

Received August 11, 2020, accepted August 24, 2020, date of publication August 26, 2020, date of current version September 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3019734

A Method for Sleep Quality Analysis Based on CNN Ensemble With Implementation in a Portable Wireless Device

FÁBIO MENDONÇA^{1,2}, SHEIKH SHANAWAZ MOSTAFA^{1,2}, (Member, IEEE),
FERNANDO MORGADO-DIAS^{2,3}, (Member, IEEE), GABRIEL JULIÁ-SERDÁ⁴,
AND ANTONIO G. RAVELO-GARCÍA^{2,5}

¹Instituto Superior Técnico, University of Lisbon, 1049-001 Lisbon, Portugal

²Madeira Interactive Technologies Institute (ITI/Larsys/M-ITI), 9020-105 Funchal, Portugal

³Faculty of Exact Sciences and Engineering, University of Madeira, 9000-082 Funchal, Portugal

⁴Pulmonary Medicine Department, Hospital Universitario de Gran Canaria Dr. Negrín, 35010 Las Palmas de Gran Canaria, Spain

⁵Institute for Technological Development and Innovation in Communications, Universidad de Las Palmas de Gran Canaria, 35001 Las Palmas de Gran Canaria, Spain

Corresponding author: Fábio Mendonça (fabio.mendonca@tecnico.ulisboa.pt)


This work was supported in part by the Portuguese Foundation for Science and Technology through the Projeto Estratégico under Grant LA 9—UID/EEA/50009/2019, in part by the Agência Regional para o Desenvolvimento da Investigação Tecnologia e Inovação (ARDITI) through the Scope of the Project under Grant M1420-09-5369-FSE-000001—PhD Studentship, and in part by the Regional Funding from the MITIExcell—Excelencia Internacional de IDT&I NAS TIC, provided by the Regional Government of Madeira, under Project M1420-01-0145-FEDER-000002.

ABSTRACT The quality of sleep can be affected by the occurrence of a sleep related disorder and, among these disorders, obstructive sleep apnea is commonly undiagnosed. Polysomnography is considered to be the gold standard for sleep analysis. However, it is an expensive and labor-intensive exam that is unavailable to a large group of the world population. To address these issues, the main goal of this work was to develop an automatic scoring algorithm to analyze the single-lead electrocardiogram signal, performing a minute-by-minute and an overall estimation of both quality of sleep and obstructive sleep apnea. The method employs a cross-spectral coherence technique which produces a spectrographic image that fed three one-dimensional convolutional neural networks for the classification ensemble. The predicted quality of sleep was based on the electroencephalogram cyclic alternating pattern rate, a sleep stability metric. Two methods were developed to indirectly evaluate this metric, creating two sleep quality predictions that were combined with the sleep apnea diagnosis to achieve the final global sleep quality estimation. It was verified that the quality of sleep of the nineteen tested subjects was correctly identified by the proposed model, advocating the significance of clinical analysis. The model was implemented in a non-invasive and simple to self-assemble device, producing a tool that can estimate the quality of sleep and diagnose the obstructive sleep apnea at the patient's home without requiring the attendance of a specialized technician. Therefore, increasing the accessibility of the population to sleep analysis.

INDEX TERMS 1DCNN, CAP, ECG, OSA, sleep quality.

I. INTRODUCTION

The quality of sleep is one of the most important aspects that can affect physical and mental health since sleep related complaints are the second most usual causes for pursuing medical care, only superseded by the feel of pain [1]. Another relevant factor is the prevalence of poor sleep quality in older adults,

The associate editor coordinating the review of this manuscript and approving it for publication was Zhanpeng Jin .

where it was projected that it affects approximately half of the population [2]. Consistent growth in the prevalence of sleep disturbances and neurodegenerative disorders is expected by considering the significant increase in the world aged population. Thus, it is likely that sleep quality assessment will become a relevant indicator in clinical diagnosis [3].

In most cases, poor quality of sleep is directly connected to the presence of a sleep related disorder. Sleep-related breathing disorders are the most prevalent and, among them,

Obstructive Sleep Apnea (OSA), characterized by a complete or partial obstruction of the upper airway that affects the ventilation during sleep [4], is the most common in the adult population. It was estimated to affect 10% of the 30 to 49 year-old men and 3% of the 30 to 49 year-old women, considering an Apnea-Hypopnea Index (AHI) greater than or equal to 15 events/hour. This prevalence increases with age, and it was projected to affect 17% and 9% of the 50 to 70 year-olds men and women, respectively [5]. However, most cases (estimated to be around 80%) are undiagnosed, frequently due to the shortage of resources to perform the analysis or the lack of knowledge about the disorder [6]. Taking into consideration that the present threshold for the AHI, regarding the OSA diagnosis, was changed to 5 or more events/hour [7], a significant increase in the number of undiagnosed cases is expected. Polysomnography (PSG) is considered the gold standard for sleep analysis and records multiple sensors, including the electroencephalogram (EEG) and electrocardiogram (ECG) [8], [9]. This information is then analyzed by a specialist to perform multiple clinical diagnoses [10]. However, the exam is expensive [11], labor-intensive [12], and is unavailable to a large group of the world population with a usual long waiting time [13]. Therefore, a non-invasive and easy to use Home Monitoring Device (HMD) capable of performing the estimation of both sleep quality and OSA could allow to overcome these disadvantages and became a significant tool for future clinical diagnosis.

The EEG signals are used as a reference to define the sleep structure that is conventionally divided into the macrostructure, defined by repetitive variations of Rapid Eye Movement (REM) and Non-REM (NREM) epochs that are scored every 30 second [14], and the microstructure (characterized by phasic and transient events, in the brain electrical activity, that are scored by 1 second epochs [15]). Most of the available HMD that provide an estimation of the sleep quality are based on the assessment of sleep duration metrics (such as total sleep time) that are related to the macrostructure. Among the available devices, actigraphs are the most common in the consumer market, possibly due to the fact that these devices are easy to use [3]. However, the validity of the predictions made by these devices still requires a systematic examination [16].

It was also verified that duration, continuity and intensity metrics have a minor correlation with the patient's subjective ratings of prior-night sleep quality [17], possibly indicating that sleep stability metrics could be more significant for medical diagnosis [18]. Among these stability metrics, the ones related to the analysis of the EEG Cyclic Alternating Pattern (CAP) present the best correlation with the subjective prior-night sleep quality appraisal [3]. A CAP cycle is composed of an activation period (named A phase) that is followed by a quiescent phase (known as B phase). The A phase comprises multiple microstructure patterns, and both A and B phases should last between 2 and 60 seconds to be considered valid CAP phases [15].

The ratio of the total duration of the CAP cycles to the total duration of the NREM sleep is known as the CAP rate. It is

a stability measure that reflects the mechanisms that regulate the arousals during sleep and it is an indicator of the quality of sleep of the total sleep period since the occurrence of disturbances during sleep lead to a higher CAP rate [19]. However, this metric is defined with the analysis of the EEG signals that are measured by a sensor that is difficult for self-assembly. This issue led to a major difficulty for HMD to employ the CAP rate as a metric to define the sleep quality. Nevertheless, the concept of CAP can be extended in a broader context, so it could be assessed using other less complex sensors (fewer channels), such as the single-lead ECG [20]. This indirect approach is followed in this work, with the analysis of the signal measured by a single-lead ECG since this sensor is considerably easier for self-assembly than the EEG sensor. This is particularly significant for a HMD that is intended to be assembled by the user.

The single-lead ECG signal was evaluated to compute a Spectrographic Image (SI), which was examined by both a minute-by-minute and an average based models to attain two sleep quality prediction. Taking into consideration the strong relation between the CAP and the OSA (one of the most prevalent sleep related disorders) [21], a second minute-by-minute classification of the SI was employed to estimate the occurrence of OSA events whose output feeds to a threshold based classifier to perform the diagnosis of this disorder. The global quality of sleep was then assessed by combining the two sleep quality predictions and the OSA diagnosis. Several methods for sleep quality analysis [3] and OSA detection [22]–[24] were previously reviewed. The methods that evaluated the single-lead ECG signal were further examined in the discussion section.

The algorithms were implemented in an HMD, producing an important tool for home sleep analysis, capable of predicting the quality of sleep of the general population, assess the presence of OSA and perform the diagnosis of this disorder, and help in the follow-up treatment for OSA by indicating the number of events and the overall quality of sleep of every night. This tool could also be used for scheduling the PSG analysis by eliminating the need for the possible negative cases (for example, a subject with scarce OSA events and good quality of sleep) to have to perform the PSG exam and prioritize the possible positive cases (for example, a subject with a high number OSA events and poor quality of sleep) for the exam.

Therefore, the objectives of this work were: develop an algorithm to estimate both the presence of OSA and the quality of sleep, using only one sensor that is simple to self-assembly and non-invasive; implement the algorithms in a prototype to produce a low-cost and easy to use HMD. Hence, the sleep test can be performed at the patient's home without the attendance of a specialized technician to help assembly the sensor or monitor the subject. All signals and results of the analysis are stored in files that can later be inspected by a physician for further validation.

A total of four works were found in the literature with reference to the estimation of CAP from ECG signals and

all employed Cardiopulmonary Coupling (CPC) analysis. Thomas *et al.* [20] used tuned thresholds to classify CAP and REM or wake periods, while Mendonça *et al.* [25] fed the CPC signal to a Deep Stacked Autoencoder (DSAE) and tested multiple thresholds to define the CAP minute concept. A feed-forward neural network was also tested as a base comparison with the deep learning algorithm and it was verified that it reaches a performance that was on average 10% lower, supporting the relevance of using a deep learning algorithm for this analysis. An expansion of the work was later presented [26], analyzing two ways of estimating the CPC signal, entropy features, and a causality metric. The features were then ranked by a Minimum Redundance Maximum Relevance (mRMR) procedure and the more relevant were fed to the DSAE for classification. Mendonça *et al.* [27] proposed a tool for time series analysis, named matrix of lags, that evaluated the connection between the Normal-to-Normal sinus interbeat intervals (N-N series) and the ECG Derived Respiration (EDR) by feeding the information regarding the energy of lags to a Support Vector Machine (SVM).

Several methods have been proposed for OSA detection based on multiple source signals [22]. However, only the works based on the ECG signal analysis, reporting both the performance metrics in a minute-by-minute approach and the global accuracy, are significant for the comparison with the results attained in this work.

A linear discriminate discriminant analysis was employed by Ravelo-García *et al.* [28] to examine 15 features (time and frequency based) selected by a proposed feature selection process, for OSA diagnosis. The same classifier was also used by Chazal *et al.* [29] that have examined multiple features from the EDR and Heart Rate Variability (HRV) signals. A combination of Cepstrum features, from the HRV, was tested by Ravelo-García *et al.* [30] to feed a Quadratic Discriminant Analysis (QDA). Cepstrum features were also examined by Martín-González *et al.* [8] with a QDA classification.

A discriminative Hidden Markov Model (HMM) was proposed by Song *et al.* [31] to evaluate the HRV, considering each ECG segment has a hidden state. The distribution of features was subject independent while the transition probabilities, among states, are subject-specific. Rachim *et al.* [32] employed a wavelet decomposition (Debauches 4 wavelet) to obtain statistical features that were fed to a SVM.

Several approaches have been proposed in the state of the art for sleep quality estimation [3]. Wu *et al.* [33] based the analysis in sound events that were clustered by a self-organizing map (using the Kullback-Leibler kernel) and categorized by hierarchical clustering. The quality of sleep was classified as either good or bad by a multinomial HMM with five hidden states. Sathyanarayana *et al.* [34] estimated the sleep efficiency by feeding the actigraphy signal, during sleep, to a Convolutional Neural Network (CNN).

A method for the CAP cycles detection was proposed by Mendonça *et al.* [35], classifying features (Teager energy operator, Shannon entropy, power spectral density in the

theta and beta bands, and autocovariance) produced from one EEG monopolar derivation signal through a Feed-Forward Neural Network (FFNN). A large number of features from the five characteristic EEG bands (62 power spectral density features, one for each electrode, for each band) were analyzed by Wang *et al.* [36], and the mRMR algorithm was used for feature selection. The most relevant features were then fed to a discriminative Graph regularized Extreme Learning Machine (GELM) to assess the quality of sleep by analyzing the total sleep time.

An estimation of the CAP rate was proposed by Mendonça *et al.* [25], analyzing an index of the CPC signal, computed from the EDR and N-N series, to feed two DSAE that predicted the minutes of NREM sleep and CAP. Time and frequency features from the HRV and EDR signals (60 from HRV and 52 from EDR) were used by Bsoul *et al.* [37] to feed a multi-stage SVM for the deep sleep efficiency estimation and assessment of the quality of sleep. Choi *et al.* [38] have developed a set of rules that analyzed nineteen attributes (such as the age and body mass index) to qualify the quality of sleep as either good, normal or bad.

The paper is organized as follows: materials and methods are presented in Section II; the performance of the developed algorithms is assessed in Section III; the developed home monitoring device is presented in Section IV; a discussion of the results is carried out in Section V; the paper is concluded in the final section.

II. MATERIALS AND METHODS

A portable solution that performs a combined assessment of sleep quality and apnea was developed using the algorithm whose block diagram is presented in Fig. 1. The algorithm is capable of predicting the quality of sleep by analyzing the single-lead ECG signal. In addition to that, a minute-by-minute based OSA prediction and a global assessment of this disorder (“OSA-positive” or “OSA-negative”) was also produced by examining the same signal. The developed models were implemented in an HMD.

The algorithm can be interpreted as having two main steps. The first performs the analysis of the preprocessed (resampled and standardized) ECG signal to assess the connection between the variability of the respiratory volume, using the EDR, and the heart rate, through the N-N series, by employing a CPC technique. Specifically, the Cross-Spectral Coherence (CSC) was computed for each epoch, producing a spectrographic measure. The results for all epochs were grouped (in sequence) to create a SI which was subsequently examined in the second step by three One-Dimensional CNN (1D-CNN). CNN was chosen since it was identified as one of the best networks for automatic feature extraction [39], [40]. Afterwards, the classifiers’ outputs were combined to form the final classification. The method tries to replicate the process executed by physicians when they are performing the analysis of a biomedical image, carrying out a visual examination to extract information. In this case, the visual analysis and subsequent evaluation was performed by the CNNs.

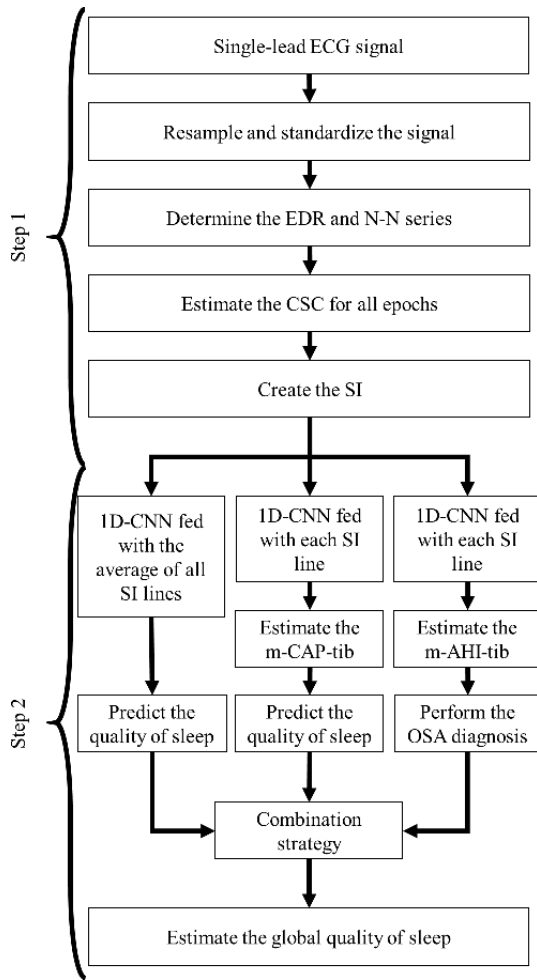


FIGURE 1. Block diagram of the developed algorithm.

A seven minute window, with one minute of displacement between windows, was used to create the epochs. Therefore, the CSC (computed for each epoch) evaluated two signals (EDR and N-N series) with a duration of seven minutes. Yet, the first and last three minutes are overlapping with, respectively, the previous and next epochs while the central minute relates to the label used for the classification. Therefore, the evaluation of the epochs lead to a minute-by-minute analysis. The created SI is a time-frequency matrix representation, where each line (row of the matrix) contained the frequency based information of the epoch and was fed to two classifiers which performed the minute-by-minute assessment of CAP and OSA. The average of all lines was fed to the average based classifier. Therefore, the classifiers are evaluating the spectral information that composed the SI. This image displays the pattern that is created by a full night recording and can also be used as a reference for the specialized physician that is analyzing the results. An example of the SI creation is presented in Fig. 2.

The ratio of the number of minutes classified as CAP (m-CAP) to the time in bed (tib) in minutes is indicated as m-CAP-tib. It is a metric proposed in this work to evaluate the

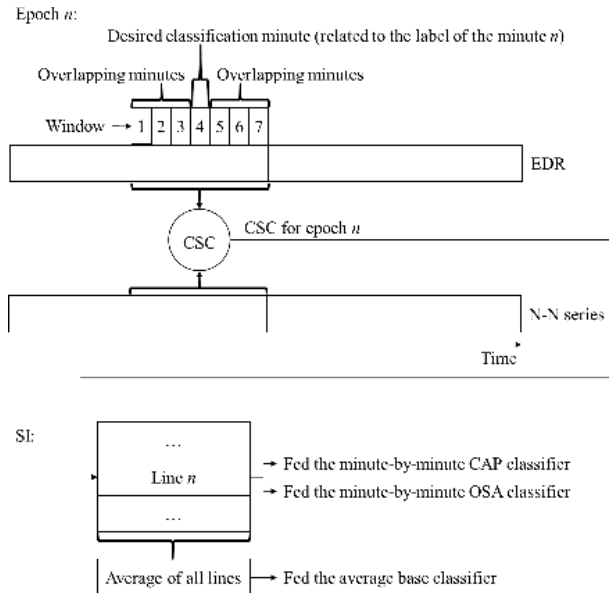


FIGURE 2. Example of the SI creation.

quality of sleep that is related to the CAP rate which does not require the estimation of the sleep macrostructure. After that, this ratio was compared with a threshold, tuned by comparing with the CAP rate (predicted by PSG), to determine the first estimation for the quality of sleep (SQ-m). The total number of minutes classified as an OSA event was then used to perform the OSA diagnosis by comparing the ratio of the number of minutes classified as OSA to the time in bed in minutes (m-AHI-tib) with a threshold. The average of all lines from the SI feeds to the last 1D-CNN to perform the second estimation for the quality of sleep (SQ-ave). The two sleep quality estimations and the OSA diagnosis were combined to create a single output for the assessment of the global quality of sleep (SQ-g). The trained models were employed by the algorithm (developed in Python 3) that was implemented in the HMD. The device is composed of two units, a sensing unit that acquires the signals that are wirelessly sent to a processing unit (second unit) that performs the analysis and generates the results.

A. DATABASES

Nineteen full-night recordings (from eleven men and eight women), were selected from the CAP Sleep Database (CAPSD) from Physionet [15], [41] to train and test the 1D-CNN for the CAP and SQ-ave assessment. This database was also used to test the capability of the models for the sleep quality estimation (SQ-m, SQ-ave and SQ-g). Fifteen of the subjects were free of any neurological disorders, and four have been diagnosed with sleep-disordered breathing. The subjects' average age is 39.95 years old (ranging from 23 and 78 years old), with a normal age-related CAP rate percentage of 55 for four subjects, 32 for five subjects, and 38 for the remaining subjects. The average time in bed was 492.11 minutes and only the single-lead ECG

signals were analyzed, recorded with a sampling frequency that ranged from 128 to 512 Hz.

This database was chosen due to the availability of the sleep macrostructure and CAP phase's annotations, named Database_Label, provided by specialized physicians. These allowed to determine the CAP cycles at every second using the scoring rules presented by Terzano *et al.* [15], which were implemented in a finite state machine. The estimated CAP cycles were stored in a vector, named CAP_cycle, that indicated if each second corresponds either to a CAP (marked as "1") or a non-CAP (marked as "0") cycle. The vector was then reshaped to form a matrix with sixty columns (corresponding to a minute of data) and the number of rows corresponds to the number of minutes. Afterward, the CAP ratio (C_r) of each minute was computed by adding all elements of the columns and divide it by sixty. This ratio was compared with a threshold to define if the minute corresponds to CAP or non-CAP, thus producing the label for the minute-by-minute CAP (CAPm) assessment used for the SQ-m estimation. This threshold was chosen to be 35% since it was indicated as the more suitable for CAP analysis based on the ECG signal [25], [27]. It considers the CAP periods that are longer than 21 s, filtering the short duration cycles (that may not significantly manifest in the ECG signal) but still covering the majority of the events since the average CAP cycle duration is 26.9 ± 4.1 s [42].

The CAP rate for each subject was estimated by analyzing the CAP cycles and the sleep macrostructure (employing the scoring rules defined by Terzano *et al.* [15]). Afterwards, the estimated CAP rate was compared with the CAP rate percentages in healthy subjects [19], [43] (the information of the gender and age of each subject is available in the database) to produce the label for the classifier that examined the average CSC signal, indicating the sleep quality (SQ-ave) as either "1" if the determine CAP rate was lower than the CAP rate percentage in healthy subject for the subject, designating a good sleep quality, or "0" otherwise. These labels were also used as ground truth for the global sleep quality assessment.

The flow diagrams of the algorithm used to generate the labels of the first database are presented in Fig. 3. This analysis is validated by the fact that the CAP rate is characterized by a low night-to-night intra-individual variability in normal subjects [43].

A database recorded by the sleep unit of the Hospital Universitario de Gran Canaria Dr. Negrín (HUGCDN) [44] was used to develop the OSA detection algorithm since both the single-lead ECG signal and a minute-by-minute OSA annotation was provided (a minute-by-minute OSA annotation is not available in the CAPSD). It is composed of seventy suspected OSA patients (nineteen females and fifty-one males) with an age variation between 18 and 82 years old. The recording's length ranged from 230 to 486 minutes, and the respiratory events were annotated every minute by a specialized expert. Forty-six subjects have an AHI of 10 or more with, at least, 70 minutes of OSA, and four recordings have an AHI of 5

or more (the OSA minutes ranged from 5 and 69). The remaining subjects have an AHI lower than 5. The single-lead ECG signals were recorded by a computerized system from VIASYS Healthcare Inc. (Wilmington, MA, USA), digitized at 200 Hz with 16-bit resolution [28]. The respiratory events were scored according to the American Academy of Sleep Medicine criteria [7].

No balancing operation (modify the datasets to obtain an equal number of positive and negative examples) was performed in any dataset since it could change the expected distribution of the data. However, it was verified that the classifier's performance can be improved using cost-sensitive learning (attribute a higher cost to misclassifying a minority class element compared to a majority class element). Therefore, this approach was used since the data distribution is preserved and can significantly improve the performance of the developed models [45].

B. PREPROCESSING

Since the sampling frequency of the single-lead ECG signals, available at the CAPSD, varied between 128 to 512 Hz then, all the records were resampled at 200 Hz (a normally used frequency by ECG sensors [46] that is also employed in the HUGCDN dataset), either by decimation or interpolation [47], providing a uniform database. The signals were then standardized by subtracting the average and dividing the result by the standard deviation. In the developed HMD, both signals were acquired with a sampling frequency of 100 Hz (the device can support either 1, 10, 100 or 1000 Hz). Therefore, the ECG signal measured by the developed HMD was resampled to 200 Hz by interpolation, allowing the employment of the trained models.

For the minute-by-minute classifications, a seven minute window, with one minute displacement between adjacent windows (first and last three minutes overlap and the central minute corresponds to the database label), was used to create the epochs since it was previously identified to be the more suitable window for CAP analysis based on the ECG signal [25], [27].

As recommended by the task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, the spectral analysis of the ECG signal should preferably use a signal with a minimum duration of five minutes [48]. This condition is in line with the seven minute window (with six minutes overlapping) that was used. It is also relevant that the OSA event duration can range from 10.0 to 62.4 seconds [49]; thus, the one minute annotation can cover both the longer and short duration events. However, shorter windows (as for example the evaluation based on 15 seconds [50]) can also be used for OSA evaluation.

C. FEATURE CONSTRUCTION

The QRS complex was attained from the ECG signal using the algorithm developed by Pan and Tompkins [46]. The R-peaks were then used to estimate the interbeat intervals. Atypical peaks were corrected by replacing the peak by the

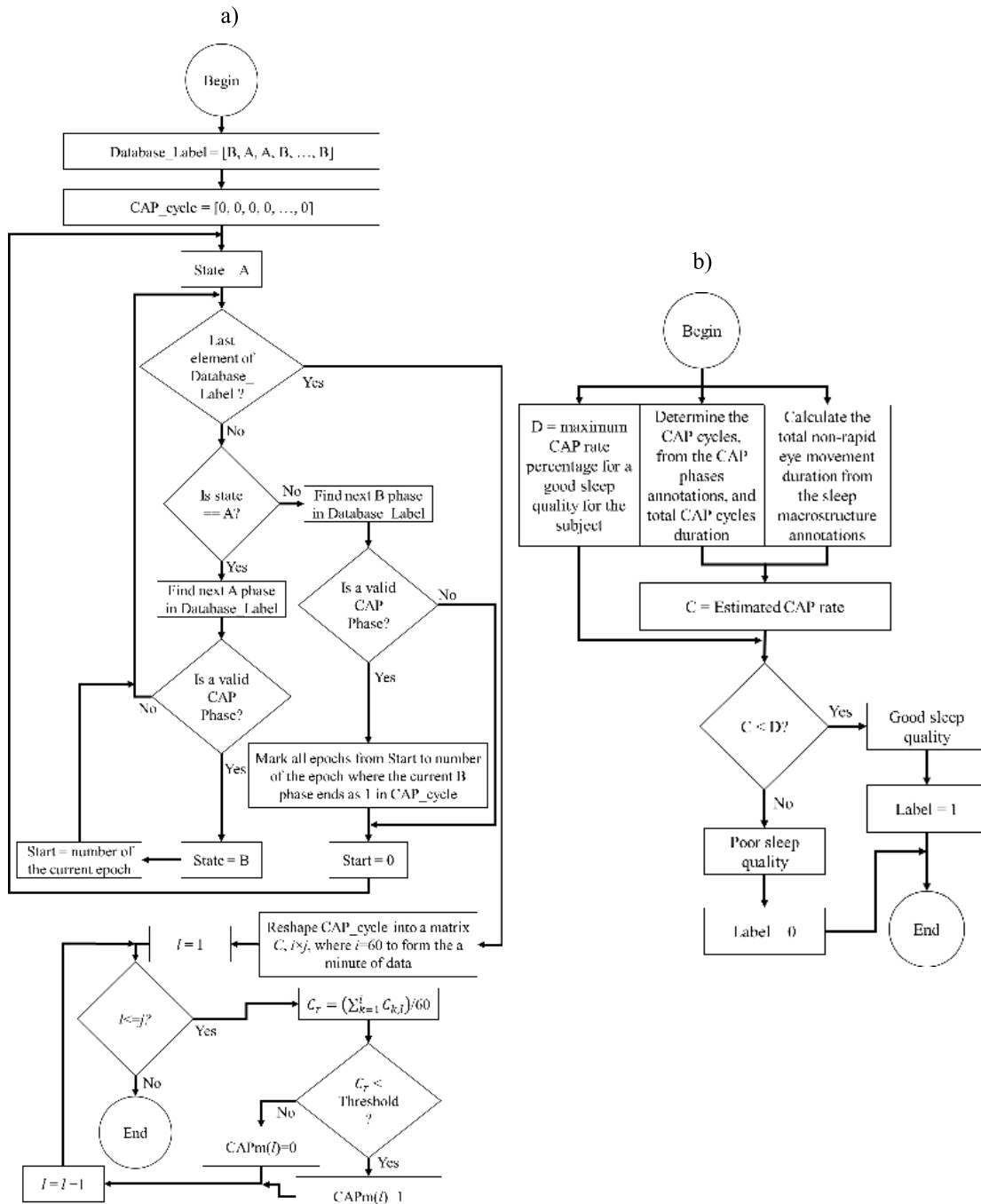


FIGURE 3. Flow diagram of the algorithms used to generate the labels, from the CAPSD, for a) the minute-by-minute CAP assessment (used for the SQ-m estimation) and for b) the classifier that examined the average of the SI lines (used for the SQ-ave assessment).

average of the previous and next peak, thus producing the N-N series.

Since the respiratory cycle modulates the QRS morphology, the EDR signal can be produced by analyzing the modulation [51]. This analysis was performed by employing the algorithm developed by Arunachalam and Brown [52] that estimates the respiratory amplitude modulation factor as the ratio of the current R-peak amplitude and a running

average of the R-peaks amplitudes. Subsequently, a cubic spline interpolation was applied to produce a continuous estimation of the EDR signal. An elliptic filter was used at the end to smooth the signal and reduce the high frequency noise.

The connection between the EDR and N-N series was assessed by the CSC, that considers the cross spectral power and coherence of the signals [25]. The correlation between

the N-N series and EDR was evaluated by the cross-spectrum between the discrete Fourier transform components of the signals, respectively, N and E , that was computed by [53]

$$\Gamma_f(N, E) = \alpha_{N,f} \alpha_{E,f} e^{j(\Phi_{E,f} - \Phi_{N,f})} \quad (1)$$

where Φ and α are, respectively, the phase and amplitude of the Fourier components. The consistency of the phase difference, between the signals, was determined by the Magnitude Squared Coherence (MSC), defined as [53]

$$H_f(N, E) = \frac{C_f(N, E)^2}{(A_{N,f} e^{j\Phi_{N,f}})^2 (A_{E,f} e^{j\Phi_{E,f}})^2} \quad (2)$$

The Welch averaged periodogram method was then applied to estimate the cross-correlation matrices [54]. Finally, the MSC was multiplied by the square of the cross spectral power to estimate the CSC of the epoch [20].

This procedure was applied to all epochs, producing the lines of the SI (one line for each epoch), which were fed to the minute-by-minute classifiers (for the m-CAP-tib and m-AHI-tib estimation). Thus, the classifiers are evaluating the CSC signal of the epoch. The average of all lines (average of all CSC signals) of the SI feeds the 1D-CNN that estimates the SQ-ave.

D. CLASSIFICATION

Three 1D-CNN were combined to form the classification ensemble. The 1D-CNN employ convolution kernels to implement a transformation of the inputs, allowing to detect patterns, and it reduces the redundancy through pooling processes [55]. The convolution operation, executed in the convolution layers, can be defined as [56]

$$c_d = \varphi(K_d \otimes X + B_d) \quad (3)$$

where φ is the activation function, selected to be the Rectified Linear Unit (ReLU) [57], $1 \leq d \leq nK_d$, given that nK_d is the number of convolution kernels, n is the dimension of the input, K is the kernel and \otimes is the n dimensional convolution operation. X are the layer inputs and B is the bias vector. A batch normalization layer was used after the activations to maintain the mean activation close to zero with a nearly unitary standard deviation. This layer allows to decrease the network's initialization sensitivity and increases the training speed. The maximum pooling operation was performed after the convolution and normalization layers to reduce the dimensionality of the data [56].

Fully connected (dense) layers were used at the end of the network to improve the learning capability of the nonlinear parameters and perform the classification by [56]

$$O = \varphi(W \times X + B) \quad (4)$$

where W is the weights matrix and φ is the activation function, selected to be the ReLU in the first dense layer and the softmax function [57] in the output layer, providing a final probabilistic classification.

The 1D-CNN was composed of a sequence of an input layer, groups of layers and the classification layers (chosen to be fully connected layers with the final classification performed by the Softmax function). A grid search method was used to select the hyperparameters of the convolution and pooling layers. In each iteration the searching algorithm decided if another group should be added. Each group was composed of a sequence of a convolution layer that was followed by a batch normalization layer which, in turn, was followed by the activation and pooling layers. A dropout of 20% was used after each group of layers to reduce the possibility of overfitting.

Two 1D-CNNs performed the minute-by-minute classification of the SI lines. The outputs of the classifiers were then used to compute the ratio of minutes classified (either as CAP or OSA) to the time in bed in minutes (m-CAP-tib or m-AHI-tib). Afterwards, the m-CAP-tib was compared with a threshold (tuned during the training) to determine the SQ-m. The m-AHI-tib ratio was then compared with a threshold (in this case 8% was used since it is correlated to the ratio of 5/60 \approx 0.083 that is given by an AHI of 5, the minimum value to diagnose OSA, in 60 minutes) to perform the OSA diagnosis. This approach was verified to be highly correlated with the AHI obtained with polygraphy [28], [58], [59] thus, validating the application. Hence, the output of the classifier (minute-by-minute quantification as either normal respiration or OSA) allows the production of a global score regarding the presence of clinically significant OSA that is equivalent to an AHI greater than or equal to 5. If the subject was diagnosed with OSA then, the OSA classifier' output was interpreted as "vote bad", otherwise it was inferred as "vote good", in the combination procedure.

The average of all SI lines was fed to the third 1D-CNN classifier to estimate the SQ-ave. The final global quality of sleep, SQ-g, was determined by considering a majority voting strategy (the output of each classifier was considered as a vote and the system chooses the output class with more votes) to perform the classifiers ensemble. Since tree classifiers were considered and all performed a binary classification, thus, SQ-g was given by the class that was chosen by either two or three classifiers. The complete process of feature creation, classification and sleep quality assessment is presented in Fig. 4.

E. PERFORMANCE METRICS

The performance assessment was based on the calculation of the True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) to estimate the average accuracy (Acc), sensitivity (Sen), and specificity (Spe) defined by [60]

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

$$Sen = \frac{TP}{TP + FN} \quad (6)$$

$$Spe = \frac{TN}{FP + TN} \quad (7)$$

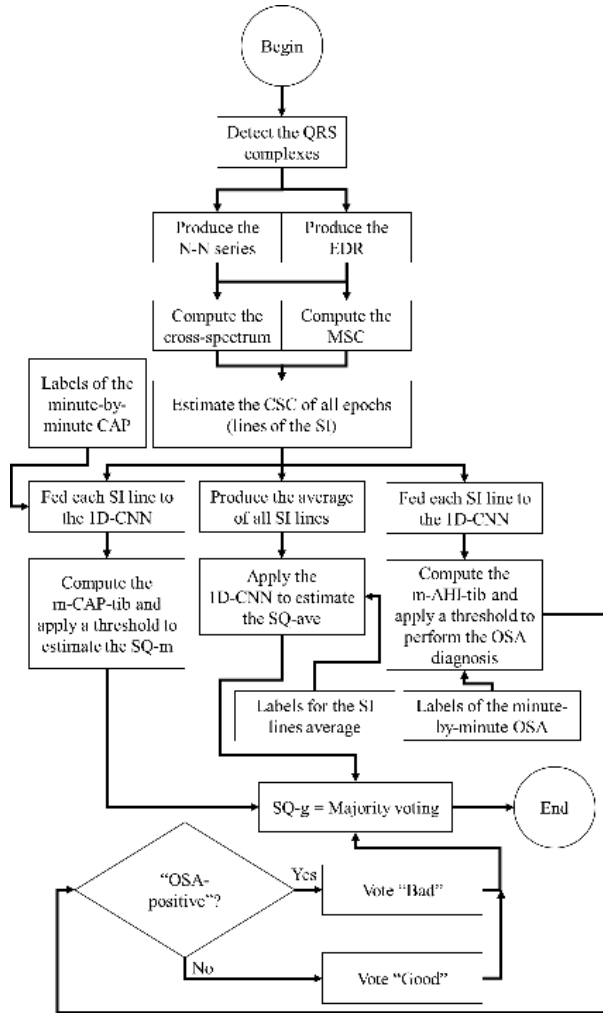


FIGURE 4. Flow diagram of the algorithm for sleep quality estimation.

For the minute-by-minute classification the diagnostic capacity of the classifier was calculated by the Area Under the receiver operating characteristic Curve (AUC) as it designates how likely the classifier is to rank a randomly selected positive instance higher than a randomly selected negative instance [61].

The average global accuracy (Acc-G) was considered as the performance metric to evaluate the sleep quality estimations and the OSA diagnosis. In addition to that, the sensitivity and specificity of the global classification (respectively, Sen-G and Spe-G) were also evaluated. Cohen’s kappa coefficient (k) was calculated for the global analysis to measure the agreement between the expert and the proposed method [62]. The statistical significance of the results was assessed by calculating the average value and the 95% confidence interval [63].

III. RESULTS

An example of the SI created by the used CPC technique is shown in Fig. 5. By analyzing the Fig. 5 c), it is possible

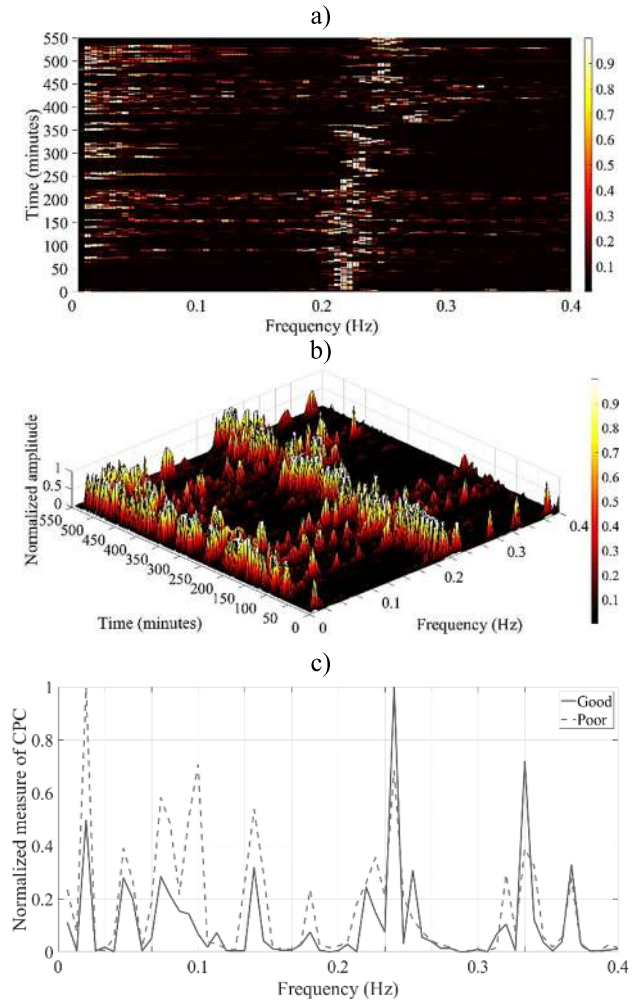


FIGURE 5. Example of a SI for the minute-by-minute analysis (each minute corresponds to a line of the image), a) top view and b) side view, and c) for the average examination (average of all lines) with good and poor sleep quality.

to verify that the subject with poor sleep quality has the highest power peaks in the very low (0-0.01 Hz) and low (0.01-0.1 Hz) frequency bands while the subject with good sleep quality has a significant amount of power in the high (0.1-0.4 Hz) frequency band. This information is in agreement with the findings reported in the state of the art where it was verified that the power in the high frequency is associated with physiologic respiratory sinus arrhythmia, deep sleep, and absence of CAP periods while the power in the very low and low frequency bands was associated with wake or REM periods, the presence of CAP (suggesting instability in sleep) and the occurrence of OSA or sleep fragmentation [20], [64]. By examining Fig. 5 a) and b) it is possible to verify the occurrence of good (as an example, from 1 s to 50 s) and poor (as an example, from 400 s to 450 s) sleep quality periods.

The 1D-CNN structure and hyperparameters were selected by a grid search method. A group of layers was added to all classifiers if the Acc-G increased. Otherwise, the previous

network configuration was chosen. Therefore, all classifiers have the same number of layers, significantly reducing the simulation time that would be required to test all possible combinations. The number of filters used by the first convolution layer was selected to be a power of two (for optimization), varying from 8 to 512, and the filter length was varied from 1 to 10. The subsequent convolution layers were chosen to have twice the number of filters used in the previous layer with the same length. The number of channels of the batch normalization layer was chosen to be the same as the number of filters used on the convolution layer and the pool size of the pooling layers was selected to be 2. A stride of one was used for the convolution layers and a stride of two for the pooling layers. The ReLU was selected as the activation function.

The network’s error optimization was implemented with the Adam algorithm [65], performing fifty runs in each iteration to attain statistically significant results. A 2-fold cross-validation scheme [57] was employed to find the optimal layer parameters for the minute-by-minute classifiers (this scheme is reasonably fast and was used due to the large number of simulations that were performed in the grid search). The training set was composed of data from half of the subjects (from the CAPSD for CAP analysis and from the HUGCDN for the OSA evaluation) while the testing set was composed of data from the remaining subjects. The subjects that composed the sets were randomly chosen at each iteration. Subject independence was assured by only using the data from a subject either on the training set or on the testing set. The performance was assessed by averaging the results of all iterations and the layer parameters that attained the best AUC were chosen.

The one hold out cross validation scheme [57] was used to find the optimal layer parameters for the classifier that evaluates the average CSC signal because it is a global classification, training with data from eighteen subjects of the CAPSD and testing with the data from the leftover subject. The process was repeated nineteen times and, each time, a different subject was selected to form the testing set. The layer parameters that attained the best Acc-G were chosen. After the optimal layer parameters for all classifiers were found, the performance of the tuned classifiers was assessed by the one hold out cross validation scheme.

It was verified that the best Acc-G was achieved using two groups in the hidden layer. The chosen layer parameters for each classifier are presented in Table 1 and an example of how the SI was evaluated by the minute-by-minute CAP classifier is presented in Fig. 6.

The performance of the algorithms for the minute-by-minute assessment is presented in Table 2. The minute-by-minute CAP estimation is in the range of the mutual agreement among physicians examining the same EEG signals (69% to 77.5%. [66]). This result is highly relevant since the developed method is based on the analysis of the ECG signal. Thus, it is an indirect estimation that achieved a result as good as a specialist physician examining the EEG signal. It is also important to take into consideration that the

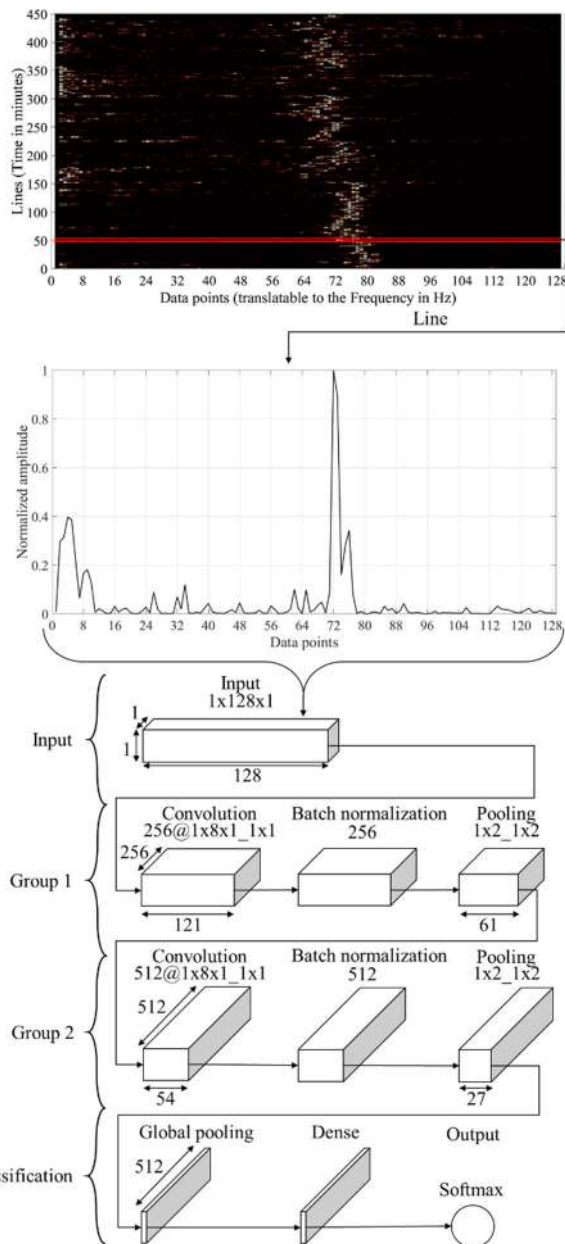


FIGURE 6. Example of the line examination by the minute-by-minute classifier for CAP assessment.

agreement gets closer to the lower bound (69%) as the number of physicians that examine the signals increases as verified by Largo et al. [67] where the overall average of the pairwise inter-scorer agreement of seven physicians was 69.9%. These observations are further substantiated by the results reported by Mendez et al. [68], where it was estimated that exist 25% of ambiguity and subjectivity in the manual classification of the CAP phases.

The results for the OSA detection are in the range of the best methods reported in the state of the art, that performed the analysis based on the ECG signal, where the Acc, Sen and Spe range, respectively, from 76% to 100%, 70% to 92% and 59% to 100% [22].

TABLE 1. Layer parameters for each classifier.

Layer	Classifier		
	Minute-by-minute – CAP	Average examination	Minute-by-minute – OSA
Input ¹	1x128x1	1x128x1	1x128x1
Convolution ²	256@1x8x1_1x1	64@1x4x1_1x1	128@1x5x1_1x1
Batch normalization ³	256	64	128
Pooling ⁴	1x2_1x2	1x2_1x2	1x2_1x2
Convolution ²	512@1x8x1_1x1	128@1x4x1_1x1	256@1x5x1_1x1
Batch normalization ³	512	128	256
Pooling ⁴	1x2_1x2	1x2_1x2	1x2_1x2
Global pooling	512	128	256
Dense	512	128	256
Dense	2	2	2

¹vertical width x horizontal width x number of channels
²number of filters @ vertical width of the filters x horizontal width of the filters x number of channels _ vertical stride x horizontal stride
³number of channels
⁴horizontal width of the filters x number of channels _ vertical stride x horizontal stride

TABLE 2. Performance of the minute-by-minute classifications.

Classification	Dataset	Performance metrics (average value and 95% confidence interval)			
		Acc (%)	Sen (%)	Spe (%)	AUC
CAP	CAPSD	70.5	73.2	69.8	0.77
		(69.9–71.1)	(71.7–74.6)	(69.3–70.4)	(0.76–0.79)
		76.4	72.7	77.2	0.83
OSA	HUGCDN	(75.2–76.8)	(70.9–73.0)	75.7–78.6	(0.81–0.84)

The regression plot of the predicted m-CAP-tib and the CAP rate obtained by PSG is presented in Fig. 7. The regression R² was 0.88, supporting the relevance of the proposed metric. Fig. 8 presents the regression plot of the predicted AHI (m-AHI-tib defined as the number of minutes with events per hour of time in bed [28]) and the AHI obtained by PSG. The regression R² was 0.79, further advocating the validity of the method for OSA diagnosis.

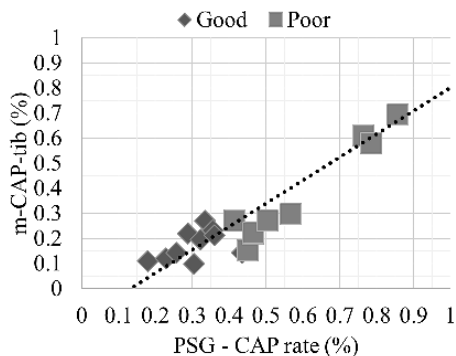


FIGURE 7. Regression plot of the predicted m-CAP-tib and the CAP rate obtained by PSG (CAPSD), using the labels for the global sleep quality assessment to dictate the sleep quality as either “Good” or “Poor”.

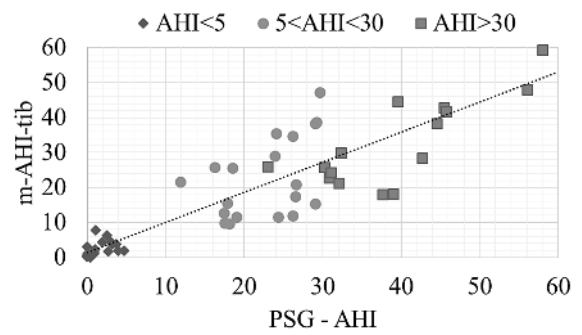


FIGURE 8. Regression plot of the predicted m-AHI-tib and the AHI obtained by PSG (HUGCDN).

TABLE 3. Average performance of the global algorithms.

Classification	SQ-ave	SQ-m	OSA diagnosis	SQ-g
Acc-G (%)	73.7	84.2	94.3	100.0
Sen-G (%)	70.0	80.0	96.0	100.0
Spe-G (%)	77.8	88.9	90.0	100.0
k (%)	47.5	68.5	86.0	100.0

The performance attained by the global classification algorithms is presented in Table 3. The accuracy for the OSA diagnosis is in the range of the best methods reported in the state of the art, where the Acc-G ranges from 72% to 100% [22], with an almost perfect agreement according to the k value. The classifier for OSA detection was trained and tested in the HUGCDN dataset and the concept of cross database analysis (use a classifier that was trained in a dataset to perform predictions in a new dataset) was employed to perform the analysis in the CAPSD (used for the SQ-g estimation). The model based on the average CSC metric achieved the lowest performance for the sleep quality estimation, with a moderate agreement according to the k value, while the combined approach for the global sleep quality estimation attained the highest possible performance, improving the results that were reached by applying a threshold to the m-CAP-tib estimation (it was verified that 0.22 was the best threshold for the developed model, optimized by performing multiple tests with different values and choosing the model that attained the highest Acc-G). The OSA diagnosis allowed to correctly predict the quality of sleep for the subjects where the estimates from SQ-m and SQ-ave were different, allowing to correctly classify the sleep quality of all analyzed subjects from the CAPSD.

IV. DEVELOPMENT OF THE HOME MONITORING DEVICE

A non-invasive HMD that could detect the presence of OSA and predict the quality of sleep, by analyzing the signals measured by the sensing unit, was developed. The device is composed of two units that wirelessly communicate via Bluetooth. The sensing unit was developed to be easily self-assembled and is responsible for collecting the sensors signals and send the data to the processing unit that is a

single-board computer where an application stores the information and performs the analysis. The employed hardware is presented in Fig. 9.

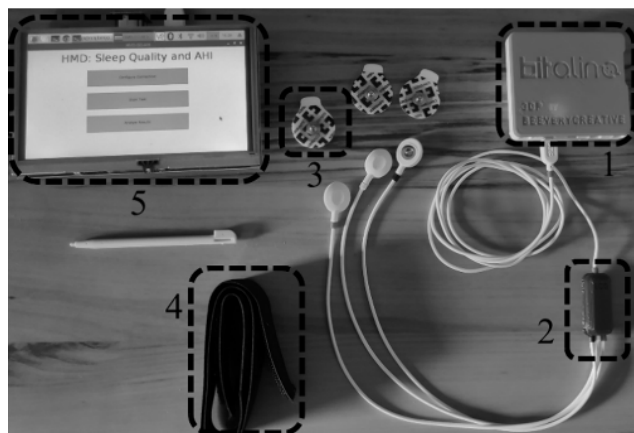


FIGURE 9. Hardware employed on the HMD. 1- Sensing unit, 2- ECG sensor, 3- Electrode, 4- Armband, 5- Processing unit.

The sensing unit was implemented by using the BITalino Core BT [69] that is composed of a microcontroller (ATmega328P), a power management module fed by a 3.7 V lithium ion battery, with an average 50 mAh load current (lasting at least seventeen hours in real-time acquisition over Bluetooth [69]), and a communication module for Bluetooth communications. The device sensing rate can be configurable by the user in the processing unit (the device can support either 1, 10, 100 or 1000 Hz, however, the default value of 100 Hz was used in this work since the measurement had less noise related artifacts than the signal measured at 1000 Hz) and the resolution of the signal is either six or ten bits, depending upon the Analog to Digital Conversion (ADC) port (only the ten bits ports were used in this work).

The ECG sensor measures the electrical potentials (through the electrodes) in the chest with respect to a ground reference and it was verified, when comparing with the output of the BioPac MP35 Student Lab Pro (an established gold standard device), that the average measurement root mean squared error was 0.049 ± 0.016 [70]. The small error advocates the practicality of the sensor for clinical diagnosis.

The processing unit is composed of a Raspberry Pi 3 B+ with a 1.4 GHz, 64-bit, ARM quad-core processor that feeds by the DC power supply, and a touch screen that displays the Graphical User Interface (GUI), allowing the user to configure the sensing unit (the two units automatically connect once the application is opened), start (and stop) the examination and produces the analysis of the signals.

By using the GUI, the user can configure the connection with the sensing unit by choosing the bit rate (the default value is 19200 bits/s) and can also change the ADCs and sampling rate that will be used. However, for the normal examination the user does not need to change any of the default configurations and the procedure can be summarized in the following steps: 1- remove the electrodes plastic cover;

2- place the electrodes on the Einthoven triangle configuration to create a single-lead ECG signal [71]; 3- tight the armband around the arm; 4- attach the sensing unit to the armband; 5- turn on the sensing unit; 6- turn on the processing unit and wait until the GUI is open; 7- press the “Start Test” button (a new window will be displayed with the “Stop Test” button and the sensing unit will start transmitting the data to the processing unit which, in turn, will store all the information in a text file with a timestamp); 8- press the “Stop Test” button when the test is finished (the communication between the units is ended); 9- press the “Analyze Results” button (the application reads the stored data, uses the developed algorithms to detect the minutes with OSA, the minutes with the occurrence of CAP, diagnose the occurrence of the sleep disorder and estimate the quality of sleep) and wait until the a message indicating that the analysis is over (the results are stored in a new text file with a timestamp for each OSA and CAP detected); 10- the user can either analyze the text files to verify the signals and results of the test or deliver the HMD to an expert to verify the files. An example of a signal stored in a text file, measured by the device when the subject was sleeping, is presented in Fig. 10.

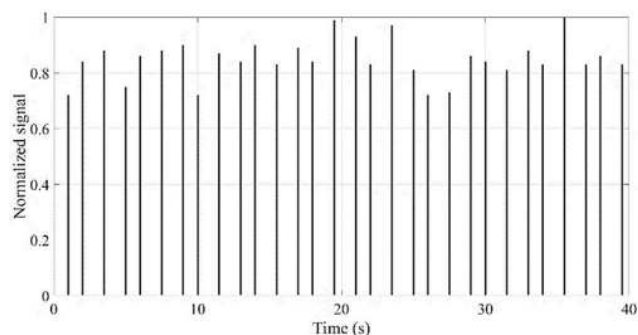


FIGURE 10. Example of a signal stored in the text file. The vertical lines are the amplitude and location of the R peaks of the ECG signal.

The sensing unit cost was 115 € (75 € for the BITalino Core and 40 € for the ECG sensor) while the processing unit cost was 60 €. Nevertheless, the total cost of a potential commercial product could be considerably reduced by developing the hardware of the sensing unit. However, the used module was already validated while this approach of developing a new unit will require a validation by performing a parallel recording with the PSG. A review of validated commercial out-of-hospital ECG devices was performed by Bansal and Joshi [72]. Among the reviewed devices, only five employed a single-lead measurement and the price ranged from 63 € to 1255 €. Hence, the cost of the developed device is in the lowest bond of the cost for validated commercial solution, supporting the choices of this work for a low cost HMD.

V. DISCUSSION

A summary of the reported results from the works that performed CAP analysis from the ECG signals is presented in Table 4. By analyzing the table it is possible to conclude

TABLE 4. Comparison of the proposed method with the work presented in the state of the art regarding the minute-by-minute CAP estimation.

Work	Population (number of subjects)	Performance metrics (rounded average value)		
		Acc (%)	Sen (%)	Spe (%)
[20]	70	-	40	84
[25]	13	71	74	66
[26]	13	76	71	79
[27]	13	77	71	82
This work	19	71	73	70

that only Mendonça et al. [26], [27] achieved better results, using a significantly more complex approach that is less suitable for hardware implementation, while Thomas et al. [20] presented the less complex approach though, the results are too unbalanced for clinical analysis.

A comparison between the results of the works that have performed the OSA detection based on the ECG signal analysis (reporting both the performance metrics in a minute-by-minute approach and the global accuracy) is presented in Table 5. By analyzing Table 5, it is possible to verify that, although other methods achieved a better minute-by-minute OSA detection accuracy, the proposed method attained the third best performance for the OSA diagnosis per-subject (Acc-G). It is also relevant to notice that the models that attained the best performance, for the minute-by-minute, have used the recordings from the PhysioNet apnea-ECG database while the developed model was trained and tested with recordings collected in a hospital. Therefore, it is difficult to establish a direct comparison with the other results due to the use of different datasets that were created with different conditions. Only the work presented by Ravelo-García et al. [28] used the same recordings as employed on this work. By comparing the results it is possible to verify that the proposed method attained a balanced performance (similar sensitivity and specificity) while Ravelo-García et al. [28] have a 52%

TABLE 5. Comparison of the proposed method with the work presented in the state of the art regarding the OSA detection.

Work	Population (number of subjects)	Performance metrics (rounded average value)			
		Acc-G (%)	Acc (%)	Sen (%)	Spe (%)
[28]	35	82	79	42	94
[30]	35*	93	-	74	86
[29]	35*	89	90	89	91
[32]	35*	94	94	95	93
[8]	35*	97	85	82	87
[31]	35*	97	86	83	88
This work	35	94	76	73	77

* Recordings from the PhysioNet apnea-ECG database [73]

difference between sensitivity and specificity. The Acc-G of the proposed method is also 12% higher. Hence, the developed model is more suitable for clinical analysis.

A summary of the works that reported the Acc-G for the sleep quality analysis is presented in Table 6, where it is possible to verify that the developed method attained the highest accuracy.

TABLE 6. Comparison of the proposed method with the work presented in the state of the art regarding the sleep quality estimation.

Work	Description	Population (number of subjects)	Dataset	Acc-G (%)
[33]	Clustered sound events classified by an HMM	36	-	70
[38]	Set of rules that analyzed attributes	280	-	73
[36]	Features from EEG to fed a GELM	8	-	76
[25]	CPC classified by two DSAE	13	CAPSD	77
[35]	Features from EEG to fed a FFNN	14	CAPSD	78
[37]	Features from EDR and HRV to fed a SVM	15	MIT-BIH [74]	78
[34]	Actigraphy signal fed to a CNN	92	-	93
This work	Combinatory scheme that examined the output of three 1D-CNN	19	CAPSD	100

VI. CONCLUSION

The main objective of this work was to develop a method capable of assessing the quality of sleep and the presence of OSA using only the signal from one sensor and implement the method in an HMD that is non-invasive and simple to self-assemble, allowing the examination to be performed at the patient’s home.

It was verified that the performance of the sleep quality estimations, produced in this work, is higher than most methods available in the state of the art. The accuracy of the indirect CAP estimation is in the range of the agreement among experts scoring the CAP events, according to the values reported by Rosa et al. [66] and Largo et al. [67]. It was also verified that the performance of the global OSA assessment is in the same range as the best works available in the state of the art. Therefore, the developed algorithms could possibly be employed for clinical analysis with the potential to increase the accessibility of the population to the OSA diagnosis and the assessment of sleep quality deficits. However, the possible benefits of replacing human analysis still need to be evaluated.

The proposed sleep quality metric (m-CAP-tib) also attained a good correlation with the CAP rate estimated by PSG. Conversely, it is simpler to be estimated; thus, it could possibly lead to further developments in the sleep quality assessment by HMDs. The SQ-g correctly predicted the

quality of sleep of all subjects from the database (CAPSD), advocating the relevance of the developed work. It is relevant to notice that only one (minute-by-minute CAP assessment) of the six classifications that are performed in this work is similar to the works preened by Mendonça et al. [25], [27] and uses a different classifier (the LSTM which is a recurrent network, an approach that is considerably different from the DSAE and SVM).

The next steps of this research are the validation of the device against a PSG to assess the performance of the implementation and perform usability tests to determine if the subjects are capable to easily use the device. It is also intended to include the ability to detect other common sleep related disorders, providing an even more valuable device for clinical applications. A study will be carried out to evaluate other classifiers with the goal of assessing if the performance of the minute-by-minute models can be improved. An extended comparison of multiple methodologies presented in the state of the art will be performed in future work.

REFERENCES

- [1] M. W. Mahowald and C. H. Schenck, "Insights from studying human sleep disorders," *Nature*, vol. 437, no. 7063, pp. 1279–1285, Oct. 2005.
- [2] K. L. Waller, E. L. Mortensen, K. Avlund, M. Osler, B. Fagerlund, M. Lauritzen, and P. Jennum, "Subjective sleep quality and daytime sleepiness in late midlife and their association with age-related changes in cognition," *Sleep Med.*, vol. 17, no. 1, pp. 165–173, Jan. 2016.
- [3] F. Mendonça, S. S. Mostafa, F. Morgado-Dias, A. G. Ravelo-García, and T. Penzel, "A review of approaches for sleep quality analysis," *IEEE Access*, vol. 7, pp. 24527–24546, 2019.
- [4] M. J. Sateia, "International classification of sleep disorders-third edition," *Chest*, vol. 146, no. 5, pp. 1387–1394, Nov. 2014.
- [5] P. E. Peppard, T. Young, J. H. Barnett, M. E. W. Palta, Hagen, and K. M. Hla, "Increased prevalence of sleep-disordered breathing in adults," *Amer. J. Epidemiol.*, vol. 177, no. 9, pp. 1006–1014, May 2013.
- [6] A. Nassir and O. Barnea, "Wireless body-area network for detection of sleep disorders," presented at the IEEE 27th Conv. Elect. Electron. Eng. Israel, Eilat, Israel, Nov. 2012.
- [7] A. T. Thornton, P. Singh, W. R. Ruehland, and P. D. Rochford, "AASM criteria for scoring respiratory events: Interaction between apnea sensor and hypopnea definition," *SLEEP*, vol. 35, no. 3, pp. 425–432, Mar. 2012.
- [8] S. Martín-González, J. L. Navarro-Mesa, G. Juliá-Serdá, J. F. Kraemer, N. Wessel, and A. G. Ravelo-García, "Heart rate variability feature selection in the presence of sleep apnea: An expert system for the characterization and detection of the disorder," *Comput. Biol. Med.*, vol. 91, no. 1, pp. 47–58, Dec. 2017.
- [9] A. G. Ravelo, C. M. Travieso, F. D. Lorenzo, J. L. Navarro, S. Martín, J. B. Alonso, and M. Ferrer, "Application of support vector machines and Gaussian mixture models for the detection of obstructive sleep apnoea based on the RR series," presented at the 8th WSEAS Int. Conf. Appl. Math., Tenerife, Spain, Dec. 2005.
- [10] L. Almazaydeh, K. Elleithy, and M. Faezipour, "Detection of obstructive sleep apnea through ECG signal features," presented at the IEEE Int. Conf. Electro Inf. Technol., Indiana, NA, USA, May 2012.
- [11] D. R. Hillman, A. S. Murphy, R. Antic, and L. Pezzullo, "The economic cost of sleep disorders," *Sleep*, vol. 29, no. 3, pp. 299–305, Mar. 2006.
- [12] R. Agarwal and J. Gotman, "Computer-assisted sleep staging," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 12, pp. 1412–1423, Dec. 2001.
- [13] A. H. Khandoker, J. Gubbi, and M. Palaniswami, "Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 1057–1067, Nov. 2009.
- [14] R. Berry, *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Darien, IL, USA: American Academy of Sleep Medicine, 2017.
- [15] M. G. Terzano, L. Parrino, A. Sherieri, R. Chervin, S. Chokroverty, C. Guilleminault, M. Hirshkowitz, M. Mahowald, H. Moldofsky, A. Rosa, R. Thomas, and A. Walters, "Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep," *Sleep Med.*, vol. 2, no. 6, pp. 537–553, Nov. 2001.
- [16] S. Khosla et al., "Consumer sleep technology: An American academy of sleep medicine position statement," *J. Clin. Sleep Med.*, vol. 14, no. 5, pp. 877–880, May 2018.
- [17] K. A. Kaplan, J. Hirshman, B. Hernandez, M. L. Stefanick, A. R. Hoffman, S. Redline, S. Ancoli-Israel, K. Stone, L. Friedman, and J. M. Zeitzer, "When a gold standard isn't so golden: Lack of prediction of subjective sleep quality from sleep polysomnography," *Biol. Psychol.*, vol. 123, no. 1, pp. 37–46, Feb. 2017.
- [18] A. D. Krystal and J. D. Edinger, "Measuring sleep quality," *Sleep Med.*, vol. 9, no. 1, pp. S10–S17, Sep. 2008.
- [19] L. Parrino, R. Ferri, O. Bruni, and M. G. Terzano, "Cyclic alternating pattern (CAP): The marker of sleep instability," *Sleep Med. Rev.*, vol. 16, no. 1, pp. 27–45, Feb. 2012.
- [20] R. J. Thomas, J. E. Mietus, C.-K. Peng, and A. L. Goldberger, "An electrocardiogram-based technique to assess cardiopulmonary coupling during sleep," *Sleep*, vol. 28, no. 9, pp. 1151–1161, Sep. 2005.
- [21] L. Chen, C. Liu, Z. Ye, B. Wang, and S. He, "Assessment of sleep quality using cardiopulmonary coupling analysis in patients with Parkinson's disease," *Brain Behav.*, vol. 8, no. 5, Art. no. e00970, doi: 10.1002/brb3.970.
- [22] F. Mendonça, S. Shanawaz Mostafa, A. G. Ravelo-García, F. Morgado-Dias, and T. Penzel, "A review of obstructive sleep apnea detection approaches," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 2, pp. 825–837, Mar. 2019.
- [23] F. Mendonça, S. S. Mostafa, A. G. Ravelo-García, F. Morgado-Dias, and T. Penzel, "Devices for home detection of obstructive sleep apnea: A review," *Sleep Med. Rev.*, vol. 41, pp. 149–160, Oct. 2018.
- [24] S. S. Mostafa, F. Mendonça, A. G. Ravelo-García, and F. Morgado-Dias, "A systematic review of detecting sleep apnea using deep learning," *Sensors*, vol. 19, no. 22, p. 4934, Nov. 2019.
- [25] F. Mendonça, S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García, "Sleep quality estimation by cardiopulmonary coupling analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 12, pp. 2233–2239, Dec. 2018.
- [26] F. Mendonça, S. Mostafa, F. Morgado-Dias, A. Ravelo-García, and T. Penzel, "Sleep quality of subjects with and without sleep-disordered breathing based on the cyclic alternating pattern rate estimation from single-lead ECG," *Physiol. Meas.*, vol. 40, pp. 1–14, Oct. 2019, doi: 10.1088/1361-6579/ab4f08.
- [27] F. Mendonça, S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García, "Matrix of lags: A tool for analysis of multiple dependent time series applied for CAP scoring," *Comput. Methods Programs Biomed.*, vol. 189, no. 1, Jun. 2020, Art. no. 105314.
- [28] A. Ravelo-García, J. Kraemer, J. Navarro-Mesa, E. Hernández-Pérez, J. Navarro-Esteve, G. Juliá-Serdá, T. Penzel, and N. Wessel, "Oxygen saturation and RR intervals feature selection for sleep apnea detection," *Entropy*, vol. 17, no. 5, pp. 2932–2957, May 2015.
- [29] P. D. Chazal, T. Penzel, and C. Heneghan, "Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram," *Physiol. Meas.*, vol. 25, no. 4, pp. 967–983, Aug. 2004.
- [30] A. G. Ravelo-García, J. L. Navarro-Mesa, E. Hernández-Pérez, S. Martín-González, P. Quintana-Morales, I. Guerra-Moreno, and Gabriel Juliá-Serdá, "Cepstrum feature selection for the classification of sleep apnea-hypopnea syndrome based on heart rate variability," presented at the Comput. Cardiol. Conf., Zaragoza, Spain, Sep. 2013.
- [31] C. Song, K. Liu, X. Zhang, L. Chen, and X. Xian, "An obstructive sleep apnea detection approach using a discriminative hidden Markov model from ECG signals," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1532–1542, Jul. 2016.
- [32] V. P. Rachim, G. Li, and W.-Y. Chung, "Sleep apnea classification using ECG-signal wavelet-PCA features," *Bio-Med. Mater. Eng.*, vol. 24, no. 6, pp. 2875–2882, 2014.
- [33] H. Wu, T. Kato, M. Numao, and K.-I. Fukui, "Statistical sleep pattern modelling for sleep quality assessment based on sound events," *Health Inf. Sci. Syst.*, vol. 5, no. 1, pp. 1–11, Dec. 2017.
- [34] A. Sathyanarayana, S. Joty, L. Fernandez-Luque, F. Ofli, J. Srivastava, A. Elmagarmid, T. Arora, and S. Taheri, "Sleep quality prediction from wearable data using deep learning," *JMIR mHealth uHealth*, vol. 4, no. 4, p. e125, Nov. 2016.

- [35] F. Mendonça, A. Fred, S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García, "Automatic detection of cyclic alternating pattern," *Neural Comput. Appl.*, Advances in Bio-Inspired Intelligent Systems, pp. 1–11, Apr. 2018. [Online]. Available: <https://link.springer.com/article/10.1007/s00521-018-3474-5>, doi: 10.1007/s00521-018-3474-5.
- [36] L.-L. Wang, W.-L. Zheng, H.-W. Ma, and B.-L. Lu, "Measuring sleep quality from EEG with machine learning approaches," presented at the Int. Joint Conf. Neural Netw. (IJCNN), Vancouver, BC, Canada, Jul. 2016.
- [37] M. Bsoul, H. Minn, M. Nourani, G. Gupta, and L. Tamil, "Real-time sleep quality assessment using single-lead ECG and multi-stage SVM classifier," presented at the Annu. Int. Conf. IEEE Eng. Med. Biol., Buenos Aires, Argentina, Sep. 2010.
- [38] R.-H. Choi, W.-S. Kang, and C.-S. Son, "Explainable sleep quality evaluation model using machine learning approach," presented at the IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI), Daegu, South Korea, Nov. 2017.
- [39] A. Craik, Y. He, and J. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, pp. 1–28, 2019.
- [40] B. Pourbabaee, M. J. Roshtkhari, and K. Khorasani, "Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 12, pp. 2095–2104, Dec. 2018.
- [41] A. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiological signals," *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.
- [42] A. Smerieri, L. Parrino, M. Agosti, R. Ferri, and M. G. Terzano, "Cyclic alternating pattern sequences and non-cyclic alternating pattern periods in human sleep," *Clin. Neurophysiol.*, vol. 118, no. 10, pp. 2305–2313, Oct. 2007.
- [43] L. Parrino, G. Milioli, A. Melpignano, and I. Trippi, "The cyclic alternating pattern and the brain-body-coupling during sleep," *Epileptologie*, vol. 33, no. 1, pp. 150–160, 2016.
- [44] F. Mendonça, S. S. Mostafa, F. Morgado-Dias, J. L. Navarro-Mesa, G. Juliá-Serdá, and A. G. Ravelo-García, "A portable wireless device based on oximetry for sleep apnea detection," *Computing*, vol. 100, no. 11, pp. 1203–1219, Nov. 2018.
- [45] N. Thai-Nghe, Z. Gantner, and L. Schmidt-Thieme, "Cost-sensitive learning methods for imbalanced data," presented at the Int. Joint Conf. Neural Netw. (IJCNN), Barcelona, Spain, Jul. 2010.
- [46] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.
- [47] I. Digital Signal Processing Committee, *Programs for Digital Signal Processing*. New York, NY, USA: IEEE Press, 1979.
- [48] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, and P. J. Schwartz, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, no. 3, pp. 354–381, Mar. 1996.
- [49] T. Leppänen, A. Kulkas, E. Mervaala, and J. Töyräs, "Increase in body mass index decreases duration of apneas and hypopneas in obstructive sleep apnea," *Respiratory Care*, vol. 64, no. 1, pp. 77–84, Jan. 2019.
- [50] N. Banluesombatkul, T. Rakthanmanon, and T. Wilaiprasitporn, "Single channel ECG for obstructive sleep apnea severity detection using a deep learning approach," presented at the IEEE Region Conf. TENCON, Jeju, South Korea, Oct. 2018.
- [51] E. Helfenbein, R. Firoozabadi, S. Chien, E. Carlson, and S. Babaeizadeh, "Development of three methods for extracting respiration from the surface ECG: A review," *J. Electrocardiol.*, vol. 47, no. 6, pp. 819–825, Nov. 2014.
- [52] S. P. Arunachalam and L. F. Brown, "Real-time estimation of the ECG-derived respiration (EDR) signal using a new algorithm for baseline wander noise removal," presented at the Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., Minneapolis, MI, USA, Sep. 2009.
- [53] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.
- [54] I. Santamaria and J. Via, "Estimation of the magnitude squared coherence spectrum based on reduced-rank canonical coordinates," presented at the IEEE Int. Conf. Acoust., Speech Signal Process., 2007, Honolulu, HI, USA, Apr. 2007.
- [55] J. Nagi, F. Ducatelle, G. A. Di Caro, D. Ciresan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber, and L. M. Gambardella, "Max-pooling convolutional neural networks for vision-based hand gesture recognition," presented at the IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA), Kuala Lumpur, Malaysia, Nov. 2011.
- [56] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [57] C. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.
- [58] A. G. Ravelo-García, J. L. Navarro-Mesa, M. J. Murillo-Díaz, and J. G. Juliá-Serdá, "Application of RR series and oximetry to a statistical classifier for the detection of sleep apnoea/hypopnoea," presented at the IEEE Comput. Cardiol. Conf. (CinC), Chicago, IL, USA, Sep. 2004.
- [59] H. M. Al-Angari and A. V. Sahakian, "Automated recognition of obstructive sleep apnea syndrome using support vector machine classifier," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 463–468, May 2012.
- [60] D. Sackett, R. Haynes, G. Guyatt, and P. Tugwell, *Clinical Epidemiology: A Basic Science for Clinical Medicine*, 2nd ed. Philadelphia, PA, USA: Lippincott Williams and Wilkins, 1991.
- [61] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.
- [62] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, Apr. 1960.
- [63] A. Hazra, "Using the confidence interval confidently," *J. Thorac. Dis.*, vol. 9, no. 10, pp. 4125–4130, 2017.
- [64] R. J. Thomas, C. Wood, and M. T. Bianchi, "Cardiopulmonary coupling spectrogram as an ambulatory clinical biomarker of sleep stability and quality in health, sleep apnea, and insomnia," *Sleep*, vol. 41, no. 2, p. zsx196, Feb. 2018.
- [65] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," presented at the 3rd Int. Conf. Learn. Representations, Los Angeles, CA, USA, May 2015.
- [66] A. Rosa, G. Alves, M. Brito, M. Lopes, and S. Tufik, "Visual and automatic cyclic alternating pattern (CAP) scoring: Inter-rater reliability study," *Arq. Neuropsiquiatr.*, vol. 64, no. 3A, pp. 578–581, 2006.
- [67] R. Largo, M. C. Lopes, K. Spruyt, C. Guilleminault, Y. P. Wang, and A. C. Rosa, "Visual and automatic classification of the cyclic alternating pattern in electroencephalography during sleep," *Brazilian J. Med. Biol. Res.*, vol. 52, no. 3, p. e8059, 2019.
- [68] M. O. Mendez, A. Alba, I. Chouvarda, G. Milioli, A. Grassi, M. G. Terzano, and L. Parrino, "On separability of A-phases during the cyclic alternating pattern," presented at the 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., Chicago, IL, USA, Aug. 2014.
- [69] H. Silva, J. Guerreiro, A. Lourenço, A. Fred, and R. Martins, "BITalino: A novel hardware framework for physiological computing," presented at the Int. Conf. Physiological Comput. Syst. (PhyCS), Lisbon, Portugal, Jan. 2014.
- [70] D. Batista, H. Plácido da Silva, A. Fred, C. Moreira, M. Reis, and H. A. Ferreira, "Benchmarking of the BITalino biomedical toolkit against an established gold standard," *Healthcare Technol. Lett.*, vol. 6, no. 2, pp. 32–36, Apr. 2019.
- [71] A. Dupre, S. Vincent, and P. Iaizzo, "Basic ECG theory, recordings, and interpretation," in *Handbook of Cardiac Anatomy, Physiology, and Devices*. New York, NJ, USA: Humana Press, 2005, pp. 191–201.
- [72] A. Bansal and R. Joshi, "Portable out-of-hospital electrocardiography: A review of current technologies," *J. Arrhythmia*, vol. 34, no. 2, pp. 129–138, Apr. 2018.
- [73] T. Penzel, G. Moody, R. Mark, A. Goldberger, and J. Peter, "The apnea-ECG database," in *Proc. Comput. Cardiol.*, Boston, MA, USA, Sep. 2000, pp. 255–258.
- [74] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/June 2001.



FÁBIO MENDONÇA was born in Funchal, Portugal, in 1990. He received the B.S. and M.S. degrees in electrical and telecommunications engineering from the University of Madeira, Portugal, in 2013 and 2016, respectively. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the Instituto Superior Técnico, University of Lisbon, Portugal.

He was a Lecturer with the University of Madeira and a Researcher with the Madeira Interactive Technologies Institute. He is also a Researcher with the Laboratory for Robotics and Engineering Systems. His research interests include sleep monitoring, machine learning, and renewable energy.



SHEIKH SHANAWAZ MOSTAFA (Member, IEEE) was born in Khulna, Bangladesh, in 1988. He received the B.S.Engg. degree from Khulna University, Bangladesh, in 2010, the M.S. degree from the Khulna University of Engineering and Technology, Bangladesh, in 2012. He is currently pursuing the Ph.D. degree with the Instituto Superior Técnico, University of Lisbon, Portugal.

He worked in different research projects related to biomedical signal and image processing and hardware implementation. His research interests include biomedical signal processing, artificial neural networks, and hardware implementations. He is a Graduate Student Member of the IEEE Signal Processing Society and the IEEE Circuits and Systems Society. He is also a member of the International Association of Engineers (IAENG) and the Bangladesh Computer Society.



GABRIEL JULIÁ-SERDÁ received the degrees in medicine and specialty of pulmonology from the Autonomous University of Barcelona and Bellvitge University Hospital, Barcelona, respectively, and the Ph.D. degree from the University of Las Palmas de Gran Canaria.

Subsequently, he spent three years as a Post-doctoral Fellow with the Toronto General Hospital and the Mount Sinai Hospital, Toronto, ON, Canada, where he investigated issues related to airway physiology. He has worked as a Pulmonologist in several hospitals in Spain, being his last destinations at the Hospital Universitario de Gran Canaria Dr. Negrín and the Perpetuo Socorro Clinic, Las Palmas de Gran Canaria, current job. He has been an Associate Professor of medicine with the University of Las Palmas de Gran Canaria. His research interests include physiology and respiratory diseases. He has made several publications both nationally and internationally on these topics.



FERNANDO MORGADO-DIAS (Member, IEEE) was born in Coimbra, Portugal, in 1971. He received the B.S. degree from the University of Aveiro, Portugal, in 1994, the M.S. degree from Joseph Fourier University, Grenoble, France, in 1995, and the Ph.D. degree from the University of Aveiro, in 2005.

He was a Lecturer with the Technical University of Setúbal, Portugal. He is currently an Assistant Professor with the Universidade da Madeira and a Researcher with the Madeira Interactive Technologies Institute. His research interests include sleep monitoring, renewable energy, artificial neural networks, and FPGA implementations. He is also a member of the Fiscal Council of the Portuguese Association of Automatic Control and the Director of the Ph.D. degree in automation and instrumentation.



ANTONIO G. RAVELO-GARCÍA was born in Las Palmas de Gran Canaria, Spain, in 1973. He received the M.S. and Ph.D. degrees in telecommunication engineering from the Universidad de Las Palmas de Gran Canaria (ULPGC), Spain.

He is currently a Docent and a Researcher with the Department of Signal and Communications (DSC) and the Institute for Technological Development and Innovation in Communications (IDeTIC), ULPGC. He is also an International Collaborator Member of ITI/LARSyS, Portugal. He has participated in multiple research projects and publications. His current research interests include biomedical signal processing, nonlinear signal analysis, data mining and intelligent systems in multimedia, communications, environment, health, and wellbeing.

...