IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 17, 2020, accepted November 1, 2020, date of publication November 5, 2020, date of current version November 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3036024

Classifier Precision Analysis for Sleep Apnea Detection Using ECG Signals

NUNO POMBO[1], (Member, IEEE), BRUNO M. C. SILVA[1,2], ANDRÉ MIGUEL PINHO[3], AND NUNO GARCIA[1]

[1]Instituto de Telecomunicações, Universidade da Beira Interior, 6200-001 Covilhã, Portugal
[2]IADE, Universidade Europeia, 1200-649 Lisboa, Portugal
[3]Department of Informatics, Universidade da Beira Interior, 6200-001 Covilhã, Portugal

Corresponding author: Nuno Pombo (ngpombo@di.ubi.pt)

This work was supported by the FCT/MCTES through National Funds and When Applicable Co-Funded EU Funds under Project UIDB/EEA/50008/2020.

ABSTRACT This article presents a study on the efficiency of implementing classifiers for the detection of sleep apnea moments based on a minute-to-minute Electrocardiogram (ECG) signal, detailing the comparison of the accuracy for different classifiers. At each ECG signal, a Sgolay filter was applied to extract the Heart Rate Variability (HRV) and the ECG-Derived Respiration (EDR) and they were used for the training, testing and validation of the classifiers. The same features were extended in a second phase in order to understand if all the classified features were important. According to the results obtained, the best accuracy was 82.12%, with a sensitivity and a specificity of 88.41% and 72.29%, respectively. This study shows the importance of choosing the right classifier for a specific problem as well as choosing and using the best features for a better accuracy. These promising early-stage results may lead to complementary studies to improve the classifiers for a possible real-world application. The performance of the proposed model was compared with other approaches used for the detection of sleep apnea.

INDEX TERMS Sleep apnea, electrocardiogram, feature extraction, feature selection, artificial neural network, support vector machine.

I. INTRODUCTION

Sleep apnea is a clinical disorder characterized by cessation of breathing during sleep that can last seconds or even minutes. Due to the fact that has direct effects on the cardiovascular system, such as systemic hypertension and sympathetic activity increment, it is considered an important cause for morbidity and mortality [1]. Since sleep is a key activity for each individual as it permits the human body to repair and maintain health [2], then is crucial to promote adequate clinical practices to mitigate its effects as evidenced when patients with sleep apnea who developed COVID-19 were considered at risk of great morbidity and mortality compared to other patients [3].

The gold-standard for sleep apnea diagnosis is the Polysomnography (PSG) that aggregates data collected from a myriad of body functions, such as: heart rhythm, eye movement, brain activity, and muscle activity, among others. However, this multi-parametric concurrent recording of physiologic data, limits its adoption. Indeed, this is a complex, cumbersome, and time-consuming activity because it requires an exhaustive test in a controlled environment; like an hospital setting, to monitor the patient's sleep, hence this diagnosis is both unfeasible for a large population and extremely expensive. So, is timely and promising the introduction of surrogate techniques that may be not only comfortably applied to the patient but also a low-cost and simpler solution. The literature in alternative models to PSG are very abundant, namely related with proposals based on either a reduced set of signals [4]–[8] or a combination of signals [9]. Thus, in this study we demonstrate a comprehensive benchmark of different classifiers and selected features based on a single signal, the Electrocardiogram (ECG). In line with this, four different classifiers to detect sleep apnea from ECG data were evaluated. In addition, these classifiers were tested on three different scenarios using distinct features (also extracted from the signal). The proposed methods could provide practitioners with a robust, simple and cost-efficient diagnosis tool compared with the classical screening schemes provided by PSG.

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

The main contributions of this article are as follows:

- Implementation of feature selection principles aiming at to determine the most relevant descriptors
- Benchmark of multiple classifiers to detect sleep apnea
- Explanatory and up-to-date state of the art on sleep apnea detection techniques.

The rest of the article is organized as follows. Section II introduces related works with special focus on physiological signals and classifiers for sleep apnea diagnosis. Section III details methods, and experimental settings. Section IV presents the results of our experiments and Section V explicates their significance. Finally, Section VI brings the article to conclusion.

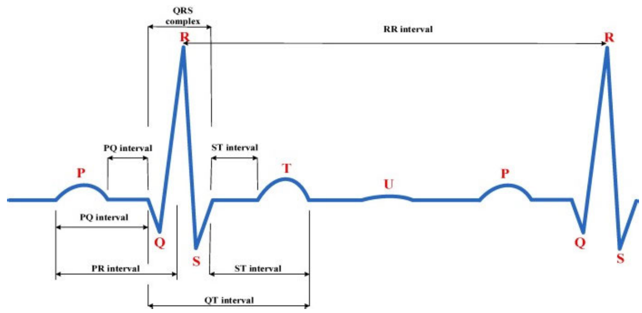

FIGURE 1. The ECG signal [14].

II. BACKGROUND

In recent years, different methods have been proposed in the literature for the diagnosis of sleep apnea disease. In, [10] authors conducted a systematic review on classification techniques used on computerised systems for sleep apnea diagnosis, identifying clusters of classifiers as follows: neural networks, regression, instance-based, Bayesian algorithms, reinforcement learning, dimensionality reduction, ensemble learning, and decision trees. On the one hand, separately of the adopted classifier its accuracy is highly dependant on an effectiveness features selection from the multitude of sources of data. On the other hand, since the PSG requires an exhaustive data collection fused by multiple sources of data such as ECG, electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG), oxygen-saturation (SpO2), among others, an observed trend in the literature is related with the adoption of a reduced number of physiological signals [11] as an alternative methodology for sleep apnea diagnosis.

Thus, is convenient to provide a brief perspective on candidate physiological signals for sleep apnea diagnosis with special focus on the ECG due to its immense adoption into these systems.

The flow of electricity generated from brain activity is measured by the EEG signal whereas the EMG signal measures the electrical activity generated from muscle motion. Furthermore, the electrocardiogram is a method of observing the heart function by measure the electric potential change related to the heartbeat resulting in the ECG signal [12], [13]. At a grassroots level, the ECG signal may be considered as a response to an impulse originated by the body. Indeed, this is an oscillatory signal due to the nature of the ECG signal. First, the ECG encompasses six features which corresponds to different stages that makes up a heartbeat which are denoted by letters P, Q, R, S, T and U as depicted in Figure 1. Second, the RR (a.k.a. RR interval) is the interval between successive heartbeats. Third, since R peaks are detected and if we measure the time between them we obtain the Heart Rate (HR). Four, the beat to beat variation in a heart-beat pattern is known as Heart Rate Variability (HRV). Five, the ECG-Derived Respiration (EDR) is the respiration signal derived from the ECG. Six, the Instantaneous Heart Rate (IHR) is the number of beats per minute.

In [15] authors presented a sleep apnea detection model using both HR, and RR signals extracted from the ECG signal. The Support Vector Machine (SVM) and the Random Forest (RF) were applied to classify normal and sleep apnea episodes. The observations revealed that both classifiers have yielded higher accuracies using features from HR signal as compared to RR signal. In addition, the 10-fold cross-validation demonstrated that the SVM has less error value than the RF. Also based on the ECG signal, authors in [16] combined the RR and the EDR signal as cornerstone of a sleep apnea system. The SVM, and the Stacked Autoencoder Based Deep Neural Network (SAE-DNN) were considered for classification. The experimental results demonstrated that SVM coupled with the Radial Basis Function (RBF) kernel performs better as compared to SAE-DNN. Similarly, authors in [17] proposed a sleep apnea detection system based on EDR and RR signals. The performance was determined using the fuzzy K-means clustering and the SVM classifier. The experiments revealed that the RBF kernel-based SVM has yielded the highest accuracy. In addition, authors recommended either the adoption of entropy features [18] or the implementation of deep learning algorithms. In [19] and [20] authors also used the EDR signal, but this time jointly the HRV. Authors in [19] implemented both Artificial Neural Network (ANN), and SVM to benchmark the system performance. The SVM classifier has yielded higher accuracy as compared to the ANN. The experimental results demonstrated that different features meet different significance in the system performance. On the contrary [20] used the Kernel Extreme Learning Machine (KELM) to distinguish between normal, and sleep apnea episodes. Main findings revealed that the polynomial kernel based KELM provided higher average accuracy as compared to linear, RBF, and cosine wavelet. Moreover, the inclusion of higher order spectral and non-linear features based on EDR and HRV signals were recommended.

On the contrary, [21] focused on the analysis of single-lead ECG signals. The classification of events either normal or apnea were performed by the following classifiers: Logistic Regression (LR), Linear Discriminant Analysis (LDA), SVM, Naive Bayes (NB), RF, and k-nearest neighbors (kNN). Authors findings included that not only RF provided higher accuracy, but also increasing the number of features led to a reduced accuracy. Congruently, [22] also used the single-lead ECG and implemented the following classifiers: kNN, Multilayer Perceptron Neural Network (MLPNN), SVM, Least-Square Support Vector Machine (LS-SVM). The experimental results demonstrated that the RBF kernel-based LS-SVM has yielded higher accuracy. Also based on a single-lead ECG, [23], and [24] proposed the Adaptive boosting (AdaBoost), and the SVM respectively to classify normal and apnea events. In addition, authors in [23] suggested the usage of time-frequency wavelet transforms to analyze oscillatory signals such as the ECG. In [25], authors extracted the EDR signal from the single-lead ECG and applied the following classifiers to detect sleep apnea episodes: ANN, SVM, kNN, Linear Discriminant (LD), and Quadratic Discriminant (QD). Main findings revealed that the ANN with two hidden layers performs better. Similarly in [26], authors used the HRV signal in they proposal for sleep apnea detection and implemented the following classifiers: ANN, BN, kNN, and SVM. The linear kerner SVM obtained the highest performance. In addition, authors highlighted that a feature extraction method has different performance in every classification method. In [27] authors proposed a sleep apnea detection system with the edge-computing principles in mind. Based on data provided by a single-channel ECG sensor, authors determined the system's performance through RF, Extremely Randomized Trees, SVM, NB, AdaBoost, kNN, and LR. It was observed that the SVM coupled with RBF kernel achieved the best accuracy in spite of the reduced number of features provided. In [28], authors proposed a microelectromechanical system (MEMS) based acceleration sensor for sleep apnea detection. The main goal was to measure diaphragm movements during the respiratory activity. The ANN was used as classifier of the proposed model. Furthermore, authors in [29] also proposed a wearable for ambulatory sleep apnea monitoring. The model used a single-lead ECG and a SVM classifier do distinguish normal, and apnea events. On the other hand, in [30] authors proposed a system based on the oronasal airflow signal. The SVM was the classifier elected to access the system's performance. Authors in [31] used the SpO2 sensor to acquire both oxygen blood rate, and heart rate. The notion behind this model is to determine a correlation between the oxygen saturation and the HRV during apnea episodes. The experimental results evidenced that the SVM provided higher accuracy as compared to KNN and ANN. In addition, it was observed that the 1-min variance demonstrated a good discriminant capacity.

Finally, [32] and [33] used deep learning methods on the sleep apnea detection. In [32], authors used the ECG signal and implemented the following deep learning models: Deep Neural Network (DNN), one-dimensional (1D) Convolutional Neural Networks (CNN), two-dimensional (2D) CNN, Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated-Recurrent Unit (GRU). In addition, authors suggested the implementation of either the 1D CNN or the GRU coupled with time series signals. On the other hand, [33] used the single-channel nasal pressure signal and applied a CNN model. Moreover, [34] combined the IHR with the SpO2 and applied the Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN).

III. METHODS AND MATERIALS

Based on lessons learned from the aforementioned literature we formulate the hypothesis that: (1) the ECG alone is a promising signal to use for sleep apnea detection. In addition, (2) an adequate feature selection is preponderant for the classifier accuracy, and (3) the SVM algorithm revealed its suitability to cope with apneic ECG signals. With those notions in mind, we developed a system to detect sleep apnea in which feature selection and classifiers were benchmarket. The flow of the proposed model is depicted in Figure 2 including: pre-process, feature extraction, classification, and feature selection. These architecture is explained in detail in the sections below.

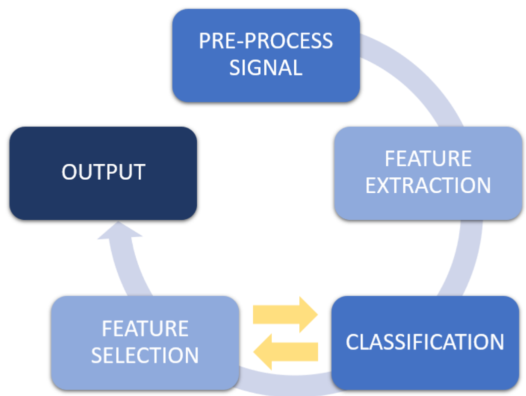

FIGURE 2. Proposed system activity model.

For the development of this study, we used GNU Octave,[1] which is software compatible with MATLAB[2] and its packages.

A. DATABASE

Our experiments were based on the PhysioNet [35] database. The considered datasource comprises 70 records, but only 35 records were used as only these were annotated. The signals during the 8h episodes were sampled with a frequency of 100 Hz, and annotated every minute by sleep disorder experts using standard criteria with respiratory signals, namely, each minute was labeled as 'A' or 'N' in case of sleep apnea moment or no-apnea, respectively.

B. PRE-PROCESS SIGNAL

The database used for evaluation has a wide variety of QRS complexes and P- and T-wave morphologies and the records

[1]https://www.gnu.org/software/octave/
[2]https://www.mathworks.com/products/matlab.html

have noise and artefacts that occur in a clinical setting. In line with this, as QRS detection is based on the time of occurrence of the QRS complex in the ECG signal it is pertinent to reduce the signal noise since it tends to decrease the classifiers' performance. Thus, the Sgolay filter [36] was used to remove the baseline wander, as it decreases the accuracy of the EDR. Afterwards, the obtained signal was subtracted from the original to yield the waveform. With an ECG signal free from noise, it is possible to detect the R-peaks and the QRS complex without missing or misclassifying a heartbeat. The TEO algorithm [37] was applied off-line over the signal on the basis of the discrete time domain. In addition, to detect the R-waves, the signal was processed in a one-second window. An adaptive threshold at 10% of the maximum R amplitude was applied due to the contrasting amplitude of the R-peaks along the signal. If the output at t0 exceeds the threshold and no greater value was observed in the next 0.25 seconds, then t0 is marked as an R-peak.

C. FEATURE EXTRACTION

The feature extraction was processed on both signals: the HRV and the EDR. In addition, features were extracted from one-minute segments congruently with the database annotations. Thus, in a first phase, 18 features were extracted from the HRV as detailed in Table 1, and 2 features were extracted from the EDR as presented in Table 2, giving a total of 20 features.

TABLE 1. Time-Domain Measure For HRV(m) Epoch Sequence.

RR(m)	$RR(m) = [rr_i]_{i=1}^{m}$	Feature Count
Mean	$\mu = \frac{\sum rr_i}{m}$	1
Standard Deviation	$\sigma = \sqrt{\frac{\sum (rr_i - \mu_{rr})^2}{m}}$	1
Sum of beats with inter-beat difference over 50 ms, variant 1	$NN50v1 = \sum_{i=2}^{m} unit[\lvert rr_i - rr_{i+1}\rvert - 50ms]$	1
Sum of beats with inter-beat difference over 50 ms, variant 2	$NN50v2 = \sum_{i=1}^{m-1} unit[\lvert rr_{i+1} - rr_i\rvert - 50ms]$	1
Ratio of NN50v1 to segment length	$pNN50v1 = \frac{NN50v1}{m}$	1
Ratio of NN50v2 to segment length	$pNN50v2 = \frac{NN50v2}{m}$	1
Mean of interbeat differentials	$\mu_{rd} = \frac{\sum rd_i}{m}$, where $rd_i = rr_{i+1} - rr_i$	1
Standard deviation of interbeat differentials	$\sigma = \sqrt{\frac{\sum (rd_i - \mu_{rd})^2}{m}}$	1
Root mean square of interbeat differentials	$RMSSD = \frac{\sqrt{\sum rd_i^2}}{m}$	1
Serial correlation coefficients (k=1,...,5)	$r_k = \frac{\sum_{i=1}^{m}(rr_i - \mu_{rd})(rr_{i+k} - \mu_{rr})}{\sum_{i=1}^{m}(rr_i - \mu_{rr})^2}$	5
Fractal Alan Factors (k=5,10,15)	$AT_k = \frac{\sum (N_{i+1}[k] - N_i[k])^2}{2 * \sum N_{i+1}[k]}$, $N_i[k]$ is the number of beats in the i-th window of k seconds	3
NEP (Number of Extreme Points)	$NEP = \frac{1}{m-2}\sum_{i=2}^{m-1}(1 - unit[(rr_i - rr_{i-1})(rr_{i+1} - rr_i)])$	1

TABLE 2. Time-Domain Measure For EDR(q) Epoch Sequence.

EDR(q)	$EDR(q) = [edr]_{i=1}^{q}$	Feature Count
Mean	$\mu_{edr} = \frac{\sum edr_i}{q}$	1
Standard Deviation	$\sigma_{edr} = \frac{\sqrt{\sum (edr_i - \mu_{edr})^2}}{q}$	1

In a second phase of the study, and to extend it with existing results from the literature, more features were added, allowing for an extended analysis of the results and of the behavior of different classifiers. In line with this, our experiments included 50 features in total for the HRV and 34 from the EDR signal, giving a total of 84 features. The additional features were extracted from the the 256-point FFT power spectral density, namely 32 points for each HRV, and EDR were considered.

D. CLASSIFICATION

With the classification in mind, all records extracted from one-minute segments were labeled as 0 or 1 representing non-apnea or apnea event respectively. The database containing 17401 records in which 46.33% are related to apnea moments, whereas non-apnea moments are observable in 53.67%. Then, database was segmented into three different vectors for training, testing, and validation purposes. The k-fold cross-evaluation method was adopted with k=10, in order to improve the training of the classifiers. Finally, sensitivity, specificity and accuracy were calculated as follows:

$$sensitivity = \frac{TP}{TP + FN} \tag{1}$$

$$specificity = \frac{TN}{TN + FP} \tag{2}$$

$$accuracy = \frac{TP + TN}{P + N} \tag{3}$$

where P: Positive. N: Negative. TP: True Positive. TN: True Negative. FP: False Positve. FN: False Negative.

In the classification phase, five classifiers (ANN, SVM, LDA, PLS, and aNBC) were implemented and its performance were comparatively evaluated. All algorithms were implemented following its default settings except the ANN and the SVM that were configured for our experiments.

1) ARTIFICIAL NEURAL NETWORK (ANN)

The ANN was implemented with both 20 and 84 input neurons (congruently with the 20 and the 84 features extracted respectively). The hyperbolic tangent sigmoid transfer function, i.e. *tansig* was used as a transfer function between the input layer and the hidden layer. Then, the linear transfer function i.e. *purelin* was used as a transfer function between the hidden layer and the output layer. The *tansig* function is defined as:

$$tansig(n) = \frac{2}{1 + e^{-2n}} - 1$$

and the *purelin* function is defined as:

$$purelin(n) = n.$$

2) SUPPORT VECTOR MACHINE (SVM)

An RBF kernel-based SVM was implemented as defined below:

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}, \gamma \geq 0 \tag{4}$$

In which γ determines the variance *i.e.* the similarity measure between two points. A large value means a small variance (two points are similar when they are close to each other). On the contrary, a lower value means a large variance (two points are similar even if are distant to each other) [19].

On the other hand, aiming at to obtain a better overall fit model [38], we tuned the SVM soft margin, namely the C parameter. Based on the experimental results the models performed best with the C parameter equal to 512.

3) LINEAR DISCRIMINANT ANALYSIS (LDA)

The LDA was introduced by [39] for dimensionality reduction. On the one hand its simple to implement since is based on generalized eigenvalue decomposition. In addition, its easy to adapt for discriminating non-linearly separable classes [39]. In other words, the LDA aims to identify a low-dimensional linear subspace whereon instances of multiple classes; at least two, are best separable [40]. Figure 3 depicts a two class-separation using the LDA by means of axes maximization.

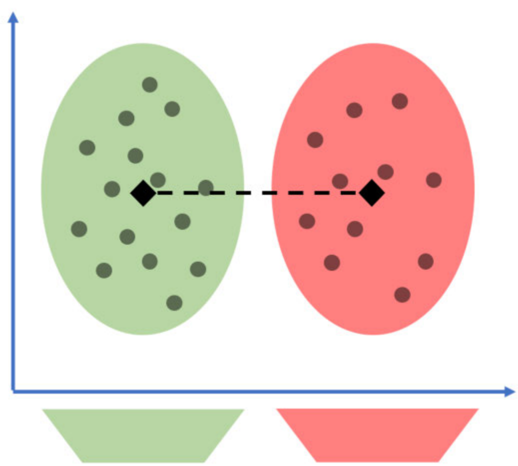

FIGURE 3. **LDA maximizing the component axes for class-separation.**

4) PARTIAL LEAST SQUARES (PLS) REGRESSION

The (PLS) regression may also be applied to reduce the data dimensionality. Indeed, the main goal of PLS regression is to determine an input vector composed by relevant and informative data according to the output [41]. As depicted in Figure 4, the notion behind PLS regression is to describe the relationship between multiple response variables and predictors through the latent variables wisely selected to provide maximum correlation with the dependant variable.

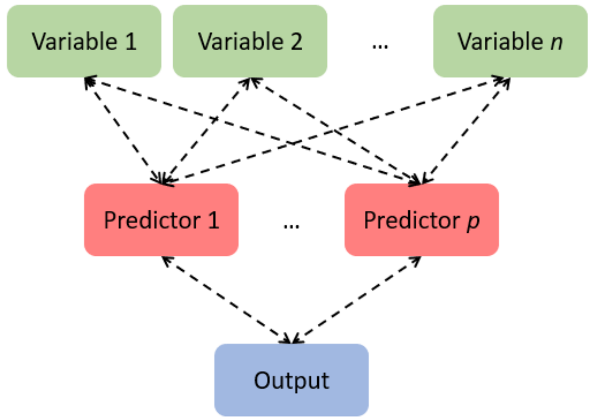

FIGURE 4. **PLS summarizing variability of variables and use it as predictor.**

5) AUGMENTED Naïve BAYES CLASSIFIER (aNBC)

The aNBC is an extension of the naive Bayes classifier, in which the class node directly points to all attribute nodes, and there exist links among attribute nodes [42]. At a grass-roots level, all attributes are independent given the value of the class variable in the naive Bayes classifier, while they are dependents in the aNBC scheme. As depicted in Figure 5, the attribute A1 is dependent on A2, and A*n* whereas A2 is dependent on A1, and A3. Finally A3 is dependant on A2.

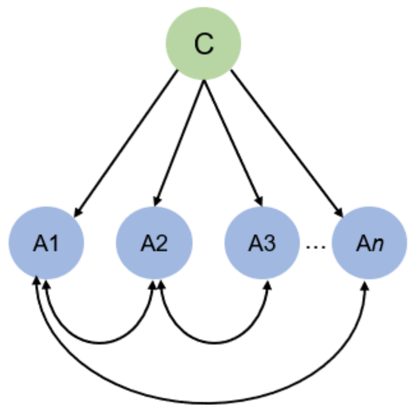

FIGURE 5. **Example of aNBC.**

IV. RESULTS ANALYSIS

Our computational experiments were based on the above mentioned classifiers. Firstly, 20 extracted features were applied to train and simulate the model. Secondly, 64 additional features obtained via PSD/FFT points were added to the initial features set (i.e., 84 features in total). All five classifiers were trained using 8507 records of features, with 2836 records used as training set and 5671 records used to evaluate the performance. The data provided to the classifiers for training and testing were divided using the k-fold cross-validation method with k=10.

The classification performance was assessed on both data sets (of 20 and 84 features) as shown in Table 3, Table 4, Figure 6, and Figure 7.

TABLE 3. Results when using 84 features with comparison between the classifiers.

Classifiers	Accuracy	Sensitivity	Specificity
ANN	59.40%	96.43%	2.67%
SVM	61.61%	**99.23%**	0.87%
LDA	60.57%	98.56%	0.95%
PLS	**63.00%**	54.93%	**65.24%**
aNBC	62.12%	0%	62.12%

TABLE 4. Results when using 20 features with comparison between the classifiers.

Classifiers	Accuracy	Sensitivity	Specificity
ANN	**82.12%**	**88.41%**	72.29%
SVM	70.94%	80.87%	54.94%
LDA	62.93%	83.98%	28.40%
PLS	64.49%	57.78%	66.05%
aNBC	41.20%	39.24%	**79.21%**

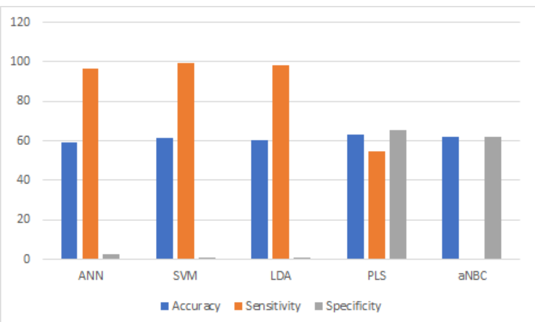

FIGURE 6. Results when using 84 features with comparison between the classifiers.

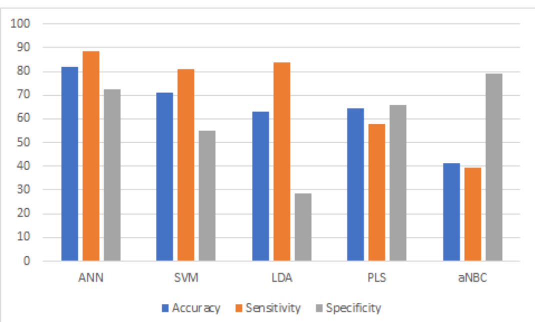

FIGURE 7. Results when using 20 features with comparison between the classifiers.

Our experiments encompassed two scenarios: (1) extraction and classification of features; (2) additional features extraction and classification. On the one hand, when 20 features were extracted and classified then we may observe that most accurate model is the ANN classifier (82.12% with a sensitivity of 88.41%). On the contrary, when 84 features were extracted the PLS classifier performed better (63.00%).

Moreover, we also observe that the higher specificity was obtained by the aNBC (79.21%), and PLS (62.24%) respectively. Finally, we note that the LDA (83.98%), and the SVM (68.36%) presented the higher sensitivity.

V. FINDINGS

The results obtained in our experiments are comparable with other studies; using the same database (PhysioNet ECG-Apnea Database), existing in the literature. This is relevant to consolidate the knowledge on the use of ECG signal for sleep apnea detection and diagnosis.

In [43] authors compared the performance of different classifiers based on a selected feature set, being that the most accurate was the Bagging.REPTree with an accuracy of 84.40%, and a specificity of 85.90%. The best sensitivity was achieved by the AdaBoost algorithm with a score of 87.03%. On the contrary, on the absence of feature selection the performance is lower, namely the observed accuracy, sensitivity, and specificity was 77.74% (Bagging.REPTree), 72.47% (AdaBoost), and 80.29% (Bagging.REPTree) respectively. In [1], only four features were selected to be considered on the classification leading to the following results: sensitivity, 88.84%, specificity, 83.29% and accuracy, 85.07%.

In addition, in [44]; that used the the ELM classifier, authors obtained results as follows: an accuracy of 87.71%, a specificity of 91.70% and a sensitivity of 81.30%. Despite this good performance, the sensitivity is lower compared with similar studies. It should be noted that, since is more relevant to detect the sleep apnea moments than the normal moments then sensitivity is more crucial than specificity. In a real-time monitor [45] a high sensitivity was obtained (96.00%), however the use of the PSG is mandatory to collect patients' data.

Finally, in [46], the LD and the QD classifiers were tested using three different methods (no optimization, feature selection, and co-variance regularization) which resulted that the QD provided the best specificity (94.60%), and LD the best sensitivity (94.00%) with no optimization. The highest accuracy (93.20%) was obtained by the QD using feature selection.

In our study two different methods were implemented using a set of 20 and 84 features respectively. In spite of the better performance of SVM compared to ANN as observed in the aforementioned literature, our experiments revealed that the best combination is achieved with the ANN coupled with 20 features. Indeed, an accuracy of 82.12%, a sensitivity of 88.41%, and a specificity of 72.29% were obtained for the ANN classifier. One explanation for this could be that these features are more correlated with the detection of sleep apnea and normal moments. Further, it provides additional evidence that when the pair sensitivity-specificity ratio is higher it may lead to an accurate detection of either normal or sleep apnea moments. Moreover, Table 4 evidences that using LDA or SVM results in a very sensitive classifier, bur very low specific. On the contrary, the aNBC results in a very specific classifier, but very low sensitive. In addition, it should be noted that higher sensitivity combined with reduced specificity may lead to poor classifier performances as evidenced in Table 3 on ANN (Sensitivity: 96.43% Specificity: 2.60% Accuracy: 59.40%) and SVM (Sensitivity: 99.23% Specificity: 0.87% Accuracy: 61.61%) classifiers.

In line with this, the proposed model revealed its suitability for sleep apnea detection and diagnosis based on a single signal, the ECG.

Finally, a particularly relevant finding of the present study is the correlation between the wisely feature selection and the accuracy. Into the context of this study, when the quantity of selected features increased lead to reduce the accuracy of the proposed model. Thus, it would be interesting to explore whether and how a wise selection of features may improve sleep apnea detection model.

VI. CONCLUSION

This study presents an ECG-based model for minute-based analysis of sleep apnea. The main goal is to implement an efficient and precise alternative method to the classical PSG, based on a single signal, the ECG. In addition, a benchmark with five classifiers are implemented, namely: ANN, SVM, LDA, PLS, and aNBC.

As expected and according with the presented results, it can be concluded that different classifiers have different behaviors to solve the same problem. Additionally, it is shown that the model proposed in this study is suitable, feasible and accurate in the detection of sleep apnea with an ECG signal. Our findings highlighted the ANN using 20 features as the most accurate model with an accuracy of 82.12%, a sensitivity of 88.41% and a specificity of 72.29%. Moreover, the experimental results revealed that is crucial determining the most relevant features with the ambition to enhance the accuracy of the model. Indeed, a same classifier may present contrasting performances as observed on the lower accuracy obtained when classifiers were evaluated with 84 features.

Future work may include the introduction of feature selection in order to determine an optimized characteristic set for the detection of sleep apnea; improving sensitivity so that all apnea moments are detected; comparing and calculating the performance of the different methods applied in the study, including evaluating the computational costs of classifiers; and simulating the same study in real patients to examine the viability of the method presented here and its implementation.

ACKNOWLEDGMENT

The authors would like to acknowledge the contribution of the COST Action IC1303–Architectures, Algorithms and Platforms for Enhanced Living Environments (AAPELE).

REFERENCES

[1] C. Varon, D. Testelmans, B. Buyse, J. A. K. Suykens, and S. Van Huffel, "Sleep apnea classification using least-squares support vector machines on single lead ECG," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 5029–5032.

[2] G. Atkinson and D. Davenne, "Relationships between sleep, physical activity and human health," *Physiol. Behav.*, vol. 90, nos. 2–3, pp. 229–235, Feb. 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031938406003957

[3] M. Thorpy, M. Figuera-Losada, I. Ahmed, R. Monderer, M. Petrisko, C. Martin, J. Akhtar, J. Thorpy, and C. Haines, "Management of sleep apnea in New York city during the COVID-19 pandemic," *Sleep Med.*, vol. 74, pp. 86–90, Oct. 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389945720303142

[4] A. Patangay, P. Vemuri, and A. Tewfik, "Monitoring of obstructive sleep apnea in heart failure patients," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 1043–1046.

[5] T. Al-ani, C. K. Karmakar, A. H. Khandoker, and M. Palaniswami, "Automatic recognition of obstructive sleep apnoea syndrome using power spectral analysis of electrocardiogram and hidden Markov models," in *Proc. Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process.*, Dec. 2008, pp. 285–290.

[6] T. Sugi, F. Kawana, and M. Nakamura, "Automatic EEG arousal detection for sleep apnea syndrome," *Biomed. Signal Process. Control*, vol. 4, no. 4, pp. 329–337, Oct. 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809409000494

[7] C.-C. Hsu and P.-T. Shih, "A novel sleep apnea detection system in electroencephalogram using frequency variation," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 6014–6024, May 2011. [Online]. Available: www.sciencedirect.com/science/article/pii/S095741741001256X

[8] N. Sezgin, "EMG classification in obstructive sleep apnea syndrome and periodic limb movement syndrome patients by using wavelet packet transform and extreme learning machine," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 23, pp. 873–884, 2015. [Online]. Available: http://sdu.dergipark.gov.tr/tbtkelektrik/144851

[9] N. Pombo and N. M. Garcia, "UbiSleep: An ubiquitous sensor system for sleep monitoring," in *Proc. IEEE 12th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2016, pp. 1–4.

[10] N. Pombo, N. Garcia, and K. Bousson, "Classification techniques on computerized systems to predict and/or to detect apnea: A systematic review," *Comput. Methods Programs Biomed.*, vol. 140, pp. 265–274, Mar. 2017. [Online]. Available: www.sciencedirect.com/science/article/pii/S0169260716306113

[11] A. M. da Silva Pinho, N. Pombo, and N. M. Garcia, "Sleep apnea detection using a feed-forward neural network on ECG signal," in *Proc. IEEE 18th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Sep. 2016, pp. 1–6.

[12] S. Saechia, J. Koseeyaporn, and P. Wardkein, "Human identification system based ECG signal," in *Proc. TENCON - IEEE Region Conf.*, Nov. 2005, pp. 1–4.

[13] J. S. Kim, G. H. Choi, and S. B. Pan, "A study on electrocardiogram based biometrics using embedded module," in *Proc. Int. Conf. Platform Technol. Service (PlatCon)*, Jan. 2019, pp. 1–4.

[14] A. Diker, E. Avci, E. Tanyildizi, and M. Gedikpinar, "A novel ECG signal classification method using DEA-ELM," *Med. Hypotheses*, vol. 136, Mar. 2020, Art. no. 109515. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306987719312381

[15] R. K. Tripathy, P. Gajbhiye, and U. R. Acharya, "Automated sleep apnea detection from cardio-pulmonary signal using bivariate fast and adaptive EMD coupled with cross time–frequency analysis," *Comput. Biol. Med.*, vol. 120, May 2020, Art. no. 103769. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482520301414

[16] H. Singh, R. K. Tripathy, and R. B. Pachori, "Detection of sleep apnea from heart beat interval and ECG derived respiration signals using sliding mode singular spectrum analysis," *Digit. Signal Process.*, vol. 104, Sep. 2020, Art. no. 102796. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S105120042030141X

[17] C. S. S. Viswabhargav, R. K. Tripathy, and U. R. Acharya, "Automated detection of sleep apnea using sparse residual entropy features with various dictionaries extracted from heart rate and EDR signals," *Comput. Biol. Med.*, vol. 108, pp. 20–30, May 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482519300897

[18] S. M. Pincus and A. L. Goldberger, "Physiological time-series analysis: What does regularity quantify?" *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 266, no. 4, pp. H1643–H1656, Apr. 1994, doi: 10.1152/ajpheart.1994.266.4.h1643.

[19] A. Pinho, N. Pombo, B. M. C. Silva, K. Bousson, and N. Garcia, "Towards an accurate sleep apnea detection based on ECG signal: The quintessential of a wise feature selection," *Appl. Soft Comput.*, vol. 83, Oct. 2019, Art. no. 105568. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1568494619303485

[20] R. K. Tripathy, "Application of intrinsic band function technique for automated detection of sleep apnea using HRV and EDR signals," *Biocybern. Biomed. Eng.*, vol. 38, no. 1, pp. 136–144, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0208521617303133

[21] A. Zarei and B. M. Asl, "Performance evaluation of the spectral autocorrelation function and autoregressive models for automated sleep apnea detection using single-lead ECG signal," *Comput. Methods Programs Biomed.*, vol. 195, Oct. 2020, Art. no. 105626. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169260720314590

[22] H. Sharma and K. K. Sharma, "An algorithm for sleep apnea detection from single-lead ECG using Hermite basis functions," *Comput. Biol. Med.*, vol. 77, pp. 116–124, Oct. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482516302086

[23] A. R. Hassan, "Computer-aided obstructive sleep apnea detection using normal inverse Gaussian parameters and adaptive boosting," *Biomed. Signal Process. Control*, vol. 29, pp. 22–30, Aug. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809416300519

[24] B. Fatimah, P. Singh, A. Singhal, and R. B. Pachori, "Detection of apnea events from ECG segments using Fourier decomposition method," *Biomed. Signal Process. Control*, vol. 61, Aug. 2020, Art. no. 102005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809420301610

[25] P. Janbakhshi and M. B. Shamsollahi, "Sleep apnea detection from single-lead ECG using features based on ECG-derived respiration (EDR) signals," *IRBM*, vol. 39, no. 3, pp. 206–218, Jun. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1959031818300794

[26] B. Sulistyo, N. Surantha, and S. M. Isa, "Sleep apnea identification using HRV features of ECG signals," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, p. 3940, Oct. 2018, doi: 10.11591/ijece.v8i5.pp3940-3948.

[27] A. Stojanovski, E. Zdravevski, S. Koceski, and V. Trajkovik, "Real-time sleep apnea detection with one-channel ecg based on edge computing paradigm," *ICT Innov.*, vol. 2018, pp. 124–134, Sep. 2018.

[28] A. H. Yüzer, H. Sümbül, M. Nour, and K. Polat, "A different sleep apnea classification system with neural network based on the acceleration signals," *Appl. Acoust.*, vol. 163, Jun. 2020, Art. no. 107225. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0003682X19314483

[29] G. Surrel, A. Aminifar, F. Rincon, S. Murali, and D. Atienza, "Online obstructive sleep apnea detection on medical wearable sensors," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 4, pp. 762–773, Aug. 2018.

[30] B. L. Koley and D. Dey, "Automatic detection of sleep apnea and hypopnea events from single channel measurement of respiration signal employing ensemble binary SVM classifiers," *Measurement*, vol. 46, no. 7, pp. 2082–2092, Aug. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0263224113000833

[31] L. Haoyu, L. Jianxing, N. Arunkumar, A. F. Hussein, and M. M. Jaber, "An IoMT cloud-based real time sleep apnea detection scheme by using the SpO_2 estimation supported by heart rate variability," *Future Gener. Comput. Syst.*, vol. 98, pp. 69–77, Sep. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167739X18326980

[32] U. Erdenebayar, Y. J. Kim, J.-U. Park, E. Y. Joo, and K.-J. Lee, "Deep learning approaches for automatic detection of sleep apnea events from an electrocardiogram," *Comput. Methods Programs Biomed.*, vol. 180, Oct. 2019, Art. no. 105001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169260719303086

[33] S. H. Choi, H. Yoon, H. S. Kim, H. B. Kim, H. B. Kwon, S. M. Oh, Y. J. Lee, and K. S. Park, "Real-time apnea-hypopnea event detection during sleep by convolutional neural networks," *Comput. Biol. Med.*, vol. 100, pp. 123–131, Sep. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482518301768

[34] R. K. Pathinarupothi, D. P. J., E. S. Rangan, G. E. A., V. R., and K. P. Soman, "Single sensor techniques for sleep apnea diagnosis using deep learning," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Aug. 2017, pp. 524–529.

[35] T. Penzel, G. B. Moody, R. G. Mark, A. L. Goldberger, and J. H. Peter, "The apnea-ECG database," in *Proc. Comput. Cardiol.*, vol. 27, Sep. 2000, pp. 255–258.

[36] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares Procedures.," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964, doi: 10.1021/ac60214a047.

[37] H. M. Teager and S. M. Teager, "Evidence for nonlinear sound production mechanisms in the vocal tract," in *Speech Production and Speech Modelling*. Dordrecht, The Netherlands: Springer, 1990, pp. 241–261, doi: 10.1007/978-94-009-2037-8_10.

[38] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1023/A:1022627411411.

[39] R. A. Fisher, "The statistical utilization of multiple measurements," *Ann. Eugenics*, vol. 8, no. 4, pp. 376–386, Aug. 1938, doi: 10.1111/j.1469-1809.1938.tb02189.x.

[40] N. Pombo, N. Garcia, and K. Bousson, "Machine learning approaches to automated medical decision support systems," in *Handbook of Research on Artificial Intelligence Techniques and Algorithms*, V. Pandian, Ed. Hershey, PA, USA: IGI Global, 2015, pp. 183–203.

[41] D. Kleinbaum, *Applied Regression Analysis and Other Multivariable Methods*. Australia Belmont, CA, USA: Brooks/Cole, 2008.

[42] H. Zhang, "The optimality of naive Bayes," in *Proc. 17th Int. Florida Artif. Intell. Res. Soc. Conf. (FLAIRS)*, V. Barr and Z. Markov, Eds. Palo Alto, CA, USA: AAAI Press, 2004.

[43] B. Xie and H. Minn, "Real-time sleep apnea detection by classifier combination," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 469–477, May 2012.

[44] N. Sadr and P. D. Chazal, "Automated detection of obstructive sleep apnoea by single-lead ECG through ELM classification," in *Proc. Comput. Cardiol.*, Sep. 2014, pp. 909–912.

[45] M. Bsoul, H. Minn, and L. Tamil, "Apnea MedAssist: Real-time sleep apnea monitor using single-lead ECG," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 3, pp. 416–427, May 2011.

[46] P. de Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 686–696, Jun. 2003.

NUNO POMBO (Member, IEEE) is currently an Assistant Professor with the University of Beira Interior (UBI), Covilhã, Portugal. He is also a Coordinator with the Assisted Living Computing and Telecommunications Laboratory (ALLab), UBI. He is also a member with the BSAFE Laboratory and the Instituto de Telecomunicações (IT), UBI. His current research interests include information systems (with a special focus on clinical decision support systems), data fusion, artificial intelligence, and software.

BRUNO M. C. SILVA received the Ph.D. degree in informatics engineering from the University of Beira Interior, in 2015.

He is currently an Assistant Professor, the Head of the Technology Department, and a Coordinator of games and apps development course. Since 2015, he has been a Senior Researcher with the Instituto de Telecomunicações, Universidade da Beira Interior, and a Co-Coordinator with the NetGNA Laboratory, integrated with the Network Protocols and Algorithms Research Group. He is also a Research Member with the Centro de Investigação em Cidades Inteligentes, Instituto Politécnico de Tomar. He has authored or coauthored several international conference papers and International Journal publications. His research interests include computer networks, mobile computing, and cloud computing, but especially: e-health, mobile health, the Internet of Things, and ambient assisted living.

Dr. Silva is a member of many international TPCs and participated in several international conferences organization.

ANDRÉ MIGUEL PINHO received the B.Sc. degree from the Universidade da Beira Interior (UBI), Covilhã, Portugal, in 2016. He is currently a member with the Assisted Living Computing and Telecommunications Laboratory (ALLab), a research group within the Instituto de Telecomunicações, UBI. His current research interests include machine learning, algorithms for bio-signal processing, e-health, and clinical decision support systems.

NUNO GARCIA received the B.Sc. degree (Hons.) in mathematics/informatics and the Ph.D. degree in computer science engineering from the University of Beira Interior (UBI), Covilhã, Portugal, in 2004 and 2008, respectively. He is currently an Assistant Professor with the Faculty of Engineering Computer Science Department, UBI, and an Invited Associate Professor with the Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal. He was the Founder and is a Coordinator with the Assisted Living Computing and Telecommunications Laboratory (ALLab), a research group within the Instituto de Telecomunicações, UBI. He was also the Co-Founder and is a Coordinator with the Executive Council of the BSAFE LAB–Law enforcement, Justice and Public Safety Research and Technology Transfer Laboratory, a multidisciplinary research laboratory in UBI. He is also a Coordinator with the Cisco Academy at UBI and the Chair of the COST Action IC1303–Architectures, Algorithms, and Platforms for Enhanced Living Environments (AAPELE), Brussels, Belgium. He is the main author of several international, European, and Portuguese patents. His main research interests include next-generation networks, algorithms for bio-signal processing, distributed and cooperative protocols. He is a member of the Non-Commercial Users Constituency, a group within GNSO is ICANN. He is also a member of ISOC.