A review of automated sleep stage scoring based on physiological signals for the new millennia

FAUST, Oliver <http://orcid.org/0000-0002-0352-6716>, RAZAGHI, H. <http://orcid.org/0000-0002-4752-216X>, BARIKA, R., CIACCIO, E.J. and ACHARYA, U.R.

Citation:

FAUST, Oliver, RAZAGHI, H., BARIKA, R., CIACCIO, E.J. and ACHARYA, U.R. (2019). A review of automated sleep stage scoring based on physiological signals for the new millennia. Computer Methods and Programs in Biomedicine, 176, 81-91. [Article]

A review of automated sleep stage scoring based on physiological signals for the new millennia

Oliver Faust[a,1,*], Hajar Razaghi[a], Ragab Barika[a], Edward J Ciaccio[b], U Rajendra Acharya[c,d,e]

[a]*Department of Engineering and Mathematics, Sheffield Hallam University, United Kingdom*
[b]*Department of Medicine – Cardiology, Columbia University, New York, New York, USA*
[c]*Department of Electronic & Computer Engineering, Ngee Ann Polytechnic, Singapore*
[d]*Department of Biomedical Engineering, School of Science and Technology, SIM University, Singapore*
[e]*Department of Biomedical Imaging, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia*

Abstract

Background and Objective: Sleep is an important part of our life. That importance is highlighted by the multitude of health problems which result from sleep disorders. Detecting these sleep disorders requires an accurate interpretation of physiological signals. Prerequisite for this interpretation is an understanding of the way in which sleep stage changes manifest themselves in the signal waveform. With that understanding it is possible to build automated sleep stage scoring systems. Apart from their practical relevance for automating sleep disorder diagnosis, these systems provide a good indication of the amount of sleep stage related information communicated by a specific physiological signal.
Methods: This article provides a comprehensive review of automated sleep stage scoring systems, which were created since the year 2000. The systems were developed for Electrocardiogram (ECG), Electroencephalogram (EEG), Electrooculogram (EOG), and a combination of signals.
Results: Our review shows that all of these signals contain information for sleep stage scoring.
Conclusions: The result is important, because it allows us to shift our research focus away from information extraction methods to systemic improvements, such as patient comfort, redundancy, safety and cost.

Keywords: Sleep stage, Deep Learning, Internet of Health Things, Decision support systems

1. Introduction

Sleep is a basic human function, which is characterized by a sequence of alterations in brain, muscle, eye, heart and respiratory activity. That active and regulated process is a prerequisite for physical and mental health. Sleep renews the body by protecting the metabolizable energy, maturing the neuronal connections, as well as consolidating learning and memory. However, when the life rhythm quickens and lifestyle changes, sleepiness and sleep structure disorders threaten people's routine activities and public safety [1]. Apart from these direct, or immediate risk factors, traumatic childhood experiences may also increase the risk for a number of sleep disorders in adulthood [2]. Demographics show that up to 24% of the adult population have regular sleep problems [3]. In a more focused study, Ohayon and Smirne found that 27.6% of the Italian population had sleep disorder symptoms [4]. The 'Sleep Heart Health Study' established that, across the world, patients experiencing difficulty initiating or maintaining sleep or daytime sleepiness have a reduced Health-Related Quality of Life (HRQoL) [5, 6]. The impact of sleep problems on health and HRQoL translates into economic consequences [7, 8]. Wickwire et al.

[*]Corresponding author
[1]These authors contributed equally to this work.

estimate that the global aggregate cost for sleep disorders exceeds $100 billion USD per year [9]. Ozminkowski et al. found that, within a six month period, the average direct and indirect costs for adults with sleep disorders were about $1,000 greater than for patients without sleep problems [10]. Several scientific studies provide evidence that there is a strong link between fatigue and occupational safety [11]. Léger demonstrated that sleep problems are statistically linked to poorer medical status and worse socio-professional indicators [12]. A French study found that employees with sleep problems missed twice as many workdays during a year when compared to normal sleepers [12]. Sleep studies help to establish the diagnosis of pathologies, such as circadian rhythm disorders, epilepsy, sleep apnea, insomnia and hypersomnia [13, 14]. Insomnia is the most common sleep problem in industrialized countries [15, 16]. For example, the prevalence of insomnia is 23% in Japan and 56% in the United States). Around 50% of insomnia patients did not seek medical attention [17]. Hence, a large number of patients suffer without treatment. To maintain public health and productivity it is of great importance to monitor sleep and analyze sleep stages.

On an abstract level, there are two main sleep stages, Non-REM (NREM) and Rapid Eye Movement (REM). REM sleep occurs 5–30 min at 90 min intervals. During REM sleep the neuronal activity is higher than during NREM sleep. During NREM sleep, metabolic rate, sympathetic activity, blood pressure, and Heart Rate (HR) decrease while parasympathetic activity increases. Sleep experts follow well-established guidelines for sleep scoring based on guidelines from standardization bodies [18, 19]. Nowadays, overnight Polysomnography (PSG) is the 'gold standard' for sleep stage evaluation [20]. It is a multi-parametric measurement apparatus that records a wide range of physiological signals in parallel, such as Electroencephalogram (EEG), Electrocardiogram (ECG), Electrooculogram (EOG), Electromyogram (EMG), blood oxygenation, airflow, and respiratory effort. In the majority of cases, the PSG data is captured in the controlled environment of a sleep laboratory. During pre-processing, the data is divided into 30 s epochs, and every epoch is categorized as either wakefulness, REM sleep or one of four states (S1, ..., S4) during NREM sleep [21, 22, 23]. In 2012, the American Academy of Sleep Medicine (AASM) published guidelines where the NREM stages S3 and S4 were combined to one stage (S3) [19], also known as Slow Wave Sleep (SWS) [24]. The guidelines for sleep staging, from Rechtschaffen & Kales, suggest the use of two EEG channels[2], two EOG electrodes and one EMG electrode [18]. Despite the efforts to standardize sleep staging, ambiguities still exist. One such ambiguity comes from the fact that the sleep stage definitions leave some space for individual interpretation [25]. Hence, expert based sleep staging is subject to bias and may therefore be unreliable [24]. For example, Danker et al. examined inter-operator variability of human expert scorers and found an interrater agreement of only 76.8% [26]. Another problem is that the physiological mechanisms, which shape the physiological signals recorded during sleep, are not well understood [6]. It is understood that sleep patterns change significantly with age, but what causes these changes is less clear [27]. The lack of well-established causality between physiological processes and the observed signals makes the data interpretation complex. Furthermore, understanding physiological processes is an active research area, hence new and refined relationships have to be learned all the time. Human experts can extract the required information from medical data and make a diagnosis. However, computational methods can be used as assistive devices to detect subtle differences in imagery, speed up the analysis process, and reduce cost. These systems can provide a wide range of results, starting from event labelling[3], over feature extraction, up to the level of suggesting a diagnosis [28]. Despite progress made in the development of diagnostic support systems, fundamental questions still remain, such as 'Which physiological signals contain sufficient information to support a particular diagnosis?'

[2]The AASM manual suggests three EEG channels while keeping all other signals the same.

[3]For example, respiratory and body movement events.

and 'How can we ensure the safety of the diagnosis?'

In this review, we establish that there is a wide range of physiological signals which contain sleep stage related information. This information can be used to support diagnosis, treatment monitoring and drug efficacy tests. However, before we can harvest these benefits, it is necessary to measure the signals and extract the information. The fact that we found a wide range of signal processing methods indicates that there is no standard method for information extraction, and indeed it is unclear which signals provide sufficient information for diagnosis. To address this uncertainty, we reviewed information extraction mechanisms for different physiological signals, to provide an indication of the information that is actually contained in the data. With respect to this focus, we recognized that automated sleep stage scoring is likely to play a leading role in future work. Computer machinery can assist to reduce inter- and intra-observer variability. Supplementing manual analysis with computerized assistance has the potential to provide cost savings. Furthermore, computer based systems can increase the quality of the extracted information including the utilization of decision support systems to assist in signal interpretation. We have discovered a large body of research on automated sleep stage scoring. This research tends to follow a traditional design methodology of feature extraction, and in some cases automated decision making. The feature extraction step must be carefully considered, because it reduces the information available for decision making, and its design process can be error prone. We recognize that automatic sleep stage classification is a starting point for sleep stage scoring. However, its diagnostic quality is usually insufficient in a practical setting, so that the sleep stage recognition technique ultimately requires manual inspection of the polysomnograms by expert human scorers. To improve outcome, we propose a general sleep stage scoring systems design based on deep learning and Internet of Health Things (IoHT) technology, described in more detail herein.

2. Review

This section presents a review of relevant scientific literature on automated sleep stage scoring. We have structured the review such that the results can be used to support our position on computer assisted sleep stage scoring and to justify our vision for future sleep scoring systems. To be specific, we have structured the review in accordance with the physiological signals that underpin sleep stage scoring. An initial analysis of the available literature showed that EEG, ECG, and EOG data were most often used in automated sleep stage scoring systems. The next three sections provide the review results for sleep scoring systems based on these signals. Individual physiological signals can represent only one aspect of sleep stages. Measuring multiple signals provides the benefit of redundant information as well as possibly providing additional uncorrelated information. Hence, a number of scientific studies have investigated automated sleep stage scoring based on multiple signals. Section 2.4 provides the review results for these systems.

2.1. Electroencephalogram

The EEG is a recording of electrical activity of the brain. EEG patterns show different characteristics during sleep stages. These features have been used for development of numerous sleep stage classification systems [29, 30, 31, 32]. A wide variety of signal processing techniques have been used to extract sleep- related information from EEG signals including: time-domain features [33, 34, 35, 36], spectral features [37, 38, 39], time-frequency features [40, 41, 35] and non-linear features [42, 43]. To provide adequate decision support for medical practitioners, several classification methods have been utilised in the reviewed sleep classification studies including: K-means [33], Support Vector Machine (SVM) [41], Ensemble Classification, such as Random Forest [44], Bootstrap Aggregating [45] and Artificial Neural Network (ANN) [46].

Table 1: A summary of the review results for selected research work that used a *EEG signals* to support sleep stage scoring.

Author	Data	Feature extraction method	Classification method	Classification results
Mousavi et al., 2019 [53]	The benchmark Sleep-European Data Format (EDF) dataset	time and frequency-domain as well as sequence to sequence features	Deep learning	84.26% accuracy for a two class problem
Michielli et al., 2019 [52]	The benchmark Sleep-EDF dataset	55 time and frequency-domain features	Deep learning	83.6% accuracy for a two class problem
Sharma et al., 2018 [29]	The benchmark Sleep-EDF dataset	Three-band time-frequency localized wavelet filter bank followed by log-energy, signal-fractal-dimensions, and signal-sample-entropy	SVM	Up to 98.3% accuracy for a two class problem
Seifpour et al., 2018 [34]	The benchmark Sleep-EDF dataset	Novel time domain feature named Statistical Behaviour of Local Extrema	Multi class SVM	Up to 97.9% accuracy
Chriskos et al., 2018 [54]	23 healthy male adults between the ages of 23 and 45 (mean: 29 ± 6 years).	Two novel methods offunctional connectivity estimation: Synchronization Likelihood and Relative Wavelet Entropy	SVM [Highest accuracy], K-nearest parameters, Neural network	Accuracy: 92.93%
Memar and Faradji 2018 [44]	Sleep-EDF database (Pz-Oz channel), St. Vincent's University Hospital and University College Dublin (UCDDB), the Expanded Sleep-EDF database (XSEDFDB)	Nested 5-fold cross validation, subject cross-validation	Random Forest	Accuracy: 95.31% for nested 5-fold and 86.64% for subject cross-validation

Hassan and Subasi, 2017 [45]	Sleep-EDF database -DREAMS	Tunable-Q wavelet transform	Bootstrap aggregating (Bagging)	Accuracy: 92.43% for 6-classes from the Sleep-EDF database
Pillay et al., 2018 [35]	16 preterm and term born newborns of 27–41 weeks gestational age (their age at birth)	Multiple features from the time- and frequency-domain	Hidden Markov Models (HMMs), Gaussian Mixture Models (GMMs)	HMM: mean kappa : 0.62 ($\pm$0.16) GMM: mean kappa : 0.55 ($\pm$0.15).
Supratak et al., 2017 [51]	Montreal Archive of Sleep Studies, Sleep-EDF database	a deep learning model, named DeepSleepNets on raw single-channel data	No classifier	Sleep EDF: Kappa: 0.76 MASS: Kappa: 0.80
da Silveira et al., 2017 [39]	Sleep-EDF (Pz-Oz channel)	Discrete Fourier Transform (DFT)	Random Forest	Accuracy for 6-state sleep stages: 90.5%
Hassan and Bhuiyan, 2016 [40]	Sleep-EDF database	Tunable-Q factor wavelet transform	Random forest	Accuracy: 90.38%, for 6-classes
Hassan and Bhuiyan, 2017 [47]	Sleep-EDF database	Ensemble Empirical Mode Decom- position	Random under sampling boosting (RUSBoost)	Accuracy of up to 98%
Bajaj and Pachori, 2013 [48]	Sleep-EDF database	time-frequency image based on the Wigner-Ville distribution (WVD)	multiclass least squares sup-port SVM.	Accuracy: 88.47
Diykh et al., 2016 [33]	Sleep-EDF database, Sleep Spindles database	The time domain features and structural graph similarity	The K-means clustering algorithm	Accuracy: 95.93% for Sleep-EDF dataset

Dimitriadis et al., 2018 [38]	Sleep-EDF database	Wavelet decomposition and cross-frequency coupling techniques	multi-class Naive Bayes classifier	Accuracy: 94.4%
Čić et al., 2013 [41]	Twenty healthy Croatian babies aged 3 months	intrinsic mode functions decomposition and generalised zero-crossing methods	SVM	Accuracy: 90%
Shi et al., 2015 [32]	25 adult subjects; Sleep Apnea Dataset provided by St. Vincent's University Hospital and University College Dublin.	A two-stage multi-view learning algorithm based on a joint collaborative representation	K-means clustering	Accuracy: 81.10%
Vural and Yildiz, 2010 [50]	International Database PhysioNet Sleep Records	Principle component analysis of time domain and frequency domain	no classifier	41.1, 33.7, 92.6, 76.4, 96.4, 79.7% success rates for 6-classes
Koley and Dey, 2012 [55]	28 subjects aged between 35 and 56 suspected to have sleep apnea	SVM based recursive feature elimination technique	Binary SVMs were combined with a one-against-all strategy.	Accuracy: 85%
Şen et al., 2014 [56]	25 individuals aged 50 $\pm$10 years; Data set provided by provided by St. Vincent's University Hospital and University College Dublin	Hybrid approach	Random Forest	Accuracy: 98.02%
Hsu et al., 2013 [49]	Sleep-EDF (Fpz_Cz channel)	Energy feature extraction using FIR bandpass filters	Elman recurrent neural classifier	Accuracy: 87.2%
Acharya et al., 2005 [57]	Sleep-EDF database	Nonlinear measures	Feature statistics through Analysis Of Variance (ANOVA) test	–

Hassan and Bhuiyan decomposed EEG signals and developed a sleep classification system using the Ensemble Empirical Mode decomposition technique and the RUSBoost classifier with an average accuracy of 88.1% for a six class problem [47]. The accuracy is increased to 90.4% for the six class problem using a tunable-Q factor wavelet transform technique together with Random forest classifier [40]. Diykh, Li and Wen decomposed time domain features of EEG signals and employed and identified six sleep stages using the K-means algorithm with 95.9% accuracy [33]. Bajaj and Pachori [48] used time-frequency features of EEG signals and a multiclass least square SVM classifier to solve a six class problem. The classification accuracy was 88.5%. Hsu et al. proposed a system to classify sleep stages using EEG signal energy features and recurrent neural classifier, resulting in 87.2% accuracy [49]. Seifpour et al. [34] proposed a novel approach for multi-class sleep stage classification by using the symbolic analysis concepts to develop a new time domain feature termed Statistical Behaviour of Local Extrema (SBLE). They achieved 90.6% and 97.9% accuracy for six-stage and two-stage classification respectively. Principal component analysis [50] and Deep Learning methods [51, 52, 53] have also been employed to construct an EEG-based sleep staging system with reasonable accuracy. Table 1 provides a summary of the review results. The table columns are Author, Data, Feature extraction method, Classification method, and Classification results. The columns of the subsequent three tables have the same content. This allows us to contrast and compare automated sleep stage scoring systems that were based on different physiological signals.

2.2. Electrocardiogram

ECG signals are recordings of the electrical activity of the human heart. In the absence of heart diseases, ECG signals are highly structured and individual signal components can be identified through visual inspection [58]. Individual sleep stages manifest themselves in subtle changes in the ECG signal. Yücelbaş et al., Xiao et al., and Kesper et al. proposed that sleep staging with ECG is less complex, but equally accurate, when compared to PSG analysis [59, 60, 61]. Redmond et al. provide further support for the validity of ECG based sleep staging by comparing it with EEG based sleep staging [62, 63]. Fell et al. made the case for nonlinear analysis of ECG signals for sleep staging [64, 65].

Sleep stages are associated with activities of the Autonomic Nervous System (ANS) [66]. To be specific, during REM sleep, the lung tidal volume decreases, and the respiratory rate exhibits a frequent and irregular pattern compared with that in NREM sleep [67]. Therefore, the characteristics of the related physiological information, such as the respiratory rate and HR vary according to the sleep stages. In a clinical setting, the HR is established by measuring successive beat to beat (RR) intervals from ECG signals [68, 69]. Heart Rate Variability (HRV) provides one or multiple measures that help to establish the regularity of HR signals [70, 71]. These measures provide meaningful information for clinical intervention [72, 73], because they reflect the ANS condition [74, 75, 76]. During REM sleep, the HR and its variability are increased due to fluctuations between sympathetic and parasympathetic activities [77, 78]. Various HRV parameters, calculated with time-, frequency-domain, and nonlinear analyses, revealed significant differences between NREM and REM sleep [79, 80]. Trinder et al. analyzed the autonomic activity during sleep with HRV measures [81]. de Zambotti et al. documented the effects of alcohol on sleep by analysing the cardiac autonomic function [82]. Penzel et al. used detrended fluctuation and spectral analysis for sleep stage information extraction [79]. Respiratory sinus arrhythmia, a periodic variation in the HR according to the respiratory cycle, also exhibits different patterns for REM versus NREM sleep [19]. Liu et al. compared HR and pulse rate variability [83]. They found that pulse rate variability contained similar information as HRV. That is of practical importance, because pulse rate is easier to measure than HR. Virtanen et al. analyzed sleep stage changes in postmenopausal women [84]. Crasset et al. and Faust et al. established that HRV changes with age and gender [85, 86]. Mendez et al. proposed a real-time Decision Support System (DSS) for sleep stage scoring based on HR signals [87]. Table

Table 2: A summary of the review results for selected research work that used a *ECG signals* to support sleep stage scoring.

Author	Data	Feature extraction method	Classification method	Classification results
Yücelbaş et al., 2018 [59]	Sleep laboratory of Necmettin Erbakan University database and PhyisioNet	Morphological methods	Random Forest, Wake, Non-REM, REM (WNR)	Up to 87.11% accuracy
Fell et al., 2015 [64]	Data from 12 healthy male	Embedding dimension estimation	–	–
Fell et al., 2015 [65]	Data from 12 healthy male	Correlation Dimension (CD), Kolmogorov entropy, and Lyapunov exponent	–	–
Yoon et al., 2017 [78]	Twenty-one healthy subjects (male: 12, female: 9) and 30 subjects (male: 25, female: 5) with Obstructive Sleep Apnea (OSA) recorded at Seoul National University Hospital	HR Statistical parameters, Spectral power, variability measurements	Threshold, REM duration	Accuracy: 87.54%
Liu et al., 2017 [83]	Seventy-five sleep apnea patients. Data recorded as Shandong Province of Traditional Chinese Medicine Hospital	HR, Time domain statistical parameters, Spectral power, nonlinear measurements	Statistical analysis	Not reported
Kesper et al., 2012 [60]	Apnea-ECG and SIESTA Database	HR, Spectral power evaluated by ANOVA	threshold	Accuracy: 57.8%
Virtanen et al., 2007 [84]	71 healthy postmenopausal women	HR, linear, geometric and nonlinear	Statistical analysis	Not reported
Xiao et al., 2013 [61]	Public database Sleep and Stroke Volume Data Bank	HR, linear statistics, spectral power, nonlinear	WNR, random forest	Accuracy: 88.67%
Redmond and Heneghan, 2006 [62]	37 subjects	ECG derived respiration and HR statistics. EEG sleep staging for comparison	Evaluation of HR parameters during different sleep stages	Not reported

Redmond et al., 2007 [63]	31 male subjects	ECG derived respiration and HR statistics.	WNR, Linear Discriminant Analysis (LDA) and a quadratic LDA.	up to 76.1%
Mendez et al., 2010 [87]	24 subjects	HR statistics Spectral power.	REM-NREM, HMM.	Accuracy: 79.3%
Penzel et al., 2003 [79]	64 patients with symptoms of excessive daytime sleepiness and arterial hypertension	HR statistics, spectral power	Wake, light-, deep-sleep REM. ANOVA	Not reported
Crasset et al., 2001 [85]	26 subjects, 18 normal, 4 with heart transplants	HR Statistical analysis	ANOVA	Not reported
Trinder et al., 2001 [81]	14 healthy subjects	HR spectral power, blood pressure	Statistical analysis	Not reported
de Zambotti et al., 2015 [82]	17 healthy subjects	Statistical analysis of labeled data to find sleep stage transitions	Not reported	Not reported

2 provides a summary of the review results on automated sleep scoring systems based on ECG signals. That table includes work on HR, because for all relevant studies the HR signal was extracted from ECG signals with appropriate algorithms.

2.3. Electrooculography and respiratory effort

EOG results from the continuous measurement of the corneo-retinal standing potential which can be used to track eye movements. Hence, this signal provides important information for REM stage detection. According to the AASM rules [19], the EOG electrodes are positioned 1 cm lateral to the left and right outer canthi. That placement is straight forward, indeed it can be undertaken by patients [88]. The user led signal acquisition is an important factor for long term monitoring and continuous sleep stage assessment. From this perspective, the work by Virkkala et al. [89] is important, because they demonstrated that EOG signals contain information about NREM sleep stages. Rahman et al. could significantly improve the classification accuracy of EOG based sleep scoring [90].

Respiratory information has been widely used to assess human nocturnal sleep objectively [91, 92, 93]. Long et al. used respiratory effort amplitude to establish an automated sleep stage classification system [94]. To improve the classification accuracy, they performed subject specific feature normalization. Such subject specific interventions are an important topic when it comes to long term sleep health monitoring, because of age-related changes to physiological parameters. Table 3 summarizes the review results for sleep studies based on both EOG and respiratory effort.

2.4. Combination of signals

The combination of multiple physiological signals provides redundant information. That is important for human scorers, because a particular bit of information might be overlooked in one signal, but that same information might be detected in another signal. Therefore, PSG incorporates a wide range of physiological signals. As such, it is the standard method to diagnose sleep disorders [95]. Typically, PSG recordings include the EEG, EOG, EMG and ECG. In many cases, these signals are recorded during the entire night [23, 96]. With PSG, sleep stage is manually scored on each 30 s epoch throughout the night by trained sleep experts, forming a sleep hypnogram [22].

EEG signals in combination with other physiological signals, such as ECG, EOG and EMG have also been used to design automatic sleep stage scoring systems [97, 98]. Kishi et al. found that the mechanism which governs NREM sleep stage transitions is also important for the REM sleep rhythm [99].

R.S.T. Leung, studied the effects of OSA on the sleep stages by observing autonomic functions through multiple physiological signals [100]. Tracik and Ebersbach studied the sleep attack pattern of a Parkinson patient [101]. They found a very fast transition from stable wakefulness to S2 without passing through S1. Kushida et al. compared subject reports with sleep patterns extracted from PSG measurements [102]. They could not detect significant differences between the subjective case reports and the objective sleep staging. Montgomery-Downs et al. studied developmental changes of the sleep patterns in children [103]. Long et al. used actigraphy and respiratory effort to determine sleep and wake states [104]. In their study, they emphasized the nonlinear concept of dynamic warping to improve the classification results. Kirjavainen et al. fused information from both respiratory and body movement signals to determine sleep stages and wakefulness in infants and young adults [105]. The movement signals came from a novel sensor enhanced bed, which could measure body movements unobtrusively. Tripathy et al. [106] and Yildirim et al. [107] used a deep learning system to fuse information from multiple signals. Such an approach might provide better robustness in case of noisy and intermittent data. Table 4 summarizes our review findings for automated sleep stage scoring based upon a combination of signals.

Table 3: A summary of the review results for selected research work that used a *Electrooculography and respiratory effort* to support sleep stage scoring.

Author	Signals and data	Feature extraction method	Classification method	Classification results
Long et al., 2014 [94]	Respiratory effort from 48 healthy subjects participating in the SIESTA project	Respiratory amplitude, statistics, spectral power, amplitude and volume analysis	Subject specific quadratic LDA, WNR	Accuracy: 79%
Liang et al., 2015 [88]	EOG from 16 healthy experimental subjects	Spectral analysis	LDA for wake, REM, S1, S2, SWS classification	Sensitivity: 82.6%.
Virkkala et al., 2007 [89]	EOG from 265 subjects	Spectral analysis	Thresholds for REM S1, S2 and SWS	Epoch agreement: 72.9%.
Rahman et al., 2018 [90]	EOG Physionet DB	Statistics of Discrete Wavelet Transform (DWT) coefficients	6 classes problem approached with SVM, RUSboost and random forest	Accuracy of up to 91.7% with SVM

Table 4: A summary of the review results for selected research work that used a *combination of physiological signals* to support sleep stage scoring.

Yildirim et al., 2011 [107]	sleep-edf and sleep-edfx	Convolutional neural network	Up to 97.62%	
Author	Signals and data	Feature extraction method	Classification method	Classification results
Tripathy et al., 2018 [106]	MIT-BIH polysomnographic database	From RR: recurrence quantification analysis and dispersion entropy. From ECG: variance and the dispersion entropy of frequency bands	Deep neural network	95.71% accuracy for REM vs. NREM
Kishi et al., 2011 [99]	Full PSG 11 healthy subjects	Statistical analysis	State machine model for waking, REM sleep, S1, S2, and S3	Not reported
Takatani et al., 2018 [109]	74 newborns and 16 adults	EEG spectral power, HR absolute high frequency component.	Statistical analysis, REM and NREM sleep	–
Fonseca et al., 2015 [110]	Data from 48 subjects	ECG: Spectral power, variability measurements, and network analysis. PSG: Time / frequency, and network analysis	LDA, Wake, NREM and REM	Accuracy: 80%
Helland et al., 2015 [6]	EEG, ECG and respiratory signals from the SIESTA database	HR: statistics. PSG: Time / frequency, and network analysis	LDA, Wake, REM and REM	Accuracy: 80%
Kesek et al., 2009 [111]	EEG, ECG and respiratory signals from 230 habitual snorersand 170 other subjects (all female)	HR: statistics and Spectral power. PSG: manual scoring	Evaluation of HR parameters during different sleep stages	Not reported
Estévez et al., 2002 [112]	11 healthy infants	EEG sleep spindle detection EOG REM detection and EMG muscle tone	Threshold WNR	Not reported

Willemen et al., 2014 [3]	36 healthy subjects	HR statistics and spectral power, Breathing Rate (BR) statistics, and movement statistics	SVM WNR	81%
Long et al., 2014 [104]	Actigraphy and respiratory effort, 115 healthy adults	Statistical analysis of dynamic wrapping of body movement	LDA, binary problem for comparing features to a PSG study	accuracy 95.7%
Kirjavainen et al., 2018 [105]	22 infants or young children	Statistical analysis of body movements	Comparison with PSG, WNR	Not reported
R.S.T. Leung, 2015 [100]	17 healthy subjects	Statistical analysis of labeled data to find sleep stage transitions	Not reported	Not reported
Tracik and Ebersbach, 2001 [101]	Full PSG One subject with Parkinson's disease	Visual scoring of W, S1, S2, S3, and REM	Visual scoring	Not reported
Kushida et al., 2001 [102]	Full PSG 100 patients with sleep disorders	Visual scoring of wake and sleep states	Threshold	Accuracy: 77%
Montgomery-Downs et al., 2006 [103]	542 healthy children in the age range from 3.2 – 8.6 years	Visual scoring of W, S1, S2, S3, and REM	Visual scoring	Not reported

3. Discussion

Information can be defined as a measure of what we can learn from a given amount of data [108]. Hence, the idea of extracting information from physiological signals is vital to sleep stage scoring. With this information centric view, the research question can be stated as: 'How much information is needed for sleep state scoring and which signals provide that information?' In the absence of a standardized test for automated sleep stage scoring systems, this question is not readily answered, because each published study investigates a specific aspect and presents novel findings. These findings are based on a particular algorithm setup which is used to process data from specific databases. To shed some light on these questions, we have structured the review in terms of individual physiological signals. Based on this structure, we were able to establish that most of the reviewed work was concerned with EEG data. That focus is justified, because sleep and sleep stages is caused by significant changes in the brain activity [42]. Apart from EEG, all physiological signals measure symptoms of sleep stages. That makes it difficult to detect individual NREM stages. Hence, there are fewer studies which focus on these secondary signals. ECG is likely to be the most prominent secondary signal. It picks up sleep related changes in the ANS. EOG is an important signal for REM phase classification. However, NREM stages are rather complex to classify based on the EOG. Figure 1 depicts the number of studies which use a particular physiological signal. Apart from the amount of studies, another important fact is that the physiological data for all of the reviewed studies originated from clinical studies. There is as of yet no work on long term sleep stage monitoring, which would inevitably require the home recording of signals.

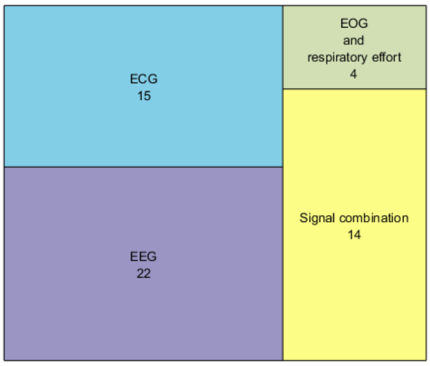

Figure 1: Treemap representation of the number of studies that used a particular physiological signal. The area of the individual rectangles is proportional to the amount of studies.

PSG studies are carried out in dedicated sleep labs, where patients are kept overnight. In the sleep lab, the cost for the individual measurement is low, compared to the overall cost of running the facility. Hence, it makes sense to measure as many physiological signals as possible during patient study. To be specific, multiple measurements add redundancy that improves the quality of the diagnosis, especially for human scorers. However, the need for redundancy implies that these systems have to address a problem which may be random in nature. Indeed it is difficult, if not impossible, to predict when a human expert will make an error. As a consequence, a prerequisite for reducing the degree of redundancy, and therefore the amount of resources required for a diagnosis, is to make the process which leads to a diagnosis more reliable. In the next section we outline a generic design of an automated sleep stage scoring system which addresses these shortcomings.

3.1. Future work

This review suggests that use of an automated DSS is one way to establish a reliable diagnosis. Trust in the DSS system should be established with traceability [113, 114], i.e. the decision process should be transparent and repeatable. Another important aspect is continuous learning. Just like a human practitioner, a DSS must also learn all the time. Furthermore, there is a need for less intrusive signal measurement systems, whereby to ensure patient comfort can be improved. In some cases long term monitoring is compromised by patients who fail to wear the sensor equipment, because wearing equipment was uncomfortable. Another important requirement for long term monitoring is real-time analysis [115], since real-time results provide an opportunity to control the therapeutic process.

To address these needs and establish the requirements, we approach the problem from a signal perspective. Recording a physiological signal over multiple sleep cycles implies that the measurement is done in the normal patient environment. EEG signals are impracticable, because the measurement setup must be done by an expert and it is difficult if not impossible for a patient to wear the recording system during the day. EOG is impractical for similar reasons, despite the fact that there are sensor masks that can be applied by patients. The masks, used to measure airflow and respiratory effort, are inconvenient to wear. Long term ECG monitoring is already a standard procedure which could be used to measure multiple sleep cycles. Even more convenient for the patient are HR measurements, because they involve only one sensor attached to a breast strap. That convenience comes from the fact that HR signals can be measured by detecting and encoding the time between two consecutive peaks (R-waves). The R-wave amplitude is rather large, usually in the rage of millivolts, when compared to the remainder of the signal. In contrast, the ECG requires constant recording[4] with a resolution of microvolts. Therefore, the data rate of the HR is much lower[5] when compared to ECG. The lower data rate implies that consumer technology can be used to communicate the HR data. With this technology an unobtrusive sleep stage monitoring, based on the IoHT, can be established [116]. The literature review in Section 2 shows that 5 studies linked HRV with sleep stages. Hence, HR signals contain the required information, i.e. they can be used for sleep staging.

Figure 2 shows an overview block diagram which captures these requirements. The data is shown flowing from the sensors to a cloud server via mobile technology. From there a deep learning system queries the measured signal in the form of data blocks. These data blocks can be used for automated sleep stage scoring and for learning. The analysis results are disseminated via social networks and other communication apps. This dissemination approach allows us to reach patient, caregivers, and medical staff in a discriminant way. The medical doctor in charge can obtain the raw data independently and review (trace) the decision process of the deep learning system. As such, sensor, mobile device, and cloud storage implement the IoHT. The deep learning system supports the medical practitioner in the process of finding a diagnosis. That diagnosis is disseminated to via the IoHT such that it reaches the correct patient.

In this design approach, deep learning takes center stage, because that method considers all of the available information content during both the training and the inference phases [117]. That is an advantage over the traditional machine learning algorithms found in most of the reviewed sleep scoring systems [118]. To be specific, traditional machine learning requires feature extraction to condense the data into a low dimensional feature vector[6], because the decision making algorithms fail to handle high dimensional data. In essence, the feature extraction step is an exercise in information reduction. Hence, traditional machine learning methods never consider all of the available information. Operating on reduced information makes them underperform

[4] The usual sampling frequency is 256 Hz.
[5] 256 times when we compare the ECG signal with a HR signal of 60 beats per minute.
[6] Typically, less than 10 dimensions

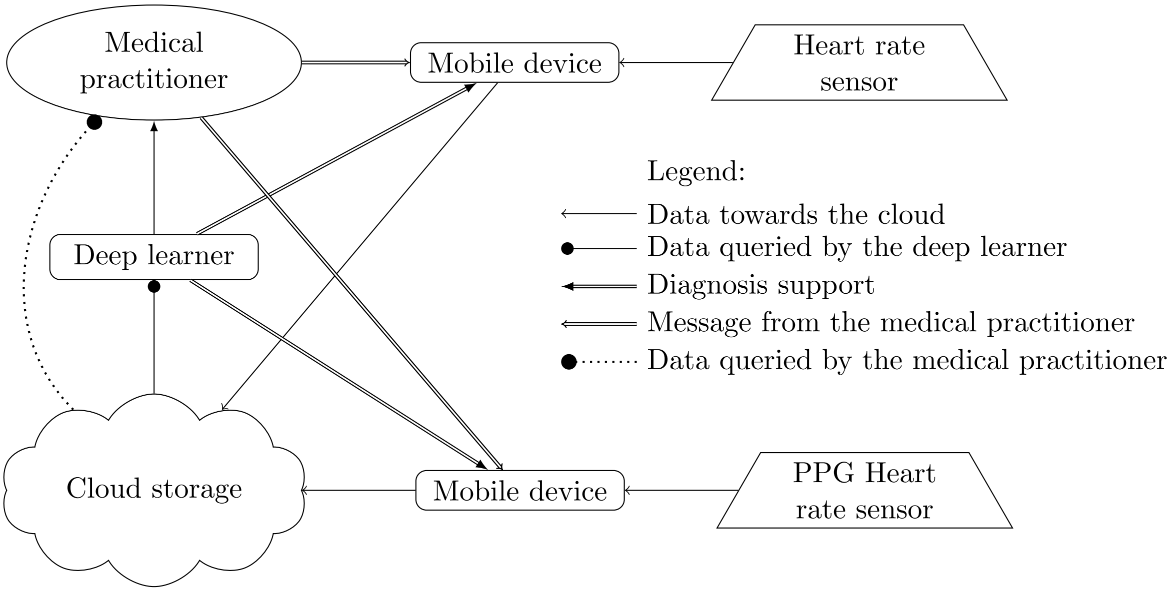

Figure 2: HR based sleep stage diagnosis support system

for unknown data. As a consequence, test result quality, as published in the scientific literature, is difficult if not impossible to achieve in a practical setting. In contrast, deep learning has the potential to excel in such blindfold validation tasks [119]. Hence, deep learning is more suitable for practical applications, such as long term sleep stage monitoring. The decision making algorithm is presented with all of the data containing all the available information. Conceptually, deep learning moves away from information reduction towards knowledge extraction. However, deep learning is computationally complex [119]. Thus, the data must travel to the processing, i.e. physiological data must travel to a data center. Depending on the physiological signal, this might create problems for the communication and storage infrastructure. Hence, we propose HR signals for automated sleep stage scoring. They have the lowest data-rate of all signals taken into consideration.

3.2. Limitations

Traditionally, HR is extracted from ECG signals by detecting the heartbeat (R wave) and subsequently calculating the beat-to-beat (RR) interval [120]. However, the instrumentation effort for measuring HR directly is significantly lower when compared with ECG measurements. In other words, we do not consider the most efficient signal measurement method for scientific studies. That efficiency comes from the fact that the R wave is a readily detected signal deflection. Sensors, which measure the HR directly are efficient, because R wave detection requires less instrumentation effort than ECG measurements. However, this presents a problem, because the heartbeat detection process is not well documented and is oftentimes proprietary to the company which manufactures the sensors. Therefore, it is difficult to establish that direct HR measurements will yield the same beat-to-beat interval sequence as HR extracted from ECG, especially for the subtle signal alterations which are indicative of sleep stage changes. However, none of the reviewed studies is based upon data from HR sensors. All of the relevant research was done by extracting the beat to beat interval from ECG signals. The signals were measured with medical equipment according to measurement standards [121]. Even with standardization, the measurement setup and indeed measurement errors influence the resulting signal [122]. The problem increases if the signal acquisition does not follow medical standards. There is no evidence that direct HR measurements have the same information content as HR extracted from ECG signals. For example, modern breast strap based HR sensors detect the R wave in hardware[7]. Such a means of detection tends to be less complex when compared to software algorithms that extract the R peaks from ECG [123]. That complexity is required to improve the peak detection quality. Pulse Pressure Variation (PPV) measures HR based on blood flow measurements. It is difficult to establish the measurement quality needed for HRV analysis, because the human circulation system acts as a filter for the heartbeat which pumps the blood. As a consequence, decision making systems that were trained with HR extracted from the ECG might have reduced accuracy when they are used to analyze directly measured HR. Sleep studies are needed which either produce labeled HR signals that are directly measured for the patient or recordings of both ECG and direct HR.

Through the review process, we found that the sleep-EDF database [124] on Physionet [125] has thus far been used in 10 studies. That database contains EEG, EOG, EMG, and respiration signals as well as body temperature. The data provides an excellent opportunity to advance sleep stage technology through cooperation and competition. A common dataset makes the sleep scoring results comparable. Unfortunately, there is no ECG database which has a similar prominence. The 'Sleep HR and SV Data Bank'[8] is also a publicly accessible database, but it has thus far been used in only one study. Granting public access to these databases is a step towards open science that leads to improved technology that can benefit a large number of individuals. However, both data amount and diversity of current databases are insufficient to create universal sleep scoring systems. A sustained effort is needed for remedy.

4. Conclusion

Physiological signals contain sleep stage related information. The task of a DSS is to extract and present this information to a human practitioner. Hence, physiological signals and their information content take center stage in sleep stage scoring. The emphasis on physiological signals is also justified by the fact that instrumentation effort, data rate, and cost differ greatly between the individual signals. In our review, we have found that all investigated physiological signals contain sleep stage related information. From this perspective, the current approach of measuring EEG, EOG, EMG and ECG in one PSG sitting makes sense – a maximal amount of information is obtained in a short period of time. However, some of this information is redundant, i.e. the ECG merely confirms information already extracted from an EEG signal. Redundancy however, is assistive in making a system reliable. For example, a human practitioner might miss a sleep stage transition in an EEG signal due to fatigue, but that expert might spot the transition in the ECG signal. However, that redundancy comes at the cost of expert labor and expensive equipment. The cost and the sheer inconvenience for the patient make recordings longer than one night impractical, even though longer recordings might reveal additional sleep disorders and therefore provide a fuller picture of the patient's sleep health. Patient led signal acquisition and DSS support can help to establish long-term unobtrusive sleep monitoring.

DSS can address issues of inter- and intra-observer variability, because an algorithm produces the same output from a given input regardless of space and time. Furthermore, these systems reduce the need for interpreting multiple signals, because they are immune to fatigue related signal misinterpretations. The need for redundancy can be addressed by observing the physiological signals during multiple sleep cycles. This has the added benefit that more sleep abnormalities can be detected. Furthermore, DSS systems can be made aware of the latest research findings via software and hardware updates, which is convenient and cost effective and can be helpful in tandem with training of human experts.

[7]Web page (last accessed 16.09.2018): `https://www.edn.com/design/analog/4442954/1/What-a-circuit-designer-needs-for-a-robust--wearable-health-sensor-system-design`

[8]Web page (last accessed 04/09/2018): `http://www.pri.kmu.lt/datbank/archiv.php`

As part of our future work, we propose to combine Artificial Intelligence (AI) and IoHT technology to create a HR based sleep scoring system. Using HR ensures patient comfort as well as a lower and therefore more manageable data rate. The signals are stored in a cloud server for traceability and continuous learning. The automated decision support is established with a deep learning system which takes account of all of the available data during the decision making process. We believe that any such a system will benefit patients in part by establishing a real-time sleep monitoring system, which provides constant feedback and emergency messages.

Acknowledgements

Funding: No external funding sources supported this review.

Competing interests: The authors declare that they have no conflict of interest.

Statements of ethical approval: For this review article, the authors did not undertake work that involved human participants or animals.

Acronyms

AASM American Academy of Sleep Medicine
ANS Autonomic Nervous System
ANN Artificial Neural Network
AI Artificial Intelligence
ANOVA Analysis Of Variance
BR Breathing Rate
CD Correlation Dimension
DFT Discrete Fourier Transform
DSS Decision Support System
DWT Discrete Wavelet Transform
ECG Electrocardiogram
EDF European Data Format
EEG Electroencephalogram
EMG Electromyogram
EOG Electrooculogram
GMM Gaussian Mixture Model
HMM Hidden Markov Model
HR Heart Rate
HRQoL Health-Related Quality of Life
HRV Heart Rate Variability
IoHT Internet of Health Things
LDA Linear Discriminant Analysis
NREM Non-REM
OSA Obstructive Sleep Apnea
PSG Polysomnography
PPV Pulse Pressure Variation
REM Rapid Eye Movement
SBLE Behaviour of Local Extrema
SVM Support Vector Machine
SWS Slow Wave Sleep
WNR Wake, Non-REM, REM

References

[1] M. M. Ohayon, Epidemiology of insomnia: what we know and what we still need to learn, Sleep medicine reviews 6 (2002) 97–111.

[2] S. Kajeepeta, B. Gelaye, C. L. Jackson, M. A. Williams, Adverse childhood experiences are associated with adult sleep disorders: a systematic review, Sleep medicine 16 (2015) 320–330.

[3] T. Willemen, D. Van Deun, V. Verhaert, M. Vandekerckhove, V. Exadaktylos, J. Verbraecken, S. Van Huffel, B. Haex, J. Vander Sloten, An evaluation of cardiorespiratory and movement features with respect to sleep-stage classification, IEEE journal of biomedical and health informatics 18 (2014) 661–669.

[4] M. M. Ohayon, S. Smirne, Prevalence and consequences of insomnia disorders in the general population of italy, Sleep Medicine 3 (2002) 115–120.

[5] C. M. Baldwin, K. A. Griffith, F. J. Nieto, G. T. O'connor, J. A. Walsleben, S. Redline, The association of sleep-disordered breathing and sleep symptoms with quality of life in the sleep heart health study, Sleep 24 (2001) 96–105.

[6] V. F. Helland, A. Gapelyuk, A. Suhrbier, M. Riedl, T. Penzel, J. Kurths, N. Wessel, Investigation of an automatic sleep stage classification by means of multiscorer hypnogram, Methods of information in medicine 49 (2010) 467–472.

[7] S. C. Bolge, J. F. Doan, H. Kannan, R. W. Baran, Association of insomnia with quality of life, work productivity, and activity impairment, Quality of life Research 18 (2009) 415.

[8] M. Daley, C. Morin, M. LeBlanc, J. Gregoire, J. Savard, L. Baillargeon, Insomnia and its relationship to health-care utilization, work absenteeism, productivity and accidents, Sleep medicine 10 (2009) 427–438.

[9] E. M. Wickwire, F. T. Shaya, S. M. Scharf, Health economics of insomnia treatments: The return on investment for a good night's sleep, Sleep medicine reviews 30 (2016) 72–82.

[10] R. J. Ozminkowski, S. Wang, J. K. Walsh, The direct and indirect costs of untreated insomnia in adults in the united states, Sleep 30 (2007) 263–273.

[11] A. Williamson, D. A. Lombardi, S. Folkard, J. Stutts, T. K. Courtney, J. L. Connor, The link between fatigue and safety, Accident Analysis & Prevention 43 (2011) 498–515.

[12] D. Léger, C. Guilleminault, G. Bader, E. Lévy, M. Paillard, Medical and socio-professional impact of insomnia, Sleep 25 (2002) 621–625.

[13] L. A. Panossian, A. Y. Avidan, Review of sleep disorders, Medical Clinics of North America 93 (2009) 407–425.

[14] O. Faust, U. R. Acharya, E. Ng, H. Fujita, A review of ecg-based diagnosis support systems for obstructive sleep apnea, Journal of Mechanics in Medicine and Biology 16 (2016) 1640004.

[15] S.-H. Wong, B.-Y. Ng, Review of sleep studies of patients with chronic insomnia at a sleep disorder unit, Singapore medical journal 56 (2015) 317.

[16] E. Lugaresi, Good and poor sleepers: an epidemiological survey of the san marino population, Sleep/wake disorders: natural history, epidemiology, and long-term evolution (1983) 1–12.

[17] D. Leger, B. Poursain, D. Neubauer, M. Uchiyama, An international survey of sleeping problems in the general population, Current medical research and opinion 24 (2008) 307–317.

[18] J. A. Hobson, A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects: A. rechtschaffen and a. kales (editors).(public health service, us government printing office, washington, dc, 1968, 58 p., $4.00), Clinical Neurophysiology 26 (1969) 644.

[19] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, C. Marcus, B. Vaughn, The aasm manual for the scoring of sleep and associated events, Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine (2012).

[20] W. H. Spriggs, Essentials of polysomnography, Jones & Bartlett Publishers, 2014.

[21] M. M. Steriade, R. W. McCarley, Brainstem control of wakefulness and sleep, Springer Science & Business Media, 2013.

[22] A. Rechtschaffen, A manual for standardized terminology, techniques and scoring system for sleep stages in human subjects, Brain information service (1968).

[23] A. Kales, A. Rechtschaffen, A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects, US Department of Health, Education and Welfare, Public Health Service, National Institutes of Health, National Institute of Neurological Diseases and Blindness, Neurological Information Network, 1968.

[24] M. H. Silber, S. Ancoli-Israel, M. H. Bonnet, S. Chokroverty, M. M. Grigg-Damberger, M. Hirshkowitz, S. Kapen, S. A. Keenan, M. H. Kryger, T. Penzel, The visual scoring of sleep in adults, Journal of Clinical Sleep Medicine 3 (2007) 22–22.

[25] T. Penzel, P.-G. Behler, M. Von Buttlar, R. Conradt, M. Meier, A. Möller, H. Danker-Hopfe, Reliablität der visuellen schlafauswertung nach rechtschaffen und kales von acht aufzeichnungen durch neun schlaflabore: Reliability of visual evaluation of sleep stages according to rechtschaffen and kales from eight polysomnographs by nine sleep centres, Somnologie 7 (2003) 49–58.

[26] H. Danker-Hopfe, D. Kunz, G. Gruber, G. Klösch, J. L. Lorenzo, S.-L. Himanen, B. Kemp, T. Penzel, J. Röschke, H. Dorn, Interrater reliability between scorers from eight european sleep laboratories in subjects with different sleep disorders, Journal of sleep research 13 (2004) 63–69.

[27] M. M. Ohayon, M. A. Carskadon, C. Guilleminault, M. V. Vitiello, Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan, Sleep 27 (2004) 1255–1273.

[28] O. Faust, R. J. Martis, L. Min, G. L. Z. Zhong, W. Yu, Automated detection of pulmonary edema and respiratory failure using physiological signals, Journal of Medical Imaging and Health Informatics 3 (2013) 424–431.

[29] M. Sharma, D. Goyal, P. Achuth, U. R. Acharya, An accurate sleep stages classification system using a new class of optimally time-frequency localized three-band wavelet filter bank, Computers in biology and medicine 98 (2018) 58–75.

[30] L. Doroshenkov, V. Konyshev, S. Selishchev, Classification of human sleep stages based on eeg processing using hidden markov models, Biomedical Engineering 41 (2007) 25–28.

[31] A. Flexer, G. Gruber, G. Dorffner, A reliable probabilistic sleep stager based on a single eeg signal, Artificial Intelligence in Medicine 33 (2005) 199–207.

[32] J. Shi, X. Liu, Y. Li, Q. Zhang, Y. Li, S. Ying, Multi-channel eeg-based sleep stage classification with joint collaborative representation and multiple kernel learning, Journal of neuroscience methods 254 (2015) 94–101.

[33] M. Diykh, Y. Li, P. Wen, Eeg sleep stages classification based on time domain features and structural graph similarity, IEEE Transactions on Neural Systems and Rehabilitation Engineering 24 (2016) 1159–1168.

[34] S. Seifpour, H. Niknazar, M. Mikaeili, A. M. Nasrabadi, A new automatic sleep staging system based on statistical behavior of local extrema using single channel eeg signal, Expert Systems with Applications 104 (2018) 277–293.

[35] K. Pillay, A. Dereymaeker, K. Jansen, G. Naulaers, S. Van Huffel, M. De Vos, Automated eeg sleep staging in the term-age baby using a generative modelling approach, Journal of neural engineering 15 (2018) 036004.

[36] F. Ebrahimi, S.-K. Setarehdan, H. Nazeran, Automatic sleep staging by simultaneous analysis of ecg and respiratory signals in long epochs, Biomedical Signal Processing and Control 18 (2015) 69–79.

[37] U. R. Acharya, E. C.-P. Chua, K. C. Chua, L. C. Min, T. Tamura, Analysis and automatic identification of sleep stages using higher order spectra, International journal of neural systems 20 (2010) 509–521.

[38] S. I. Dimitriadis, C. Salis, D. Linden, A novel, fast and efficient single-sensor automatic sleep-stage classification based on complementary cross-frequency coupling estimates, Clinical Neurophysiology 129 (2018) 815–828.

[39] T. L. da Silveira, A. J. Kozakevicius, C. R. Rodrigues, Single-channel eeg sleep stage classification based on a streamlined set of statistical features in wavelet domain, Medical & biological engineering & computing 55 (2017) 343–352.

[40] A. R. Hassan, M. I. H. Bhuiyan, A decision support system for automatic sleep staging from eeg signals using tunable q-factor wavelet transform and spectral features, Journal of neuroscience methods 271 (2016) 107–118.

[41] M. Čić, J. Šoda, M. Bonković, Automatic classification of infant sleep based on instantaneous frequencies in a single-channel eeg signal, Computers in biology and medicine 43 (2013) 2110–2117.

[42] U. R. Acharya, S. Bhat, O. Faust, H. Adeli, E. C.-P. Chua, W. J. E. Lim, J. E. W. Koh, Nonlinear dynamics measures for automated eeg-based sleep stage detection, European neurology 74 (2015) 268–287.

[43] M. Peker, An efficient sleep scoring system based on eeg signal using complex-valued machine learning algorithms, Neurocomputing 207 (2016) 165–177.

[44] P. Memar, F. Faradji, A novel multi-class eeg-based sleep stage classification system, IEEE Transactions on Neural Systems and Rehabilitation Engineering 26 (2018) 84–95.

[45] A. R. Hassan, A. Subasi, A decision support system for automated identification of sleep stages from single-channel eeg signals, Knowledge-Based Systems 128 (2017) 115–124.

[46] M. Ronzhina, O. Janoušek, J. Kolářová, M. Nováková, P. Honzík, I. Provazník, Sleep scoring using artificial neural networks, Sleep medicine reviews 16 (2012) 251–263.

[47] A. R. Hassan, M. I. H. Bhuiyan, Automated identification of sleep states from eeg signals by means of ensemble empirical mode decomposition and random under sampling boosting, Computer methods and programs in biomedicine 140 (2017) 201–210.

[48] V. Bajaj, R. B. Pachori, Automatic classification of sleep stages based on the time-frequency image of eeg signals, Computer methods and programs in biomedicine 112 (2013) 320–328.

[49] Y.-L. Hsu, Y.-T. Yang, J.-S. Wang, C.-Y. Hsu, Automatic sleep stage recurrent neural classifier using energy features of eeg signals, Neurocomputing 104 (2013) 105–114.

[50] C. Vural, M. Yildiz, Determination of sleep stage separation ability of features extracted from eeg signals using principle component analysis, Journal of medical systems 34 (2010) 83–89.

[51] A. Supratak, H. Dong, C. Wu, Y. Guo, Deepsleepnet: a model for automatic sleep stage scoring based on raw single-channel eeg, IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (2017) 1998–2008.

[52] N. Michielli, U. R. Acharya, F. Molinari, Cascaded lstm recurrent neural network for automated sleep stage classification using single-channel eeg signals, Computers in biology and medicine (2019).

[53] S. Mousavi, F. Afghah, U. R. Acharya, Sleepeegnet: Automated sleep stage scoring with sequence to sequence deep learning approach, arXiv preprint arXiv:1903.02108 (2019).

[54] P. Chriskos, C. A. Frantzidis, P. T. Gkivogkli, P. D. Bamidis, C. Kourtidou-Papadeli, Achieving accurate automatic sleep staging on manually pre-processed eeg data through synchronization feature extraction and graph metrics, Frontiers in human neuroscience 12 (2018) 110.

[55] B. Koley, D. Dey, An ensemble system for automatic sleep stage classification using single channel eeg signal, Computers in biology and medicine 42 (2012) 1186–1195.

[56] B. Şen, M. Peker, A. Çavuşoğlu, F. V. Çelebi, A comparative study on classification of sleep stage based on eeg signals using feature selection and classification algorithms, Journal of medical systems 38 (2014) 18.

[57] R. Acharya, O. Faust, N. Kannathal, T. Chua, S. Laxminarayan, Non-linear analysis of eeg signals at various sleep stages, Computer methods and programs in biomedicine 80 (2005) 37–45.

[58] U. R. Acharya, E. C.-P. Chua, O. Faust, T.-C. Lim, L. F. B. Lim, Automated detection of sleep apnea from electrocardiogram signals using nonlinear parameters, Physiological measurement 32 (2011) 287.

[59] Ş. Yücelbaş, C. Yücelbaş, G. Tezel, S. Özşen, Ş. Yosunkaya, Automatic sleep staging based on svd, vmd, hht and morphological features of single-lead ecg signal, Expert Systems with Applications 102 (2018) 193–206.

[60] K. Kesper, S. Canisius, T. Penzel, T. Ploch, W. Cassel, Ecg signal analysis for the assessment of sleep-disordered breathing and sleep pattern, Medical & biological engineering & computing 50 (2012) 135–144.

[61] M. Xiao, H. Yan, J. Song, Y. Yang, X. Yang, Sleep stages classification based on heart rate variability and random forest, Biomedical Signal Processing and Control 8 (2013) 624–633.

[62] S. J. Redmond, C. Heneghan, Cardiorespiratory-based sleep staging in subjects with obstructive sleep apnea, IEEE Transactions on Biomedical Engineering 53 (2006) 485–496.

[63] S. J. Redmond, P. de Chazal, C. O'Brien, S. Ryan, W. T. McNicholas, C. Heneghan, Sleep staging using cardiorespiratory signals, Somnologie-Schlafforschung und Schlafmedizin 11 (2007) 245–256.

[64] J. Fell, K. Mann, J. Röschke, M. Gopinathan, Nonlinear analysis of continuous ecg during sleep i. reconstruction, Biological cybernetics 82 (2000) 477–483.

[65] J. Fell, K. Mann, J. Röschke, M. Gopinathan, Nonlinear analysis of continuous ecg during sleep ii. dynamical measures, Biological cybernetics 82 (2000) 485–491.

[66] R. P. Bartsch, A. Y. Schumann, J. W. Kantelhardt, T. Penzel, P. C. Ivanov, Phase transitions in physiologic coupling, Proceedings of the National Academy of Sciences 109 (2012) 10181–10186.

[67] M. Kryger, T. Roth, W. Dement, Principles and Practice of Sleep Medicine, Saunders W.B, Elsevier, 2016.

[68] N. Chattipakorn, T. Incharoen, N. Kanlop, S. Chattipakorn, Heart rate variability in myocardial infarction and heart failure, International journal of cardiology 120 (2007) 289–296.

[69] O. Faust, V. R. Prasad, G. Swapna, S. Chattopadhyay, T.-C. Lim, Comprehensive analysis of normal and diabetic heart rate signals: A review, Journal of Mechanics in Medicine and Biology 12 (2012) 1240033.

[70] U. R. Acharya, O. Faust, S. V. Sree, D. N. Ghista, S. Dua, P. Joseph, V. T. Ahamed, N. Janarthanan, T. Tamura, An integrated diabetic index using heart rate variability signal features for diagnosis of diabetes, Computer methods in biomechanics and biomedical engineering 16 (2013) 222–234.

[71] T. S. Hock, O. Faust, T.-C. Lim, W. Yu, Automated detection of premature ventricular contraction using recurrence quantification analysis on heart rate signals, Journal of Medical Imaging and Health Informatics 3 (2013) 462–469.

[72] U. R. Acharya, O. Faust, V. Sree, G. Swapna, R. J. Martis, N. A. Kadri, J. S. Suri, Linear and nonlinear analysis of normal and cad-affected heart rate signals, Computer methods and programs in biomedicine 113 (2014) 55–68.

[73] U. R. Acharya, O. Faust, N. A. Kadri, J. S. Suri, W. Yu, Automated identification of normal and diabetes heart rate signals using nonlinear measures, Computers in biology and medicine 43 (2013) 1523–1529.

[74] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, P. J. Schwartz, Heart rate variability: Standards of measurement, physiological interpretation, and clinical use, European heart journal 17 (1996) 354–381.

[75] A. Camm, M. Malik, J. Bigger, G. Breithardt, S. Cerutti, R. Cohen, P. Coumel, E. Fallen, H. Kennedy, R. Kleiger, Heart rate variability: standards of measurement, physiological interpretation and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology, Circulation 93 (1996) 1043–1065.

[76] O. Faust, R. Acharya, S. Krishnan, L. C. Min, Analysis of cardiac signals using spatial filling index and time-frequency domain, BioMedical Engineering OnLine 3 (2004) 30.

[77] M. Kryger, T. Roth, W. Dement, Principles and Practice of Sleep Medicine - E-Book: Expert Consult Premium Edition - Enhanced Online Features, Elsevier Health Sciences, 2010.

[78] H. Yoon, S. H. Hwang, J.-W. Choi, Y. J. Lee, D.-U. Jeong, K. S. Park, Rem sleep estimation based on autonomic dynamics using r–r intervals, Physiological measurement 38 (2017) 631.

[79] T. Penzel, J. W. Kantelhardt, L. Grote, J.-H. Peter, A. Bunde, Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea, IEEE Transactions on biomedical engineering 50 (2003) 1143–1151.

[80] P. K. Stein, Y. Pu, Heart rate variability, sleep and sleep disorders, Sleep medicine reviews 16 (2012) 47–66.

[81] J. Trinder, J. Kleiman, M. Carrington, S. Smith, S. Breen, N. Tan, Y. Kim, Autonomic activity during human sleep as a function of time and sleep stage, Journal of sleep research 10 (2001) 253–264.

[82] M. de Zambotti, A. R. Willoughby, F. C. Baker, D. S. Sugarbaker, I. M. Colrain, Cardiac autonomic function during sleep: effects of alcohol dependence and evidence of partial recovery with abstinence, Alcohol 49 (2015) 409–415.

[83] S. Liu, J. Teng, X. Qi, S. Wei, C. Liu, Comparison between heart rate variability and pulse rate variability during different sleep stages for sleep apnea patients, Technology and Health Care 25 (2017) 435–445.

[84] I. Virtanen, E. Ekholm, P. Polo-Kantola, H. Huikuri, Sleep stage dependent patterns of nonlinear heart rate dynamics in postmenopausal women, Autonomic Neuroscience 134 (2007) 74–80.

[85] V. Crasset, S. Mezzetti, M. Antoine, P. Linkowski, J. P. Degaute, P. Van De Borne, Effects of aging and cardiac denervation on heart rate variability during sleep, Circulation 103 (2001) 84–88.

[86] O. Faust, L. M. Yi, L. M. Hua, Heart rate variability analysis for different age and gender, Journal of Medical Imaging and Health Informatics 3 (2013) 395–400.

[87] M. O. Mendez, M. Matteucci, V. Castronovo, L. Ferini-Strambi, S. Cerutti, A. Bianchi, Sleep staging from heart rate variability: time-varying spectral features and hidden markov models, International Journal of Biomedical Engineering and Technology 3 (2010) 246–263.

[88] S.-F. Liang, C.-E. Kuo, Y.-C. Lee, W.-C. Lin, Y.-C. Liu, P.-Y. Chen, F.-Y. Cherng, F.-Z. Shaw, Development of an eog-based automatic sleep-monitoring eye mask, IEEE Transactions on Instrumentation and Measurement 64 (2015) 2977–2985.

[89] J. Virkkala, J. Hasan, A. Värri, S.-L. Himanen, K. Müller, Automatic sleep stage classification using two-channel electro-oculography, Journal of neuroscience methods 166 (2007) 109–115.

[90] M. M. Rahman, M. I. H. Bhuiyan, A. R. Hassan, Sleep stage classification using single-channel eog, Computers in Biology and Medicine (2018).

[91] T. Penzel, N. Wessel, M. Riedl, J. W. Kantelhardt, S. Rostig, M. Glos, A. Suhrbier, H. Malberg, I. Fietze, Cardiovascular and respiratory dynamics during normal and pathological sleep, Chaos: An Interdisciplinary Journal of Nonlinear Science 17 (2007) 015116.

[92] N. Douglas, D. White, C. K. Pickett, J. Weil, C. Zwillich, Respiration during sleep in normal man., Thorax 37 (1982) 840–844.

[93] V. K. Somers, M. E. Dyken, A. L. Mark, F. M. Abboud, Sympathetic-nerve activity during sleep in normal subjects, New England Journal of Medicine 328 (1993) 303–307.

[94] X. Long, J. Foussier, P. Fonseca, R. Haakma, R. M. Aarts, Analyzing respiratory effort amplitude for automated sleep stage classification, Biomedical Signal Processing and Control 14 (2014) 197–205.

[95] S. H. Sheldon, M. H. Kryger, R. Ferber, D. Gozal, Principles and Practice of Pediatric Sleep Medicine E-Book, Elsevier Health Sciences, 2014.

[96] A. A. of Sleep Medicine Task Force, Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research, Sleep 22 (1999) 667–689.

[97] S.-F. Liang, C.-E. Kuo, Y.-H. Hu, Y.-S. Cheng, A rule-based automatic sleep staging method, Journal of neuroscience methods 205 (2012) 169–176.

[98] M. E. Tagluk, N. Sezgin, M. Akin, Estimation of sleep stages by an artificial neural network employing eeg, emg and eog, Journal of medical systems 34 (2010) 717–725.

[99] A. Kishi, H. Yasuda, T. Matsumoto, Y. Inami, J. Horiguchi, M. Tamaki, Z. R. Struzik, Y. Yamamoto, Nrem sleep stage transitions control ultradian rem sleep rhythm, Sleep 34 (2011) 1423–1432.

[100] R. S. Leung, Sleep-disordered breathing: autonomic mechanisms and arrhythmias, Progress in cardiovascular diseases 51 (2009) 324–338.

[101] F. Tracik, G. Ebersbach, Sudden daytime sleep onset in parkinson's disease: polysomnographic recordings, Movement disorders: official journal of the Movement Disorder Society 16 (2001) 500–506.

[102] C. A. Kushida, A. Chang, C. Gadkary, C. Guilleminault, O. Carrillo, W. C. Dement, Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients, Sleep medicine 2 (2001) 389–396.

[103] H. E. Montgomery-Downs, L. M. O'Brien, T. E. Gulliver, D. Gozal, Polysomnographic characteristics in normal preschool and early school-aged children, Pediatrics 117 (2006) 741–753.

[104] X. Long, P. Fonseca, J. Foussier, R. Haakma, R. M. Aarts, Sleep and wake classification with actigraphy and respiratory effort using dynamic warping, IEEE journal of biomedical and health informatics 18 (2014) 1272–1284.

[105] T. Kirjavainen, D. Cooper, O. Polo, C. SULLIVAN, Respiratory and body movements as indicators of sleep stage and wakefulness in infants and young children, Journal of sleep research 5 (1996) 186–194.

[106] R. Tripathy, U. R. Acharya, Use of features from rr-time series and eeg signals for automated classification of sleep stages in deep neural network framework, Biocybernetics and Biomedical Engineering 38 (2018) 890–902.

[107] O. Yildirim, U. B. Baloglu, U. R. Acharya, A deep learning model for automated sleep stages classification using psg signals, International journal of environmental research and public health 16 (2019) 599.

[108] C. E. Shannon, A mathematical theory of communication, ACM SIGMOBILE mobile computing and communications review 5 (2001) 3–55.

[109] T. Takatani, Y. Takahashi, R. Yoshida, R. Imai, T. Uchiike, M. Yamazaki, M. Shima, T. Nishikubo, Y. Ikada, S. Fujimoto, Relationship between frequency spectrum of heart rate variability and autonomic nervous activities during sleep in newborns, Brain and Development 40 (2018) 165–171.

[110] P. Fonseca, X. Long, M. Radha, R. Haakma, R. M. Aarts, J. Rolink, Sleep stage classification with ecg and respiratory effort, Physiological measurement 36 (2015) 2027.

[111] M. Kesek, K. A. Franklin, C. Sahlin, E. Lindberg, Heart rate variability during sleep and sleep apnoea in a population based study of 387 women, Clinical physiology and functional imaging 29 (2009) 309–315.

[112] P. Estévez, C. Held, C. Holzmann, C. Perez, J. Pérez, J. Heiss, M. Garrido, P. Peirano, Polysomnographic pattern recognition for automated classification of sleep-waking states in infants, Medical and Biological Engineering and Computing 40 (2002) 105–113.

[113] E. Piva, M. Pelloso, L. Penello, M. Plebani, Laboratory critical values: automated notification supports effective clinical decision making, Clinical biochemistry 47 (2014) 1163–1168.

[114] A. Patterson, P. Bonissone, M. Pavese, Six sigma applied throughout the lifecycle of an automated decision system, Quality and Reliability Engineering International 21 (2005) 275–292.

[115] O. Faust, W. Yu, U. R. Acharya, The role of real-time in biomedical science: A meta-analysis on computational complexity, delay and speedup, Computers in biology and medicine 58 (2015) 73–84.

[116] S. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, K.-S. Kwak, The internet of things for health care: a comprehensive survey, IEEE Access 3 (2015) 678–708.

[117] O. Faust, A. Shenfield, M. Kareem, T. R. San, H. Fujita, U. R. Acharya, Automated detection of atrial fibrillation using long short-term memory network with rr interval signals, Computers in biology and medicine (2018).

[118] O. Faust, U. R. Acharya, T. Tamura, Formal design methods for reliable computer-aided diagnosis: a review, IEEE reviews in biomedical engineering 5 (2012) 15–28.

[119] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, U. R. Acharya, Deep learning for healthcare applications based on physiological signals: a review, Computer methods and programs in biomedicine (2018).

[120] J. Pan, W. J. Tompkins, A real-time qrs detection algorithm, IEEE Trans. Biomed. Eng 32 (1985) 230–236.

[121] P. Macfarlane, E. Coleman, Resting 12-lead ecg electrode placement and associated problems, Soc Cardiological Technicians Update 2 (1995) 10.

[122] J. García-Niebla, P. Llontop-García, J. I. Valle-Racero, G. Serra-Autonell, V. N. Batchvarov, A. B. De Luna, Technical mistakes during the acquisition of the electrocardiogram, Annals of Noninvasive Electrocardiology 14 (2009) 389–403.

[123] B.-U. Kohler, C. Hennig, R. Orglmeister, The principles of software qrs detection, IEEE Engineering in Medicine and Biology Magazine 21 (2002) 42–57.

[124] B. Kemp, A. H. Zwinderman, B. Tuk, H. A. Kamphuisen, J. J. Oberye, Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the eeg, IEEE Transactions on Biomedical Engineering 47 (2000) 1185–1194.

[125] G. B. Moody, R. G. Mark, A. L. Goldberger, Physionet: a web-based resource for the study of physiologic signals, IEEE Engineering in Medicine and Biology Magazine 20 (2001) 70–75.