on each 30-second signal epoch. The frequency bands over which the above parameters were calculated are defined in the table below.

Of these, meanHR and SDNN are considered general measures of HRV; RMSSD, HF and normalisedHF are parameters influenced by the parasympathetic nervous system, while LF is an index of the autonomic nervous system in general, both sympathetic and parasympathetic [22]. Regarding the physiological significance of normalisedLF, some authors indicate a correlation only with the sympathetic nervous system, others correlate it also with the parasympathetic, and still others consider the normalised parameters to be redundant [23]. As the topic is still debated, it was decided to include all parameters in the study.

Feature name	**Meaning**
meanHR	mean heart rate
SDNN	standard deviation of the R-R interval
RMSSD	root mean square of successive R-R interval differences
VLF	spectral power in < 0.04 Hz frequency band
LF	spectral power in 0.04-0.15 Hz frequency band
HF	spectral power in 0.15-0.4 Hz frequency band
normalizedLF	$\frac{LF}{LF + HF}$
normalizedHF	$\frac{HF}{LF + HF}$
LF-HF ratio	$\frac{LF}{HF}$

Then, once the entire signal has been processed, nine arrays are obtained (one for each extracted feature), with a length equal to the number of 30-second epochs in the examined signal. On each of these, six statistical parameters were calculated: mean, standard deviation, 25th percentile, 75th percentile, kurtosis, skewness.

To summarise, six statistical parameters were calculated on nine arrays obtained from the analysis of five different concatenations, resulting in 270 features from the ECG signal.

III.2 EMG features

In the introduction to this thesis, the importance of the presence of RSWA for the correct diagnosis of RBD was explained. RSWA can be defined either as the sustained (tonic) loss of normal muscle atonia, as well as the intermittent (phasic) increase in EMG activity during REM sleep.

For the same reasons as described in the chapter on ECG feature extraction, it was also chosen for the EMG not to work on the signal in its entirety, but to create five different subsets. These, as before, are the result of the concatenation of the sleep stages considered most indicative:

- concatenation of N2 stages
- concatenation of N3 stages
- concatenation of REM stages
- concatenation of N3 + REM stages
- concatenation of N2 + N3 + REM stages

Among the several assessment methods of muscle atonia in REM sleep, developed over the years, in this work, to extract features from the EMG signal, it was chosen to follow the guidelines provided by the SINBAR group [24] by integrating them with an automatic RSWA scoring algorithm called REM Sleep Atonia Index (RAI) [25].

The method developed by the SINBAR group examines each 30-second epoch of EMG signal and defines the muscle activity as 'tonic' if it is above a threshold value for at least 50% of the epoch duration. 'phasic' EMG activity was scored into 3-second mini-epochs and was defined as any burst of EMG activity lasting 0.1 to 5 seconds with amplitude exceeding twice the background EMG activity.

The threshold referred to in the SINBAR method is defined as twice the background EMG muscle tone or 10μV. In the case of the algorithm developed in this work, it was set by calculating the 75th percentile of the entire EMG signal in order to have the most accurate estimate of background EMG activity. Three features are derived from this, namely the percentage of 'tonic' epochs within the track, the percentage of 'phasic' mini-epochs and the percentage of 30-second epochs containing at least five 'phasic' classified mini-epochs.

Turning to frequency analysis, the Spectral Edge Frequency at 50% and 95% was calculated on each 1-second epoch referred to as SEF50 and SEF95 respectively. These

show the frequency value below which 50 and 95 per cent of the signal's spectral power lies, respectively. Given this, SEF50 is nothing more than the median frequency of the analysed signal.

Following the same procedure used in the ECG features extraction, of the vectors obtained (having a length equal to the number of 1-second epochs present in the examined concatenation) the six statistical parameters were calculated: mean, standard deviation, 25th percentile, 75th percentile, kurtosis, skewness.

Finally, on the first three listed concatenations, those formed by the succession of epochs of the same stage, the REM Sleep Atonia Index (RAI) was assessed. This metric provides an index regarding the level of atonia during the REM stage. In the study, this concept was extended not only to the REM stage, but also to the N2 and N3 stages. Together with the RAI, p1, p2 and p3 were calculated, namely the percentage of epochs with an average amplitude below 1µV, between 1 and 2 µV and above 2µV, respectively.

IV. Feature Selection

To verify the actual ability of the features extracted from the ECG to reach a correct automatic identification of the RBD patients, the supervised learning algorithms were trained and tested on 3 different feature sets. The first consisted of the 87 features extracted from the EMG signal, the second of the 270 features extracted from the ECG and the third given by the combination of the two previous ones (357 features in total).

Before the tuning of feature selection and classification models parameters could proceed, it was decided to apply a z-score normalization to all extracted features, so that they would all have the same weight during subsequent steps.

In order to choose the best combination of feature selection algorithm and classification model, in this work the performance of three of the former and six of the latter was evaluated.

With regard to feature selection algorithms, the aim here is to find the smallest possible set of parameters that has the highest predictive power, which is called the

minimal-optimal approach. A brief description of the chosen feature selection algorithms follows:

- **ANOVA F-test**: Analysis of variance (ANOVA) can determine whether the means of three or more groups are different. ANOVA uses F-tests to statistically test the equality of means.
- **Mutual information**: is calculated between two variables and measures the reduction in uncertainty for one variable given a known value of the other variable.
- **Principal Component Analysis (PCA)**: uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components.

V. Classification

In this study, the classification models had the task of performing a binary classification between healthy and RBD patients. For this purpose, the performance of six models was assessed:

- **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).
- **SVM**: it aims at finding a hyperplane in an N-dimensional space (with N equal to the number of features) that distinctly classifies the data points, seeking to maximise the distance between the data points (support vectors) and the hyperplane.
- **NB**: it is a probabilistic classifier based on the Naive-Bayes theorem.
- **DT**: it is a non-parametric supervised learning method used for classification and regression. It wants to attain prediction of the target variable by learning simple decision rules inferred from the data features.
- **BAG**: it is an ensemble meta-estimator that fits base classifiers each on random subsets of the original dataset and then aggregate their individual predictions (either by voting or by averaging) to form a final prediction.
- **XGB**: is a decision tree-based ensemble learning algorithm like random forest, for classification and regression. It improves a single weak model by combining

it with several other weak models in order to generate a collectively strong model.

- **RF**: it fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting.

For each combination of a feature selection algorithm with a classification model, tuning of the hyper-parameters was carried out. This step is crucial as it tests all possible configurations looking for the one that outputs the best F1-score, a metric preferred over accuracy due to the slightly unbalanced dataset.

Finally, two different cross validation algorithms were evaluated. The first, more traditional, is called 'Repeated k-Fold Cross-Validation' and aims to improve the prediction performance of the classification model by repeating the cross-validation procedure several times. In fact, since the dataset is not particularly large, different splits of the data may result in very different results. The second takes even more account of the aspect just described, in fact the training set will consist of all but one patient who will make up the test set. Training the model many times by varying the patient that composes the test set allows a more accurate estimate of the actual performance of the model.

RESULTS

I. Classification Performance

The study compared the predictive capabilities of the models trained on three different datasets, the first consisting of the features extracted from the EMG, the second consisting of those extracted from the ECG and the third given by the union of the previous two. The number of trained models for each dataset corresponds to the total number of possible combinations of feature selection algorithms and classification models, resulting in 18 models for each dataset. In addition, the two cross-validation algorithms described in the previous chapter were evaluated, thus doubling the number of trainings performed.
In the case of the models in which the Repeated k-fold Cross-Validation algorithm was applied, a hold-out approach was also adopted, and 20% of the dataset was used to evaluate the performance, thus composing the test set. With regard to the Leave One Out cross-validation, the entire dataset was used for model training, so in order to assess its performance, a confusion matrix was constructed by evaluating the prediction of the algorithm at every single iteration.

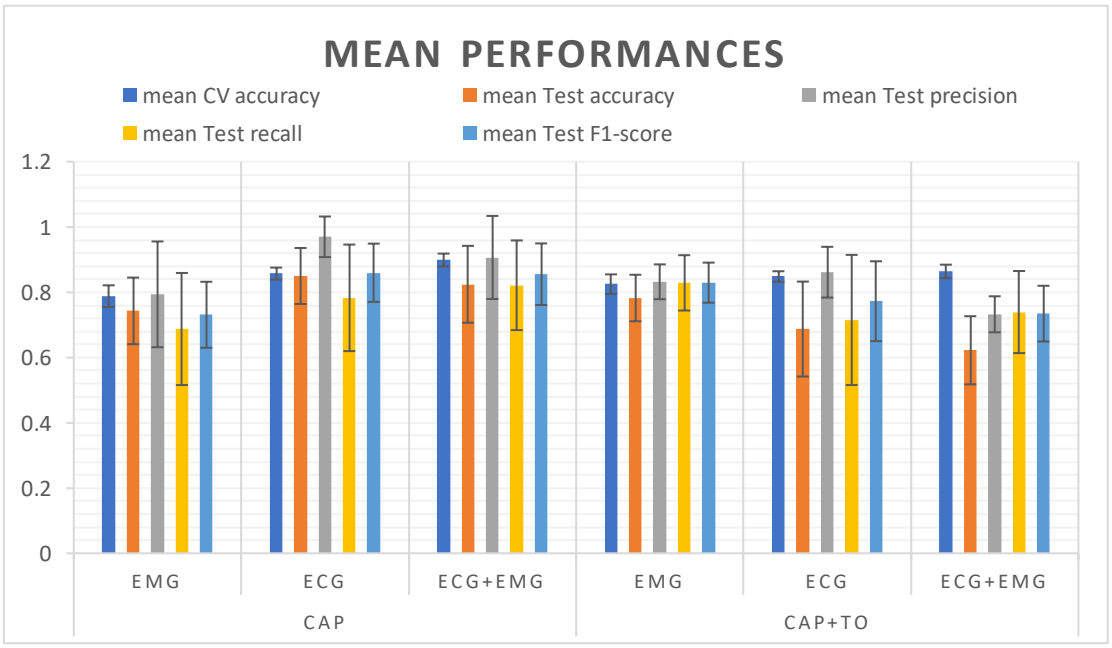

Figure 2 *Graph of performance averaged over all trained models*

DISCUSSION

Several considerations can be made with respect to the results reported in the previous chapter.

Firstly, by examining the results obtained from the training on the CAP Sleep dataset alone, it can definitely be said that the correlation between HRV and RBD is indeed present. To be able to make this assertion is in itself a very good omen for research in the field of sleep disorders, regardless of the numerical results. Going into more detail about the results that enabled the above statement to be made, one finds an average accuracy, calculated on the 20/80 cross-validation, that is 6.9% better than that obtained from the EMG features (going from 78.8±0.03% to 85.5±0.02%). In addition, the combination of features obtained from the ECG with those obtained from the EMG further improved, giving a mean result of 89.9±0.02%. The same trend can also be seen on the training performed on the dataset formed from the union of the CAP Sleep and the TuSDi dataset. This is reassuring and further confirms what has been said so far. It should, in any case, be noted that with the introduction of the patients examined in the Turin hospital, although the trand remained unchanged, the performance dropped by about 3%. This should not be a cause for concern, as the same procedure (in this case polysomnography) conducted in different clinical structures may introduce variability due to factors external to the pathology being examined.

That said, the performance of the models on patients not included in the training set should be evaluated. This subset of the dataset is called the test set. Here, as expected, there is a small decrease in performance, which nevertheless remains very satisfactory. If only the results from the training on the CAP Sleep dataset are taken into consideration, the expected values are found, in fact, with regard to the average accuracy, they remain consistent with those described above, showing a reduction of approximately 3%: 74.3±0.1% on EMG, 85.0±0.09% on ECG and 82.4±0.12% on EMG+ECG. The same does not happen when training the models on the two datasets combined, in fact in this case there is a significant deterioration in performance. Not

so much for the models trained on the features extracted only from the EMG, which indeed show a slight improvement, obtaining an average accuracy of 78.2±0.07%, but rather for those trained on the HRV features and, consequently, for those trained on the EMG+ECG features. In fact, they obtained a mean accuracy of 68.7±0.15% (16.3% worsening) and 62.2±0.1% (20.2% worsening) respectively.

This effect may be a consequence of a modification made in the dedicated HRV feature extraction algorithm of the TuSDi dataset. In it, seeking to further improve the detection of R peaks within the ECG signal, a threshold was set for which it was ensured that at least 20 R peaks were identified for each 30-second epoch. This approach, at first glance, appeared to be indeed ameliorative, but most likely reduced the specific variability of each signal by flattening its differences. Consequently, it is much more complex for the machine learning algorithm to distinguish a pathological subject from a healthy one.

In any case, going into more detail on the performance of each individual model trained, it can be seen that the one with the best performance of all those tested is the one that applies the 'mutual information' feature selection algorithm followed by the 'SVC' classification model and is trained solely on the HRV features of the dataset made up of the union of the two available datasets (CAP Sleep + TuSDi). After being validated on the test set, it performs as follows:

- Accuracy: 92.9%
- Precision: 91.7%
- Recall: 100%
- F1-score: 95.7%

CONCLUSION AND FUTURE WORK

In conclusion, considering all the results obtained, it can be said that there is indeed a correlation between HRV and RBD and that it can be exploited to distinguish healthy from pathological subjects. A major advantage of leveraging HRV-related features lies in the fact that the sequence of R-R intervals, which can also be measured with a simple wrist accessory equipped with a heart rate reader, is all you need to obtain them. In this way, it might no longer be necessary to perform a traditional polysomnography, an examination that requires spending at least one night in hospital, but it would be possible to instruct the patient to perform the examination in the comfort of his or her own home. This would lead to multiple advantages:

- The examination would more accurately describe the patient's sleep, as it is common experience that sleeping in a bed other than one's own can change the quality of sleep.
- It would be less expensive for public healthcare, which would save on hospital-based examinations.
- It would make possible a screening campaign potentially on the whole population over 50, which could lead to positive results for the early diagnosis, even by decades, of sleep-related neurodegenerative synucleinopathies such as dementia with Lewy bodies, Parkinson's disease.

Note that the algorithm developed in this work uses the classification of sleep stages done manually by a specialist in this field. Obviously, if one ever achieved the possibility of implementing a population-wide screening campaign this would not be possible, but it would be necessary to exploit algorithms for automatic sleep stage detection. Fortunately these have already been developed with also very promising performance and in some cases make use of a single EEG trace [26], potentially not hindering the home course of the examination.

Certainly, a major limitation of this study is the small size of the dataset as well as its strong bias. In fact, it is not representative of the true incidence of the disease in the population, of about 2% if only the population over 60 is taken into account. Consequently, in order to have a more realistic indication of the algorithm's performance, it would first of all be necessary to address these two flaws just described. In addition, the number of trained models should be expanded by opening up a greater variety of parameters to choose from when tuning the hyperparameters, which has not been done because of their high computational cost.

REFERENCES

[1] American Academy of Sleep Medicine. International Classification of Sleep Disorders, 3rd ed.; American Academy of Sleep Medicine: Darien, IL, USA, 2014.

[2] Ngo, H.V.V., Claassen, J., Dresler, M.: Sleep: slow wave activity predicts amyloid-β accumulation. Current Biology 30(22), R1371–R1373 (2020)

[3] Winer J.R. Mander B.A. Kumar S. Reed M. Baker S.L. Jagust W.J. Walker M.P. Sleep disturbance forecasts β-amyloid accumulation across subsequent years. Curr. Biol. 2020; 30: 4291-4298

[4] Michael H. Silber, Erik K. St. Louis, Bradley F. Boeve, Chapter 103 - Rapid Eye Movement Sleep Parasomnias. Principles and Practice of Sleep Medicine (Sixth Edition), Elsevier, 2017, Pages 993-1001

[5] Barone DA, Henchcliffe C. Rapid eye movement sleep behavior disorder and the link to alpha-synucleinopathies. Clin Neurophysiol. 2018 Aug;129(8):1551-1564.

[6] McCarter SJ, St Louis EK, Sandness DJ, Arndt K, Erickson M, Tabatabai G, Boeve BF, Silber MH. Antidepressants Increase REM Sleep Muscle Tone in Patients with and without REM Sleep Behavior Disorder. Sleep. 2015 Jun 01;38(6):907-17.

[7] Gagnon JF, Postuma RB, Montplaisir J. Update on the pharmacology of REM sleep behavior disorder. Neurology. 2006 Sep 12;67(5):742-7.

[8] Knudsen S, Gammeltoft S, Jennum PJ. Rapid eye movement sleep behaviour disorder in patients with narcolepsy is associated with hypocretin-1 deficiency. Brain. 2010 Feb;133(Pt 2):568-79.

[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 2018; 41.

[10] Howell MJ, Schenck CH. Rapid Eye Movement Sleep Behavior Disorder and Neurodegenerative Disease. JAMA Neurol. 2015 Jun;72(6):707-12.

[11] Boeve BF. REM sleep behavior disorder: Updated review of the core features, the REM sleep behavior disorder-neurodegenerative disease association, evolving concepts, controversies, and future directions. Ann N Y Acad Sci 2010; 1184:15.

[12] Ferman TJ, Boeve BF, Smith GE, et al. Inclusion of RBD improves the diagnostic classification of dementia with Lewy bodies. Neurology 2011; 77:875.

[13] Du L, Xu L, Liang T, Wing YK, Ke Y, Yung WH. Progressive Pontine-Medullary Dysfunction Leads to REM Sleep Behavior Disorder Symptoms in a Chronic Model of Parkinson's Disease. Nat Sci Sleep. 2021;13:1723-1736.

[14] Nomura T, Inoue Y, Takigawa H, Nakashima K. Comparison of REM sleep behaviour disorder variables between patients with progressive supranuclear palsy and those with Parkinson's disease. Parkinsonism Relat Disord. 2012 May;18(4):394-6.

[15] Ebben MR, Shahbazi M, Lange DJ, Krieger AC. REM behavior disorder associated with familial amyotrophic lateral sclerosis. Amyotroph Lateral Scler. 2012 Sep;13(5):473-4.

[16] Lo Coco D, Cupidi C, Mattaliano A, Baiamonte V, Realmuto S, Cannizzaro E. REM sleep behavior disorder in a patient with frontotemporal dementia. Neurol Sci. 2012 Apr;33(2):371-3.

[17] Chokroverty S, Bhat S, Rosen D, Farheen A. REM behavior disorder in myotonic dystrophy type 2. Neurology. 2012 Jun 12;78(24):2004.

[18] Fantini ML, Corona A, Clerici S, Ferini-Strambi L. Aggressive dream content without daytime aggressiveness in REM sleep behavior disorder. Neurology 2005;65:1010–15.

[19] Goldberger AL, Amaral LA, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. 2000;

[20] Zitser J, During EH, Chiaro G, Miglis MG. Autonomic impairment as a potential biomarker in idiopathic REM-sleep-behavior disorder. Auton Neurosci. 2019;220: 102553.

[21] Postuma RB, Gagnon JF, Pelletier A, Montplaisir J. Prodromal autonomic symptoms and signs in Parkinson's disease and dementia with Lewy bodies. Mov Disord. 2013;28(5):597–604.

[22] Ferini-Strambi L, Oertel W, Dauvilliers Y, et al. Autonomic symptoms in idiopathic REM behavior disorder: a multicentre case-control study. J Neurol.2014;261(6):1112–1118. Published erratum appears in J Neurol. 2015;262:249–250.

[23] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996) Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation 93(5):1043–1065

[24] Burr RL (2007) Interpretation of normalized spectral heart rate variability indices in sleep research: a critical review. Sleep 30(7):913–919

[25] Frauscher B, Iranzo A, Gaig C, et al.; SINBAR (Sleep Innsbruck Barcelona) Group. Normative EMG values during REM sleep for the diagnosis of REM sleep behavior disorder. Sleep. 2012; 35(6): 835–847.

[26] Ferri R, Rundo F, Manconi M, et al. Improved computation of the atonia index in normal controls and patients with REM sleep behavior disorder. Sleep Med. 2010; 11(9): 947–949.

[27] Rechichi I, Zibetti M, Borzì L, Olmo G, Lopiano L. Single-channel EEG classification of sleep stages based on REM microstructure. Healthc Technol Lett. 2021;8(3):58-65. Published 2021 Apr 20. doi:10.1049/htl2.12007