Received 27 June 2022, accepted 17 July 2022, date of publication 27 July 2022, date of current version 3 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3194145

SURVEY

A Systematic Review of Literature on Automated Sleep Scoring

HADEEL ALSOLAI[1], SHAHNAWAZ QURESHI[2], SYED MUHAMMAD ZEESHAN IQBAL[3], SIRIRUT VANICHAYOBON[4], LAWRENCE EDWARD HENESEY[5,6], CRAIG LINDLEY[7], AND SEPPO KARRILA[8]

[1]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[2]Department of Computer Science, National University of Computing and Emerging Sciences, Fasialabad 38000, Pakistan
[3]Research and Development, BrightWare LLC, Riyadh 13323, Saudi Arabia
[4]Department of Computer Science, Prince of Songkla University, Songkhla 90110, Thailand
[5]Blekinge Institute of Technology, 37424 Karlshamn, Sweden
[6]DP World, 37424 Karlshamn, Sweden
[7]CSIRO Data 61, Pullenvale, QLD 4069, Australia
[8]Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Makham Tia, Muang, Surat Thani 84000, Thailand

Corresponding authors: Hadeel Alsolai (haalsolai@pnu.edu.sa); Shahnawaz Qureshi (Shahnawaz.qureshi@nu.edu.pk); and Sirirut Vanichayobon (sirirut.v@psu.ac.th)

This work was supported by the Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R303), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

ABSTRACT Sleep is a period of rest that is essential for functional learning ability, mental health, and even the performance of normal activities. Insomnia, sleep apnea, and restless legs are all examples of sleep-related issues that are growing more widespread. When appropriately analyzed, the recording of bio-electric signals, such as the Electroencephalogram, can tell how well we sleep. Improved analyses are possible due to recent improvements in machine learning and feature extraction, and they are commonly referred to as automatic sleep analysis to distinguish them from sleep data analysis by a human sleep expert. This study outlines a Systematic Literature Review and the results it provided to assess the present state-of-the-art in automatic analysis of sleep data. A search string was organized according to the PICO (Population, Intervention, Comparison, and Outcome) strategy in order to determine what machine learning and feature extraction approaches are used to generate an Automatic Sleep Scoring System. The American Academy of Sleep Medicine and Rechtschaffen & Kales are the two main scoring standards used in contemporary research, according to the report. Other types of sensors, such as Electrooculography, are employed in addition to Electroencephalography to automatically score sleep. Furthermore, the existing research on parameter tuning for machine learning models that was examined proved to be incomplete. Based on our findings, different sleep scoring standards, as well as numerous feature extraction and machine learning algorithms with parameter tuning, have a high potential for developing a reliable and robust automatic sleep scoring system for supporting physicians. In the context of the sleep scoring problem, there are evident gaps that need to be investigated in terms of automatic feature engineering techniques and parameter tuning in machine learning algorithms.

INDEX TERMS Artificial neural network, deep learning, automatic sleep scoring system, big data, feature extraction, inter-rater variability, machine learning, sleep stages.

I. INTRODUCTION

Sleep is defined as the absence of alertness and is regarded essential for a person's ability to learn, mental health, and even the everyday activities. During sleep, the body's major organs closely coordinate with one another, and this impacts the sleep at any given time. Sleep-related issues, including insomnia, sleep apnea, and restless legs, are becoming more widespread in the society, despite the fact that humans spend one-third of their lives sleeping.

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

Bio-signals are used to track electrical activity in the human brains. Electromyography (EMG), Electroencephalography (EEG), Electrooculography (EOG), and Electrocardiography (ECG) are the most widely used types of such signals. Large volumes of data are generated by these recordings, and are stored in data archives, whether public or private. A signal trace summarizes changes in the signal's properties, analyzed from representative features that collect and process data on key characteristics of the signal. Statistical and machine learning tools can learn to discern complicated patterns in the data, and to assist in making informed choices.

This study began by reviewing sleep-related publications published since 1968, in order to acquire insights for future research on automatic sleep grading. Sleep stages, computational approaches, machine learning, and in particular selection or extraction of features were all discussed. The study's systematic search process, inclusion and exclusion criteria, data extraction and synthesis, and data analysis and synthesis, are all detailed in a framework that allows replication of this part of the study.

The review's findings provide new insights into the dimensions that are frequently used in the development of automatic sleep grading systems. The American Academy of Sleep Medicine and Rechtschaffen & Kales are two sleep scoring standards that have been published in the literature. The following domains were investigated using various feature extraction techniques: (I)Time, (ii) Frequency, (iii) Time and Frequency, and (iv) Non-Linear and Entropy Domain. Finally, various machine learning methods are assembled based on their purpose, features used, number of Sleep Stages, dataset, and data accessibility, and their prediction accuracy is evaluated.

A. MOTIVATION

1) AIMS AND OBJECTIVES

The advances in signal processing, computer science, and statistical techniques incorporated in open source and simple data analysis tools have the potential to revolutionize the neuroscience field, particularly the understanding of sleep signal data. Rapid advances are ongoing in data mining, machine learning, artificial intelligence, and digital signal processing. However, the fields of signal processing and machine learning are diverse; therefore, many different algorithms, theories, and methods are available. This appears to be an obstacle in the adoption of these sophisticated tools by many sleep data professionals, which could limit the use of the large amounts of data accessible. According to the above arguments, this paper aims to:

- Present feature extraction techniques for bio signals; propose an overall structure for them; and discuss their applications to diagnosing sleep related problems.
- Present different machine learning techniques and provide advantages and disadvantages in the context of automated sleep scoring.
- Discuss from automatic sleep scoring perspective
  - the clinical acceptance of automated methods
  - understanding the inter-rater reliability of human scoring
- Discuss challenges or critical issues in using automated methods in clinical practice, with further extended use in a home environment

The next section discusses the current issues that automatic sleep scoring encounters. This lays the foundation for a later discussion of the importance of sleep disorders, sleep scoring standards, automated feature extraction approaches, and machine-learning as a tool for sleep data analysts to confront such difficulties.

2) CHALLENGES OF AUTOMATIC SLEEP SCORING

Sleep medicine is among the well-established fields; however, the importance of automatic sleep scoring is not rated high enough. It has grown in importance and is now part of the standard of care in the field of health sciences. It enabled the establishment of a small clinical unit to monitor patients with various cardiac, respiratory, and metabolic problems while they slept [1]. The discovery of electroencephalogram and sleep stages is directly linked with modern sleep medicine. Several educational programs to revamp the sleep medicine study were started, e.g., a survey to assess the current (2013) state of sleep medicine educational resources offered in the US [2].

Key challenges have been identified as follows.

- Sleep disorder analysis and scoring standards.
- Utilization of advanced knowledge in interdisciplinary fields
- Collaboration between Academia and Industry to adopt new technologies.
- Large and complex datasets, and difficult problems in their analysis
- Application of signal-processing methods (i.e., feature extraction techniques) in bio-signal analysis
- New methods or models combining (advanced) statistical, signal processing, and machine learning approaches.

These key challenges highlight the trend of increasing dynamic complexity. Adding to the challenges, there is relative lack of scientific experimentation on sleep, although data sets are available, but the high dimensionality and variety of data, as well as the NP completeness in model training, present challenges.

To overcome the major challenges in this complex domain, candidate aspects with high potential include sleep scoring standards, feature extraction alternatives, and machine learning tools. These techniques support finding highly complex and non-linear patterns in data of various types. Further, the raw data need to be converted to features for classification, prediction, regression or forecasting.

B. CONTRIBUTION AND SIGNIFICANCE OF THE STUDY

The major goal of this report is to present a systematic literature review that explores scientific machine learning, feature extraction and selection, and big data published in the context of Automatic Sleep Scoring Systems, given the importance of sleep and related difficulties, obstacles faced by automatic sleep scoring, and driven by data availability on bio-electric signals, as well as promising developments in the computational artificial intelligence techniques.

II. RESEARCH METHODOLOGY

This article uses Systematic Literature Review (SLR) technique, in order to ensure impartial search and study selection.

It is characterized as a research method that aims to gather all empirical evidence in a certain topic, evaluate the material, and synthesize new results. This SLR adheres to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) quality reporting criteria [3]. The SLR process is presented in Figure 1.

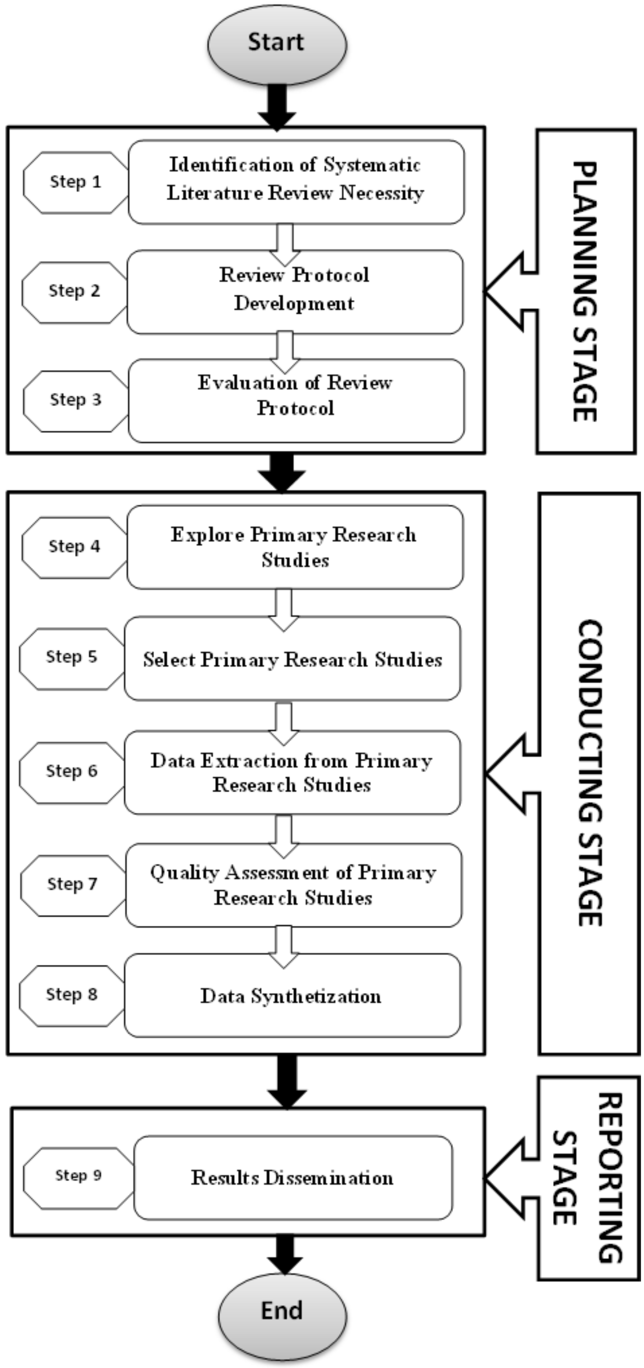

FIGURE 1. Steps in Systematic Literature Review Process.

A. RESEARCH QUESTIONS

A set of research questions addressed in this study is listed in Table 1.

B. SEARCH STRATEGY

1) SEARCH TERMS

The search string was generated following the PICO approach [3], by decomposing the scope of review to its population of interest, intervention, comparison, and outcome, see Table 2.

The following keywords were used in these searches: Human Sleep Stages, Sleep Scoring Standard, Sleep EEG, Automatic Sleep Stages, Feature Extraction Techniques, Sleep EEG, Classification of Sleep Stages, Machine Learning Techniques for Sleep Stages. Search Query Language expressions were used to find the relevant articles, for example EEG features extraction OR machine learning AND sleep stages'.

2) SEARCH PROCESS

We found 295 articles matching the search parameters, with 223 of them being non-duplicates that were chosen for full-text inspection. Twenty-five articles were discarded following title/abstract screening, another 22 were discarded in screening based on full text due to insignificance, and 19 were eliminated during data extraction. We found 157 articles that were chosen for full-text examination, with 130 articles included based on qualitative evaluation and 27 on quantitative synthesis, see Figure 2.

C. INCLUSION/EXCLUSION IN STUDY

The quality of this new group of studies was evaluated by qualitative review of the article title, abstract, and keywords as per established inclusion and exclusion criteria. As a result, the papers that met the assessment criteria are included in the study sets, see Table 3. Further, the quality evaluation is further guided by a set of questions, see Table 4

D. FRAMEWORK FOR DATA RETRIEVAL AND SYNTHESIS

Following the selection of studies, this is an important step in which the study's assessment criteria are developed. We provide a separate quality assessment methodology for each

TABLE 1. Research Questions.

RQ#	Research Questions	Motivation
RQ1	Which standards are used for sleep stage scoring?	Identify standards commonly used for sleep scoring
RQ2	What are the types of sleep stages?	Identify the reported sleep stages in the literature.
RQ3	What are the types of feature extraction techniques?	Explore the types of feature extraction techniques to analyze the signals.
RQ3.1	Which feature extraction techniques are used for ASS?	Identify the feature extraction techniques used for sleep scoring.
RQ3.2	What are the domains of hand engineering features?	Investigate the reported domains of hand engineering features.
RQ3.3	What is the 'without hand engineering feature' technique?	Investigate the automatic feature learning techniques
RQ3.4	What are the pros, cons or challenges of time and frequency, non-linear entropy based features?	Performance analysis of each feature extraction technique in context of ASS.
RQ4	Which machine learning techniques are used for ASS?	Explore the types of machine learning techniques used for ASS.
RQ4.1	What are the advantages or disadvantages or challenged of machine learning techniques?	Investigate the advantages / disadvantage of each mostly used machine learning techniques
RQ4.2	What are the application areas of machine learning techniques in ASS?	Determine the machine learning application areas in ASS
RQ5	What kind of the deep learning techniques used in ASS?	Identify the deep learning techniques used in ASS.
RQ5.1	What are the strengths and weaknesses of deep learning techniques?	Investigate the advantages and disadvantages of deep learning techniques.
RQ5.2	What are the challenges deep learning techniques?	Determine the challenges faced by deep learning techniques.
RQ6	What is the role of big data in sleep science?	Investigate the role of big data in sleep science.
RQ7	What is the inter-rater variability of manual and automated sleep scoring?	Investigate the inter-rater variability of manual and ASS.

TABLE 2. Decomposition of Search Keywords using PICO approach.

PICO	
Population of Interest	The studies focused on sleep **EEG** analysis, Sleep Stages and automatic sleep scoring systems
Intervention	Machine learning techniques, Feature Extraction Techniques
Comparison	Accuracies of used classifiers. Parameter Tuning, Dataset size
Outcome	Classification of sleep stages for an individual or subject
Population Search	The studies focused on sleep **EEG** analysis, Sleep Stages and automatic sleep scoring systems
Population Search: Sleep Scoring	
Subject Headings	Human sleep scoring, Automatic sleep stages/scoring
Key words	(Sleep standards* **OR** Sleep disorders OR Sleep Scoring* **OR** Sleep stages* **OR** Automatic sleep **OR** Sleep Brain **OR** Human sleep) **OR** (Sleep EEG* **OR** Sleep signal **OR** Signal processing techniques **OR** Feature extraction techniques* **OR** EEG signal processing* **OR** sleep EEG features) **OR** (Machine learning techniques ***OR** Machine learning methods **OR** Sleep expert system* **OR** Sleep Neural network *)
Intervention Search: Sleep disorders identification	
Subject Headings	Sleep disorders
Key words	Somnipathy* **OR** Insomnia* **OR** narcolepsy* **OR** sleep apnea* **OR** hypersomnia* **OR** sleep walking* **OR** night terror * **OR** bed wetting*

research topic in order to answer it, which could be useful for a new researcher starting out in this subject.

Because the human brain goes through various stages of sleep, it's important to keep track of the criteria for each stage as well as the related brain signals. As a result, information regarding sleep staging criteria and EEG signal qualities is extracted, see Figure 3.

A number of computerized analysis approaches are based on the concept that the EEG signal is generated by a highly sophisticated linear system, resulting in non-stationary or unpredictable features. The signal, on the other hand, could alternatively stem from a deterministic system with a low level of complexity but a lot of nonlinear features. As a result, in order to answer the second question, we must first identify the different types of feature extraction methodologies. The data synthesis and processing framework is shown in Figure 3.

Machine learning algorithms can learn to perform a task from a series of examples and after such training the equivalent actions can be applied to a new data set. The problem of sleep EEG has been addressed using a variety of ways. A data synthesis and processing framework is created to address the third research question, see Figure 3.

E. THREATS TO VALIDITY

A systematic literature review starts with a complete literature search of all the relevant studies from the major bibliographic databases, after identifying the research questions. The searches in this study were formulated using the PICO

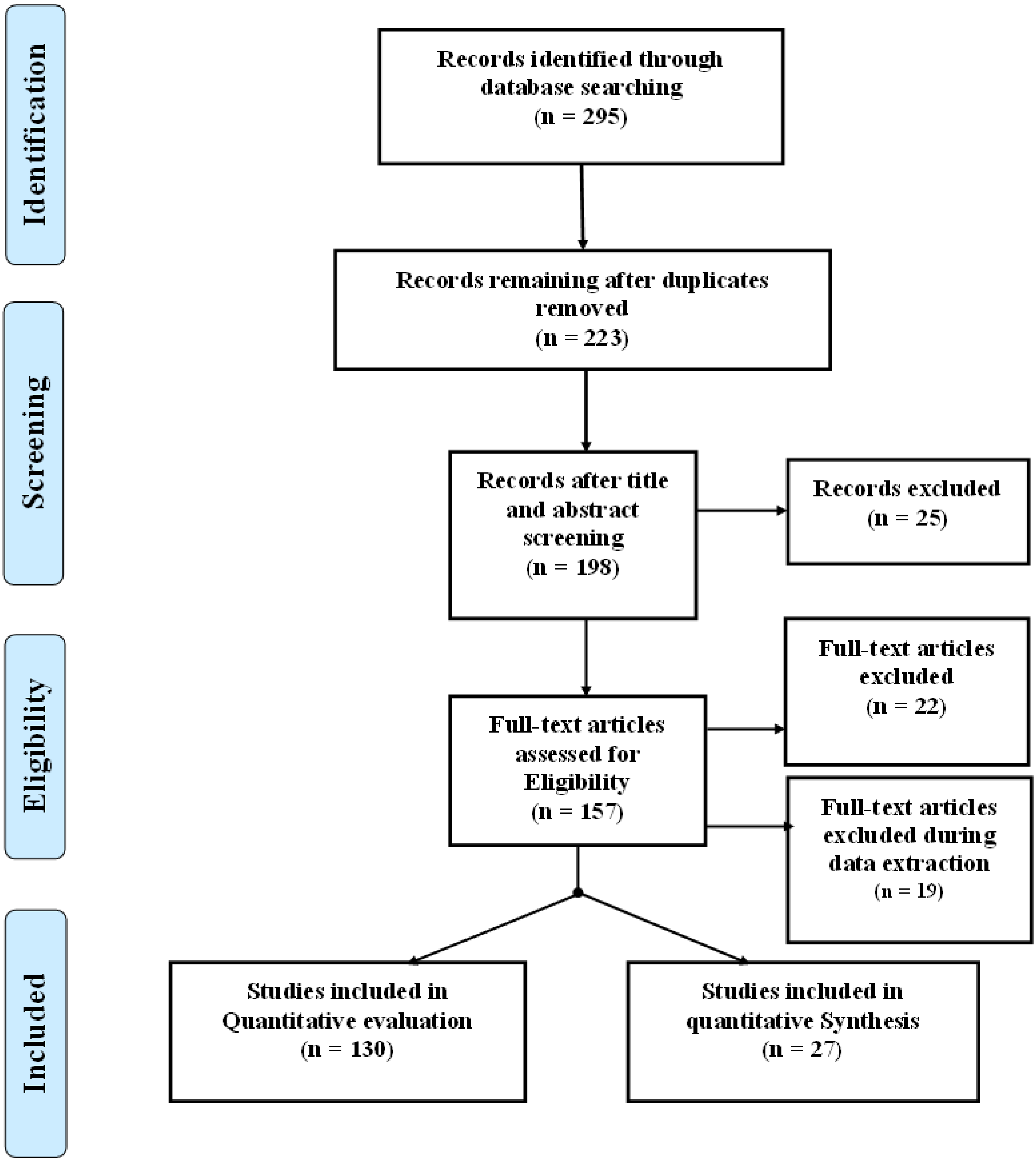

FIGURE 2. PRISMA Flow Diagram for the selection of studies on automated sleep scoring.

technique, which includes specifying search keywords and examining the numbers of results returned.

The eligibility criteria for study selection must be defined. To address this danger to validity, a set of inclusion and exclusion criteria has been established.

What criteria were used to evaluate the quality of each study? Data extraction and synthesis frameworks are defined to address the research issues in order to answer this query. The discrepancies between studies should be described.

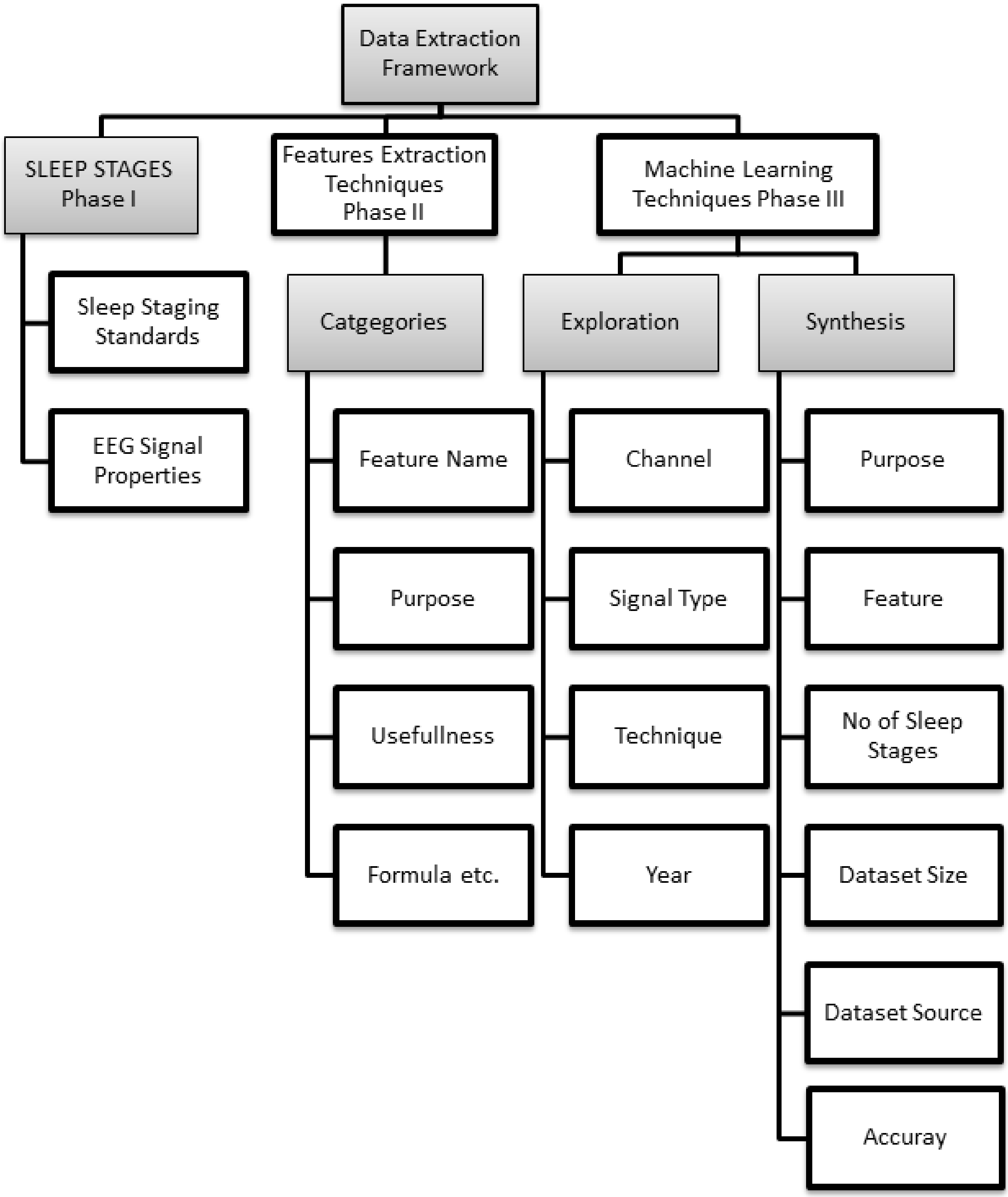

FIGURE 3. Quality Assessment Framework and Data Synthesis.

III. STRUCTURING AND EXPLANATION OF SLEEP SCORING STANDARDS AND SLEEP STAGES

The SLR identified the sleep scoring standards, Rechtschaffen & Kales (R & K) [4] and American Academy of Sleep Medicine (AASM) [5]. The former has six sleep stages while the latter has five of them, defined by brain activity characteristics. Neuroscientists have identified various brain signals associated with these stages, named as Alpha, Beta, Delta, and Theta waves that can distinguish the stages (see Figure 4).

According to R&K rule [4] these stages are classified as:

- **Wake (W)** define feeling relaxed, fall asleep quickly or in less than 10 minutes.

TABLE 3. Criteria for Inclusion or Exclusion from Study.

Criteria	Description
Inclusion	Only English language articles were considered.
	Book / Journals / Conference / Report related to automatic sleep scoring were included.
	Citation factor was involved in filter article.
	How to collect the raw sleep EEG data?
	How many subjects/participants that were involved in the experiment?
	Whether publicly available on-line data set is used or data collected through in-house experiment?
	Data interpretation methods were also considered.
	Parameter tuning was also considered
Exclusion	Any article related to EEG but not specific for sleep EEG data.
	Any article title/abstract that did not match sleep EEG data
	Papers with incomplete information about sleep EEG data acquiring.
	Fake publisher's studies according to Beall's list.

TABLE 4. A set of questions for quality evaluation.

Q#	Quality evaluation question
Q1	Are the research aims and objectives expressed clearly?
Q2	Is the size of the data set appropriate?
Q3	Is the technique for gathering data well-defined?
Q4	Is the use of machine learning techniques justified?
Q5	Is the machine learning techniques well-defined?
Q6	Are the performance metrics for evaluating ASS models well-defined?
Q7	Are the outcomes and conclusions conveyed clearly?
Q8	Have the study's shortcomings been stated?
Q9	Is the research technique repeatable?
Q10	Is the research repeatable?
Q11	Has a comparison analysis been done?
Q12	Has a comparative analysis of ML techniques been done?
Q13	Does the study add to or contribute to the existing literature?

- **STAGE 1: Non-rapid eye movement (N-REM)** Refers to very light sleep (feeling like in a cloud, when hearing a noise in the house or room not feeling like responding to it but can still understand overheard conversation). The quick beta waves of awareness are replaced by slower alpha waves, and the slower theta wave emerges after a period of falling asleep.
- **STAGE 2: N-REM** state light sleep indicates that the subject can still hear but cannot understand speech. The EEG signals continue to decrease in frequency while increasing in amplitude during this period of light sleep. Burst activity known as sleep spindles disrupts theta waves, which have a frequency of 8-14 Hz. During sleep stage 2, the K-complex, or fast and high amplitude waves, can be seen.
- **STAGE 3: N-REM** belongs to deep stage (subject no longer hears anything, cuts off the world). During the third stage of N-REM, delta waves occur on the EEG. Sleep spindle and K-complex waves do appear, but they are less frequent than in stage 2.

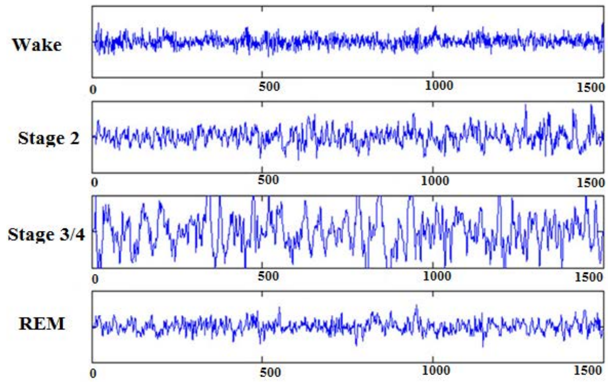

FIGURE 4. Signals related to sleep stages.

- **STAGE 4: N-REM** shows sleeping deeply. In stage 4, delta waves are influential, and overall neural activity is at its lowest. The range of frequencies is less than 2Hz.
- **Rapid eye movement (REM):** REM refers to dream sleep, in which the brain recharges its battery and records what it has learned during the day. It is distinguished by theta, beta, and gamma frequencies of 4-8, 16-32, and >32 Hertz, respectively.

However, AASM [5] defined sleep stages as W, S1, S2, S3 and S4 instead of N-REM stage1, N-REM stage 2, N-REM stage 3 and N-REM stage 4 respectively. Meanwhile the representations of stages N-REM stage 3 and N-REM stage 4 are identical; the AASM merged stage 3 and stage 4 into deep sleep or to the slow wave sleep (S3) stage. The standards and characteristics of EEG signals associated with each sleep stage are given in Table 5.

Besides the simplification of the sleep stage classification problem, the following challenges are faced [6]: healthy to unhealthy subjects ratio, test or validation test dataset size, class imbalance problem, visual inspection time, human error in the manual annotation, and inter-rater reliability, etc.

One noteworthy discovery is that the R&K sleep grading standards have been used in research for decades. Another significant conclusion is that, following the release of the revised sleep scoring standard by AASM in 2007, the research community is split between the two standards. As a result, fewer sleep datasets are rated with AASM than with R&K. The use of AASM in automatic sleep scoring research is highlighted in the publications evaluated for the literature analysis. This observation is explained by the fact that studying human brain events is a difficult undertaking that necessitates a large amount of work in order to get relevant insights from brain signal data. The bulk of the investigations used EEG signals instead of EMG or EOG signals, which is a noteworthy finding.

IV. TECHNIQUES FOR EXTRACTING FEATURES

A set of selected features must be retrieved from EEG in order to obtain meaningful information. The waveform of the EEG signal changes over time (i.e., it has different

TABLE 5. Sleep stages in each standard with description.

Stages	Rechtschaffen and Kales (R & K	American Academy of Sleep Medicine (AASM)	Description
WAKE (W)	W	W	
NREM (Non–Rapid Eye Movement)	N1	S1	Alpha
	N2	S2	Theta
	N3	S3	Delta
	N4		
REM (Rapid Eye Movement)			

frequencies). As a result, extracting information from two domains namely time and frequency could be beneficial. Frequency-domain features are generated from frequency spectra, while time-domain features are derived from EEG signals in time. The entire purpose of signal processing is lost if either domain is ignored, so evaluation should utilize the time-frequency domain. A list of feature extraction strategies has been produced from the finalized set of previous tests.

A. STRUCTURING OF FEATURE EXTRACTION TECHNIQUES WITH ADVANTAGES

Raw brain signals do not give enough information for effective analysis due to noise. An important step in signal processing is the feature extraction that converts noisy signals into meaningful values. A number of methods, each with its specific advantages and disadvantages, exists. They have potential for use in diagnoses of sleep related problems. The main purpose of this section is to provide insight into the feature extraction techniques.

Before looking into the feature extraction techniques, the terms used are briefly introduced. These techniques are known for their ability to solve problems that often appear in the automatic sleep scoring domain. Obstructive Sleep Apnea (OSA) is a critical sleep disorder in which breathing stops and starts periodically while sleeping. Neonatal EEG is used to study sleep staging in newly born babies. Nocturnal Oximetry is an oxygen test that is used to evaluate oxygen need during sleep. Sleep spindles are rhythmic oscillations with a frequency range of 10 to 14 Hz. A seizure is uncontrolled electrical disturbance in the brain.

As explained before, feature extraction techniques have been developed for research on human sleep related problems. This SLR identified four types of feature extraction techniques for the EEG signals (see Figure 5) that most studies have utilized.

1) TIME DOMAIN FEATURES

- **Zero Crossing:** An event is counted whenever a zero crossing of the signal occurs i.e. a point at which the wave form performs a crossing of the time axis.

$$Zero\ Crossing = \frac{1}{T-1}\sum_{t=1}^{T-1} 1_{R<0}(S_t S_t - 1)$$

where S is a signal of length T. This is useful as an indicator of noise, but the value also varies by sleep stage. For automatic sleep scoring, the delta wave (0.5-2Hz) in infants is detected. This feature is used to measure the number of baseline zero crossing in a fixed period interval [7].

- **Hjorth Parameters:** Derivatives of signals are used to calculate the Hjorth parameters.

$$Activity = \frac{var(y(t))}{Mobility} = \sqrt{\frac{var(\frac{dy}{dt}y(t))}{var(y(t))}}$$

$$Complexity = \frac{Mobility(\frac{dy}{dt}y(t))}{Mobility(y(t))}$$

where: y(t) represents the signal and var takes the variance of function. It has been used to detect OSA. It is sensitive to noise and is one of the candidates to construct automatic sleep staging [8].

- **Arithmetic Mean:** This summarizing feature is used to extract information from a signal.

$$AM = ArithmeticMean = \frac{1}{N}\sum_{n=1}^{N} x_n.$$

Its use as a feature is obvious, e.g. in sleep staging in neonatal, especially REM state detection [9].

- **Median:** It is used to extract information from a signal.

$$Median = \frac{(N+1)^{th}}{2}$$

This feature is used in sleep apnea diagnosis from nocturnal oximetry [10].

- **Variance:** It is calculated from squared differences of each number in a data set from the mean.

$$Var = \frac{\sum_{i=1}^{N}(x_i - AM)^2}{N-1}$$

It is used to classify neonatal sleep states [11].

- **Standard Deviation:** An alternative to amplitude for characterising strength of signal.

$$S = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - AM)^2}$$

It is used as a feature for topographic study of neonatal biosignals [11].

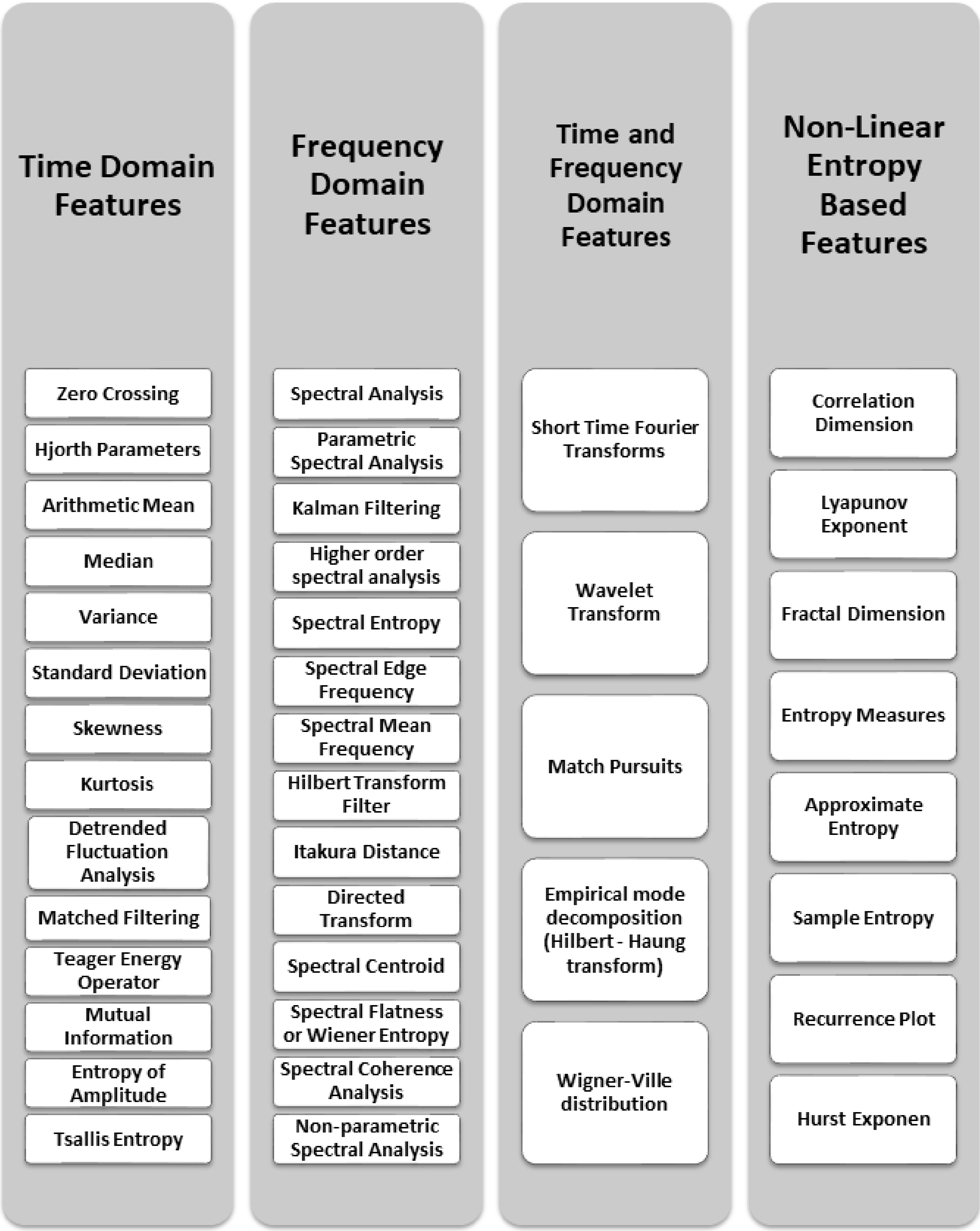

FIGURE 5. **Classification of Feature Extraction Techniques for Automatic Sleep Scoring.**

- **Skewness:** It is used to define an irregularity from a probability distribution in a set of real data. It can be positive or negative depending on how the data are skewed.
$$S = \frac{\sum_{i=1}^{N}(x_i - AM)^3}{(N-1)SD^3}$$
In a study on automatic sleep stage classification, skewness was used as a feature [12].
- **Kurtosis:** This numerical measure is used to describe the shape of the data, indicating whether the data are heavy or light tailed.
$$K = \frac{\sum_{i=1}^{N}(x_i - AM)^4}{(N-1)SD^4}$$
It is widely used as feature from a biosignal, also in computer based sleep staging [12], and in an automated sleep stage classification system using an ensemble technique [13], among others.
- **Detrended Fluctuation Analysis:** This method is helpful to detect a long-range correlation in a noisy signal.
$$F(n) = \sqrt{\frac{1}{N}\sum_{K=1}^{N}[y(k) - y_n(k)]^2}$$
It is used in different sleep EEG related studies, e.g., on the relation of sleep stages and sleep apnea with heart rate variability [14].
- **Matched Filtering:** This method is helpful to find a template matching. Matched filtering is used for perceiving a signal that contains strong noise. The filter increases the signal to noise ratio. It has been used to detect cyclic alternating patterns in sleep. Three types of matched filter were used to detect k-complex in sleep stages. It has been used for sleep spindle detection. The disadvantage of this method is that the frequency deviation within the spindle may be problematic because matched filter output is relying on the spindle template [15].
- **Teager Energy Operator:** It measures the energy of the input signal in a particular frequency band [16].
$$\psi_{Ts}(n) = \psi_s^2(n) - \psi_s(n-1)\psi_s(n+1)$$
where Ψ s (n) and Ψ Ts is the n^{th} sample value of signal and Teager Energy Operator as output. Automatic sleep spindle detection was done in biosignals during NREM sleep stage [15]. It is used to detect K-complex signals in sleep EEG automatically [16]. It's been used successfully in a variety of signal processing applications.
- **Mutual Information (MI):** measures the mutual information among two random variables. Two steps are needed to estimate the MI: calculating the joint distribution and computing the MI from the joint distribution.
$$I(S;R) = I[p(S;R)] = \sum_{s,r} p(s,r) log(\frac{p(s,r)}{p(s)p(r)})$$
It is applied in EEG to measure the effects of total sleep deprivation [17].
- **Tsallis Entropy:** It is used for diagnosis based on the entropic index of biosignals. This method provides better accuracy than Shannon entropy, since it maximizes the probabilities of the events by using entropic index [10].
$$S_q(p_i) = \frac{k}{q-1}(1 - \sum_i p_i^q)$$
Improving sleep stage separation by using Markov model was based on Tsallis entropy [18] and its use to analyse sleep stages [19] has been reported.

2) FREQUENCY DOMAIN FEATURES
Investigation of signals in a frequency domain gives new insights for sleep data analysis. A set of commonly used spectral features and associated signal processing techniques is described in this section.

- **Fast Fourier Transformation** This algorithm helps to decrease the number of computations in non-stationary tests for power spectrum analysis. There are two classes: Parametric and Non-parametric methods [20]. Fast Fourier Transformation is advantageous for the detection of sleep spindles [21]. It was also applied in the analysis of preterm infant signal spectrum [22].
- **Parametric Spectral Analysis:** It finds the parameters of a signal. For parametric analysis based on the occurrence of poles, the following models are used: (i) Auto Recursive, (ii) Yule-Walker and Burg's, (iii) Moving Average, and (iv) Prony's Auto Regressive Moving Average. AR modelling is a popular technique for analysis, because of its advantage in finding positions in signals with low noise levels and determining the short data record [23].
It is appropriate for signals that do not change with time, for spectral estimation and stability assessed in human biosignals analysis [24].
- **Kalman Filtering:** It is an optimal estimator for a large class of problems. It follows two stages (i) predict the state of system (ii) refinement to estimation using noisy measurements.
The major advantage of this method is that it does signal parameterization. It is applied for sleep dynamics analysis and automatic arousal detection with an AR model of the signals [25].
- **Higher order spectral analysis:** This uses higher-order moment spectra for deterministic signals. Cumulant spectra are defined for random processes. It is used in signal processing to: (i) contain Gaussian Noise of unknown data (ii) reform the phase and magnitude response of signals and (iii) to identify and distinguish non-linearity in the data. Sleep EEG of healthy neonates is an important topic, and parametric bi-spectrum analysis [26] has been used for this.

- **Spectral Entropy:** It measures the irregularity or complexity levels of signals.

$$E = -\sum_{f=-\frac{fs}{2}}^{f=\frac{fs}{2}} PSD_n(f) log_2[PSD_n(f)]$$

It is used as spectral information during OSA diagnosis and in Automatic REM sleep stage detection [27].
- **Spectral Edge Frequency:** It is defined as a frequency below which X% of total signal power is located. It differentiates between different sleep stages. In neonates, it is used to distinguish between active and quiet sleep [28].
- **Spectral Mean Frequency:** It is a mean value from power spectrum of signal.

$$f_{mean} = \sum_{i=0}^{n} I_i \cdot f_i \,/ \sum_{i=0}^{n} I_{i.}$$

Quiet sleep is investigated in premature and full term infants and spectral moment has been suggested for automatic sleep stage detection [29].
- **Hilbert Transform Filter:** It derives the analytic representation (filter negative frequency component) of a signal and is useful for envelope detection.

$$H(u)(t) = p.\int_{-\infty}^{\infty} u(\tau)h(t-\tau)d\tau = \frac{1}{\pi}\, p.\int_{-\infty}^{\infty} \frac{u(\tau)}{t-\tau}\, d\tau$$

It is useful for detecting spike activity in newborns as well as calculating variance between two quite sleep patterns in preterm and full-term infants. The instantaneous envelope and frequency waveform have been considered for micro-structure of sleep spindles [30], [31].
- **Itakura Distance:** It measures the degree of similarity between EEG and EOG with different sleep stages [32].

$$d_t(A, \widehat{A}) = ln[\frac{1}{2\pi}\int_{-\pi}^{\pi} |A(e^j)|^{\,2} \,/\, |A(e^j)|^{\,2}\, d\,] \\ -ln\,[d_{LR}(A, \hat{A}) + 1].$$

- **Directed Transform Function:** determines the relationship between channels as a function of frequency and time.

$$DTF^2{}_{j\to i}(f) = |H_{ij}(f)|^{\,2} / \sum_{m=1}^{k} |H_{im}(f)|^{2}$$

It recognizes the main centres' of EEG activity during sleep and wakefulness, and the direction of information flow is estimated through presleep wake and early sleep stages [33].
- **Spectral Centroid:** It measures the power spectrum "centre of mass" by employing Fourier transform frequency and magnitude information [34].

$$SC = \frac{\sum_{m=0}^{N-1} m|X(m)|}{\sum_{m=0}^{N-1} |X(m)|}$$

It is useful to detect and classify human stress and for automatic classification of healthy and sick neonates [35].
- **Spectral Flatness or Wiener Entropy:** A method to quantify the noise of spectrum known as Wiener entropy.

$$SF = \prod_{m=0}^{N-1} |X(m)|^{\frac{1}{N}} \Big/ \ \frac{1}{N}\sum_{m=0}^{n-1} |X(m)|$$

It is a feature parameter and used for automatic detection of snoring in studies conducted on sleep [36].
- **Spectral Coherence Analysis:** It depicts the relationship between two signals as a function of frequency. It represents the degree of integration among frequency components of two signals and may indicate a large scale functional connectivity in the brain.

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy(f)}}$$

It is widely used in various studies and for the purpose of sleep related oscillation (slow-wave and spindle), and temporal evolution in human sleep brain signals. It is applied during the wake-sleep transition period [37], [38].
- **Non-parametric Spectral Analysis:** Wolfowitz coined the term non-parametric in 1942 for cases when the parameters of variables are unknown and they do not rely on the estimation of parameter's mean and standard deviation. This is also known as parameter-free or distribution free. It has been used to investigate the relationship between fitness, behaviour and sleep [20]. Spectral analysis is proposed to identify the inter-dependencies among heart rate and sleep recording [21].

3) TIME AND FREQUENCY DOMAIN FEATURES

This representation is used to analyse non-stationary signals (sleep EEG) in both time and frequency domains [20]. The methods used in time-frequency analysis are given in this section.

- **Short Time Fourier Transforms:** The signal is divided into segments by using a window function (i.e., in terms of time and frequency), defined as:

$$STFT_x^w t, f = \int_{-\infty}^{\infty} [x(t).w^*(t-t^{\prime})].e^{-j2\pi ft} dt$$

Here x(t) is signal, w(t) is window function, and * is the complex conjugation. It is used for analysis of respiratory cycle related EEG changes in sleep. Further, human sleep onset estimation is achieved by this feature and it is found useful for sleep spindle detection. Furthermore, this feature was proposed to visualize both macro and micro levels of human sleep [39].
- **Wavelet Transform:** This feature can provide the time-frequency of signals, and can be expressed as:

$$F(a, b) = \int_{-\infty}^{\infty} f(x)\psi^*_{(a,b)}(x)dx$$

where * indicates complex conjugation and ψ is the generating function. It can detect automatic arousals and classify sleep/wake stages. It is proposed to capture sleep spindle activity. Because they are borderline in the time-frequency domain, certain spindles are difficult to recognize. Wavelet transform was applied to calculate features for coefficients of decomposition scale during EEG sleep in neonates [40].

- **Match Pursuits:** is used with dictionaries of Gabor functions in time-frequency analysis of signals and has following advantages (i) explicit parametrization of transients, (ii) robust time frequency estimate [41].

$$x \approx \sum_{n=0}^{M-1} (R^n x, g_{\gamma_n}) g_{\gamma_n}$$

This can be suitable for finding and parametrizing delta waves and sleep spindles. The MP algorithm is used for EEG structures like Slow Wave Activity (SWA) with time-frequency parameters [42].

- **Empirical mode decomposition (Hilbert - Haung transform):** based on the function known as Intrinsic Mode Function (IMF), decomposes the signal into its component IMFs along with trends and extracts instantaneous frequency data [43]. Since a signal is decomposed in the time-domain and is of the same original signal length, it allows maintaining varying frequency. It can be compared with other transformations such as Fourier transform and wavelet decomposition. It is calculated easily and yields high time and frequency resolution.
It is employed to analyze sleep stages and automatically detect sleep spindle. It is used for automatic sleep staging with the nearest neighbor algorithm [44].
- **Wigner-Ville distribution:** This is also a good choice for extracting features from a signal that comprises only a single component [45], defined as:

$$WV(t, w) = \int_{-\infty}^{\infty} f(t + \frac{t_0}{2}) f^*(t - \frac{t_0}{2}) e^{-jt_0 w} dw_0$$

Here WV(t,w) is energy distribution of signal, * the complex conjugate of signal and w is the frequency. It is utilized to locate the sleep spindle's structural position and to solve the difficulty of detecting K-Complexes and Delta waves [46].

4) NON-LINEAR ENTROPY BASED FEATURES

Non-Linear Entropy Based Features provide complementary information in sleep EEG analysis. Although reliability and interpretability of results are important issues, a good understanding of these techniques helps in their application and interpretation. A list of commonly used features is given below:

- **Correlation Dimension:** The Grassberger-Procassia algorithm is a fast and simple numerical method for calculating a fractal measure's Correlation Dimension. It successfully identified sleep stages and considered slow wave activity both in adults and infants [47]. Automatic REM detection is based on the spectral measure [48].
- **Lyapunov Exponent:** measures the convergence or divergence rate of trajectories and describes the performance of a dynamical system. The exponent can be positive, negative, or zero and this reveals the behavior it implies [20]. Positive value indicates that the system is chaotic; a negative value relates to converging trajectories; and a zero indicates the system maintains its relative position.

$$\lambda = lim_{t\to\infty} \frac{1}{t} ln|\Delta x(X_0\, t)|/|\Delta X_0$$

where λ is the Lyanpunov exponent and X_0 and $X_0 + \Delta X_0$ are two EEG data points in space. This is utilized in studies done on EEG sleep analysis, for predictability of different sleep stages [49], and it provides information regarding the neural process of brain during sleep [50]. EEG signal characterization in different sleep stages has been calculated by positive LE [51]. Automatic REM stage detection is based on non-linear measures such as LE and correlation dimension [48].
- **Fractal Dimension:** It is a scaling parameter that describes how patterns change with the scale, and this is associated with signal complexity. It can be employed for short segments of EEG signals [20] and to classify physiological function of a state [23]. Kats and Higuchi's algorithms are used to calculate the FD. Behavior of fractal dimension has been studied during the different sleep stages in infants [52] and adults [53]. Higuchi fractal dimension has been calculated on comparing sleep spindle and anesthesia [54]. Multifractal analysis of sleep EEG characterization has been investigated by using wavelet transforms [55].
- **Entropy Measures:** Those measure the disorder in a signal.
- **Approximate Entropy:** It measures the anomalies in a time series' variation. A low value suggests strong regularity and predictability, whereas a large value indicates unpredictability and random data variances.

$$ApEn(S_n, m) = ln[\frac{C_m(r)}{C_{m+1}(r)}]$$

where S_n is approximate entropy for length m and similarity criterion r.
A low ApEn indicates predictability and high regularity of time series data and high ApEn shows unpredictability and random deviation. ApEN performed well in the classification of sleep EEG signals to sleep stages [56]. It is used to compare the sleep spindle and anesthesia in EEG signals [54].
- **Sample Entropy:** This is an improved version of the approximate entropy. It is based on the negative logarithm of the probability. It is more reliable, unbiased, and ideal for brief data segments. It is also unaffected

by sample size.

$$SampEn(K, r, N) = \frac{-ln(A(K))}{B(K-1)}$$

It is used to record and indicate the characteristics of sleep [57]. It was used to analyze different sleep stages and deeper sleep was associated with a lesser SampEn value [23].

- **Recurrence Plot:** This tool helps visualize the recurrence state in the phase space, when the distance between two points on a trajectory is smaller than the threshold. It is represented by two-dimensional matrices of black and white dots with time axes. It can help find interrelations and visualize time dependencies in the data [58].
  Recurrence is typically a visual aid for the analysis of dynamical systems. It has been used in the analysis of EEG signals at different sleep stages [59]. Recurrence analysis of sleep EEG data has been studied to obtain information regarding treatment effects in patients with depression [60].
- **Hurst Exponent:** It is used to measure the self-similarity in a time series. It evaluates the presence or absence of long range dependencies and irregularity in a time series [61]. Its value ranges between 0 and 1; higher values show a smoother signal with less roughness.

$$H = \frac{log(\frac{R}{S})}{log(T)}$$

where T is duration of a sample of the data and R/S is the corresponding value of the re-scaled range. It has been used for characterizing non-stationary behaviour in sleep EEG data [62].

B. CHALLENGES OF FEATURE EXTRACTION TECHNIQUES

A very common challenge is feature extraction from single or multiple channels. For automatic sleep stage classification it remains to be determined whether single or multiple channels would perform better. Nevertheless, attempts to diagnose sleep related problems and stage classification highlight the need for a set of feature extraction techniques. Especially, due to the increased attention of clinical practitioners, and researchers, different types of feature extraction techniques are available. Adding to the signal variation and complexity, combinations of different feature extraction techniques are becoming more common in diagnosing various sleep related problems: sleep apnea, automatic spindle detection, etc.

Another challenge is usefulness of single or multiple feature(s) to diagnose sleep related problem(s). It has to be taken into account that not only single features are used to diagnose the problem but multiple ones are employed for diagnostic purposes, e.g. in Matched Filtering and Teager Energy Operator for sleep spindle detection, etc. Secondly, many alternative or complementary feature extraction techniques are used in the sleep context. One motivation for this is that the EEG signal is difficult to understand as it involves changes in frequency and in amplitude. Further, one must consider effects of the human subject's age and mental state, disease, etc. Therefore the study of neuron activity can benefit both linear and non-linear signal processing techniques and needs to be considered along with physiological aspects. Hence, several types of extracted features can help with gaining insights into sleep EEG data.

V. FEATURE SELECTION TECHNIQUES

All features present in available data are not useful for a specific classification task. Instead, some features can reduce the classifier performance due to irrelevance and/or noise. Feature selection refers to the process of selecting a discriminating subset of features in order to avoid over-fitting. In order to increase the classification accuracy, feature selection methods or algorithms remove the unnecessary or redundant features from the given dataset. Hence, the dimensionality of the dataset is reduced, and learning accuracy increased with improved results. Sequential forward and backward selection are two common techniques for selecting features. Both are greedy approaches to a combinatorial optimization problem, and the subset of features obtained might be far from ideal. A summary of pros and cons of the feature selection methods [63] is given as follows:

Euclidean distance is linear in computational cost having O(n) time complexity. It is sensitive to noise and outliers. If time series similarity measurement is required then it requires extensive data preprocessing.

The T-test does not necessitate a huge dataset and eliminates subject to subject variation. Its drawback is that it is unconcerned about feature dependencies and disregards the classifier's interaction.

Information gain is an entropy that aids in the elimination of redundancy and ensures feature relevance with other features.

When compared to other methods, the Correlation Based Feature (CBF) feature selection method has a lower computing complexity and is less prone to over-fitting. However, it is very reliant on the model, which may fail to fit the data.

The Markov blanket filter (NBF) approach can handle large datasets. It is independent of the classification technique and computationally simple. Its disadvantage is that it ignores feature dependencies and does not take into account communication with the classifier, resulting in poorer classification especially in comparison to other feature selection methods.

The feature goodness for classification is measured using the fast correlation based feature (FCBF) selection method. It removes the feature of a class with a near-zero linear correlation. It eliminates repetition among certain characteristics. It is slow and less scalable than univariate approaches, and it ignores classifier interaction.

In x^2 feature selection method, over-fitting is reduced, and learning precision is improved. In terms of time and space complexity, it is effective. However, it ignores the classifier's individual heuristics and biasing, which could lead to reduced classification accuracy.

Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) reduce dimensionality, i.e. the count of attributes or features for an item to be labelled by the classifier. Instead of selecting from existing features, PCA and LDA create new ones that usefully summarize the given ones.

Feature selection is carried out using meta-heuristic methods such as the Genetic Algorithm (GA) or Particle Swarm Optimization. FCBF, t-test, ReliefF, and Fisher score are all examples of fast correlation based filters used for feature selection [64], [65].

VI. TYPES OF MACHINE LEARNING TECHNIQUES

A feature may be classified based on a measure of its similarity to each class. A set of classification algorithms has been studied for sleep EEG signal analysis. This section discusses different categories of machine learning algorithms available in the literature.

Artificial Neural Network: It is a frequently used in different applications in aerospace, military, electronics, signal processing, and medical field, etc., due to giving non-linear models with computational efficiency [66].

It was used in an investigation of sleep EEG for automated k-complex detection [67], [68], for sleep stages and apnea in children [69], diagnosis of OSA [70], [71] and in a study of drug effects [72]. Back-propagation trained ANN has been used for REM, sleep spindles, and waking state in their automatic detection [71]. Identification of arousal [73] and spike detection in neonates [50] from Polysomnography recordings was based on ANN. Finally, a system based on ANN for micro and macro-structure of sleep is presented in [74].

Multilayer Perceptron: This type of neural network comprises three categories of layers: input, (multiple) hidden, and output layers. Any continuous function can be arbitrarily well approximated, provided the hidden layer is made large enough [41]. It may be flexible for classification but too sensitive to over-fitting [75].

They have been used to examine EEG recordings taken during sleep [76], automatic sleep spindle detection [77]–[79], automatic REM detection [80], OSA diagnosis [81], and automatic sleep staging [82], [83].

Self-Organizing Maps: This allows visualization of multi-dimensional data [41] and clusters data into several classes. It is suited for applications with a small amount of input data and no output available [41]. In the cooperative learning the neurons not only adjust themselves to the data but also to the neighbouring neurons as well.

The SMO have been applied in automatic sleep stage detection [84], [85], and in the classification of patterns of k-complexes during sleep [86].

Linear Discriminant Analysis: Fisher linear discriminant analysis is another name for this method. The basic idea is to search for a linear combination of variables (predictors) which distinguishes the data into various classes separated by hyperplanes. This requires that the classes should be linearly separable, and the method is numerically robust, but cannot handle strongly nonlinear class boundaries [87].

It has been used for artefact detection in sleep EEG [88] and for the classification of newborn baby's brain state and burst suppression pattern [55]. Further, it has been used for automatic sleep state recognition between preterm and full-term infants [89]. It was applied to low and high voltage pattern discrimination of infants in sleep stages [90].

Support Vector Machine: This technique is useful both for classification as well as for regression problems. It is based on the design of an optimal hyperplane which classifies all training vectors into two classes and this optimal hyperplane leaves the maximum margins to the two classes [41]. It can be used to classify data in both linear and non-liner classifications. It is a useful tool for non-linear classification since it uses a kernel function to map the feature space for classification [41].

It's commonly utilized in sleep EEG analysis including automatic spindle recognition [78], arousal detection [73], [91], [92], sleep staging [93], [94], and automatic REM detection [95]. It has been applied in recognition of behavioural sleep states in infants [62].

Hidden Markov Model: At each time step, a system's alternative states and transition probabilities between them are given [66]. This works well because of simplicity and the parameters can be estimated for various real world applications. It has been employed in automatic sleep staging schemes [96], [97].

Based on probabilistic principles, it's been used to create sleep staging algorithms in place of the R & K rules [96], [97]. However, this method has been proposed for sleep EEG including automatic sleep stage classification in infants and adults [98], k-complex detection [99], demonstrated for sleep EEG dynamic activity [100], and for automatic sleep stages and for sleep apnea diagnosis [101].

Naïve Bayes: The probabilities of each class are predicted first using this probabilistic classifier and then call is made of that class which has the highest probability, based on a set of observations. It allows you to calculate the posterior probability p (c|x) of a class using the prior probability, predictor (p (c). p (x)) and likelihood (p (x|c)). NB performs well in classification speed and accuracy for large training datasets [102].

NB classifier has not commonly been used but has been employed in sleep EEG analysis including sleep stage discrimination [76] and for neonatal state discrimination [62].

K-nearest Neighbors: It's a nonlinear lazy learning technique that can be used to solve regression and classification problems. It makes predictions based on known labels of the K closest neighbours [102], [103] according to some distance function, often the Euclidean distance.

KNN has been recommended for the analysis of EEG signals [104]. It has been used to detect sleep apnea occurrences using ECG signals during the night [105], [106]. It performed

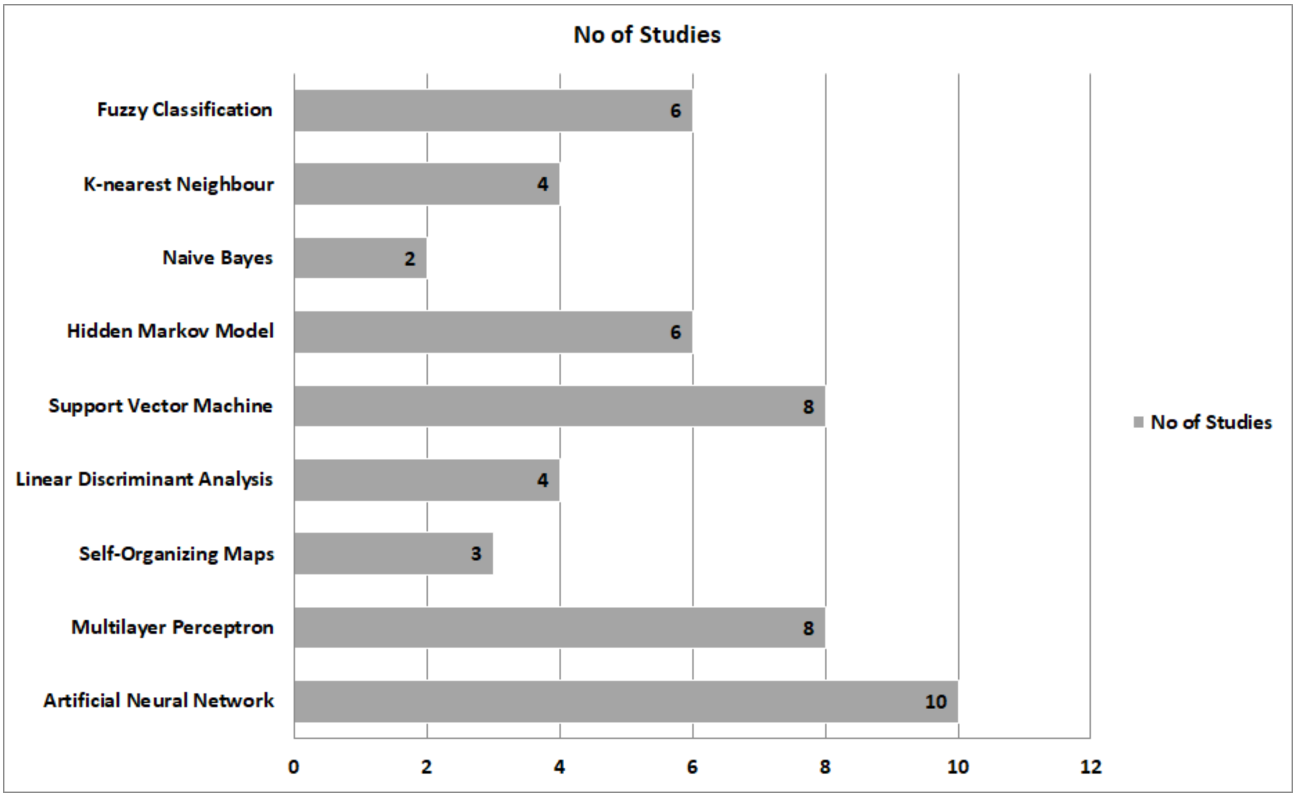

FIGURE 6. Most used ML techniques for Automatic Sleep Scoring.

well with low computation complexity in automatic sleep stage classification [107].

Fuzzy Classification: It can work with other classification systems, such as neuro-fuzzy classifiers and fuzzy decision trees, to improve performance [41].

Fuzzy reasoning-based classifier (FRBC) is a reliable tool for automatic sleep EEG staging [85]. Fuzzy ganglionic lattices have been applied to classify the sleep/wake states in newborns. Fuzzy based methods have been applied for the detection of alpha activity [108] and automatic cyclic alternating pattern detection [109] in sleep EEG analysis. The classification of sleep stages in newborns has been done using neuro-fuzzy classifiers (NFC) [110], [111] and used to detect the k complex in EEG signals from sleep [112].

A list of feature extraction techniques in each category was extracted from the finalized set of the studies, and the extracted information was synthesized, see Table 6.

A. ADVANTAGES OF MACHINE LEARNING TECHNIQUES

ML approaches have a number of general advantages, including the ability to tackle (although imperfectly) NP complete problems like bio-signal analysis. The following discussion organizes the advantages of machine learning, with a focus on capacity to handle high-dimensional multi-variate data and extract implicit associations. The classification and analysis of sleep bio-signal data is complicated by the nature and necessitates multidisciplinary expertise.

ML approach reduces cycle time, and improves execution time and resource utilization in sleep stage classification. Moreover, it provides powerful tools for performance improvement in diagnosis of sleep related problems, such as sleep apnea or spindle detection.

The ability to handle high-dimensional issues is one of the benefits of machine learning. As sleep EEG data availability is increasing, it is becoming more important to utilize ML techniques, but it's also true that the majority of the advantages and disadvantages of individual algorithms aren't generalizable. Support Vector Machines and Artificial Neural Networks are two approaches that excel in dealing with large dimensionality [117], [127]. As stated before, most ML techniques are applicable to sleep bio-signal analysis, and the ability to handle high-dimensional data is considered an advantage.

Another benefit of machine learning approaches is the accessibility and usefulness of open source algorithm packages like WEKA.

In data mining, machine learning techniques are used to identify unknown knowledge and relationships in data sets. The requirements for training data vary depending on the properties of ML algorithms. ML algorithms have been successfully demonstrated in applications to sleep data analysis [107], [117], [119], [120], [124], [125].

While sleep bio-signals are complicated and dynamic, machine learning algorithms can learn from them and adapt

TABLE 6. Synthesis of Machine Learning Techniques with Data Extraction Framework.

Reference	Purpose	Features	# of Sleep Stages	Dataset Size	Data set Source Available at:	Accuracy in (%) : Features
K-Means Clustering						
[34]	Using structural graph similarity and the K-means method together	Statistical Features and Structural Graph properties	Six	48,555	[113], [114]	95.93 : 12
[115]	Feature weighting based on k-means clustering	Statistical	Six	4196	Sleep Lab of Meram Medicine Faculty of Selcuk University Turkey	82.15 : 4
[116]	K-means clustering based wavelet de-noising, EEG data feature extraction and spectrum analysis, k=10	Frequency & Time Domain	Five	N/A	[113]	81 : 4
[70]	K-means based on time and spectral	Time & Spectral	Five	N/A	Lab of Neuro-physiology Oscar Moscoso Arisa Autonoma University of Manizales UAM	74 : 5
Artificial Neural Network (ANN)						
[117]	Single layer neural network	Spectral Density	NREM, REM, and WAKE	N/A	4 nights home based experiment	63
[118]	The dual-tree complex wavelet transform and a complex-valued neural network based	Statistical	Six	20,774	[113]	93.84 (R & K) 95.42 (AASM)
[119]	Convolutional neural network (CNN) based on five layer (One output layer, two conventional and pooling layers)	Time-Frequency	4 Awake, light, deep and REM	N/A	[113]	88.83
[120]	Mixed Neural Network	Time-domain	Five	N/A	[113]	83.35
[121]	Three feed-forward ANNs	Time-Frequency	Five	13650	[113]	81.1
[122]	Bidirectional Long Short-Term Memory based Convolutional Neural Network (CNN)	Time invariants features	Five	N/A	[113], [123]	86.2
K-Nearest Neighbours (KNN)						
[124]	KNN based on iterative filtering	Time and frequency domain	Six	N/A	[113]	83 to 95
[107]	KNN with Separability & Correlation (SEPCOR) analysis	Spectral entropy	Two wake and stage 1	N/A		92
[125]	KNN	Time Domain and frequency domain	wake and sleep	N/A		73.36 : 36
[126]	KNN based on EOG signal	Time and Frequency	Five	N/A	Belgian sleep hospital	80

TABLE 6. *(Continued.)* Synthesis of Machine Learning Techniques with Data Extraction Framework.

Support Vector Machine (SVM))						
[127]	SVM	Time and Frequency domain	Five	N/A		94 : 102
[128]	Based on time frequency images and multiclass least squares support vector machines	Time-Frequency	Six	N/A	[113]	92.93
[129]	SVM	Graph domain features	Six	14,963	[113]	87.5
[130]	SVM	Frequency Domain	Five	N/A	[113]	93.8
[131]	SVM	Time and Frequency Domain and nonlinear features	Five	N/A		87 : 39
Naïve Bayes (NB)						
[132]	Naïve Bayes based on dynamic time wrapping method(DTW)	Time domain feature	Two sleep and wake states	N/A	National Heart Lung & Blood institute	84.19
[133]	NBC	Time domain	Three wake, NREM and REM	N/A	[113]	70
[134]	NB based on bed sheet system	Time Domain	Wake, NREM and REM	N/A	UCLA School of Nursing	70.3

to changing environments (depending on the ML methodology) very quickly, and in almost all situations faster than traditional approaches [117], [119], [120].

Machine learning approaches assist in the discovery of patterns in existing data sets, which can be used to develop approximations. Clinical decision-making could be aided by the knowledge acquired. As a result, some machine learning algorithms aim to find patterns, regularities, or abnormalities.

Generally, ML performance varies in prediction speed, memory usage, and interpretability. It is not suggested to base the selection of the ML technique on previously reported comparisons. Each algorithm's performance is determined by the type of problem and data provided, as well as pre-processing and parameter choices.

B. MACHINE LEARNING TECHNIQUES' CHALLENGES

In sleep data analysis, a typical challenge is the acquisition and availability of relevant data. There are also issues with the quality and the composition of data that affect the performance of an ML algorithm. An example of the challenges is the high dimensionality of data, as it can contain irrelevant and redundant variables. Several factors impact the result, including the algorithm and its parameter settings. Obtaining any data is a general challenge. Though machine learning allows for the extraction of information and produces better outcomes than most traditional approaches with fewer requirements for training data, certain characteristics of data must still be considered. Overall, this emphasizes the increased requirement to comprehend data in order to use a machine learning algorithm. In contrast to traditional approaches, which spend a lot of time extracting information, ML spends a lot of effort preparing the data.

Following the collecting of data, the next step is to preprocess it according to the algorithm's specifications. The results are heavily influenced by data preprocessing. Standardized tools are frequently used to normalize and filter data. It is also tested to see if the data is balanced, as this can provide a problem when training some algorithms. Missing values is a common problem in industrial or medical databases and the use of different types of sensors may cause varying quality of data, even within one dataset. These issues may bias the classification results, and in a clinical application a patient's problem may be misdiagnosed.

A key decision to make is the choice of ML algorithm. Generally, some strengths and weaknesses of ML algorithms are well-known. Due to the increasing use by practitioners and academics, a huge variety of diverse ML algorithms and their modifications are available in medical areas. Literature available can show effective applications of machine learning techniques for specific issues. On the other hand, in most situations, the test dataset is not publicly available, making an unbiased evaluation of the results impossible. The steps in selecting a suitable ML algorithm for a certain problem (type) are as follows:

1) Screening the available data, i.e., labelled, unlabelled, etc. Make a choice between supervised, and unsupervised approaches.
2) The structure, data categories, and overall volume of available data must all be considered when evaluating the general applicability of existing algorithms.
3) Previous applications of ML algorithms to solve similar problems should be explored.

The analysis of the findings poses another challenge. Attention has to be given to ensure that the output format and its interpretation are relevant. Points to be considered: algorithm specification, parameter settings, planned outcome, and also data preprocessing. Over-fitting, bias, and variance are typical issues that need attention. In order to address the overfitting, following techniques can be used in the context of sleep data. Regularization techniques include early stopping, batch normalization, weight decay, dropout, particle swarm optimization, max-norm regularization, data augmentation and cross validation [135]–[137].

TABLE 7. Exploration of Machine Learning and Feature Extraction Techniques over the years: single or multiple EEG channels along with choice of classifier.

Machine Learning Techniques or Classifiers:
Feed Forward Neural Network (**FNN**), Probabilistic Neural Network (**PNN**), Support Vector Machine (**SVM**), K-Nearest Neighbour Search (**KNN**), Artificial Neural Network (**ANN**), Decision Tree (**DT**), Naïve Bayes (**NB**), Radial Basis Probabilistic Neural Network (**RBPNN**), Fisher's linear discriminant (**FLD**), Fuzzy, Adaptive Boosting (**AB**), Decision-tree multi-class Support Vector (**DTMCSV**), Random Forest Classifier (**RFC**)

Feature (s): Energy, Power Spectrum (**PS**), Hjorth Complexity Parameters (**HCP**), Frequency Domain (**FD**), Wavelet Packet Decomposition (**WPD**), Time Domain (**TD**), Non-Linear Analysis (**NLA**), Spectral Measure (SM), Fast Fourier transform (**FFT**), Adaptive Auto-regressive (**AAR**), Spectral Features (**SF**), Teager Energy Operator (**TEO**), Wavelet Packets (**WP**), Autocorrelation Function (**ACF**), Wavelet Transform (**WT**), Choi-Williams distribution (**CWD**), Continuous wavelet transform (**CWT**), Hilbert–Huang Transform (**HHT**), Renyi's Entropy (**RE**), Discrete wavelet transform (**DWT**), Hjorth features (**HF**), Itakura Distance **(ID)**

SET_1 : {EEG, ECG, EMG, EOG}

Channel		**Classifier**	Year	Reference
Single	Multiple			
	Signal			
EEG	SET_1			
X	X	ANN	2010	[138]
		FNN & PNN	2013	[139]
	X	SVM & KNN	2005	[140]
X		SVM	2009	[141]
X		DT, Fuzzy, GMM, KNN, NB, RBPNN	2012	[142]
X		SVM	2012	[131]
X		ANN	2001	[48]
X		ANN	1997	[143]
X		Fuzzy	2002	[144]
X		SVM	2004	[145]
	X	-	2006	[146]
X		FLD	2010	[35]
X		AB	2015	[36]
X		SVM	2006	[147]
X		-	2009	[15]
	X	ANN	2008	[148]
X		ANN	2006	[40]
X		RFC	2012	[149]
	X	DTMCSV	2015	[127]
X		SVM	2013	[150]
X		RFC, FFNN, DT, SVM	2014	[7]
X		SVM	2016	[34]
X		K-means & SVM	2016	[151]
X		RFC	2016	[152]
X		Deep learning	2017	[119]
X		DBN	2017	[122]
X		K-NN	2018	[103]

VII. THEORETICAL APPLICABILITY OF MACHINE LEARNING TECHNIQUES TO AUTOMATE SLEEP SCORING CHALLENGES

Before looking into the applicability of machine learning (ML) for sleep stage classification, a quick review of the terminologies is in order. ML is known for its potential to handle problems of an NP-complete nature.

See Table 7 for a summary of sleep research over the previous decade. As sleep disorders become more common in today's culture, developments in machine learning and data analysis techniques provide academics new opportunity to investigate the problem and build automatic sleep scoring systems. As previously stated, the EEG is utilized to record human brain activity, and even a single signal channel can yield good findings. Electrical monitoring systems have been used in a few investigations to target numerous channels.

Focus on a single channel simplifies the problem, at the risk of potentially poor performance. K-Nearest Neighbours (KNN), Support Vector Machine (SVM), Naive Bayes (NB), and Artificial Neural Network (ANN), have all been used in reported studies of sleep disorders, together with Time and/or Frequency-domain characteristics.

ANNs are frequently used because they allow non-linear activation functions combined with computational efficiency. For both linear and non-linear classification tasks, SVMs can be useful. Non-linear classification with an SVM is based on a kernel function that maps the feature space non-linearly for classification, but does it implicitly via manipulating the inner product. The KNN method is a non-linear lazy learning method that can be used to solve classification and regression problems. When it comes to classification speed and accuracy, NB does better with large training datasets. Figure 6 shows a summary of the state-of-the-art machine learning algorithms employed.

Machine Learning (ML) approaches have been used in a number of studies on automatic sleep rating. The reason for this is possibly the rise in number of people suffering sleep disorders world-wide. Another conclusion is that due to developments in ML approaches in other biomedical domains, researchers and/or health practitioners are working on automatic sleep scoring systems.

Deep neural networks were used in a small number of studies that did not use feature engineering approaches. The reported levels of accuracy range around 84 to 88 percent, although feature engineering methods have yielded more accurate results at the cost of more human labor in the implementation of classification.

There is no discernible trend in the techniques' performance. The technique used, the size of the data collection, the target (i.e., the number of sleep stages to detect), and the feature extraction techniques used all influence the accuracy of the results. The performance of machine learning systems for automatic sleep scoring is unclear, which supports this observation. The absence of evident advances in classifier performance could be due to the fact that this discipline is very young, and when more trials are published in scientific journals, such trends may develop.

Given the challenge of high dimensional data, and possible changes in measurements and thereby in sets of features, ML has a distinct advantage because of its adaptability to changes. The designer of automatic sleep scoring does not need to provide solutions for all possible situations, instead the developed tools can adapt by re-training to new types of data. Adaptability to and learning from a changing environment are major strengths of ML.

Data are analyzed using machine learning algorithms to extract patterns and information. We seek to crystallize the information in accumulated data by learning from it a classifier of sleep stages. A more thorough examination of the existing machine learning approaches, as well as their advantages and disadvantages, is required, and the ML perspective on automatic sleep scoring has to be further elaborated and pursued.

VIII. DEEP LEARNING TECHNIQUES

A. HAND ENGINEERED FEATURES

Hand engineering features for sleep stage classification affect (1) signal preprocessing and dataset preparation, (2) extraction of features, (3) classification, and (4) performance evaluation [153]. Filtering and normalizing of signals are included in the preprocessing phase. Signal features can be extracted using time, frequency, or time-frequency domain features. The classifiers in the third phase represent typical machine learning techniques. Finally, the last phase is performance evaluation.

B. WITHOUT HAND ENGINEERED FEATURES

Deep learning processes data hierarchically in multiple layers, extracting highly non-linear and complex features. Computer vision and natural language processing applications are the key drivers in this domain. In sleep data analysis, it involves preprocessing and labelling the dataset for 30-second epochs. The network's ability to extract a near-optimal collection of features without human bias is advantageous. Deep Belief Nets (DBNs) and Convolutional Neural Networks (DeepSleepNet) are examples of deep learning [119], [122]. Although the methods can be used on raw data, their blackbox nature is disadvantageous [154].

C. STRUCTURING OF DEEP LEARNING TECHNIQUES

1) AUTOENCODER

These expand linear dimension reduction (commonly done with principal component analysis) to nonlinear dimension reduction, using a neural network with a bottleneck layer that encodes, while this encoding is expanded to (approximately) the input as the target output.

Does not require labelled data for training and has many variants, e.g., k-sparse, de-noising, contractive, and separable deep auto-encoder. The vanishing gradient problem affects the trained model as well.

2) RESTRICTED BOLTZMANN MACHINE

It's a bidirectionally trained stochastic neural network. Contrasting divergence is used to speed up the sampling procedure. Because it is trained without supervision, there is no guarantee that the features extracted from Restricted Boltzmann Machine hidden layer will be helpful for supervised work in the future.

3) DEEP BELIEF NET

It combines Restricted Boltzmann Machine and sigmoid belief networks to provide a deep hierarchical representation of the training data. Pre-training and discriminative fine-tuning are the two stages in the training process. The advantages include: high-dimensional raw data is transformed into a homogeneous representation. It is good in learning features, processing unlabeled data, and avoiding problems with over-fitting, but run-time complexity is high.

4) RECURRENT NEURAL NETWORK

These are used for time sequence data, with the output depending on previous computations, and the same weights are shared by sets of nodes. It maintains data in the form of activations and is utilized in natural language processing. It has problems with gradient vanishing and exploding, and it can't be layered for really deep models.

5) CONVOLUTIONAL NEURAL NETWORK

In at least one of its layers, it is a Neural Network that uses convolution operations instead of basic matrix multiplication. It can handle sparsity in data and share the parameters in different functions. A limitation is that a large amount of training data is required.

6) GENERATIVE ADVERSARIAL NETWORK

It's made up of two models: a generative G model and a distribution D model. This can be used in any domain, including music, images, and speech. It is not required to have Mote

Carlo approximation in training of generative adversarial network and it is faster than completely transparent belief nets at generating samples. It is unable to generate discrete data, such as text.

D. CHALLENGES OF DEEP LEARNING TECHNIQUES

Data Volume: The polysomnography signals are more complex than many other data types, because each patient recording spans overnight (8 hours). The training time also increases with this huge amount of data, or big data.

Data Quality: Heterogeneity, noise, improper recording devices, fluctuations in voltage, faults in instruments, blinks of the eye, movements of the eyes, muscular movement, and missing values due to other reasons, pose challenges to DL. The DL model needs to tolerate sparsity, missing values, and data redundancy.

Temporality: The static vector based models cannot deal with dynamic changes happening as time passes.

Domain Complexity: Sickness of heterogeneous type, unavailability of information about majority of ailments, and limited number of patients add complexity to the domain.

Interpretability: In biomedicine, quantitative algorithms as well as significance estimates are also important. Model interpretability is vital for gaining expert confidence to ML based calls.

IX. BIG DATA IN SLEEP SCIENCE AND MEDICINE

''Big data'' as a modern term refers to a situation with a large amount of data that is complex and heterogeneous so that conventional techniques are unable to analyze it. It demands large computational resources for processing and analysis, therefore, the term ''Big Data Analytics'' has been coined. Typically one wants to extract significant patterns, trends, interactions, and associations. Three V's are used to characterize many big data situations: Velocity (Data Acquisition speed), Volume (Amount of Data), and Variety (Number of sources to create big data sets) [155], [156].

A wealth of physiological information is available in polysomnography, which is helpful in clinical research and decision making. The databases available to public include those at NSRR (National Sleep Research Resource), NHLBI (a new National Heart, Lung, and Blood Institute), PhysioNet (accessible at www.physionet.com), and the MASS (Montreal Archive of Sleep Studies). Clinical databases support big data research goals and can support heterogeneity (a heterogeneous dataset possibly suitable for clustering and other exploratory methods) and diversity.

However: ''Academic centres may have different referral biases, for example, being enriched for complicated cases. Although most clinical laboratories have standardized physiological recording protocols, the collection of self-reported clinical information may not be standardized. Variation across recording and scoring technologists may contribute heterogeneity despite quality efforts required in accredited laboratories. Centralized scoring common to large clinical trials may not be practical for clinical databases'' [155].

X. INTERRATER VARIABILITY OF MANUAL AND AUTOMATED SLEEP SCORING

In sleep stage scoring, interrater variability is well-known and requires clarification. The degree of (dis)agreement between experienced sleep scoring experts is called interrater reliability [157]. This variability exists due to different rules used to score events and their interpretation. The variability exist because it is difficult to determine whether in transitional epochs, the wake stage lasts longer than 15 seconds, spindles of sleep are present, and delta waves last longer than 6 seconds in a 30-second epoch.

It is critical to assess the reliability of human-assisted manual sleep scoring. Credibility necessitates a high level of trustworthiness. The AASM Inter-scorer Reliability Program, which began in April 2010, was created for this aim. The evaluation of inter-scorer dependability can be done based on a very large number of scorers due to the experience gathered through this program. The sleep stage R has a high level of reliability, with 90.5 percent agreement, while the sleep stage N1 has the lowest level of agreement, just 63.0 percent [158].

Manual or visual scoring in sleep medicine involves rules, such as those given by R&K (Rechtschaffen and Kales) almost half a century ago [159]. Manual scoring by an expert is expensive and time consuming by its nature [160], and has many limitations: (I) Sleep depth is thought to progress in stages from light to moderate to profound. (II) In stages 1 to 3 of non-REM sleep, there is a lot of inter-scorer variability on this small scale. (III) Variability among EEG features: sleep spindles and K complexes, arousal intensity, alpha intrusion amount and frequency.

Digital analysis may reduce the variability in labelling stages, and solve the above mentioned problems, but it is challenging to develop that system. It is argued that such systems have implemented the R&K rules efficiently but do not explore the micro structure of sleep that contains clinically important information. In the past decades, for automated sleep scoring, numerous systems have been developed, with some of them clinically proven. However, their use in clinical practice faces resistance. The main criticism leveled against these systems is that they are unreliable and require human assistance. Therefore, the apparent advantages in economy, speed and consistency are lost partly due to the lack of human trust [161].

The study [162] presents arguments favoring an automated system. The evidence for benefits: the digital system can reproduce the R&K staging, and more information is obtained than from manual scoring; further, the automated system demonstrated yields similar calls as those by experienced technologists. It is concluded that laboratory efficiency may increase if manual editing is supplemented.

It seems appealing to develop a home based sleep monitoring system or device, but lack of adequate monitoring of EEG and quantification of sleep time objectively are major obstacles. Manual scoring of a home sleep testing system would add considerable costs. Further, the manual scoring has both inter-scorer and intra-scorer variability. Therefore, its reliability and reproducibility are questionable. Numerous attempts have been made for automation but the available systems have moderate accuracy and are perceived as expensive.

To address the above considerations, some studies have assessed interrater variability and reliability. Further, there is no acceptable tolerance limit for accuracy of an automated sleep scoring system. The various sleep research centres would need to collaborate to measure the variability among their systems. However, while the clinical acceptance of automated systems is low, they are superior to manual systems because they require less labour, and several attempts have been made to reduce the criticism. As far as the home based clinical systems are concerned, the above arguments show that a lot of research would be required to develop a handy portable device for novice users.

It is obvious from the databases investigated that studies on automatic sleep scoring have intensified in the preceding decade. We believe this is due to the availability of published sleep data in the modern IT systems. Another aspect is the high expense of manual sleep stage analysis by sleep technologists or experts, which has fueled the demand for automated sleep scoring. Brainstorming, feature selection or creation, appropriateness evaluation, feature improvement, and repeating as needed are all common feature engineering processes. A further factor contributing to this trend is the overall drive toward automation, particularly in measurement, in order to eliminate operator-dependent outcomes.

XI. CONCLUSION

The current study's major purpose was to assess state-of-the-art in sleep scoring standards, bio-electric signal feature extraction methodologies, and sleep data classification using machine learning approaches. The study's second goal was to figure out how the components above are combined to create an autonomous sleep rating system. The project aims to learn more about the human sleep problem while also providing a foundation for young researchers.

Two sleep scoring standards have been identified in this study: (i) R&K, and (ii) AASM. The second finding is that the features fall into four categories: (i) time domain, (ii) frequency domain, (iii) time-frequency domain, and (iv) nonlinear domain. Finally, sleep data are frequently collected using both single and multiple channel signals. Additionally, sleep data are classified using a set of machine learning approaches.

It is obvious from the databases examined that automatic sleep scoring studies have been strongly pursued over the past decade. A small number of feature extraction approaches were used. Automated techniques often dominate manual feature engineering. The study's main finding is that machine learning techniques have been deployed without fully utilizing their adjustable parameters. Furthermore, based on prediction speed, memory utilization, and call interpretability or traceability, or which categorization algorithms would be the most effective for sleep data analysis, have yet to be determined. Furthermore, it was discovered that classifier performance is affected not only by the size of the data set, but also by the feature extraction choices.

These findings suggest that the AASM standard should be used and that the collection of datasets should be expanded. Other sorts of signals outside EEG can also be investigated, according to the review. Multiple aspects must be considered since each one contributes to understanding sleep in the diagnostic context.

Automatic feature engineering techniques and parameter choices for machine learning algorithms, in the context of sleep scoring, are two areas that need to be researched more in the future. Parameter adjustment will not only allow for a fair comparison of machine learning options, but it may also enhance accuracy to a level comparable to sleep expert calls. As a result, this study indicates that using an alternate sleep scoring standard, as well as numerous feature extraction with selection approaches, machine learning algorithms with parameter tweaking, and big data analytics, physicians can produce a practically useful automatic sleep scoring system.

ACKNOWLEDGMENT

The authors would like to thank Gideon Mbiydzenyuy, a Senior Lecturer of the University of Borås, for comments that helped significantly improve the manuscript.

REFERENCES

[1] R. Pelayo, N. Hodgson, and C. Guilleminault, "The history of sleep medicine," in *History of Neurology* (Handbook of Clinical Neurology), vol. 95, M. J. Aminoff, F. Boller, and D. F. Swaab, Eds. Amsterdam, The Netherlands: Elsevier, 2009, pp. 547–556. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0072975208021349

[2] A. Y. Avidan, B. V. Vaughn, and M. H. Silber, "The current state of sleep medicine education in u.s. neurology residency training programs: Where do we go from here?" *J. Clin. Sleep Med.*, vol. 9, no. 3, pp. 281–286, Mar. 2013.

[3] A. Liberati, D. G. Altman, J. Tetzlaff, C. Mulrow, P. C. Gotzsche, J. P. A. Ioannidis, M. Clarke, P. J. Devereaux, J. Kleijnen, and D. Moher, "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration," *BMJ*, vol. 339, Dec. 2009, Art. no. b2700. [Online]. Available: http://www.bmj.com/content/339/bmj.b2700

[4] J. A. Hobson, "A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects," *Electroencephalogr. Clin. Neurophysiol.*, vol. 26, no. 6, p. 644, Jun. 1969.

[5] C. Iber, "Respiratory rules," in *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Darien, IL, USA: American Academy of Sleep Medicine, 2007, pp. 45–50.

[6] S. Fallmann and L. Chen, "Computational sleep behavior analysis: A survey," *IEEE Access*, vol. 7, pp. 142421–142440, 2019.

[7] B. Şen, M. Peker, A. Çavuşoğlu, and F. V. Çelebi, "A comparative study on classification of sleep stage based on EEG signals using feature selection and classification algorithms," *J. Med. Syst.*, vol. 38, no. 3, pp. 1–21, Mar. 2014.

[8] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.

[9] V. Krajca, S. Petranek, K. Paul, M. Matousek, J. Mohylova, and L. Lhotska, "Automatic detection of sleep stages in neonatal EEG using the structural time profiles," in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Sep. 2005, pp. 1–4.

[10] J. V. Marcos, R. Hornero, D. Álvarez, F. del Campo, and C. Zamarrón, "Assessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry," *Med. Eng. Phys.*, vol. 31, no. 8, pp. 971–978, Oct. 2009.

[11] K. Paul, V. Krajča, Z. Roth, J. Melichar, and S. Petránek, ''Quantitative topographic differentiation of the neonatal EEG,'' *Clin. Neurophysiol.*, vol. 117, no. 9, pp. 2050–2058, Sep. 2006.

[12] S. Khalighi, T. Sousa, and U. Nunes, ''Adaptive automatic sleep stage classification under covariate shift,'' in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 2259–2262.

[13] S. T.-B. Hamida and B. Ahmed, ''Computer based sleep staging: Challenges for the future,'' in *Proc. 7th IEEE GCC Conf. Exhib. (GCC)*, Nov. 2013, pp. 280–285.

[14] T. Penzel, J. W. Kantelhardt, L. Grote, J. H. Peter, and A. Bunde, ''Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea,'' *IEEE Trans. Biomed. Eng.*, vol. 50, no. 10, pp. 1143–1151, Oct. 2003.

[15] B. Ahmed, A. Redissi, and R. Tafreshi, ''An automatic sleep spindle detector based on wavelets and the teager energy operator,'' in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 2596–2599.

[16] A. Erdamar, F. Duman, and S. Yetkin, ''A wavelet and teager energy operator based method for automatic detection of K-complex in sleep EEG,'' *Expert Syst. Appl.*, vol. 39, no. 1, pp. 1284–1290, Jan. 2012.

[17] S. H. Na, S.-H. Jin, and S. Y. Kim, ''The effects of total sleep deprivation on brain functional organization: Mutual information analysis of waking human EEG,'' *Int. J. Psychophysiol.*, vol. 62, no. 2, pp. 238–242, Nov. 2006.

[18] O. A. Rosso, M. T. Martin, and A. Plastino, ''Brain electrical activity analysis using wavelet-based informational tools (II): Tsallis non-extensivity and complexity measures,'' *Phys. A, Stat. Mech. Appl.*, vol. 320, pp. 497–511, Mar. 2003.

[19] N. A. Lofgren, N. Outram, and M. Thordstein, ''EEG entropy estimation using a Markov model of the EEG for sleep stage separation in human neonates,'' in *Proc. 3rd Int. IEEE/EMBS Conf. Neural Eng.*, May 2007, pp. 634–637.

[20] S. Motamedi-Fakhr, M. Moshrefi-Torbati, M. Hill, C. M. Hill, and P. R. White, ''Signal processing techniques applied to human sleep EEG signals—A review,'' *Biomed. Signal Process. Control*, vol. 10, pp. 21–33, Mar. 2014.

[21] E. Huupponen, G. Gómez-Herrero, A. Saastamoinen, A. Värri, J. Hasan, and S.-L. Himanen, ''Development and comparison of four sleep spindle detection methods,'' *Artif. Intell. Med.*, vol. 40, no. 3, pp. 157–170, Jul. 2007.

[22] A. Okumura, T. Kubota, N. Toyota, H. Kidokoro, K. Maruyama, T. Kato, F. Hayakawa, and K. Watanabe, ''Amplitude spectral analysis of maturational changes of delta waves in preterm infants,'' *Brain Develop.*, vol. 25, no. 6, pp. 406–410, Sep. 2003.

[23] D. P. Subha, P. K. Joseph, R. Acharya, and C. M. Lim, ''EEG signal analysis: A survey,'' *J. Med. Syst.*, vol. 34, no. 2, pp. 195–212, 2010.

[24] J. Pardey, S. Roberts, and L. Tarassenko, ''A review of parametric modelling techniques for EEG analysis,'' *Med. Eng. Phys.*, vol. 18, no. 1, pp. 2–11, Jan. 1996.

[25] A. Schlögl, P. Anderer, S. Roberts, M. Pregenzer, and G. Pfurtscheller, ''Artefact detection in sleep EEG by the use of Kalman filtering,'' in *Proc. EMBEC*, vol. 99. Princeton, NJ, USA: Citeseer, 1999, pp. 1648–1649.

[26] K. Schwab, M. Eiselt, P. Putsche, M. Helbig, and H. Witte, ''Time-variant parametric estimation of transient quadratic phase couplings between heart rate components in healthy neonates,'' *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1077–1083, Dec. 2006.

[27] J. Fell, J. Röschke, K. Mann, and C. Schäffner, ''Discrimination of sleep stages: A comparison between spectral and nonlinear EEG measures,'' *Electroencephalogr. Clin. Neurophysiol.*, vol. 98, no. 5, pp. 401–410, 1996.

[28] I. Korotchikova, S. Connolly, C. Ryan, D. Murray, A. Temko, B. Greene, and G. Boylan, ''EEG in the healthy term newborn within 12hours of birth,'' *Clin. Neurophysiol.*, vol. 120, no. 6, pp. 1046–1053, 2009.

[29] M. Eiselt, M. Schendel, H. Witte, J. Dörschel, L. Curzi-Dascalova, A. M. D'Allest, and U. Zwiener, ''Quantitative analysis of discontinuous EEG in premature and full-term newborns during quiet sleep,'' *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 5, pp. 528–534, Nov. 1997.

[30] D. B. Melges, A. Infantosi, F. Ferreira, and D. Rosas, ''Using the discrete Hilbert transform for the comparison between tracé alternant and high voltage slow patterns extracted from full-term neonatal EEG,'' in *World Congress on Medical Physics and Biomedical Engineering*. Cham, Switzerland: Springer, 2007, pp. 1111–1114.

[31] P. Xanthopoulos, S. Golemati, V. Sakkalis, P. Y. Ktonas, M. Zervakis, and C. R. Soldatos, ''Modeling the time-varying microstructure of simulated sleep EEG spindles using time-frequency analysis methods,'' in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 2438–2441.

[32] I. Zhovna and I. D. Shallom, ''Automatic detection and classification of sleep stages by multichannel EEG signal modeling,'' in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 2665–2668.

[33] M. J. Kamiński, K. Blinowska, and W. Szelenberger, ''Topographic analysis of coherence and propagation of EEG activity during sleep and wakefulness,'' *Electroencephalogr. Clin. Neurophysiol.*, vol. 102, no. 3, pp. 216–227, Mar. 1997.

[34] M. Diykh, Y. Li, and P. Wen, ''EEG sleep stages classification based on time domain features and structural graph similarity,'' *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 11, pp. 1159–1168, Nov. 2016.

[35] J. Löfhede, M. Thordstein, N. Löfgren, A. Flisberg, M. Rosa-Zurera, I. Kjellmer, and K. Lindecrantz, ''Automatic classification of background EEG activity in healthy and sick neonates,'' *J. Neural Eng.*, vol. 7, no. 1, Feb. 2010, Art. no. 016007.

[36] A. R. Hassan, S. K. Bashar, and M. I. H. Bhuiyan, ''Automatic classification of sleep stages from single-channel electroencephalogram,'' in *Proc. Annu. IEEE India Conf. (INDICON)*, Dec. 2015, pp. 1–6.

[37] R. Armitage, R. F. Hoffmann, and A. J. Rush, ''Biological rhythm disturbance in depression: Temporal coherence of ultradian sleep EEG rhythms,'' *Psychol. Med.*, vol. 29, no. 6, pp. 1435–1448, Nov. 1999.

[38] P. G. Grieve, J. R. Isler, A. Izraelit, B. S. Peterson, W. P. Fifer, M. M. Myers, and R. I. Stark, ''EEG functional connectivity in term age extremely low birth weight infants,'' *Clin. Neurophysiol.*, vol. 119, no. 12, pp. 2712–2720, Dec. 2008.

[39] V. Kokkinos, A. Koupparis, M. L. Stavrinou, and G. K. Kostopoulos, ''The hypnospectrogram: An EEG power spectrum based means to concurrently overview the macroscopic and microscopic architecture of human sleep,'' *J. Neurosci. Methods*, vol. 185, no. 1, pp. 29–38, Dec. 2009.

[40] R. Lin, R. Lee, C. Tseng, H. Zhou, C. Chao, and J. Jiang, ''A new approach for identifying sleep apnea syndrome using wavelet transform and neural networks,'' *Biomed. Eng. Appl. Basis Commun.*, vol. 18, no. 3, pp. 138–143, 2006.

[41] E. F. Estrada, H. Nazeran, and H. Ochoa, ''HRV and EEG signal features for computer-aided detection of sleep apnea,'' in *Proc. 25th Southern Biomed. Eng. Conf.* Miami, FL, USA: Springer, May 2009, pp. 265–266.

[42] P. J. Durka and K. J. Blinowska, ''Analysis of EEG transients by means of matching pursuit,'' *Ann. Biomed. Eng.*, vol. 23, no. 5, pp. 608–611, Sep. 1995.

[43] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, ''The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,'' *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, pp. 903–995, Mar. 1971.

[44] Z. Yang, L. Yang, and D. Qi, ''Detection of spindles in sleep EEGs using a novel algorithm based on the Hilbert–Huang transform,'' in *Wavelet Analysis and Applications*. Cham, Switzerland: Springer, 2006, pp. 543–559.

[45] N. Baydar and A. Ball, ''A comparative study of acoustic and vibration signals in detection of gear failures using Wigner–Ville distribution,'' *Mech. Syst. Signal Process.*, vol. 15, no. 6, pp. 1091–1107, Nov. 2001.

[46] M. Sun, S. Qian, X. Yan, S. B. Baumann, X.-G. Xia, R. E. Dahl, N. D. Ryan, and R. J. Sclabassi, ''Localizing functional activity in the brain through time-frequency analysis and synthesis of the EEG,'' *Proc. IEEE*, vol. 84, no. 9, pp. 1302–1311, Sep. 1996.

[47] T. Kobayashi, K. Misaki, H. Nakagawa, S. Madokoro, H. Ihara, K. Tsuda, Y. Umezawa, J. Murayama, and K. Isaki, ''Non-linear analysis of the sleep EEG,'' *Psychiatry Clin. Neurosci.*, vol. 53, no. 2, pp. 159–161, 1999.

[48] M. Grözinger, J. Fell, and J. Röschke, ''Neural net classification of REM sleep based on spectral measures as compared to nonlinear measures,'' *Biol. Cybern.*, vol. 85, no. 5, pp. 335–341, Nov. 2001.

[49] D. Gallez and A. Babloyantz, ''Predictability of human EEG: A dynamical approach,'' *Biol. Cybern.*, vol. 64, no. 5, pp. 381–391, Mar. 1991.

[50] N. Pradhan and P. K. Sadasivan, ''The nature of dominant Lyapunov exponent and attractor dimension curves of EEG in sleep,'' *Comput. Biol. Med.*, vol. 26, no. 5, pp. 419–428, Sep. 1996.

[51] J. Fell, J. Röschke, and P. Beckmann, ''Deterministic chaos and the first positive Lyapunov exponent: A nonlinear analysis of the human electroencephalogram during sleep,'' *Biol. Cybern.*, vol. 69, no. 2, pp. 139–146, Jun. 1993.

[52] M. Carrozzi, A. Accardo, and F. Bouquet, ''Analysis of sleep-stage characteristics in full-term newborns by means of spectral and fractal parameters,'' *Sleep*, vol. 27, no. 7, pp. 1384–1393, Oct. 2004.

[53] B. S. Raghavendra and D. N. Dutt, ''Multiresolution area-based fractal dimension estimation of signals applied to EEG data,'' in *Proc. TENCON IEEE Region Conf.*, Nov. 2008, pp. 1–5.

[54] R. Ferenets, T. Lipping, P. Suominen, J. Turunen, P. Puumala, V. Jantti, S.-L. Himanen, and A.-M. Huotari, ''Comparison of the properties of EEG spindles in sleep and propofol anesthesia,'' in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 6356–6359.

[55] I. H. Song, Y. S. Ji, B. K. Cho, J. H. Ku, Y. J. Chee, J. S. Lee, S. M. Lee, I. Y. Kim, and S. I. Kim, ''Multifractal analysis of sleep EEG dynamics in humans,'' in *Proc. 3rd Int. IEEE/EMBS Conf. Neural Eng.*, May 2007, pp. 546–549.

[56] W.-X. He, X.-G. Yan, X.-P. Chen, and H. Liu, ''Nonlinear feature extraction of sleeping EEG signals,'' in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Jan. 2005, pp. 4614–4617.

[57] G. Jiayi, Z. Peng, Z. Xin, and W. Mingshi, ''Sample entropy analysis of sleep EEG under different stages,'' in *Proc. IEEE/ICME Int. Conf. Complex Med. Eng.*, May 2007, pp. 1499–1502.

[58] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, ''Recurrence plots of dynamical systems,'' *EPL Europhys. Lett.*, vol. 4, no. 9, p. 973, 1987.

[59] R. Acharya, O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan, ''Non-linear analysis of EEG signals at various sleep stages,'' *Comput. Methods Programs Biomed.*, vol. 80, pp. 37–45, Oct. 2005.

[60] V. C. Figueroa Helland, S. Postnova, U. Schwarz, J. Kurths, B. Kundermann, U. Hemmeter, and H. A. Braun, ''Comparison of different methods for the evaluation of treatment effects from the sleep EEG of patients with major depression,'' *J. Biol. Phys.*, vol. 34, nos. 3–4, pp. 393–404, Aug. 2008.

[61] A. Accardo, M. Affinito, M. Carrozzi, S. Cisint, and F. Bouquet, ''Comparison between spectral and fractal EEG analyses of sleeping newborns,'' in *Proc. 20th Annu. Int. Conf. Eng. Med. Biol. Soc.*, vol. 3, Oct. 1998, pp. 1569–1571.

[62] S. Dangel, P. Meier, H. Moser, S. Plibersek, and Y. Shen, ''Time series analysis of sleep EEG,'' *Comput. Assist. Phys.*, pp. 93–95, 1999.

[63] P. Kavipriya and K. Karthikeyan, ''A comparative study of feature selection algorithms in data mining,'' *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 6, no. 11, pp. 1–7, 2017.

[64] B. Şen, M. Peker, A. Çavuşoğlu, and F. V. Çelebi, ''A comparative study on classification of sleep stage based on EEG signals using feature selection and classification algorithms,'' *J. Med. Syst.*, vol. 38, no. 3, p. 18, Mar. 2014.

[65] R. Boostani, F. Karimzadeh, and M. Nami, ''A comparative review on sleep stage classification methods in patients and healthy individuals,'' *Comput. Methods Programs Biomed.*, vol. 140, pp. 77–91, Mar. 2017.

[66] S. Motamedi-Fakhr, M. Moshrefi-Torbati, M. Hill, C. M. Hill, and P. R. White, ''Signal processing techniques applied to human sleep EEG signals—A review,'' *Biomed. Signal Process. Control*, vol. 10, pp. 21–33, Mar. 2014.

[67] I. N. Bankman, V. G. Sigillito, R. A. Wise, and P. L. Smith, ''Feature-based detection of the K-complex wave in the human electroencephalogram using neural networks,'' *IEEE Trans. Biomed. Eng.*, vol. 39, no. 12, pp. 1305–1310, Dec. 1992.

[68] C. Strungaru and M. S. Popescu, ''Neural network for sleep EEG K-complex detection,'' *Biomedizinische Technik/Biomed. Eng.*, vol. 43, no. s3, pp. 113–116, 1998.

[69] D. P. Baraglia, M. J. Berryman, S. W. Coussens, Y. Pamula, D. Kennedy, A. J. Martin, and D. Abbott, ''Automated sleep scoring and sleep apnea detection in children,'' *Proc. SPIE*, vol. 6039, Jan. 2006, Art. no. 60390T.

[70] J. Rodríguez-Sotelo, A. Osorio-Forero, A. Jiménez-Rodríguez, F. Restrepo-de Mejía, D. Peluffo-Ordonez, and J. Serrano, ''Sleep stages clustering using time and spectral features of EEG signals,'' in *Proc. Int. Work-Conf. Interplay Between Natural Artif. Comput.* Cham, Switzerland: Springer, 2017, pp. 444–455.

[71] R. K. Sinha, ''Artificial neural network and wavelet based automated detection of sleep spindles, REM sleep and wake states,'' *J. Med. Syst.*, vol. 32, no. 4, pp. 291–299, Aug. 2008.

[72] A. S. Gevins, R. K. Stone, and S. D. Ragsdale, ''Differentiating the effects of three benzodiazepines on non-REM sleep EEG spectra,'' *Neuropsychobiology*, vol. 19, no. 2, pp. 108–115, 1988.

[73] D. Álvarez-Estévez and V. Moret-Bonillo, ''Model comparison for the detection of EEG arousals in sleep apnea patients,'' in *Proc. Int. Work-Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, 2009, pp. 997–1004.

[74] N. McGrogan, E. Braithwaite, and L. Tarassenko, ''BioSleep: A comprehensive sleep analysis system,'' in *Proc. Conf. Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Oct. 2001, pp. 1608–1611.

[75] D. Balakrishnan and S. Puthusserypady, ''Multilayer perceptrons for the classification of brain computer interface data,'' in *Proc. IEEE 31st Annu. Northeast Bioeng. Conf.*, Apr. 2005, pp. 118–119.

[76] L. Zoubek, S. Lesecq, F. Chapotot, S. Charbonnier, and A. G. C. Buguet, ''Feature selection for sleep/wake stages classification using data driven methods,'' *Biomed. Signal Process. Control*, vol. 2, no. 3, pp. 171–179, Jul. 2007.

[77] E. Huupponen, A. Värri, S. Himanen, J. Hasan, M. Lehtokangas, and J. Saarinen, ''Autoassociative mlp in sleep spindle detection,'' *J. Med. Syst.*, vol. 24, no. 3, pp. 183–193, 2000.

[78] M. Corsi-Cabrera, M. A. Guevara, and Y. del Río-Portilla, ''Brain activity and temporal coupling related to eye movements during REM sleep: EEG and MEG results,'' *Brain Res.*, vol. 1235, pp. 82–91, Oct. 2008.

[79] N. Acır and C. S. Güzeliş, ''Automatic recognition of sleep spindles in EEG by using artificial neural networks,'' *Expert Syst. Appl.*, vol. 27, no. 3, pp. 451–458, 2004.

[80] M. Grözinger, C. Wolf, T. Uhl, C. Schäffner, and J. Röschke, ''Online detection of rem sleep based on the comprehensive evaluation of short adjacent EEG segments by artificial neural networks,'' *Prog. Neuro-Psychopharmacol. Biol. Psychiatry*, vol. 21, no. 6, pp. 951–963, Aug. 1997.

[81] N. Sezgin and M. E. Tagluk, ''Energy based feature extraction for classification of sleep apnea syndrome,'' *Comput. Biol. Med.*, vol. 39, no. 11, pp. 1043–1050, Nov. 2009.

[82] F. Ebrahimi, M. Mikaeili, E. Estrada, and H. Nazeran, ''Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients,'' in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 1151–1154.

[83] M. H. Schwaibold, T. Penzel, J. Schochlin, and A. Bolz, ''Combination of AI components for biosignal processing application to sleep stage recognition,'' in *Proc. Conf. Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Oct. 2001, pp. 1692–1694.

[84] A. Berger, D. P. Möller, and M. Reuter, ''Detection of sleep with new preprocessing methods for EEG analysing,'' in *Proc. Int. Conf. Comput. Intell.* Cham, Switzerland: Springer, 1997, pp. 304–310.

[85] J. Y. Tian and J. Q. Liu, ''Automated sleep staging by a hybrid system comprising neural network and fuzzy rule-based reasoning,'' in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Jan. 2005, pp. 4115–4118.

[86] M. Golz, D. Sommer, T. Lembcke, B. Kurella, and F. Schmalkalden, ''Classification of pre-stimulus EEG of k-complexes using competitive learning networks,'' *Proc. Eufit*, vol. 98, 1998, pp. 1767–1771.

[87] D. Michie, *Machine Learning: Neural and Statistical Classification*. London, U.K.: Overseas Press, 2009.

[88] P. Lima, J. Leitao, and T. Paiva, ''Artifact detection in sleep EEG recording,'' in *Proc. Electrotech. Conf. Integrating Res., Ind. Educ. Energy Commun. Eng.*, Apr. 1989, pp. 273–277.

[89] M. S. Scher, S. G. Dokianakis, M. Sun, D. A. Steppe, R. D. Guthrie, and R. J. Sclabassi, ''Computer classification of sleep in preterm and full-term neonates at similar postconceptional term ages,'' *Sleep*, vol. 19, no. 1, pp. 18–25, Jan. 1996.

[90] P. R. Galhanone, D. M. Simpson, A. F. C. Infantosi, E. Faveret, M. A. Genofre, H. Bello, and L. de Azevedo, ''Multivariate analysis of neonatal EEG in different sleep stages: Methods and preliminary results,'' in *Proc. 17th Int. Conf. Eng. Med. Biol. Soc.*, Sep. 1995, pp. 1021–1022.

[91] I. A. Rezek and S. J. Roberts, ''Stochastic complexity measures for physiological signal analysis,'' *IEEE Trans. Biomed. Eng.*, vol. 45, no. 9, pp. 1186–1191, Sep. 1998.

[92] S. Cho, H. Choi, H. Lee, and K. Lee, ''Detection of EEG arousals in patients with respiratory sleep disorder,'' in *World Congress on Medical Physics and Biomedical Engineering 2006*. Cham, Switzerland: Springer, 2007, pp. 1131–1134.

[93] S. Gudmundsson, T. P. Runarsson, and S. Sigurdsson, ''Automatic sleep staging using support vector machines with posterior probability estimates,'' in *Proc. Int. Conf. Comput. Intell. Modelling, Control Automat. Int. Conf. Intell. Agents, Web Technol. Internet Commerce (CIMCA-IAWTIC)*, vol. 2, Nov2005, pp. 366–372.

[94] W.-C. Tang, S.-W. Lu, C.-M. Tsai, C.-Y. Kao, and H.-H. Lee, "Harmonic parameters with HHT and wavelet transform for automatic sleep stages scoring," *REM*, vol. 365, pp. 6–8, Sep. 2007.

[95] Y.-H. Lee, Y.-S. Chen, and L.-F. Chen, "Automated sleep staging using single EEG channel for REM sleep deprivation," in *Proc. 9th IEEE Int. Conf. Bioinf. BioEng.*, Jun. 2009, pp. 439–442.

[96] A. Flexerand, G. Dorffner, P. Sykacekand, and I. Rezek, "An automatic, continuous and probabilistic sleep stager based on a hidden Markov model," *Appl. Artif. Intell.*, vol. 16, no. 3, pp. 199–207, Mar. 2002.

[97] A. Flexer, G. Gruber, and G. Dorffner, "A reliable probabilistic sleep stager based on a single EEG signal," *Artif. Intell. Med.*, vol. 33, no. 3, pp. 199–207, 2005.

[98] V. Gerla, V. Djordjevic, L. Lhotska, and V. Krajca, "System approach to complex signal processing task," in *Proc. Int. Conf. Comput. Aided Syst. Theory*. Cham, Switzerland: Springer, 2009, pp. 579–586.

[99] A. Kam, A. Cohen, A. Geva, and A. Tarasiuk, "Detection of k-complexes in sleep EEG using CD-HMM," in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 1, Sep. 2004, pp. 33–36.

[100] B. R. Kosanovic, L. F. Chaparro, and R. J. Sclabassi, "Signal modeling with dynamic fuzzy sets," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. Conf.*, May 1996, pp. 2829–2832.

[101] D. Novak, D. Cuesta-Frau, T. A. ani, M. Aboy, R. Mico, and L. Lhotska, "Speech recognition methods applied to biomedical signals processing," in *Proc. 26th Annu. Int. Conf. Eng. Med. Biol. Soc. (IEMBS)*, vol. 1, Sep. 2004, pp. 118–121.

[102] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.

[103] S. Qureshi, S. Karrila, and S. Vanichayobon, "Human sleep scoring based on K-nearest neighbors," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 26, no. 6, pp. 2803–2819, Nov. 2018.

[104] S. Güneş, K. Polat, and Ş. Yosunkaya, "Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 7922–7928, 2010.

[105] M. O. Mendez, D. D. Ruini, O. P. Villantieri, M. Matteucci, T. Penzel, S. Cerutti, and A. M. Bianchi, "Detection of sleep apnea from surface ECG based on features extracted by an autoregressive model," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 6105–6108.

[106] B. Yılmaz, M. H. Asyalı, E. Arıkan, S. Yetkin, and F. Özgen, "Sleep stage and obstructive apneaic epoch classification using single-lead ECG," *Biomed. Eng. OnLine*, vol. 9, no. 1, p. 39, 2010.

[107] T. K. Padma Shri and N. Sriraam, "Comparison of T-test ranking with PCA and SEPCOR feature selection for wake and stage 1 sleep pattern recognition in multichannel electroencephalograms," *Biomed. Signal Process. Control*, vol. 31, pp. 499–512, Jan. 2017.

[108] V. Kokkinos, A. Koupparis, M. L. Stavrinou, and G. K. Kostopoulos, "The hypnospectrogram: An EEG power spectrum based means to concurrently overview the macroscopic and microscopic architecture of human sleep," *J. Neurosci. Methods*, vol. 185, no. 1, pp. 29–38, Dec. 2009.

[109] A. C. Rosa and L. J. Allen, "Fuzzy classification of microstructural dynamics of human sleep," in *Proc. IEEE Int. Conf. Syst., Man Cybern. Inf. Intell. Syst.*, Oct. 1996, pp. 1108–1113.

[110] C. M. Held, J. E. Heiss, P. A. Estévez, C. A. Perez, M. Garrido, C. Algarín, and P. Peirano, "Extracting fuzzy rules from polysomnographic recordings for infant sleep classification," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 10, pp. 1954–1962, Oct. 2006.

[111] J. E. Heiss, C. M. Held, P. A. Estevez, C. A. Perez, C. A. Holzmann, and J. P. Perez, "Classification of sleep stages in infants: A neuro fuzzy approach," *IEEE Eng. Med. Biol. Mag.*, vol. 21, no. 5, pp. 147–151, Sep. 2002.

[112] V. Pohl and E. Fahr, "Neuro-fuzzy recognition of K-complexes in sleep EEG signals," in *Proc. 17th Int. Conf. Eng. Med. Biol. Soc.*, Sep. 1995, pp. 789–790.

[113] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000. [Online]. Available: http://circ.ahajournals.org/content/101/23/e215

[114] S. P. Devuyst. (Feb. 2017). *The Dreams Subjects Database*. [Online]. Available: http://www.tcts.fpms.ac.be/~devuyst/Databases/DatabaseSubjects/

[115] S. Güneş, K. Polat, and Ş. Yosunkaya, "Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 7922–7928, 2010.

[116] X. Shao, B. Hu, and X. Zheng, "A study on automatic sleep stage classification based on clustering algorithm," in *Proc. Int. Conf. Brain Informat.* Cham, Switzerland: Springer, 2017, pp. 139–148.

[117] G. Garcia-Molina, K. Baehr, B. Steele, T. Tsoneva, S. Pfundtner, B. Riedner, D. P. White, and G. Tononi, "Automatic characterization of sleep need dissipation using a single hidden layer neural network," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2017, pp. 1305–1308.

[118] M. Peker, "A new approach for automatic sleep scoring: Combining Taguchi based complex-valued neural network and complex wavelet transform," *Comput. Methods Programs Biomed.*, vol. 129, pp. 203–216, Jun. 2016.

[119] S. Xie, Y. Li, X. Xie, W. Wang, and X. Duan, "The analysis and classify of sleep stage using deep learning network from single-channel EEG signal," in *Proc. Int. Conf. Neural Inf. Process.* Cham, Switzerland: Springer, 2017, pp. 752–758.

[120] H. Dong, A. Supratak, W. Pan, C. Wu, P. M. Matthews, and Y. Guo, "Mixed neural network approach for temporal sleep stage classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 324–333, Feb. 2018.

[121] M. Prucnal and A. G. Polak, "Effect of feature extraction on automatic sleep stage classification by artificial neural network," *Metrol. Meas. Syst.*, vol. 24, no. 2, pp. 229–240, Jun. 2017.

[122] A. Supratak, H. Dong, C. Wu, and Y. Guo, "DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 11, pp. 1998–2008, Nov. 2017.

[123] C. O'Reilly, N. Gosselin, J. Carrier, and T. Nielsen, "Montreal archive of sleep studies: An open-access resource for instrument benchmarking and exploratory research," *J. Sleep Res.*, vol. 23, no. 6, pp. 628–635, 2014, doi: 10.1111/jsr.12169.

[124] R. Sharma, R. B. Pachori, and A. Upadhyay, "Automatic sleep stages classification based on iterative filtering of electroencephalogram signals," *Neural Comput. Appl.*, vol. 28, no. 10, pp. 2959–2978, Oct. 2017.

[125] M. K. Uçar, M. R. Bozkurt, C. Bilgin, and K. Polat, "Automatic sleep staging in obstructive sleep apnea patients using photoplethysmography, heart rate variability signal and machine learning techniques," *Neural Comput. Appl.*, vol. 29, no. 8, pp. 1–16, Apr. 2018.

[126] E. Malaekah, H. Abdullah, and D. Cvetkovic, "Automatic sleep stage detection based on electrooculography," in *Proc. Int. Conf. Innov. Biomed. Eng. Life Sci.* Singapore: Springer, 2015, pp. 193–197.

[127] T. Lajnef, S. Chaibi, P. Ruby, P.-E. Aguera, J.-B. Eichenlaub, M. Samet, A. Kachouri, and K. Jerbi, "Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines," *J. Neurosci. Methods*, vol. 250, pp. 94–105, Nov. 2015.

[128] V. Bajaj and R. B. Pachori, "Automatic classification of sleep stages based on the time-frequency image of EEG signals," *Comput. Methods Programs Biomed.*, vol. 112, no. 3, pp. 320–328, 2013.

[129] G. Zhu, Y. Li, and P. Wen, "Analysis and classification of sleep stages based on difference visibility graphs from a single-channel EEG signal," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 6, pp. 1813–1821, Nov. 2014.

[130] T. Nakamura, T. Adjei, Y. Alqurashi, D. Looney, M. J. Morrell, and D. P. Mandic, "Complexity science for sleep stage classification from EEG," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 4387–4394.

[131] B. Koley and D. Dey, "An ensemble system for automatic sleep stage classification using single channel EEG signal," *Comput. Biol. Med.*, vol. 42, no. 12, pp. 1186–1195, 2012.

[132] C. Fu, P. Zhang, J. Jiang, K. Yang, and Z. Lv, "A Bayesian approach for sleep and wake classification based on dynamic time warping method," *Multimedia Tools Appl.*, vol. 76, no. 17, pp. 17765–17784, Sep. 2017.

[133] J. Sloboda and M. Das, "A simple sleep stage identification technique for incorporation in inexpensive electronic sleep screening devices," in *Proc. IEEE Nat. Aerosp. Electron. Conf. (NAECON)*, Jul. 2011, pp. 21–24.

[134] L. Samy, M.-C. Huang, J. J. Liu, W. Xu, and M. Sarrafzadeh, "Unobtrusive sleep stage identification using a pressure-sensitive bed sheet," *IEEE Sensors J.*, vol. 14, no. 7, pp. 2092–2101, Jul. 2014.

[135] J. B. Stephansen, "Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy," *Nature Commun.*, vol. 9, no. 1, pp. 1–15, 2018.

[136] O. Yildirim, U. Baloglu, and U. Acharya, "A deep learning model for automated sleep stages classification using PSG signals," *Int. J. Environ. Res. Public Health*, vol. 16, no. 4, p. 599, Feb. 2019.

[137] N. Surantha, T. F. Lesmana, and S. M. Isa, "Sleep stage classification using extreme learning machine and particle swarm optimization for healthcare big data," *J. Big Data*, vol. 8, no. 1, pp. 1–17, Dec. 2021.

[138] M. E. Tagluk, N. Sezgin, and M. Akin, "Estimation of sleep stages by an artificial neural network employing EEG, EMG and EOG," *J. Med. Syst.*, vol. 34, no. 4, pp. 717–725, 2010.

[139] Y.-L. Hsu, Y.-T. Yang, J.-S. Wang, and C.-Y. Hsu, "Automatic sleep stage recurrent neural classifier using energy features of EEG signals," *Neurocomputing*, vol. 104, pp. 105–114, Mar. 2013.

[140] S. Gudmundsson, T. P. Runarsson, and S. Sigurdsson, "Automatic sleep staging using support vector machines with posterior probability estimates," in *Proc. Int. Conf. Comput. Intell. Modelling, Control Automat. Int. Conf. Intell. Agents, Web Technol. Internet Commerce (CIMCA-IAWTIC)*, vol. 2, Nov. 2005, pp. 366–372.

[141] Y.-H. Lee, Y.-S. Chen, and L.-F. Chen, "Automated sleep staging using single EEG channel for REM sleep deprivation," in *Proc. 9th IEEE Int. Conf. Bioinf. BioEng.*, Jun. 2009, pp. 439–442.

[142] U. R. Acharya, S. V. Sree, A. P. C. Alvin, and J. S. Suri, "Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9072–9078, Aug. 2012.

[143] M. Grözinger, C. Wolf, T. Uhl, C. Schäffner, and J. Röschke, "Online detection of rem sleep based on the comprehensive evaluation of short adjacent EEG segments by artificial neural networks," *Prog. Neuro-Psychopharmacol. Biol. Psychiatry*, vol. 21, no. 6, pp. 951–963, Aug. 1997.

[144] E. Huupponen, S.-L. Himanen, A. Värri, J. Hasan, A. Saastamoinen, M. Lehtokangas, and J. Saarinen, "Fuzzy detection of EEG alpha without amplitude thresholding," *Artif. Intell. Med.*, vol. 24, no. 2, pp. 133–147, Feb. 2002.

[145] N. Acır and C. S. Güzeliş, "Automatic recognition of sleep spindles in EEG by using artificial neural networks," *Expert Syst. Appl.*, vol. 27, no. 3, pp. 451–458, 2004.

[146] E. Estrada, P. Nava, H. Nazeran, K. Behbehani, J. Burk, and E. Lucas, "Itakura distance: A useful similarity measure between EEG and EOG signals in computer-aided classification of sleep stages," in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Jan. 2005, p. 7.

[147] S. P. Cho, J. Lee, H. D. Park, and K. J. Lee, "Detection of arousals in patients with respiratory sleep disorders using a single channel EEG," in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Jan. 2005, pp. 2733–2735.

[148] F. Ebrahimi, M. Mikaeili, E. Estrada, and H. Nazeran, "Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 1151–1154.

[149] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, and H. Dickhaus, "Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier," *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 10–19, 2012.

[150] L. J. Herrera, C. M. Fernandes, A. M. Mora, D. Migotina, R. Largo, A. Guillen, and A. C. Rosa, "Combination of heterogeneous EEG feature extraction methods and stacked sequential learning for sleep stage classification," *Int. J. Neural Syst.*, vol. 23, no. 03, Jun. 2013, Art. no. 1350012.

[151] M. Diykh and Y. Li, "Complex networks approach for EEG signal sleep stages classification," *Expert Syst. Appl.*, vol. 63, pp. 241–248, Nov. 2016.

[152] T. L. T. da Silveira, A. J. Kozakevicius, and C. R. Rodrigues, "Single-channel EEG sleep stage classification based on a streamlined set of statistical features in wavelet domain," *Med. Biol. Eng. Comput.*, vol. 55, no. 2, pp. 343–352, Feb. 2017.

[153] S. K. Pandey and R. R. Janghel, "Recent deep learning techniques, challenges and its applications for medical healthcare system: A review," *Neural Process. Lett.*, vol. 50, pp. 1907–1935, Oct. 2019.

[154] L. Fiorillo, A. Puiatti, M. Papandrea, P.-L. Ratti, P. Favaro, C. Roth, P. Bargiotas, C. L. Bassetti, and F. D. Faraci, "Automated sleep scoring: A review of the latest approaches," *Sleep Med. Rev.*, vol. 48, Dec. 2019, Art. no. 101204. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1087079218301746

[155] M. T. Bianchi, K. Russo, H. Gabbidon, T. Smith, B. Goparaju, and M. B. Westover, "Big data in sleep medicine: Prospects and pitfalls in phenotyping," *Nature Sci. Sleep*, vol. 9, p. 11, Jan. 2017.

[156] N. Bragazzi, O. Guglielmi, and A. Garbarino, "SleepOMICS: How big data can revolutionize sleep science," *Int. J. Environ. Res. Public Health*, vol. 16, no. 2, p. 291, 2019.

[157] H. Danker-Hopfe, D. Kunz, G. Gruber, G. Klösch, J. L. Lorenzo, S. L. Himanen, B. Kemp, T. Penzel, J. Röschke, H. Dorn, A. Schlögl, E. Trenker, and G. Dorffner, "Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders," *J. Sleep Res.*, vol. 13, no. 1, pp. 63–69, Mar. 2004, doi: 10.1046/j.1365-2869.2003.00375.x.

[158] T. Penzel, X. Zhang, and I. Fietze, "Inter-scorer reliability between sleep centers can teach us what to improve in the scoring rules," *J. Clin. Sleep Med.*, vol. 9, no. 1, pp. 89–91, Jan. 2013.

[159] A. Malhotra, M. Younes, S. T. Kuna, R. Benca, C. A. Kushida, J. Walsh, A. Hanlon, B. Staley, A. I. Pack, and G. W. Pien, "Performance of an automated polysomnography scoring system versus computer-assisted manual scoring," *Sleep*, vol. 36, no. 4, pp. 573–582, Apr. 2013.

[160] K. A. I. Aboalayon, M. Faezipour, W. S. Almuhammadi, and S. Moslehpour, "Sleep stage classification using EEG signal analysis: A comprehensive survey and new investigation," *Entropy*, vol. 18, no. 9, p. 272, 2016.

[161] N. A. Collop, "Scoring variability between polysomnography technologists in different sleep laboratories," *Sleep Med.*, vol. 3, no. 1, pp. 43–47, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389945701001150

[162] M. Younes, "The case for using digital EEG analysis in clinical sleep medicine," *Sleep Sci. Pract.*, vol. 1, no. 1, pp. 1–15, Feb. 2017, doi: 10.1186/s41606-016-0005-0.

HADEEL ALSOLAI received the Ph.D. degree in computer and information sciences from the University of Strathclyde, Glasgow, U.K., in March 2021. She works as an Assistant Professor with the Department of Information System, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University. Also, she has been appointed as the Manager of the Office Business and Project Management at the College of Computer and Information Sciences. She teaches several courses with the Information System Department, such as information system, system security, and database system. Her research interests include machine learning, data mining, data science, software quality, and open-source systems.

SHAHNAWAZ QURESHI received the M.S. degree in computer science from the Blekinge Institute of Technology, Sweden, and the Ph.D. degree in computer science in the research area of machine learning and deep learning. He is currently an Associate Professor with the FAST National University of Computer and Emerging Sciences (NUCES). He won the Thailand Education Hub for Southern Region of ASEAN countries scholarship, Prince of Songkla University for Ph.D. position, from 2016 to 2019. His master's research project was relevant to European project known as PsyIntEC—Psychophysiological interaction and empathetic cognition for human-robot cooperative work. He has five years of successful professional teaching and research experience in national and international level institutions around the world. He has been playing a role as a member of the International Association of Engineers, including Society of Artificial Intelligence, Society of Data Mining, Society of Scientific Computing, and Society of Bioinformatics, since 2015.

SYED MUHAMMAD ZEESHAN IQBAL received the M.S. degree in computer science from the Blekinge Institute of Technology, Karlskrona, Sweden, in 2010.

From 2010 to 2014, he was a Project Assistant with the Blekinge Institute of Technology. He is currently a Research and Development Engineer at BrightWare LLC, Riyadh, Saudi Arabia. His research interests include the development of parallel processing systems, intelligent decision support systems, machine learning, artificial intelligence, and optimization techniques. He has ten year's of successful professional software development, teaching, and research experience in national and international level institutions around the world.

SIRIRUT VANICHAYOBON received the Ph.D. degree in computer science from the University of Oklahoma, USA. She is currently an Assistant Professor with the Department of Computer Science, Faculty of Science, Prince of Songkla University, Hat Yai, Thailand. Her research interests include database, data warehouse, data mining, data science, parallel computing, and GIS. She was the Chairman of the Academic and Examination Department at the 13th National Computer Olympiad, the Secretary of the Academic and Examination Committee at the 12th National Computer Olympiad, and the Vice President for Academic Affairs and Examination Department at the 11th National Computer Olympiad.

LAWRENCE EDWARD HENESEY has been an Assistant Professor in the area of applied artificial intelligence, since 2000, working in Sweden (http://www.henesey.eu). He posseses more than 30 year's of "hands-on experience" in commercial and operational aspects in international maritime transportation and logistics industries. Expertise in shipping, container terminal operations, sales (project and product), business development, analysis, simulation, and development of expert systems and knowledge of IT services and technologies. He has developed several innovations implemented globally for top terminal operators. Concurrently with industrial experience and work, his research area is on the application of techniques from Distributed Artificial Intelligence in Container Terminals, Logistics and Transport, which has culminated into more than 100 published articles and two books.

CRAIG LINDLEY works with the Computational Modelling Team, CSIRO's Data61 Business Unit. He has been recently involved in the design and development of technology platforms for volumetric data management and process tracking, providing 3D visualization, simulation, analytics, situation monitoring, diagnostics, alerting, optimized rescheduling (applied in mining operations, mineral processing, unmanned aerial vehicle (UAV) operations, and aerospace manufacturing). His roles at CSIRO have included the Manager of the $50 million Data61 ICT Laboratory in Hobart, from 2012 to 2016, the Research Director of Intelligent Sensing and Systems, the Group Leader of Cognitive Systems, the Team Leader of Cognitive Informatics, and the Project Manager for numerous research and development projects. Prior to this, he has been a Professor and the Research Leader at universities in Sweden, working with cognitive systems, AI, interactive 3D visualization, simulation, human-machine interaction, and robotics. Before this, he initially worked in military software development as an RAAF Air Defence Officer and a Software Developer for radar tracking, command, and control software systems development. He subsequently worked on the Australian Submarine Project, and has held senior research positions in Australia (CSIRO), Belgium (STARLAB NV/SA), and Sweden (The Interactive Institute, Uppsala and BTH). He has negotiated and led numerous collaborative research projects with industrial partners in aerospace, defence, ICT, mining, and government. He has developed several rapid prototyping, Robotics and Psychophysiology Laboratories, largely-based upon additive and subtractive prototyping technologies and open source electronics systems. He has created and taught numerous undergraduate and master's degree courses in 3D systems development, interaction design, and rapid prototyping of novel devices and robotic systems.

SEPPO KARRILA received the Ph.D. degree in engineering chemistry from the University of Wisconsin–Madison, USA. He is currently an Associate Professor with the Prince of Songkla University, Thailand. His research interests include computational fluid mechanics, engineering for papermaking, machine learning, bioinformatics, optimization, mathematical modeling, and neural networks. He has a widely diverse research background in goal oriented engineering; as an Academic Faculty Member; and in a large pharma MNC performing hands-on biomarker discovery.