Computer Methods and Programs in Biomedicine 140 (2017) 265–274

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Classification techniques on computerized systems to predict and/or to detect Apnea: A systematic review

Nuno Pombo[a,*], Nuno Garcia[a], Kouamana Bousson[b]

[a] Research Units: Instituto de Telecomunicações and ALLab Assisted Living Computing and Telecommunications Laboratory, Department of Informatics, Universidade da Beira Interior, Covilhã, Portugal and Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal
[b] Research Unit: LAETA/UBI-AEROG, Department of Aerospace Sciences, Universidade da Beira Interior, Covilhã, Portugal

ARTICLE INFO

Article history:
Received 13 June 2016
Revised 28 December 2016
Accepted 3 January 2017

Keywords:
Sleep apnea
Machine learning
Classification
Threshold-based classification
Systematic review

ABSTRACT

Background and objective: Sleep apnea syndrome (SAS), which can significantly decrease the quality of life is associated with a major risk factor of health implications such as increased cardiovascular disease, sudden death, depression, irritability, hypertension, and learning difficulties. Thus, it is relevant and timely to present a systematic review describing significant applications in the framework of computational intelligence-based SAS, including its performance, beneficial and challenging effects, and modeling for the decision-making on multiple scenarios.
Methods: This study aims to systematically review the literature on systems for the detection and/or prediction of apnea events using a classification model.
Results: Forty-five included studies revealed a combination of classification techniques for the diagnosis of apnea, such as threshold-based (14.75%) and machine learning (ML) models (85.25%). In addition, the ML models, were clustered in a mind map, include neural networks (44.26%), regression (4.91%), instance-based (11.47%), Bayesian algorithms (1.63%), reinforcement learning (4.91%), dimensionality reduction (8.19%), ensemble learning (6.55%), and decision trees (3.27%).
Conclusions: A classification model should provide an auto-adaptive and no external-human action dependency. In addition, the accuracy of the classification models is related with the effective features selection. New high-quality studies based on randomized controlled trials and validation of models using a large and multiple sample of data are recommended.

1. Introduction

Sleep apnea syndrome (SAS) is defined as a temporary closure of the upper airway during sleep when air is prevented from entering lungs which may results on the complete cessation of breathing for more than 10 s in adults [1]. This is typically, accompanied by a reduction in blood oxygen saturation and leads to arousal from sleep in order to breathe. In addition, repetitive obstructive events during sleep are hypothesized to cause intermittent hypoxia, resulting in activation of oxygen free radicals and an oxidative stress response.

As SAS events are classified according to whether the patient exhibits respiratory effort, then the presence of abdominal and thoracic effort for continuing breathing while air flow completely stops, is called Obstructive Sleep Apnea (OSA) representing the most common pattern of SAS. On the contrary, when a complete cessation of both respiratory movements and airflow during at least 10 s, is considered as Central Sleep Apnea (CSA). Finally, the combination of these two symptoms, defined by a central respiratory pause followed, in a relatively short interval of time, by an

Abbreviations: AI, Apnea Index; AHI, Apnea and Hypopnea Index; AIRS, Artificial Immune Recognition System; ANN, Artificial Neural Network; ANFIS, Adaptive Neuro-Fuzzy Inference System; AUC, Area Under receiver operating characteristic Curve; BHC, Binary Hierarchical Classification; BNN, Bayesian Neural Network; CAS, Central Sleep Apnea; ECOC, Error Correcting Output Code; ECG, Electrocardiogram; EEG, Electroencephalogram; EMG, Electromyography; EOG, electrooculography; FP, False Positive; FN, False Negative; HI, Hypopnea Index; HMM, Hidden Markov Model; KNN, K-Nearest Neighbor; LDA, Linear Discriminant Analysis; LS-SVM, Least Squares Support Vector Machine; LR, Logistic Regression; LVQ, Learning Vector Quantization; ML, Machine Learning; MLR, Multi-Linear Regression; MSA, Mixed Sleep Apnea; NARX, Nonlinear AutoRegressive network with eXogenous; OSA, Obstructive Sleep Apnea; PNN, Probabilistic Neural Network; NPV, Negative Predictive Value; PPG, Photoplethysmogram; PPV, Positive Predictive Value; PSG, Polysomnogram; RBFNN, Radial Basis Function Neural Network; RCT, Randomized Controlled Trial; RDI, Respiratory Disturbance Index; ROC, receiver operating characteristic; SAS, Sleep Apnea Syndrome; Sp02, Oxygen Saturation; SRN, Simple Recurrent Network; SVM, Support Vector Machine; TP, True Positive; TN, True Negative; VDA, Voice Activity Detection.

* Corresponding author.
E-mail addresses: ngpombo@ubi.pt (N. Pombo), ngarcia@di.ubi.pt (N. Garcia), bousson@ubi.pt (K. Bousson).

http://dx.doi.org/10.1016/j.cmpb.2017.01.001
0169-2607/© 2017 Published by Elsevier Ireland Ltd.

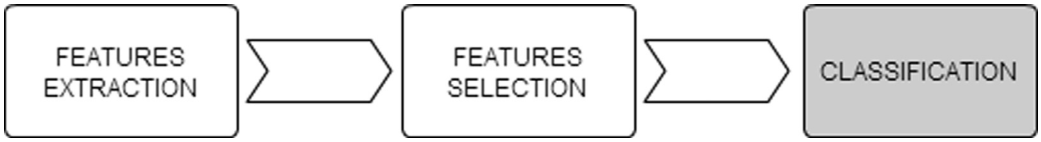

Fig. 1. Conceptual workflow of decision support systems, highlighting the focus of this systematic review.

obstructive ventilator effort is called Mixed Sleep Apnea (MSA). Moreover, hypopnea is a condition wherein the breathing is slow and shallow reducing the oxygen supply to lungs due to partial obstruction to the airway path.

The severity of apnea and hypopnea is measured by the number of episodes per hour, by means of an assessment tool, such as Apnea Index (AI), Hypopnea Index (HI), Apnea and Hypopnea Index (AHI) or Respiratory Disturbance Index (RDI). Thus, a mild case of apnea/hypopnea if observed when it happens between 5 and 15 episodes per hour, while a moderate case is verified when it occurs between 15 and 30 episodes per hour, and a severe case is a result of the occurrence of 30 or more episodes per hour.

These respiratory disturbances may lead to hypoxia and hypercapnia, which can trigger arousal from sleep by increasing ventilatory drive [2]. As a result of such sleep disruption, excessive daytime sleepiness is the most common presenting complaint [3]. Other symptoms of sleep apnea include snoring, asthma, sleep talking, sweating, chronic fatigue, falling asleep at inappropriate times of the day, morning headaches, weight gain, limited attention span, and memory loss. These symptoms can significantly decrease the quality of life and are associated with a major risk factor of health implications such as increased cardiovascular disease, sudden death, depression, irritability, hypertension, and learning difficulties [4]. Moreover, unfortunately, because of person's unawareness, SAS may go undiagnosed for years [5]. In fact, statistics show that around 100 million people worldwide are suspected to have OSA of which the vast majority undiagnosed [6]. In line with this, SAS is an important concern for public health that raises several challenges regarding it diagnosis, assessment and treatment.

The gold-standard method in SAS diagnosis is the nocturnal polysomnography (PSG). This diagnosis includes the monitoring of the breath airflow [7], snore [8], midsagittal jaw movement [9], respiratory events [10], oxygen saturation (SpO2), body position, electroencephalography (EEG), electromyography (EMG), electrooculography (EOG), and electrocardiography (ECG) which is inconvenient for a patient since it requires his connection to numerous sensors for one night, usually at the hospital under the supervision of a sleep technician. Therefore, new simplified methods and techniques for diagnosis and screening are desirable and timely. However, due to the complexity to deal with SAS symptoms, as evidenced by the multi-parameter monitoring provided by the PSG, it is advisable that new approaches could offer the ability to summarize all the collected information aiming at to extract the most relevant data, reducing the complexity of the model, before the classification process. For this reason, SAS diagnosis and screening often includes a three stages methodology as depicted in Fig. 1: feature extraction, selection of features and pattern classification. Firstly, the main purpose is to achieve a reduced set of features, extracted usually by means of algorithms from the observed data. Secondly, these features should be prioritized aiming to provide an adequate selection, which is meaningful since classification algorithms are unable to achieve high accuracy when a large number of weakly relevant and/or redundant features are managed. Thirdly, a classification method that should be wisely selected aiming to provide reasonable, reliable and consistent decisions.

In this paper, we provide a systematic review of classification methods for the decision-making on the multiple SAS scenarios, including their concepts, models, performance, plus beneficial and challenging effects.

2. Methods

2.1. Research questions

The primary research questions of this review were as follows: (RQ1) Which classification techniques have been used to support physician's decision-making on SAS? (RQ2) What is the beneficial and challenging effects stemming from the included case studies?

2.2. Inclusion criteria

Studies were included in this review if they met the following criteria: (1) presented a method to detect and/or to predict apnea, (2) were based on computerized systems, (3) included data about systems' evaluation, (4) presented preliminary or definitive results and (5) were written in English. These criteria were also applied to studies obtained from reference tracking. There were no age or disease restrictions; participants could be either adults or children, consisting of sleep disorders patients, whose data were collected either into a study or in a scientific database (e.g. Physionet).

2.3. Search strategy

To determine the state-of-the-art related with classification techniques on apnea prediction and/or detection, a search was conducted on the following databases: ACM Digital Library, and ScienceDirect. Only the studies published from the year 2005 until December 31th, 2015 and meeting the inclusion criteria were considered for this study. Two reviewers independently evaluated every study, and their suitability was determined by the agreement of both parties. A third reviewer was considered to adjudicate on differences of opinion but was not required because a consensus was reached.

2.4. Extraction of study characteristics

The following data were extracted from the studies and tabulated (see Table 2): year of publication, population, database, main decision, classification method, and metrics to evaluate it performance as presented on Table 1. When a study compared several classification models the presented metrics are related to the most accurate model.

3. Results

The combined electronic searches identified 1537 records. Out of the 1460 titles and abstracts that were then screened to test eligibility using the PICOS screening tool (population or participant, intervention or indicator, comparator or control, outcome, and study design), 77 full-text papers were eligible for inclusion. Thus, the full text evaluation of these papers resulted in the exclusion of 37 records that did not match the defined criteria. Many of the excluded papers reported studies focus on the sleep quality and/or classification of sleep/awake events. Further hand searching