REVIEW ARTICLE

Current status and prospects of automatic sleep stages scoring: Review

Maksym Gaiduk[1] · Ángel Serrano Alarcón[2] · Ralf Seepold[1] · Natividad Martínez Madrid[2]

Abstract

The scoring of sleep stages is one of the essential tasks in sleep analysis. Since a manual procedure requires considerable human and financial resources, and incorporates some subjectivity, an automated approach could result in several advantages. There have been many developments in this area, and in order to provide a comprehensive overview, it is essential to review relevant recent works and summarise the characteristics of the approaches, which is the main aim of this article. To achieve it, we examined articles published between 2018 and 2022 that dealt with the automated scoring of sleep stages. In the final selection for in-depth analysis, 125 articles were included after reviewing a total of 515 publications. The results revealed that automatic scoring demonstrates good quality (with Cohen's kappa up to over 0.80 and accuracy up to over 90%) in analysing EEG/EEG + EOG + EMG signals. At the same time, it should be noted that there has been no breakthrough in the quality of results using these signals in recent years. Systems involving other signals that could potentially be acquired more conveniently for the user (e.g. respiratory, cardiac or movement signals) remain more challenging in the implementation with a high level of reliability but have considerable innovation capability. In general, automatic sleep stage scoring has excellent potential to assist medical professionals while providing an objective assessment.

Keywords Automatic sleep assessment · Sleep scoring · Physiological signals · Sleep stages

1 Introduction

Sleep plays a vital role in human life. Its influence on physical and psychological health and human well-being is enormous, which has already been demonstrated and underlined in numerous publications [1–4]. Accordingly, it is crucial to be able to analyse sleep to have the possibility to intervene in time in case of disorders, to eliminate them if possible, or at least to reduce their negative influence on health.

As early as the 1950s, William Dement observed that the signal of the electroencephalogram (EEG), the frequency of eye movements and the frequency of body movements were subject to regular cyclical fluctuations during the night [5].

Further research into this phenomenon led over time to the recognition of typical sleep patterns and ultimately to the creation of the concept of sleep stages.

In 1968, the first standardised categorisation of sleep into sleep stages was conducted using the method of Rechtschaffen and Kales (R&K) [6]. It divided sleep into 30-s intervals—"epochs"—and assigned one of the following sleep stages to each epoch:

- Stage W—Wakefulness/Wakefulness
- Stage 1 (S1)
- Stage 2 (S2)
- Stage 3 (S3)
- Stage 4 (S4)
- Stage REM—Rapid eye movement

Furthermore, some epochs could be labelled "movement time" if movement prevents accurate identification of sleep stages.

Over time, as new evidence about sleep emerged, there was a need to establish a new guide to terminology, recording methodology and scoring rules for sleep. This was done

✉ Maksym Gaiduk
  maksym.gaiduk@htwg-konstanz.de

✉ Ángel Serrano Alarcón
  angel.serrano_alarcon@reutlingen-university.de

1   HTWG Konstanz – University of Applied Sciences,
    Alfred-Wachtel-Str.8, 78462 Konstanz, Germany

2   Reutlingen University, Alteburgstraße 150,
    72762 Reutlingen, Germany

in 2007 with the release of a new guideline by the American Academy of Sleep Medicine (AASM) [7], which has been continuously updated since its release. The most recent version is 2.6, published in 2020 [8]. It proposed, among other things, a new classification of sleep stages:

- Stage W—wakefulness
- Stage N1/NREM1 (formerly S1)
- Stage N2/NREM2 (formerly S2)
- Stage N3/NREM3 (formerly S3 + S4)
- Stage R—Rapid eye movement (REM)

In fact, inter-standard comparison of sleep stage scoring has been the focus of several scientific publications, among them [9, 10].

In sleep analysis practice, sleep stages N1-N3 are, in some cases, combined into one NREM stage. Another subdivision found in scientific papers is Wake/Light Sleep (N1 + N2)/ Deep Sleep (N3)/REM sleep (R). Table 1 summarises these sleep stage classification manners.

The most widely utilised method for assessing sleep behaviour, which has been in use for many years, is polysomnography (PSG) [11]. This approach involves measuring several signals, which are then evaluated:

- Electroencephalography (EEG) records brain activity.
- Electrocardiography (ECG) is a method that reflects the electrical activity of the heart over time.
- Electromyography (EMG) is used to record muscle activity.
- Electrooculography (EOG), on the other hand, uses electrodes to detect and measure the potential between the human eye's back and front to record eye movements.
- The oxygen saturation of the blood is measured with pulse oximetry.
- Often, both respiratory flow and respiratory effort are measured.
- Moreover, other signals can be recorded, such as the person's position or a video recording for a detailed analysis.

The recorded signals are then stored and finally analysed manually by trained sleep experts, scoring the stages of sleep. Mainly EEG, EOG, and EMG signals provide the required input.

This manual evaluation of the recordings enables accurate scoring of the sleep stages. However, it also entails a high expenditure of time and financial resources to complete this task [12]. It should also be noted that despite following the established scoring guidelines, manual analysis introduces a certain level of subjectivity, also known as interscorer/ interrater reliability/variability, which has been described in several scientific publications [13–15]. Thus, the results of the evaluation of the same sleep recording by different experts may vary. As the meta-analysis of interrater reliability conducted in [16] has demonstrated, Cohen's kappa for manual, overall sleep scoring reaches the value of 0.76.

To summarise the above situation, automatic sleep scoring could result in several benefits. Among other things, it could reduce the financial and human resources needed, thereby supporting the work of sleep physicians and making them available to provide treatment to a larger number of individuals requiring it. A high number of scientific works in the field of automatic sleep stage detection indicate that new developments are constantly being made in this area, and only a thorough analysis can provide a comprehensive overview of the status quo.

The conducted analysis of current research has indicated that the subject of a state-of-the-art review in automatic sleep stage identification has been addressed several times, highlighting the importance of this topic. However, studies in recent years have focused on specific aspects of the problem. For example, in [17], only the algorithms that work with the EEG signal were selected. In [18], consumer sleep technologies (CSTs) were investigated to analyse sleep in combination with artificial intelligence. [19] examined a more extended period of 19 years but with a relatively small number of selected articles (55 in total) for this larger time frame. EEG signals in combination with deep learning algorithms applied directly to raw signals or spectrograms, were the subject of investigation in [20], resulting in a more refined but smaller sample of 14 studies. Deep learning techniques for sleep phase detection were investigated in [21], with 36 studies selected from the period between 2010 and 2020. The number of studies reviewed in [22] is also relatively small, below 30, which is in part caused by the research question in which EEG signal-based algorithms were investigated. Also, if one looks further into the past, it can be seen that the matter of automatic detection of sleep phases was addressed considerably earlier, and was, for example, already investigated in a review article in 2012 [23]. In general, the principles of automatic sleep scoring were even reported in 2000 [24], indicating the topic's importance and demonstrating a long history of research in this scientific area.

Table 1 Sleep stages classification

Number of sleep stages	Sleep stages				
5	W	N1	N2	N3	R
4	Wake	Light		Deep	R
3	Wake	NREM			R
2	Wake	Sleep			

Due to the major importance of the topic, knowing the enormous number of recent developments in the field of automatic sleep scoring and considering the gaps in existing review articles in recent years, the decision was made to conduct a detailed investigation of state-of-the-art and to report the results in a review article. The aim is to prepare a comprehensive overview of the state-of-the-art in the field of automatic sleep stage identification and provide researchers with a consolidated summary of the current developments, which should increase the efficiency of their scientific research.

Several approaches to literature reviews exist. One of the classifications proposes the following groups: systematic, semi-systematic (or narrative) and integrative, according to [25]. In our case, it seems appropriate to use the semi-systematic approach, as it facilitates more flexibility and is commonly used for overview publications. In addition to aiming to provide an overview of a topic, a semi-systematic review often examines how research in a particular area has developed over time [25]. This type of review can be a highly effective way to cover more areas and broader topics than a systematic review can address, especially when multiple subject domains (such as computer science and medicine) are in the field of interest [26].

Three main research questions are being addressed in our article:

RQ1. Which level of quality is achieved at the current state of development of automatic sleep scoring approaches?

RQ2. How has the focus on particular signal analysis methods evolved over time?

RQ3. How has the choice of signals used for analysis varied over time?

2 Materials and methods

2.1 Eligibility criteria

A list of inclusion and exclusion criteria was established based on the analysis of the current state of the art and considering the defined research questions.

Inclusion criteria

- Automatic approach for sleep scoring is described in a manuscript
- Peer-reviewed journals
- Publications considering adults (> 18) as a group of interest
- English language
- Manuscripts published between 2018 and 2022
- Ground truth labels came from AASM rules

Exclusion criteria

- Article size of fewer than four pages
- Test dataset is smaller than 20 overnight recordings
- Animals being a subject of investigation
- Only Wake/Sleep identification without more detailed scoring
- Review article/Editorial/Book chapters (however, these arts of publications are being checked for eligible articles)
- Epochs of longer than 30 s are being analysed
- Missing significant information on some of the points from the exclusion criteria list

When a research group has published several consecutive articles that represent a further development of particular methods, only the most recent article has been included in this review. If the reported approaches represent significantly different procedures or use diverse input signals, all of them are included in the selection.

A more detailed explanation of the inclusion criterion "Automatic approach for sleep scoring is described in a manuscript" appears to be reasonable. As one of the important aspects for the developed sleep stage identification system is its applicability in practice and usability, only those publications were included in the review that allow fully automatic scoring, i.e. the systems that require manual intervention in use, such as partial manual scoring, were not considered in this work.

In 2007, the AASM guidelines [7] replaced the Rechtschaffen & Kales rules [6] as the new standard for sleep assessment. Therefore, in the selection of articles, the publications were included in which the sleep phase detection of the reference measurement is done according to AASM guidelines, as only the manuscripts since 2018 were considered. However, in order to be able to evaluate the methods that were using the older databases already available, the decision was made also to consider the articles that were using the existing recordings according to R&K rules, but adapting them to the sleep stages defined in AASM.

2.2 Information collection strategy

After analysing the requirements for articles to be included in the review and considering the interdisciplinary nature of the topic, the following selection of databases was made to ensure the comprehensive collection of information:

- IEEE Xplore
- PubMed
- Scopus
- Web of Science

The following query (with necessary adaptions for specific databases) was determined to search for articles that match the defined criteria:

(((automatic OR automated OR automatically) AND ("sleep stage" OR "sleep stages" OR "sleep phase" OR "sleep phases")) AND (scoring OR classification OR determination OR identification OR recognition OR detection)). The search using this defined query took place on the articles' titles, abstracts and keywords in databases. In addition, a restriction of years (2018–2022) was used, either directly in the query, if the database provided this option, or through the later application of an appropriate filter.

2.3 Data refining procedure

The next step after applying the designed query in the selected databases was to refine the collected data. This involved several steps. First, duplicates were removed, and all non-journal articles were excluded from the list of articles. This was followed by three phases of initial screening: title, abstract and diagonal reading. During this process, the eligibility of the manuscripts was checked according to the inclusion/exclusion criteria, and if they did not meet them, they were eliminated. At last, the final screening was applied, where the full reading of the papers was performed. In this step, the last articles that did not meet the established criteria were removed to generate the final list of manuscripts for further analysis.

2.4 Synthesis methods

A set of Python packages (Pandas, NumPy, Seaborn, Matplotlib and SciPy) were used for the statistical analysis carried out in this scientific work. Statistical data visualisations were presented to aid the understanding of the results. Where possible, statistics were proposed to explain relevant information gathered during the review of the articles.

3 Results

The conducted in-depth analysis of the current research outcomes over the last five years has led us to create a comprehensive overview, which is presented in the following. It is important to point out that not all descriptions of the implementation and evaluation approaches provide sufficient information to estimate the results' quality and correctness. Nevertheless, this review has included these articles according to inclusion/exclusion criteria.

3.1 Study selection

The article selection process carried out throughout this review of scientific papers dealing with the classification of sleep stages is shown in Fig. 1. The number of publications included and excluded during each review stage is also shown.

3.2 Which level of quality is achieved at the current state of development of automatic sleep scoring approaches?

A summary table was prepared to provide a comprehensive overview of the available approaches for automatic sleep stage scoring, including all articles selected according to inclusion criteria and after refining the set according to exclusion criteria. The results can be seen in Table 2, attached to this manuscript. Due to its comprehensive content, inserting it directly into the text would significantly decrease the manuscript's readability, which should be avoided.

The best accuracies reported in the reviewed manuscripts achieve over 90% when analysing EEG signal, being an excellent result. As reported in several research papers, the best conducted Cohen's kappa values are over 0.80, and F1 values achieve up to 85–90%.

The differences in performance of some algorithms can vary significantly depending on the composition of the test group/dataset. In particular, if the training was done on healthy subjects only, testing on a group with sleep disorders (e.g., sleep apnoea) would be associated with a high probability of a reduction in sleep stage detection accuracy, as indicated in, e.g. [27]. Therefore, the provided table includes information on a targeted population of the developed algorithms.

In the column "Algorithm/Method" in the table, only the primary used approach for sleep scoring is provided. Feature extraction/filtering procedures, being another relevant characteristic of the method, are not described due to the large variability and individuality of utilised approaches which complicates their classification and would excessively enlarge the table, decreasing its clarity.

3.3 How has the focus on particular signal analysis methods evolved over time?

Trends in research are changing over time, among other things, as new knowledge is gained and new priorities may be set. This raises the question of how the choice of methods for sleep stage estimation has evolved over the last few years. Figure 2 depicts the list of publications that met the inclusion/exclusion criteria and were selected for statistical analysis separately for the years 2018–2022, as well as

Fig. 1 Flow chart for the selection of the entire set of the included publications

Identification of studies via databases

Identification

Records identified from:

Databases (n = 1067)

PubMed (n = 235)
Scopus (n = 430)
IEEE (n = 62)
Web of Science (n = 340)

Records removed *before screening*:

Duplicate records
&
Non-journal sources
(n = 552)

Screening

Records screened
(n = 515)

Records excluded (n = 376)
Title reading (n = 190)
Abstract reading (n = 96)
Diagonal reading (n = 90)

Reports assessed for eligibility
(n = 139)

Reports excluded (n = 14)
Lack of information for reproducibility
Did not meet inclusion / exclusion criteria

Included

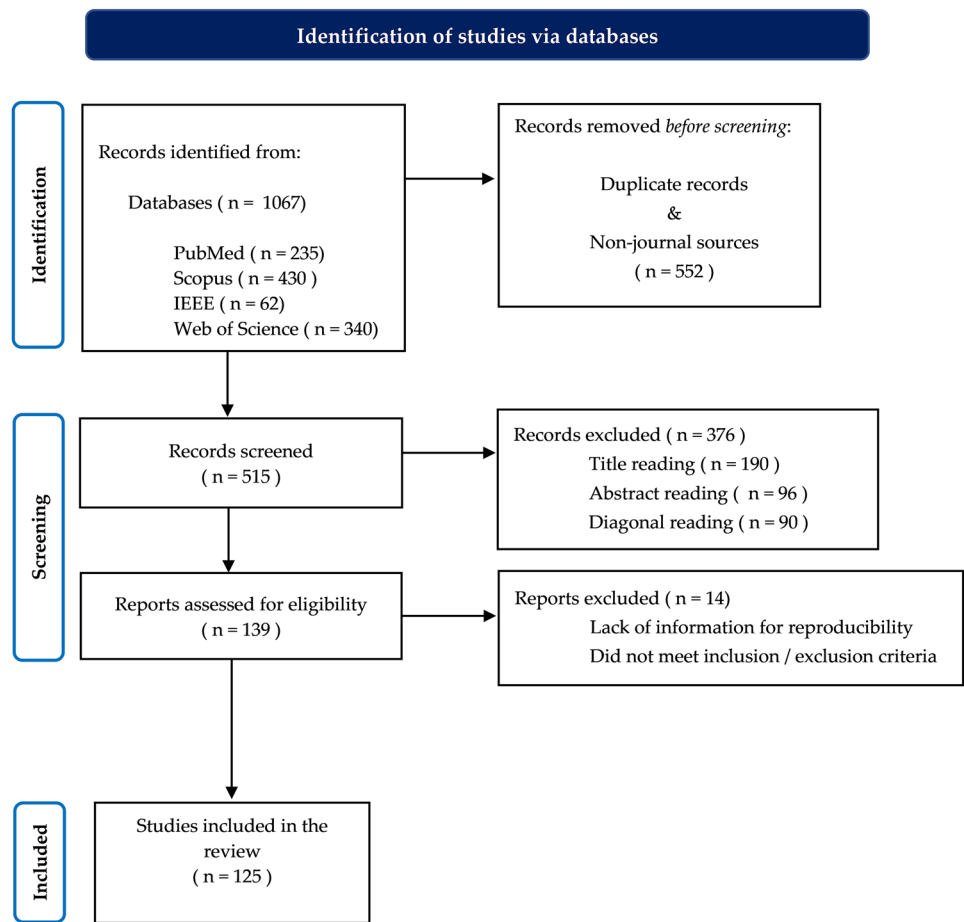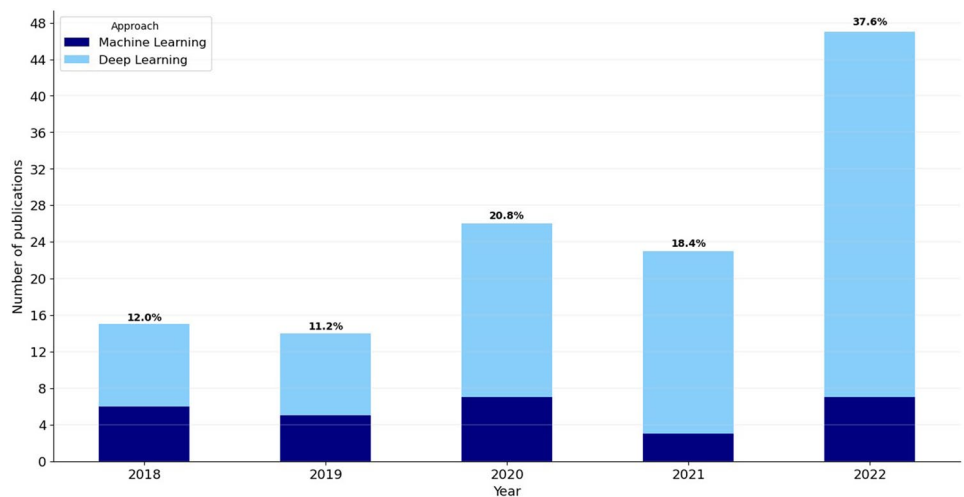Studies included in the
review
(n = 125)

Fig. 2 Bar plots with the number of publications that met the inclusion/exclusion criteria. The terms "Machine Learning" (with Deep Learning excluded) and "Deep Learning" refer to the approach followed for sleep stage classification

the breakdown by primary approach for every single year. The highest number of selected articles was published in 2022 (37.6% of all publications), and the lowest number of selected papers was in 2018 and 2019. It can also be seen in Fig. 2 that the use of deep learning techniques is much higher than other machine learning approaches in any year from 2018 to 2022.

3.4 How has the choice of signals used for analysis varied over time?

The selection of signals for analysis with the subsequent scoring of sleep stages and their variation over time also provides interest for the research. Therefore, we have graphically presented the selection of signals by year in Fig. 3 to facilitate an overview and to illustrate possible trends.

With the aim of providing more insight into the selection of signals used to score sleep stages in the articles included in the review, the detailed representation per input signal was generated and can be seen in Fig. 4.

As shown in Fig. 4, the EEG signal is the most widely used as opposed to other signal sets, with 69 scientific publications using standalone EEG and being a part of numerous other combinations with other signals.

4 Discussion

The findings of the conducted review represent, to the best of our knowledge, the most comprehensive research overview in the field of automatic sleep stage scoring in recent years. This allows a thorough analysis of current developments in this domain and can serve as a basis for further research.

Most of the information for the analysis can be found in Table 2. For example, it can be seen that the majority of publications have considered the detection of five sleep stages, which corresponds to the AASM standard. Only occasionally did the methods target fewer sleep stages.

The most common method of validation, as indicated in Table 2, is cross-validation. More specifically, tenfold and 20-fold cross-validation are the most commonly used techniques in the reviewed articles. In addition, other x-fold cross-validation and leave-one-subject-out approaches

Fig. 3 Bar plots with the number of publications per year depending on the signal used. Four main groups of signals analysed for sleep stage classification are represented: "EEG", "EEG + EOG", "EEG + EOG + EMG", and "Others"
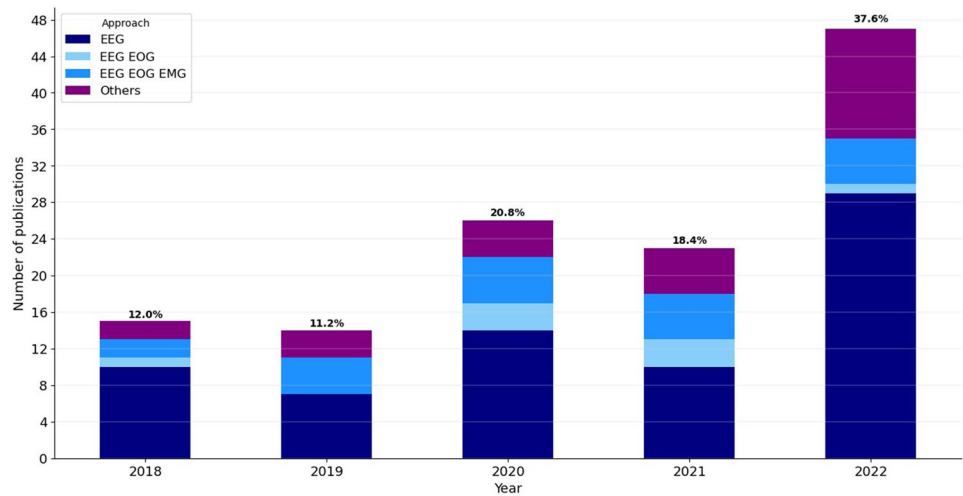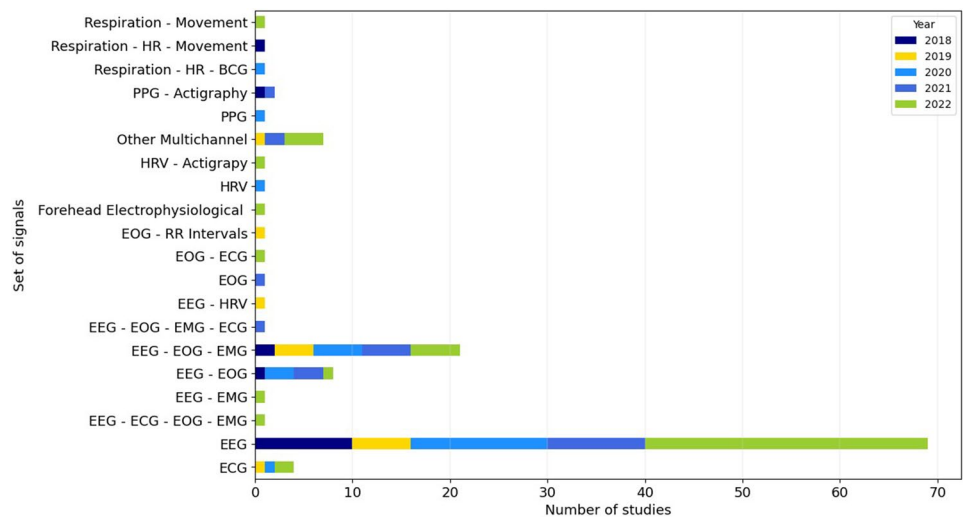

Fig. 4 Set of signals used by the reviewed scientific works for classifying sleep stages from 2018 to 2022

can be found in publications. Direct strict separation of training and validation/test datasets is also reported in several articles. It is important to note that using different epochs of the same recording for both training and evaluation would affect the algorithm's performance in terms of increasing accuracy. However, using epochs from the same recording for both training and evaluation does not allow us to assess whether the algorithm used would perform similarly if the subjects for the training and test datasets were strictly separated. Therefore, we have tried to extract this relevant information from the reviewed articles. Unfortunately, far from all articles provided this important detail. For those articles where this information was provided, or where the method of validation (e.g. leave-one-subject-out) allowed us to obtain the required data, we can say that in the majority of cases there was a strict separation of subjects into training and test datasets, and therefore the epochs of the same recording were mostly not used for both training and evaluation.

We can highlight some important points by looking at the quality parameters reported in the research papers analysed. By far, the most widely used signal for sleep stage scoring is the EEG, and the combination of EEG, EOG and EMG is the second most frequently used, as can be recognised from Table 2 and Fig. 3. Together, these two sets account for more than two-thirds of all publications. They also yield the best results in scoring—Cohen's kappa up to over 0.80 (substantial to almost perfect according to [28]) and accuracy up to over 90%, which is a very good performance, especially considering that even when evaluating the same recording by different scorers, a kappa of 0.71–0.81 is obtained, as studied in [16]. In general, the use of automated sleep phase detection methods could address the problem of interrater reliability when the same approach is used for analysis in different sleep laboratories (or even in one sleep laboratory instead of multiple experts). This could help to free up clinicians' time for other clinical tasks.

At the same time, it should be noted that there has been no breakthrough in the quality of the results when analysing these signals (EEG, EOG, EMG) over the last five years in terms of performance. This can be explained, among other things, by the high scores already achieved, as mentioned above. If one is looking for research topics with a high innovation potential, other alternatives should perhaps be considered. An example of this could be systems that work with other signals. These could be, for example, cardiac, respiratory and movement signals, as they have the potential to be recorded with more comfort for the user and possibly without contact [29, 30]. The number of works using these alternative signals for sleep scoring is significantly lower, according to the research conducted, and there is noticeable room for improvement. Nevertheless, looking at Fig. 3, we can see that the number of publications using other than

classical signal sets as input to the algorithm has increased over the years.

Another important issue in the development of sleep scoring systems is usability and suitability for practical use. Unfortunately, these aspects are not always explicitly considered in publications, although they are of great importance for the transfer of research into practical implementation. For example, it may be advantageous in terms of usability if the algorithms work fully automatically, i.e. do not require manual pre-classification or pre-processing. Another point that could improve the usability of the systems in a practical application is the determination of the measurement uncertainty during sleep stage detection. This would make it possible to see where the results of the automatic evaluation might need to be re-examined because the classification is not unambiguous, which has been recognised by the software, also known as confidence estimation or prediction certainities. Some work has already been done towards this functionality and usability improvement, e.g. [31–34]. In summary, further research aimed at improving the usability of systems in real-world environments can only be welcomed, as the potential is not yet exhausted, as this review has shown.

The explainability/interpretability of algorithms and their results is also a topic with great potential for further research, as it could at least partially help to solve the "black box" problem, especially in deep learning applications. This issue has already been addressed in some of the reviewed articles [32, 35–38], but there is still much room for further investigation to overcome the current challenges. In general, due to the fact that there is an impressive number of articles on sleep scoring, but at the same time a significantly smaller number of practical applications, the question of usability and practicality seems to be one of the relevant ones and should be further investigated.

Another area of research we observed in some of the manuscripts was the use of algorithms that adapt to the signal or recalibrate the features used [39], or incorporate the probabilities of transitions between sleep stages [40–42]. The inclusion of these additional steps in the algorithms to analyse the signal in an individualised way, while taking into account the whole sleep structure, may also lead to more accurate and higher quality sleep stage detection, and we encourage further research in this direction.

It is pretty common that more than one dataset is used for evaluation in the published articles, as can be seen in Table 2. In this case, however, there are often only two datasets used, which are additionally not always large. Though, the essential point to consider is that the algorithms developed should be evaluated on several, preferable large datasets to ensure a high-quality general scoring that is not only dataset specific, as there may be peculiarities in the signals (due to different equipment, pre-processing, etc.) in other series of recordings.

This has been addressed in a number of the papers included in the selection but not in all cases. In order to provide a reliable solution that can be applied to a wide range of situations, it is crucial to consider the universality mentioned above and evaluate with multiple data sets from diverse sleep laboratories and recorded with different devices.

An analysis of Fig. 2 reveals that there is a tendency for the number of developments in the area of sleep stage classification to increase steadily, although there are minor fluctuations in this trend from year to year. Proportionally, deep learning methods stand out from other machine learning techniques in this topic area from year to year, taking an increasingly larger share. This may be partly caused by a general increase in the popularity of deep learning methods in research and partly by progress in developing new, more effective techniques that may produce higher-quality results.

While Fig. 3 gives a rather general picture regarding the signals used in literature, Fig. 4 provides more detail. Here it can be seen that, in addition to the two most frequently used signal sets already mentioned (EEG and EEG + EOG + EMG), a third combination occurs significantly more frequently than other signals—EEG + EOG, although significantly less frequently than the first two. In general, it can be said that a combination with the EEG signal occurs more often than possible combinations with other signals in the research. This can be explained by the fact that the EEG traditionally plays a key role in the detection of sleep stages and also provides crucial information. Apart from these three signals, only the ECG is used significantly more often than the remaining signals. Besides that, other signals were only sporadically included in the approaches reported in the reviewed articles. Another interesting finding from the information presented in Table 2 and Fig. 4 is that the majority of the required signals are recorded with electrodes (EEG, EOG, EMG, ECG), which requires direct contact with the body and is not necessarily the most comfortable for users. This shows that methods that allow non-obtrusive measurement of physiological signals are not yet widely used in sleep stage assessment.

Another important point is that while automated analysis can free up physicians' time for other clinical tasks, even with fully automated sleep evaluation there is still a need for the clinician to review the signals to avoid missing potentially important points. It is known that some specific pathologies can only be detected by looking for specific patterns in the signals that may currently be missed by automated scoring systems.

5 Conclusions

Analysing the review findings, in the last years, no breakthrough progress in the quality of sleep scoring approaches analysing EEG, EOG and EMG signals could be observed

in terms of perfomance, although the most advanced of the current studies demonstrate a good level of quality in the detection of sleep stages. Nonetheless, a number of topics still remain with some gaps and have excellent innovation potential. These include, among others, the application of algorithms to categorise sleep stages that utilise less traditional signals, such as, for instance, breathing, body movement or heart signals. These signals could possibly be recorded in a contactless and more accessible way than the classical PSG approach and with more comfort for users/patients. Therefore, if an acceptable level of measurement quality could be achieved, it could simplify the measurement process and promote the widespread use of sleep scoring systems. This, in turn, would favour the detection of sleep disorders at an early stage.

Another critical issue that still needs to be further addressed is the matter of interoperability of the developed algorithms with different datasets in order to have the capability to apply the designed sleep stage identification systems universally and not only to one specific dataset. This point was addressed in several of reviewed articles (e.g. [43, 44]), where a significant number of datasets was used for the evaluation. Nevertheless, this topic requires further investigation and assessment.

The issues of explainability, usability, practicality and the use of adaptive algorithms are other areas of research that have been addressed in recent years, but are still not sufficiently explored and have significant potential.

In general, automatic sleep scoring has the potential to create an objective approach devoid of some level of subjectivity and, consequently, variance in scoring, present by manual scorers, considered in the literature as interrater reliability. This point was addressed in [45]. The analysis of current research in the area of sleep scoring presented in our article leads us to the conclusion that automated scoring of sleep stages could become a powerful tool supporting physicians in their work and helping to reduce sleep scoring ambiguity by decreasing the level of subjectivity in the analysis process. It is also noteworthy that the use of automated sleep scoring systems could lead to a saving of resources (both human and financial) that could be allocated to provide a more comprehensive medical service to the general public by medical professionals or to treat a larger number of patients with increased time capacity.

Appendix 1

See Table 2.

Table 2 Articles included in the review according to inclusion/exclusion criteria with detailed description

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Abdollahpour et al. [46]	2020	Multichannel	Transfer Learning Convolutional Neural Network Data Fusion (TLCNN-DF)	Sleep-EDF Expanded Sleep-EDF	8 subjects 20 subjects	18 Healthy subjects 10 effects of temazepam-insomnio	fivefold CV	5 stages	93.58%		0.89
Abdollahpour et al. [47]	2022	EEG (1 channel)	2×CNN	Sleep EDF Sleep EDFx DREAMS	4 subjects 20 subjects 20 subjects	4 Healthy subjects 20 effects of temazepam-insomnio 20 Healthy subjects	80% Training 20% Test	5 stages	93.48% 93.14% 83.55%		0.89 0.89 0.77
Alickovicet al. [48]	2018	EEG	Denoising with Multiscale principal component analysis (MSPCA) Feature extraction by DWT, Rotational support vector machine (RotSVM)	Sleep-EDF	20 subjects (30 PSG) 10 subject × 2 PSG	20 Healthy subjects 10 effects of temazepam-insomnio	Grand-Subject.Specific Approach (GSSA) 10-Fold CV	5 stages	91.10%		0.8
Alvarez-Estevez Diego et al. [49]	2021	EEG (2 channels) EOG EMG	CNN LSTM	HMC Dublin SHHS Telemetry DREAMS ISRUC	154 subjects 25 PSG 100 PSG 44 PSG 20 PSG 100 PSG	Mixed (healthy and SDB) subjects	80% Training 20% Test	5 stages			0.79 0.70 0.82 0.80 0.83 0.78
An et al. [50]	2022	EEGs	Amplitude-Time Dual-View signal Representation 1D-CNN Attention-based Feature Fusion	Sleep-EDF Expanded Sleep-EDF	8 subjects 20 subjects	20 Healthy subjects 10 effects of temazepam-insomnio	tenfold CV	6 stages	89.94% 92.07%	84.61% 83.44%	
Arslan et al. [51]	2022	Multichannel	DNN	Own data	50 subjects	50 Healthy subjects	50-fold CV	5 stages	91.6%	90.7%	
Baek et al. [35]	2022	EEG (1 channel)	NA-BEMD (Noise assisted emprirical bivariate mode decomposition) 1D-CNN Bi-LSTM	Sleep-EDF-2013 Sleep-EDF-2018 WSC	20 subjects 78 subjects 116 subjects	214 Healthy subjects	20-fold CV	5 stages	86.22%	80.79%	0.81
Barua et al. [52]	2022	EEG	L-tetrolet TSRFINCA (ReliefF and INCA)	CAP	108 subjects	16 Healthy subjects 92 Pathological (insomnia, SDB, PLM)	tenfold CV	6 stages	91.05%	90.01%	

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Biswal et al. [53]	2018	EEG	RCNN	MGH SHHS	9000 subjects for training (MGH) 5224 subjects for training (SHHS) 1000 subjects for test (MGH) (SHHS) 580 subjects for test (SHHS)	Healthy—Mixed	9000 subjects for training; 1000 subjects for test 1000 subjects	5 stages	87.5% 77.7%		0.80 0.73
Cai et al. [54]	2021	EEG (1 channel)	Limited Penetrable Visibility Graph (LPVG) Graph-temporal fused dual-input CNN	Sleep-EDF	20 subjects	10 Healthy subjects 10 Subjects with mild sleep disturbance	Tenfold CV	2- 6 stages	91.97%		0.86
Chambonet al. [55]	2018	EEG EOG EMG	Spatial Filtering CNN	MASS	61 subjects	Healthy subjects	Fivefold CV Split: 41 Train -10 Val- 10 Test	5 stages	78.00%	72%	0.70
Chen et al. [56]	2020	EEG EOG EMG	Bi-LSTM-CNN	Sleep -EDF	39 subjects	Healthy subjects	Tenfold CV	5 stages	89.4%		
Cho et al. [57]	2022	EEG (6 channels) EOG (2 channels) EMG (1 ch)	StageNet (CNN + RNN)	SCHBC 2018–12-007	602 subjects	397 SDB (sleep apnea) subjects 205 Healthy subjects	482 subjects—training 48 subjects—validation 72 subjects—test	5 stages	90.4%		0.84
Choi et al. [58]	2020	Respiration, HR, BCG	LSTM	PSG recordings and PVDFdata (Seoul National University Hospital)	60 PSG	60 Healthy subjects	LOOCV 30 Training 10 Validation 20 Test	4 stages	73.9%		0.55
Delimayanti et al. [59]	2020	EEG (from single to multichannel)	High-Dimensional FFT Features SVM	Sleep-EDF	61 recordings for training (42 subjects)	61 Healthy subjects	Fivefold CV/tenfold CV	2- 6 stages	87.84% (avg)		
Dimitriadis Stavros et al. [60]	2018	EEG	Single-sensor ASSC Estimations of cross-frequency coupling with a multi-class Naïve Bayesian classifier	Sleep-EDF External Database	20 PSGs (Training) 77 PSG (Test)		LOOCV 20-fold CV—100 times	5 stages	94.40%		
Diykh et al. [61]	2020	EEG	Weighted undirected network. Least square support vector machine (LS-SVM)	Sleep-EDF ISRUC-Sleep	13 subjects 18 subjects	Mixed (healthy and SDB) subjects	Tenfold CV	6 stages	96.0% 96.74%		0.87 0.90
Dong et al. [62]	2018	EEG	Mixed Neural Network (MNN) Rectify Neural Network + LSTM	MASS	62 subjects	Healthy subjects	31-fold CV 30 subsets train 1 subset test	5 stages	85.92%	77.23%	

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Dutt et al. [36]	2022	EEG (1 channel)	CNN CRFs (Conditional Random Fields)	Expanded Sleep-EDF	38 recordings (20 subjects)	Healthy subjects	20-fold CV	5 stages	86.8%		0.78
El Moaqet, Hisham et al. [63]	2022	EEG	TF images from wavelet transform GoogLeNet CNN, BiLSTM	DS-1 (Charité) DS-2(Umich)	20 subjects 61 subjects		20-fold CV	5 stages	91.2%		
Eldele et al. [39]	2021	EEG (1 channel)	Multi-resolution CNN Adaptive feature recalibration (AFR) Temporal context encoder (TCE) AttnSleep	Sleep-EDF-20 Sleep-EDF-78 SHHS	20 subjects 78 subjects 329 subjects	Healthy subjects	20-fold CV	5 stages	84.4% 81.3% 84.2%	78.1% 75.1% 75.3%	0.79 0.74 0.78
Eldele et al. [64]	2022	EEG	ADAST: Adversarial Domain Adaptation framework based on preserving attention mechanism and iterative Self-Training strategy	Sleep-EDF SHHS-1 SHHS2	20 subjects 42 subjects 44 subjects	Healthy subjects	60% training 20% validation 20% test	5 stages	74.0%	60.39%	
Erdenebayar et al. [65]	2020	EEG	SleepGRU	PSG recordings (Samsung Medical Center, Seoul, Korea)	89 subjects for training 23 subjects for testing	42 Healthy subjects/47 SDB (apnea) subjects 10 Healthy subject/12 SDB (apnea) subjects	89 Training 23 Test	3 stages	80.4%	80.7%	
Fan et al. [66]	2021	EOG	2-scale CNN RNN	MASS Sleep-EDF	62 PSG 39 PSG(20subjects)	Healthy subjects	LOOCV	5 stages	81.2% 76.3%	73.7% 69.3%	0.71 0.67
Feng et al. [67]	2021	EEGs	Time Attention Mechanism: Feature vector BiGRU	Sleep-EDF Self-Collected Dataset	8 subjects 83 recordings (21 subjects)	4 Healthy subjects/4 effects of temazepam—insomnio Healthy subjects	Fivefold CV	5 stages	92.18% 90.06%	91.77% 89.91%	0.87 0.86
Fonseca et al. [68]	2020	Heart rate variability (HRV)	LSTM	SOMNIA SIESTA Computer-Assisted Automated Scoring of Polysomnograms Using the Sonnolyzer System	389 584 97	Mixed (healthy and SDB) subjects	Hold out 584 Training 1 (SIESTA) 97 Training 2 Validation 389 (SOMNIA)	4 stages (W/ N1-N2/ N3/REM)	75.9%		0.60

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Fraiwan et al. [69]	2020	EEG (2 channels)	Bi-LSTM	Sleep-EDF	39 PSG recordings (20 subjects)	Healthy subjects	Tenfold CV	6 stages	97.28%	74.64%	0.72
Frilot et al. [70]	2018	EEG	CNN Fine-grained segments succeeding 30 s	ISRUC-Sleep	116 subjects	Healthy subjects	Tenfold CV	5 stages	92.00%		
Fu et al. [71]	2021	EEG (1 channel)	Attention Mechanism Bidirectional LSTM (AT-BiLSTM)	DRM-SUB PSEE Expanded Sleep-EDF (Test) DREAMS (Test)	20 subjects 20 subjects 153 subjects 20subjects	Healthy subjects	Fivefold CV	5 stages	83.78% 81.72%	82.14% 80.74%	0.77 0.75
Gaiduk et al. [72]	2018	Respiratory HR Movement	Multinomial Logistic Regression (MLR)	Charité Dataset	35 subjects	Healthy subjects	5 subjects for training 30 subjects for testing	3 stages 4 stages	72.00% 58.00%		0.67 0.50
Gaiduk et al. [73]	2022	Breathing, Movement	Multinomial Logistic Regression (MLR)	Charité Dataset	35 subjects	Healthy subjects	15 subjects for training 20 subjects for test	3 stages	73%		0.44
Ghasemzadeh et al. [74]	2019	EEG	D3TDWT Logistic smooth transition autoregressive (LSTAR) SVM	Sleep-EDF ISRUC-Sleep	13 subjects 10 subjects	Healthy subjects	Tenfold CV	5 stages (Sleep-EDF) 6 stages (Sleep-EDF) 5 stages (ISRUC-Sleep)	94.83% 93.92% 82.25%		0.89 0.87 0.77
Ghimatgar et al. [75]	2019	EEG	Random Forest Hidden Markov Model (HMM)	DRM-SUB ISRUC Sleep EDF Expanded Sleep-EDF	20 subjects + 100 subjects 8 subjects + 61 subjects	Healthy subjects	LOOCV p=50%	5 stages 5 stages 5 stages/6 stages 5 stages/6 stages	79.08% 74.3% 87.59%/85.64% 80.55%/76.88%		0.73 0.71 0.81/0.78 0.73/0.68
Gopan et al. [76]	2020	EEG (5 channels)	Inter-channel covariance matrices Riemannian Manifold Ensamble with bagging	DREAMS	20 subjects	Healthy subjects	Tenfold CV		83.05%		0.79
Guillot et al. [77]	2020	EEG	SimpleSleepNet (Spectogram Signals and Frecuencies Reduction Gru with Attention Positional Embeddings Sequence Encoder)	DOD-H DOD-O	25 subjects 55 subjects	25 healthy subjects 55 SDB (OSA) subjects	DOD-H: LOOCV 18 Training 6 Validation 1Test DOD-O: tenfold CV 37 Training 12 Validation 6 Test		83.3% 89.0%	89.9% 88.3%	0.84 0.82
Guo et al. [78]	2022	Forehead electrophysiological (3 channels)	Light gradient boosting machine (LGB)	Expanded Sleep-EDF Own data	61 subjects 28 subjects	Healthy subjects	LOSO (leave one subject out)	5 stages	90.25%		0.85

259

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Hasan, Md. et al. [79]	2020	EEG (Prefrontal,central,occipital)	PSD ResNet	DREAMS	20 subjects	Healthy subjects	Fourfold CV 75% Training and validation, 25 Test	5 stages	85.8%		
Hei et al. [80]	2022	EOG ECG	Feature Extraction Module, Feature Fusion Module, XGBoost	SHHS2	69 subjects (PSG)	Healthy subjects	Fourfold CV	5 stages	83%		0.7
Hong et al. [31]	2021	EEG (1 channel)	SeqConfidNet Dropout Correct Rate (DCR)	SNUBH SHHS	702 2804	Mixed (healthy and SDB) subjects	70% Training 15% Validation 15% Test	5 stages	76% 82%	76% 82%	0.67 0.75
Horie et al. [37]	2022	EEGs, EOGs, EMG	Sleep-CAM: CNN	MASS+own data	109 subjects	Healthy subjects	Sevenfold CV	5 stages	86.9%		0.81
Huang et al. [40]	2022	EEG (1 channel)	CNN HMM	Sleep-EDFx	20 subjects	Healthy subjects	Tenfold CV and paired t-test	5 stages	84.6%		0.79
Huang et al. [81]	2020	EEG (2 channels)	ReliefF SVM	Expanded Sleep-EDF	30 subjects	Mixed (healthy and SDB) subjects	Tenfold CV 7:2:1	5 stages 6 stages	93.08% 92.34%		0.84
Huttunen et al. [82]	2022	Multichannel	Model 1(U-Net based) (PPG+SpO2) Model 2(U-Net based) (PPG+SpO2+nasal pressure signal) Model 3 (U-Net based) (SpO2+nasal pressure, EEG, oronasal thermocouple+RIP sum)	Princess Alexandra Hospital's Sleep Disorders Centre	877 subjects	Healthy subjects (suspected OSA)	80% Training 10% Validation 10% Test	5 stages	69% 70% 79%	68% 68% 79%	
Jadhav et al. [83]	2022	EEG (1 channel)	1D-CNN SWT-CNN STFT-CNN	Expanded Sleep-EDF (v1) Expanded Sleep-EDF(v2)	39 PSG recordings (20 subjects) 153 PSG recordings (78 subjects)	Healthy subjects	20-fold CV (Sleep-EDF-v1) tenfold CV (Sleep-EDF-v2)	5 stages	83.7%/82.4% 85.6%/82.5% 85.7%/82.7%		0.78/0.75 0.80/0.75 0.80/0.76
Jaoude et al. [84]	2020	EEG (4 channels)	CNN RNN	MGH-PSG scalpEEG	6431 subjects 112 subjects	Mixed (healthy and SDB) subjects 40 Healthy subjects/53 subjects with other pathologies (dementia..)	Fivefold CV	5 stages	81.15% 88.9%		0.74 0.78
Jia et al. [85]	2021	EEG EOG EMG	SleepPrintNet (a hierarchical recurrent neural network for the sequence-to-sequence classification problem)	MASS-SS3	62 subjects	Healthy subjects	31-fold CV	5 stages	88.8%	84.3%	0.83

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Jiang et al. [87]	2019	EEG EOG EMG	Covariance Features Riemannian manifolds Bootstrap aggregation (Bagging) classifier	MASS	61 PSG	Healthy subjects	LOOCV	5 stages	81.20%	80.9%	0.72
Jiang et al. [86]	2019	EEG	Random Forest Hidden Markov Model (HMM)	MASS Sleep EDF Expanded Sleep EDF	89 subjects	Mixed (healthy and SDB) subjects	LOOCV	5 stages	80.8% 93.0% 92.6%	79.3% 93.0% 90.8%	0.71 0.86 0.86
Kang Dae et al. [88]	2018	EEG	Multitaper Spectral Estimation EEG Spectral Feature Straction Stage-Specific Kernel Density Estimation Hidden Markov Model	UC San Diego Sleep-EDF	65 (15 HN-50 OSA 15 (HN)	30 Healthy subjects 50 SDB (OSA) subjects	Fivefold CV	5 Stages			0.70
Karimzadeh et al. [89]	2018	EEG	Time/Frecuency Measure of Phase & Envelope Frequency Domain Segmentation Entropy Calculation (Shannon Entropy) Distributed decision tree classifier, KNN	Sleep -EDF SHHS	20 subjects 140 subject	Healthy subjects SDB subjects	LOO 60% validation/40% test 70% validation/30% test	5 stages	88.97% 83.17%		
Khalili et al. [90]	2021	EEG (1 channel)	CNN TCNN onditional Random Field layer Data augmentation	Sleep-EDF-2013 Sleep-EDF-2018	39 recordings 153 subjects	Mixed (healthy and SDB) subjects	LOOCV 39PSG SleepEDF-2013 tenfold CV 153 PSG SleepEDF-2018	5 stages	85.39% 82.46%		0.80 0.76
Khojandi A et al. [91]	2019	EEG (2 channels)	Random Forest	Sleep-EDF	20 subjects	Healthy subjects	20-fold CV	5 stages	74.00%		
Kim et al. [92]	2022	EEG (1 channel)	CNN Transformers	Own dataset (Hallym University College of Medicine (Chuncheon, Republic of Korea:No. 2021-03-005)	2274 subjects	Mixed (healthy and SDB) subjects	1590 subjects for training (70%) 341 subjects for validation (15%) 343 subjects for test (15%)	3 stages	91.4%	89%	0.84
Kim et al. [93]	2022	EEG EMG	CNN Feature-level fusion	Expanded Sleep-EDF	20 subjects	Healthy subjects	20-fold CV	5 stages	87.2%		
Kim et al. [94]	2022	HRV, Actigraphy	Random Forest	PSG recordings (Samsung Medical Center, Seoul, Korea)	30 subjects	Mixed (healthy and SDB) subjects	Paired t-test	5 stages	86.19%		0.97

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Korkalainen et al. [95]	2020	PPG	CNN RNN	Clinical Dataset (Princess Alexandra Hospital, Australia) using Compumedics Grael acquisition system)	894 subjects	Mixed (healthy and SDB) subjects	Hold out 715 Training 90 Validation 89 Test	5 stages	64.1%		0.51
Korkalainen et al. [96]	2020	EEG EOG	CNN+LSTM	Expanded Sleep-EDF Clinical dataset	153 PSG 891 PSG	Healthy subjects Mixed (healthy and SDB) subjects	tenfold CV LOOCV 80% Training 10% Validation 10% Validation	5 stages	83.9% 83.8%		0.78 0.78
Kwon et al. [97]	2022	EEG (2 channels) EOG (2 channels)	CNN	ISRUC-Sleep Lab Dataset	100 subjects 4 subjects	Healthy subjects	60% training 20% validation 20% test	5 stages	88.85% 87.05%		0.85 0.83
Lee et al. [98]	2022	EEG (1 channel)	SleepExpertNet: Embedding vector generation (temporal spectra CNN) Temporal context modeling (multi-head attention+BiLSTM)	Sleep-EDFX-2018	153 subjects	Healthy subjects	Tenfold CV	5 stages	90.8%	86.7%	0.87
Li et al. [99]	2022	EEG	EEGSNet: CNNs+Bi-LSTM	Sleep-EDFX-8 Sleep-EDFX-20 Sleep-EDFX-78 SHHS	8 subjects 20 subjects 78 subjects 329 subjects	Healthy subjects	20-fold CV LOSO (leave-one-subject-out) 20-fold CV 20-fold CV	5 stages	94.17% 86.82% 83.02% 85.12%	87.78% 81.57% 77.26% 78.54%	0.91 0.82 0.77 0.79
Li et al. [100]	2022	EEGs	GCN (Graph neural network) CNN	Sleep-EDF ISRUC-SLEEP	20 subjects 10 subjects	Healthy subjects	20-fold CV tenfold CV	5 stages	91.0% 87.4%	89.0% 86.5%	0.88 0.84
Li et al. [101]	2021	PPG Actigraphy	CNN SVM	Emory Twin Study Follow-Up (ETSF)	105 subjects	No information given	tenfold CV 105 subjects ETSF	4 stages	68.62%		0.44
Li et al. [102]	2022	EEG(1 channel)	CAttSleepNet (Attention Module+CNN+BiLSTM)	Sleep-EDFX-2013 Sleep-EDFX-2018	39 recordings (20 subjects) 159 subjects (79 subjects)	Healthy subjects	20-fold CV & tenfold CV +I53	5 stages	84.14% 80.81%	84.14% 80.81%	0.78 0.735
Li Xiaojin et al. [103]	2018	EEG	HyCLASSS: Feature extraction+Random Forest+Correction Rule	CCSHS CFS	116 subjects 82 subjects		Split:96 Train—102 Test	5 stages	85.95%	71.6%	0.8046
Li et al. [104]	2022	EEG ECG EOG EMG	MVF-SleepNet (VGG-16+GRU…)	ISRUC-S3 ISRUC-S1	10 recordings 114 recordings	10 Healthy recordings 114 SDB recordings	Tenfold CV (ISRUC-S3) 90% Training—10% Test (ISRUC-S1)	5 stages	84.1% 82.1%	82.8% 80.2%	0.79 0.76

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Ling et al. [105]	2022	EEG (1 channel)	RDB-DCGAN (Residual Dense Block and DeepConvolutional Generative Adversarial Network) CNN	Expanded Sleep-EDF	40 subjects	Healthy subjects	70% Training 30% Test	5 stages	76.0%		
Liu et al. [106]	2021	EEG (1 channel)	Ensemble Empirical Mode Decomposition (EEMD) eXtreme Gradient Boosting (XGBoost) algorithm	Sleep-EDF DREAMS SHHS	16 recordings (8 subjects) 20 subjects 111 subjects	Healthy subjects	Fivefold CV	4/5 stages	93.1%/91.9% 86.4%/83.4% 87.5%/85.8%		0.87 0.77 0.79
Liu et al. [107]	2022	EEG	MAResnet (Multi-scale Attention Residual Net) BiGRU	Sleep-EDF UCD SHHS	20 subjects 20 PSG 5445 subjects	20 Healthy subjects 20 subjects with suspected SDB 5445 Mixed (healthy and SDB) subjects	Tenfold CV	5 stages	84.24% 79.34% 81.6%		0.78 (Sleep-EDF)
Liu et al. [108]	2020	EEG (1 channel)	Intrinsic geometry of sleep dynamics SVM	Expanded Sleep-EDF	83 recordings (42 subjects)	20 Healthy subjects 22 SDB subjects	LOOCV	5 stages	82.72%	75.91%	0.76
Liu et al. [109]	2022	EEG (single channel)	Deep neural network combining MSE(Multi-Scale Extraction) based U-structure and CBAM (Convolutional Block Attention Module)	Sleep-EDF-39 Sleep-EDF-153 SHHS	20 subjects/39PSGs 78 subjects/153PSGs 329 subjects/PSGs	Healthy Healthy With regular sleep	20-fold CV	5 stages	90.3% 89.7% 86.8%	86.2% 85.2% 83.5%	
Lu et al. [110]	2022	EEG (2 channels), EOGs (2 channels), EMG (1 channel)	CNN-RNN, weighted loss function (WLF)	Sleep-EDF-13 SHHS1	20 subjects/39 PSGs 5793 subjects/PSGs	Healthy Mixed	Nested CV (20-fold, tenfold) Split: 81%/9%/10%	5 stages	87.2% 89.1%	82.1% 81.4%	0.82 0.85
Luo et al. [111]	2022	Body movement ECG Abdominal breathing	Multi-head self-attention blocks coupled with CNN	MESA SHHS1	2237 subjects 5793 subjects/PSGs	Mixed Mixed	Fivefold CV	3 stages	82.9% 84.3%	79.2% 80.38%	0.70 0.70
Lv et al. [112]	2022	EEG (1 channel)	Multilevel temporal context network (MLTCN): CNN+TCN+HMM	Sleep-EDF-2013 Sleep-EDF-2018	20 subjects/39PSGs 7 subjects/153PSGs		20-fold CV tenfold CV	5 stages	84.2% 81.0%	77.0% 74.9%	0.78 0.74
Malafeev et al. [27]	2018	EEG (1 channel) EOG (2 channels) EMG (1 channel)	LSTM CNN-LSTM	Sleep laboratory, University of Zurich Sleep disorders center, Warsaw	18 subjects/54 PSGs 22 night PSG	Healthy Mixed	Split: 70%/15%/15% All dataset—test	5 stages			Healthy: ca. 0.8 for all except for N1 (<0.5)

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Matsumori et al. [113]	2022	Patch-type wearable EEG sensor (3 ch)	CNN-LSTM	Own dataset, PSG-1100 (NIHON KOH-DEN)	25 subjects/26PSGs	Healthy, 6 suspected to have SDB at low level	Leave-one-out CV	5 stages	78.6%	73.4%	0.70
Mousavi et al. [114]	2019	EEG (1 channel)	CNN-BiRNN	Sleep-EDF-2013 Sleep-EDF-2018	61 PSGs 197 PSGs		20-fold CV tenfold CV	5 stages	84.3% 80.0%	79.7% 73.6%	0.79 0.73
Neng et al. [115]	2021	EEG (1 channel)	CCRRSleepNet: frame-level CNN and epoch-level RNN	Sleep-EDF-2013	20 subjects/39PSGs	Healthy	20-fold CV	5 stages	84.3%	79.8%	0.78
Olesen et al. [43]	2021	EEG (1 channel) EOGs (2 channels) EMG (1 channel)	Deep residual neural network	ISRUC MrOS SHHS WSC SSC	118 subjects/126PSGs 3932 PSGs 8444 PSGs 767 PSGs 2401 PSGs	Mixed	Split(ISRUC): 33%/33%/34% Split(other): 87.5%/2.5%/10%	5 stages	86.9%		0.80
Paisarnsrisomsuk et al. [116]	2021	EEG (2 channels) EOG (1 ch)	CNN	Sleep-EDF	20 subjects/39PSGs	Healthy	Nested 5×4 CV	5 stages	84.5%	78.1%	
Pan et al. [117]	2021	EEG+EOG+ EMG+Resp+Temp	SVM	Sleep-EDF Hospital sleep data	20 subjects/38 PSG 31 subjects	Mixed Depressed/Healthy	Fourfold CV	4 stages	90.1% 78.6%/69.2%		0.87 0.71/0.58
Patanaik et al. [118]	2018	EEG (2 channels) EOG (2 channels)	dCNN	Own datasets (DS1-DS4) from Singapore and San Diego, USA	DS1: 120 subjects/1046 PSGs DS2: 52 subjects/284 PSGs DS3: 210 subjects/ PSGs DS4: 77 subj/PSGs	Healthy Healthy Sleep disorders Parkinson´s disease	DS1, DS2—training DS3, DS4—test	5 stages	DS3: 81.4% DS4: 72.1%		DS3: 0.74 DS4: 0.60
Pathak et al. [38]	2021	EEG (2–8 channels) EOG (2–3 ch) EMG (1 ch)	CNN BiLSTM	SHHS1 MST	5793 PSGs 1418 PSGs	Mixed	Split: 81%/9%/10%	5 stages	85% 77%	79.0% 73.0%	
Pei et al. [119]	2022	EEG (1 channel) EOG (2 ch) EMG (1 ch)	CNN with GRU softmax classifier	UCDDB SHHS1	25 subjects/25 PSGs 700 PSGs		Split: 50%/20%/30%	5 stages	68.3%/83.2%		0.58 0.76
Perslev t al. [120]	2021	EEG (1 channel) EOG (1 channel)	U-Sleep, deep-learning-based system	16 independent clinical studies	15,660 subjects/19924 PSGs	Mixed	Split: training (at least 75%), validation (up to 10%, at most 50 subjects), testing (up to 15%, at most 100 subjects)	5 stages		79.0%	
Phan et al. [121]	2019	EEG+EOG/ EEG+EOG+EMG	Multitask 1-max CNN	Sleep-EDF (expanded) MASS	20 subjects/39 PSGs 200 subjects	Suspected sleep disorder	Leave-one-subject-out CV 20-fold CV	5 stages	82.3% 83.6%	74.7% 77.9%	0.75 0.77
Phan et al. [122]	2019	EEG+EOG+EMG	SeqSleepNet: End-to-End Hierarchical RNN	MASS	200 subjects		20-fold CV	5 stages	87.1%	83.3%	0.82

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Phan et al. [32]	2022	EEG (single channel)	SleepTransformer, a sequence-to-sequence sleep staging model	SHHS1, Sleep-EDF-78	5791 subjects, 78 subjects/163 PSGs	Mixed	Split: 70%/30% tenfold CV	5 stages	87.7%, 81.4%	80.1%, 74.3%	0.83, 0.74
Pini et al. [123]	2022	ECG	Deep-learning architecture: Neurobit-HRV	Physionet CinC, Proprietary dataset	988 subjects/PSGs, 52 subjects/112 PSGs	Mixed	Previously trained on another dataset. Here all data is for test	3/4 stages	3: 82.1%/80.3%, 4:73.0/70.3%		3:0.60/0.61, 4:0.51/0.53
Qu et al. [124]	2020	EEG	ResNet-CNN	MASS Sleep-EDF	62 subjects/PSGs 20 subjects		31-fold CV 20-fold CV	5 stages	86.5%, 84.3%	81.0%, 79.0%	0.80, 0.78
Qureshi et al. [125]	2019	EEG (1 channel)	CNN: GACNN Sleep-TuneNet	CAP Sleep, Sleep-EDF	20 subjects, 20 subjects		Tenfold CV	5 stages	95.6%, 92.5%		0.94, 0.90
Radha et al. [126]	2019	ECG, HRV	LSTM	SIESTA	292 subjects/584 PSGs	Mixed	Fourfold CV	5 stages	77.0%		0.61
Seo et al. [127]	2020	EEG (1 channel)	CRNN: IITNet	Sleep-EDF-2013, MASS, SHHS	20 subjects, 62 subjects, 5791 subjects	Mixed	20-fold CV, 31-fold CV, Split: 50%/20%/30%	5 stages	83.9%, 86.5%, 86.7%	77.6%, 80.7%, 79.8%	0.78, 0.80, 0.81
Sharma et al. [128]	2022	EEG (2 channels), EMG (1 ch), EOG (2 ch)	Wavelet-based Tsallis entropy features, EBT classifier	SHHS-1, SHHS-2	5793 subjects, 2651 subjects	Mixed	10% holdout validation	3/5 stages	90.7%/84.3%, 91.8%/86.3%		0.84/0.77, 0.86/0.80
Shen et al. [129]	2020	EEG (1 channel)	Improved model based essence features (IMBEFs), Locality energy (LE), Dual state space models (DSSMs)	Dreams, ISRUC	20 subjects/PSGs, 10 subjects/PSGs		Tenfold CV	5 stages	79.9%, 81.7%		0.76
Sokolovsky et al. [130]	2020	EEG (2 channels), EOG (1 channel)	CNN	Sleep-EDF (expanded)	20 subjects/39PSGs	Healthy	Fourfold CV	5 stages	81.0%	72%	
Sors et al. [131]	2018	EEG (1 channel)	CNN	SHHS-1	5793 subjects	Mixed	Split: 50%/20%/30%	5 stages	87.0%	78%	0.81
Sridhar et al. [132]	2020	Instantaneous heart rate (IHR), ECG	CNN	SHHS, MESA, CinC	561 subjects/800 PSGs, 194 PSGs, 993 subjects/PSGs	Mixed	Only test recordings are listed in the table	4 stages	77.0%, 80.0%, 72.0%		0.66, 0.69, 0.55
Sun et al. [133]	2019	EOG, RR-intervals	CNN, RNN	MASS (SS3), SA	62 PSGs, 24 subjects/PSGs	Suspected sleep disorder	Leave-one-subject-out	5 stages/4 stages	84.4%/88.2%, 74.3%/81.3%	78.1%/85.8%, 74.0%/80.2%	0.77/0.80, 0.66/0.71
Sun et al. [134]	2020	EEG, EMG, EOG	CNN+BLSTM	MASS (SS2-SS5)	147 subjects/PSGs	Suspected sleep disorder	Leave-one-subject-out	5 stages	87.8%	81.8%	0.82
Sun et al. [135]	2021	EEG, EOG	CNN	Sleep-EDF	20 subjects/39 PSGs	Healthy	20-fold CV	5 stages	83.4%	77.3%	0.77
Tao et al. [136]	2022	EEG (1 channel)	Cascaded CNN LSTM	Sleep-EDF-2013	20 subjects/39 PSGs	Healthy	Fivefold CV	5 stages	82.7%	76.8%	

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Urtnasan et al. [137]	2022	Single-Lead ECG	Deep convolutional recurrent (DCR)	Proprietary dataset, Samsung Medical Center, Seoul, Korea	112 subjects	Control: 52 subjects, apnoea: 60 subjects	Training set: 89 subjects, test set: 23 subjects	5 stages 3 stages	74.2% 86.4%		
Vallat et al. [138]	2021	EEG (1 channel) EOG (1 channel) EMG (1 ch)	LightGBM	MESA, CFS, CCSHS, SHHS, MrOS, CHAT, HomePAP, DOD (healthy and obstructive)	3798 PSGs	Mixed	Training: 3163 PSGs Test: 585 PSGs/50 PSGs (DOD-obstructive)	5 stages	87.5%/84.3%	74.0%	0.82/
van der Plas et al. [33]	2021	EEG EMG EOG ECG SaO2 Position Snore sound	Emission models and the random forest	1: Dataset from 10 medical centers in Belgium, the Netherlands and Austria 2: Proprietary dataset from a single center	1: 5884 PSGs 2: 50 PSGs	1: each PSG is scored by one expert 2: scored by four, five, six experts or seven	1: Split: 33.3%/33.3%/33.3% 2: Test only	5 stages	1: 80.5% 2: 82.7%		1: 0.73
Vanbuis et al. [34]	2020	EEG EMG EOG	Random forest, interpretable	Proprietary dataset, Institut de Recherche en Santé Respiratoire	400 PSGs	Suspected having OSA subjects	Split: 75%/25%	5 stages	77.8%		0.69
Vanbuis et al. [139]	2022	Pulse oximetry Respiration flow and effort Actigraphy Ambient light Tracheal sound	Multilayer perceptron classifier	Proprietary dataset, Institut de Recherche en Santé Respiratoire	400 PSGs	Suspected having OSA subjects	Split: 75%/25%	5 stages 3 stages	62.4% 78.5%		0.48 0.60
Wang et al. [140]	2022	EEG	Deepsleepnet network (CNN, bi-LSTM), LightGBM	MASS Sleep-EDF (expanded)	200 subjects 20 subjects	Suspected sleep disorder	Training: MASS, Sleep-EDF Test: Sleep-EDF Leave-one-out CV	5 stages	87.8%		
Wang et al. [141]	2022	EEG (1 channel)	Multiscale dual attention network (MSDAN)	Sleep-EDF Sleep-EDFx	8 PSGs 197 PSGs		20-fold CV	5 stages	91.7% 90.4%	82.3% 79.5%	0.87 0.83
Wei et al. [142]	2019	Single-Lead ECG	LSTM	Proprietary dataset, Xijing Hospital, Fourth Military Medical University	373 subjects/PSGs	depression or schizophrenia	Split: 238/60/75 recordings	3 stages 4 stages 5 stages	84.1% 77.8% 71.2%		0.58 0.55 0.52
Wongsirichot et al. [143]	2019	Multichannel	k-Nearest Neighbors (the best results)	SHHS	2535 PSGs	Mixed	Tenfold CV	5 stages	83.8%	69.4%	
Yan et al. [144]	2021	EOG, EEG, EMG, ECG	2d CNN LSTM	SHHS-1 ISRUC Sleep-EDF	100 PSGs 99 PSGs 19 PSGs	RDI3P <5 mixed health state 2nd night	Fivefold CV	5 stages	87.0% 86.0% 86.0%		0.81 0.82 0.81

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Yang et al. [41]	2022	EEG	Conditional random fields (CRF), CNN-LSTM	Sleep-EDF-20 DRM-SUB SVUH-UCD	20 subjects/39 PSGs 20 subjects/PSGs 25 subjects	Healthy Healthy With sleep apnoea	20-fold CV 20-fold CV 25-fold CV	5 stages	85.2% 83.2% 74.8%	78.1% 76.9% 69.9%	0.79 0.77 0.66
You et al. [145]	2022	EEG (1 channel)	Bi-LSTM, Fractional Fourier transform (FRFT)	Sleep-EDF MASS	20 subjects 62 subjects	Healthy subjects	20-fold CV	5 stages	81.6% 84.3%	74.7% 77.8%	0.75 0.77
Yu et al. [146]	2022	EEG (1 channel)	Deep neural network—MRASleepNet: a feature extraction (FE) module, an MRA module, and a gMLP module	Sleep-EDF-20 Sleep-EDF-78 CAP	20 subjects 78 subjects 63 subjects	Healthy and mixed subjects	20-fold CV tenfold CV Split: 70%/10%/20%	5 stages	84.5% 81.4% 74.3%	78.9% 75.4% 65.6%	0.79 0.74 0.65
Yulita et al. [147]	2018	EEG EMG EOG	The fast convolutional method	Physionet Proprietary dataset: Mitra Keluarga Kemayoran Hospital	25 subjects 10 subjects		Leave-one-out CV	5 stages		73.5% 56.3%	
Zhang et al. [148]	2022	EEG (dual channel)	Random forest, rule-set creation based on data-mining, Ontology-Based Modeling (OBM)	Sleep-EDF (expanded)	42 subjects/61 PSGs	Healthy and with mild difficulty in falling asleep	Split: 50%/50%	5 stages	89.1%	85.9%	0.81
Zhang et al. [149]	2022	EEG	Co-attention meta sleep staging network (CMS2-net), semi-supervised	Proprietary: Peking University Sixth Hospital DOD-H DOD-O	48 subjects 25 subjects 55 subjects	Mixed (5 subsets) Healthy OSA patients	Split: 60%/20%/20%	5 stages	74%-81% 80% 78%	53%-69% 69% 68%	0.62–0.73 0.71 0.68
Zhang et al. [150]	2021	EEG (1 channel)	Competition convolutional neural network (C-CNN)	UCD Sleep-EDF(expanded)	25 subjects 42 subjects	With sleep-respiration disorder	Leave-one-out CV 20-fold CV	5 stages	77.2% 83.5%	84.0%	0.73
Zhang et al. [44]	2019	EEG EMG EOG	Deep neural network combined recurrent and convolutional structures	SHHS-1 MrOS SOF	5793 subjects 2907 subjects 461 subjects	Mixed	Training (SHHS): 5213 subjects Test (SHHS): 580 subjects, MrOS, SOF	5 stages	87.0%	87.0% 79.0% 77.0%	0.82 0.70 0.68
Zhang et al. [151]	2020	EEG (2 channels) EOG (2 channels) EMG (1 channel)	CNN	Proprietary dataset: Beijing Tongren Hospital Sleep-EDF-2013	294 PSGs 20 subjects	Snoring	Split: 122/20/152 Sleep-EDF—test	5 stages	81.8% 83.6%	81.5% 78.1%	0.73 0.77
Zhang et al. [152]	2018	Heart rate (from PPG), Actigraphy	Multi-level feature learning RNN (BiLSTM)	Proprietary dataset: General Hospital of the Air Force, PLA, Beijing, China	39 subjects	Healthy	Leave-one-out CV	5 stages		60.5%	

Table 2 (continued)

Author	Year	Signals	Algorithm/Method	Database	N° of subjects	Population characteristics	Evaluation	Sleep Stages	Accuracy	F1	Kappa
Zhong et al. [153]	2021	EEG (1 channel)	Multiscale residual convolutional neural network (MRCNN)	Sleep-EDF(expanded) CinC2018	197 PSGs / 500 PSGs	Mixed	Fivefold CV / Subject CV, Split 400/100	5 stages	92.1% / 89.2%		0.74 / 0.62
Zhou et al. [154]	2022	EEG (1 channel)	SingleChannelNet (deep neural network)	CCSHS (children) Sleep-EDF (expanded)	515 subjects / 78 subjects/163PSGs		Fivefold CV	5 stages	90.2% / 86.1%		0.87 / 0.81
Zhou et al. [155]	2020	EEG (1 channel)	Random forest (RF) and LightGBM (LGB)	Sleep-EDF Sleep-EDF (expanded)	8 subjects/PSGs / 153 PSGs		Twofold CV (subject-independent)	5 stages	82.4% / 85.3%	75.1% / 83.4%	0.72 / 0.69
Zhu et al. [42]	2020	EEG (2 channels) EOG (1 channel) EMG (1 channel)	Multi-branch convolutional neural network Residual attention method	Sleep-EDFx UCDDB	20 subjects/39 PSGs / 25 subjects	Healthy Sleep disordered breathing	Leave-one-subject-out	5 stages	85.8% / 79.4%	81.2% / 78.8%	0.80 / 0.73

Declarations

References

1. Worley SL. The extraordinary importance of sleep: the detrimental effects of inadequate sleep on health and public safety drive an explosion of sleep research. P T. 2018;43:758–63.

2. Walker MP. Cognitive consequences of sleep and sleep loss. Sleep Med. 2008;9(Suppl 1):S29-34. https://doi.org/10.1016/S1389-9457(08)70014-5.

3. Mukherjee S, Patel SR, Kales SN, Ayas NT, Strohl KP, Gozal D, Malhotra A. An official american thoracic society statement: the importance of healthy sleep: recommendations and future priorities. Am J Respir Crit Care Med. 2015;191:1450–8. https://doi.org/10.1164/rccm.201504-0767ST.

4. Perry GS, Patil SP, Presley-Cantrell LR. Raising awareness of sleep as a healthy behavior. Prev Chronic Dis. 2013. https://doi.org/10.5888/pcd10.130081.

5. Dement W, Kleitman N. Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. Electroencephalogr Clin Neurophysiol. 1957;9:673–90. https://doi.org/10.1016/0013-4694(57)90088-3.

6. Rechtschaffen A, Kales A. A Manual of Standardised Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Bethesda, Maryland: U.S. Department of Health, Education, and Welfare; 1968.

7. Iber C, Ancoli-Israel S, Chesson AL, Quan S. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. Westchester, IL: American Academy of Sleep Medicine. 2007.

8. Berry RB, Quan SF, Abreu A.R., et.al. The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications: Version 2.6. Darien, Illinois: American Academy of Sleep Medicine; 2020.

9. Danker-Hopfe H, Anderer P, Zeitlhofer J, Boeck M, Dorn H, Gruber G, et al. Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. J Sleep Res. 2009;18:74–84. https://doi.org/10.1111/j.1365-2869.2008.00700.x.

10. Moser D, Anderer P, Gruber G, Parapatics S, Loretz E, Boeck M, et al. Sleep classification according to AASM and

Rechtschaffen & Kales: effects on sleep scoring parameters. Sleep. 2009;32:139–49. https://doi.org/10.1093/sleep/32.2.139.

11. Douglas NJ, Thomas S, Jan MA. Clinical value of polysomnography. Lancet. 1992;339:347–50. https://doi.org/10.1016/0140-6736(92)91660-Z.

12. Hirshkowitz M. Polysomnography challenges. Sleep Med Clin. 2016;11:403–11. https://doi.org/10.1016/j.jsmc.2016.07.002.

13. Deng S, Zhang X, Zhang Y, Gao H, Chang EI-C, Fan Y, Xu Y. Interrater agreement between American and Chinese sleep centers according to the 2014 AASM standard. Sleep Breath. 2019;23:719–28. https://doi.org/10.1007/s11325-019-01801-x.

14. Shambroom JR, Fábregas SE, Johnstone J. Validation of an automated wireless system to monitor sleep in healthy adults. J Sleep Res. 2012;21:221–30. https://doi.org/10.1111/j.1365-2869.2011.00944.x.

15. Zhang X, Dong X, Kantelhardt JW, Li J, Zhao L, Garcia C, et al. Process and outcome for international reliability in sleep scoring. Sleep Breath. 2015;19:191–5. https://doi.org/10.1007/s11325-014-0990-0.

16. Lee YJ, Lee JY, Cho JH, Choi JH. Interrater reliability of sleep stage scoring: a meta-analysis. J Clin Sleep Med. 2022;18:193–202. https://doi.org/10.5664/jcsm.9538.

17. Alsolai H, Qureshi S, Iqbal SMZ, Vanichayobon S, Henesey LE, Lindley C, Karrila S. A systematic review of literature on automated sleep scoring. IEEE Access. 2022;10:79419–43. https://doi.org/10.1109/ACCESS.2022.3194145.

18. Djanian S, Bruun A, Nielsen TD. Sleep classification using Consumer Sleep Technologies and AI: a review of the current landscape. Sleep Med. 2022;100:390–403. https://doi.org/10.1016/j.sleep.2022.09.004.

19. Faust O, Razaghi H, Barika R, Ciaccio EJ, Acharya UR. A review of automated sleep stage scoring based on physiological signals for the new millennia. Comput Methods Programs Biomed. 2019;176:81–91. https://doi.org/10.1016/j.cmpb.2019.04.032.

20. Fiorillo L, Puiatti A, Papandrea M, Ratti P-L, Favaro P, Roth C, et al. Automated sleep scoring: a review of the latest approaches. Sleep Med Rev. 2019;48:101204. https://doi.org/10.1016/j.smrv.2019.07.007.

21. Loh HW, Ooi CP, Vicnesh J, Oh SL, Faust O, Gertych A, Acharya UR. Automated detection of sleep stages using deep learning techniques: a systematic review of the last decade (2010–2020). Appl Sci. 2020;10:8963. https://doi.org/10.3390/app10248963.

22. Phan H, Mikkelsen K. Automatic sleep staging of EEG signals: recent development, challenges, and future directions. Physiol Meas. 2022. https://doi.org/10.1088/1361-6579/ac6049.

23. Ronzhina M, Janoušek O, Kolářová J, Nováková M, Honzík P, Provazník I. Sleep scoring using artificial neural networks. Sleep Med Rev. 2012;16:251–63. https://doi.org/10.1016/j.smrv.2011.06.003.

24. Penzel T, Conradt R. Computer based sleep recording and analysis. Sleep Med Rev. 2000;4:131–48. https://doi.org/10.1053/smrv.1999.0087.

25. Snyder H. Literature review as a research methodology: an overview and guidelines. J Bus Res. 2019;104:333–9. https://doi.org/10.1016/j.jbusres.2019.07.039.

26. Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: meta-narrative reviews. BMC Med. 2013;11:20. https://doi.org/10.1186/1741-7015-11-20.

27. Malafeev A, Laptev D, Bauer S, Omlin X, Wierzbicka A, Wichniak A, et al. Automatic human sleep stage scoring using deep neural networks. Front Neurosci. 2018;12:781. https://doi.org/10.3389/fnins.2018.00781.

28. Blackman NJM, Koval JJ. Interval estimation for Cohen's kappa as a measure of agreement. Stat Med. 2000;19:723–41. https://doi.org/10.1002/(SICI)1097-0258(20000315)19:5%3C723::AID-SIM379%3E3.0.CO;2-A.

29. Xu H, Ebrahim MP, Hasan K, Heydari F, Howley P, Yuce MR. Accurate heart rate and respiration rate detection based on a higher-order harmonics peak selection method using radar non-contact sensors. Sensors (Basel). 2021. https://doi.org/10.3390/s22010083.

30. Massaroni C, Nicolo A, Sacchetti M, Schena E. Contactless methods for measuring respiratory rate: a review. IEEE Sensors J. 2021;21:12821–39. https://doi.org/10.1109/JSEN.2020.3023486.

31. Hong JK, Lee T, Reyes RDD, Hong J, Tran HH, Lee D, et al. Confidence-based framework using deep learning for automated sleep stage scoring. Nat Sci Sleep. 2021;13:2239–50. https://doi.org/10.2147/NSS.S333566.

32. Phan H, Mikkelsen K, Chén OY, Koch P, Mertins A, de Vos M. SleepTransformer: automatic sleep staging with interpretability and uncertainty quantification. IEEE Trans Biomed Eng. 2022;69:2456–67. https://doi.org/10.1109/TBME.2022.3147187.

33. van der Plas D, Verbraecken J, Willemen M, Meert W, Davis J. Evaluation of automated hypnogram analysis on multi-scored polysomnographies. Front Digit Health. 2021;3:707589. https://doi.org/10.3389/fdgth.2021.707589.

34. Vanbuis J, Feuilloy M, Baffet G, Meslier N, Gagnadoux F, Girault J-M. Towards a user-friendly sleep staging system for polysomnography part I: automatic classification based on medical knowledge. Inform Med Unlocked. 2020. https://doi.org/10.1016/j.imu.2020.100454.

35. Baek J, Lee C, Yu H, Baek S, Lee S, Lee S, Park C. Automatic sleep scoring using intrinsic mode based on interpretable deep neural networks. IEEE Access. 2022;10:36895–906. https://doi.org/10.1109/ACCESS.2022.3163250.

36. Dutt M, Redhu S, Goodwin M, Omlin CW. SleepXAI: an explainable deep learning approach for multi-class sleep stage identification. Appl Intell. 2022. https://doi.org/10.1007/s10489-022-04357-8.

37. Horie K, Ota L, Miyamoto R, Abe T, Suzuki Y, Kawana F, et al. Automated sleep stage scoring employing a reasoning mechanism and evaluation of its explainability. Sci Rep. 2022;12:12799. https://doi.org/10.1038/s41598-022-16334-9.

38. Pathak S, Lu C, Nagaraj SB, van Putten M, Seifert C. STQS: Interpretable multi-modal Spatial-Temporal-seQuential model for automatic Sleep scoring. Artif Intell Med. 2021;114:102038. https://doi.org/10.1016/j.artmed.2021.102038.

39. Eldele E, Chen Z, Liu C, Wu M, Kwoh CK, Li X, Guan C. An attention-based deep learning approach for sleep stage classification with single-channel EEG. IEEE Trans Neural Syst Rehabil Eng. 2021;29:809–18. https://doi.org/10.1109/TNSRE.2021.3076234.

40. Huang J, Ren L, Zhou X, Yan K. An improved neural network based on senet for sleep stage classification. IEEE J Biomed Health Inform. 2022;26:4948–56. https://doi.org/10.1109/JBHI.2022.3157262.

41. Yang B, Wu W, Liu Y, Liu H. A novel sleep stage contextual refinement algorithm leveraging conditional random fields. IEEE Trans Instrum Meas. 2022;71:1–13. https://doi.org/10.1109/TIM.2022.3154838.

42. Zhu T, Luo W, Yu F. Multi-branch convolutional neural network for automatic sleep stage classification with embedded stage refinement and residual attention channel Fusion. Sensors (Basel). 2020. https://doi.org/10.3390/s20226592.

43. Olesen AN, Jørgen Jennum P, Mignot E, Sorensen HBD. Automatic sleep stage classification with deep residual networks in a mixed-cohort setting. Sleep. 2021. https://doi.org/10.1093/sleep/zsaa161.

44. Zhang L, Fabbri D, Upender R, Kent D. Automated sleep stage scoring of the Sleep Heart Health Study using deep neural networks. Sleep. 2019. https://doi.org/10.1093/sleep/zsz159.

45. Penzel T. Sleep scoring moving from visual scoring towards automated scoring. Sleep. 2022. https://doi.org/10.1093/sleep/zsac190.

46. Abdollahpour M, Rezaii TY, Farzamnia A, Saad I. Transfer learning convolutional neural network for sleep stage classification using two-stage data fusion framework. IEEE Access. 2020;8:180618–32. https://doi.org/10.1109/ACCESS.2020.3027289.

47. Abdollahpour M, Rezaii TY, Farzamnia A, Saad I. A two-stage learning convolutional neural network for sleep stage classification using a filterbank and single feature. IEEE Access. 2022;10:60597–609. https://doi.org/10.1109/ACCESS.2022.3180730.

48. Alickovic E, Subasi A. Ensemble SVM method for automatic sleep stage classification. IEEE Trans Instrum Meas. 2018;67:1258–65. https://doi.org/10.1109/TIM.2018.2799059.

49. Alvarez-Estevez D, Rijsman RM. Inter-database validation of a deep learning approach for automatic sleep scoring. PLoS ONE. 2021;16:e0256111. https://doi.org/10.1371/journal.pone.0256111.

50. An P, Zhao J, Du B, Zhao W, Zhang T, Yuan Z. Amplitude-time dual-view fused EEG temporal feature learning for automatic sleep staging. IEEE Trans Neural Netw Learn Syst. 2022. https://doi.org/10.1109/TNNLS.2022.3210384.

51. Arslan RS, Ulutas H, Köksal AS, Bakir M, Çiftçi B. Sensitive deep learning application on sleep stage scoring by using all PSG data. Neural Comput Appl. 2022. https://doi.org/10.1007/s00521-022-08037-z.

52. Barua PD, Tuncer I, Aydemir E, Faust O, Chakraborty S, Subbhuraam V, et al. L-Tetrolet pattern-based sleep stage classification model using balanced EEG datasets. Diagnostics. 2022. https://doi.org/10.3390/diagnostics12102510.

53. Biswal S, Sun H, Goparaju B, Brandon Westover M, Sun J, Bianchi MT, Westover MB. Expert-level sleep scoring with deep neural networks. J Am Med Inform Assoc. 2018;25:1643–50. https://doi.org/10.1093/jamia/ocy131.

54. Cai Q, Gao Z, Member S, An J, Gao S, Grebogi C. A graph-temporal fused dual-input convolutional neural network for detecting sleep stages from EEG signals. IEEE Trans Circuits Syst II Express Briefs. 2021;68:777–81. https://doi.org/10.1109/TCSII.2020.3014514.

55. Chambon S, Galtier MN, Arnal PJ, Wainrib G, Gramfort A. A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. IEEE Trans Neural Syst Rehabil Eng. 2018;26:758–69. https://doi.org/10.1109/TNSRE.2018.2813138.

56. Chen X, He J, Wu X, Yan W, Wei W. Sleep staging by bidirectional long short-term memory convolution neural network. Fut Gen Comput Syst. 2020;109:188–96. https://doi.org/10.1016/j.future.2020.03.019.

57. Cho JH, Choi JH, Moon JE, Lee YJ, Lee HD, Ha TK. Validation study on automated sleep stage scoring using a deep learning algorithmÔÇá. Medicina (Lithuania). 2022. https://doi.org/10.3390/medicina58060779.

58. Choi SH, Kwon HB, Jin HW, Yoon H, Lee MH, Lee YJ, Park KS. Long short-term memory networks for unconstrained sleep stage classification using polyvinylidene fluoride film sensor. IEEE J Biomed Health Inform. 2020;24:3606–15. https://doi.org/10.1109/JBHI.2020.2979168.

59. Delimayanti MK, Purnama B, Nguyen NG, Faisal MR, Mahmudah KR, Indriani F, et al. Classification of brainwaves for sleep stages by high-dimensional FFT features from EEG signals. Appl Sci (Switzerland). 2020. https://doi.org/10.3390/app10051797.

60. Dimitriadis SI, Salis C, Linden D. A novel, fast and efficient single-sensor automatic sleep-stage classification based on complementary cross-frequency coupling estimates. Clin Neurophysiol. 2018;129:815–28. https://doi.org/10.1016/j.clinph.2017.12.039.

61. Diykh M, Li Y, Abdulla S. EEG sleep stages identification based on weighted undirected complex networks. Comput Methods Prog Biomed. 2020;184:105116. https://doi.org/10.1016/j.cmpb.2019.105116.

62. Dong H, Supratak A, Pan W, Wu C, Matthews PM, Guo Y. Mixed neural network approach for temporal sleep stage classification. IEEE Trans Neural Syst Rehabil Eng. 2018;26:324–33. https://doi.org/10.1109/TNSRE.2017.2733220.

63. El Moaqet H, Eid M, Ryalat M, Penzel T. A deep transfer learning framework for sleep stage classification with single-channel EEG signals. Sensors (Basel). 2022. https://doi.org/10.3390/s22228826.

64. Eldele E, Ragab M, Chen Z, Wu M, Kwoh CK, Li X, Guan C. ADAST: attentive cross-domain EEG-based sleep staging framework with iterative self-training. IEEE Trans Emerg Top Comput Intell. 2022. https://doi.org/10.1109/TETCI.2022.3189695.

65. Erdenebayar U, Kim Y, Park JU, Lee SY, Lee KJ. Automatic classification of sleep stage from an ECG signal using a gated-recurrent unit. Int J Fuzzy Logic Intell Syst. 2020;20:181–7. https://doi.org/10.5391/IJFIS.2020.20.3.181.

66. Fan J, Sun C, Long M, Chen C, Chen W. EOGNET: a novel deep learning model for sleep stage classification based on single-channel EOG signal. Front Neurosci. 2021;15:573194. https://doi.org/10.3389/fnins.2021.573194.

67. Feng LX, Li X, Wang HY, Zheng WY, Zhang YQ, Gao DR, Wang MQ. Automatic sleep staging algorithm based on time attention mechanism. Front Hum Neurosci. 2021;15:692054. https://doi.org/10.3389/fnhum.2021.692054.

68. Fonseca P, van Gilst MM, Radha M, Ross M, Moreau A, Cerny A, et al. Automatic sleep staging using heart rate variability, body movements, and recurrent neural networks in a sleep disordered population. Sleep. 2020;43:1–10. https://doi.org/10.1093/sleep/zsaa048.

69. Fraiwan L, Alkhodari M. Investigating the use of uni-directional and bi-directional long short-term memory models for automatic sleep stage scoring. Inform Med Unlocked. 2020;20:100370. https://doi.org/10.1016/j.imu.2020.100370.

70. Frilot C II, McCarty DE, Marino AA. An original method for staging sleep based on dynamical analysis of a single EEG signal. J Neurosci Methods. 2018;308:135–41. https://doi.org/10.1016/j.jneumeth.2018.07.017.

71. Fu M, Wang Y, Chen Z, Li J, Xu F, Liu X, Hou F. Deep learning in automatic sleep staging with a single channel electroencephalography. Front Physiol. 2021;12:628502. https://doi.org/10.3389/fphys.2021.628502.

72. Gaiduk M, Penzel T, Ortega JA, Seepold R. Automatic sleep stages classification using respiratory, heart rate and movement signals. Physiol Meas. 2018;39:124008. https://doi.org/10.1088/1361-6579/aaf5d4.

73. Gaiduk M, Perea Rodríguez JJ, Seepold R, Martínez Madrid N, Penzel T, Glos M, Ortega JA. Estimation of sleep stages analyzing respiratory and movement signals. IEEE J Biomed Health Inform. 2022;26:505–14. https://doi.org/10.1109/JBHI.2021.3099295.

74. Ghasemzadeh P, Kalbkhani H, Sartipi S, Shayesteh MG. Classification of sleep stages based on LSTAR model. Appl Soft Comput J. 2019;75:523–36. https://doi.org/10.1016/j.asoc.2018.11.007.

75. Ghimatgar H, Kazemi K, Helfroush MS, Aarabi A. An automatic single-channel EEG-based sleep stage scoring method based on hidden Markov Model. J Neurosci Methods. 2019;324:108320. https://doi.org/10.1016/j.jneumeth.2019.108320.

76. Gopan KG, Prabhu SS, Sinha N. Sleep EEG analysis utilizing inter-channel covariance matrices. Biocybern Biomed Eng. 2020;40:527–45. https://doi.org/10.1016/j.bbe.2020.01.013.

77. Guillot A, Sauvet F, During EH, Thorey V. Dreem open datasets: multi-scored sleep datasets to compare human and automated sleep staging. IEEE Trans Neural Syst Rehabil Eng. 2020;28:1955–65. https://doi.org/10.1109/TNSRE.2020.3011181.

78. Guo H, Di Y, An X, Wang Z, Ming D. A novel approach to automatic sleep stage classification using forehead electrophysiological signals. Heliyon. 2022;8:e12136–e12136. https://doi.org/10.1016/j.heliyon.2022.e12136.

79. Hasan MJ, Shon D, Im K, Choi HK, Yoo DS, Kim JM. Sleep state classification using power spectral density and residual neural network with multichannel EEG signals. Appl Sci (Switzerland). 2020;10:1–13. https://doi.org/10.3390/app10217639.

80. Hei Y, Yuan T, Fan Z, Yang B, Hu J. Sleep staging classification based on a new parallel fusion method of multiple sources signals. Physiol Meas. 2022. https://doi.org/10.1088/1361-6579/ac647b.

81. Huang W, Guo B, Shen Y, Tang X, Zhang T, Li D, Jiang Z. Sleep staging algorithm based on multichannel data adding and multifeature screening. Comput Methods Prog Biomed. 2020;187:105253. https://doi.org/10.1016/j.cmpb.2019.105253.

82. Huttunen R, Leppanen T, Duce B, Arnardottir ES, Nikkonen S, Myllymaa S, et al. A comparison of signal combinations for deep learning-based simultaneous sleep staging and respiratory event detection. IEEE Trans Biomed Eng. 2022. https://doi.org/10.1109/TBME.2022.3225268.

83. Jadhav P, Mukhopadhyay S. Automated sleep stage scoring using time-frequency spectra convolution neural network. IEEE Trans Instrum Meas. 2022;71:1–9. https://doi.org/10.1109/TIM.2022.3177747.

84. Jaoude MA, Sun H, Pellerin KR, Pavlova M, Sarkis RA, Cash SS, et al. Expert-level automated sleep staging of long-term scalp electroencephalography recordings using deep learning. Sleep. 2020. https://doi.org/10.1093/sleep/zsaa112.

85. Jia Z, Cai X, Zheng G, Wang J, Lin Y. SleepPrintNet: a multivariate multimodal neural network based on physiological time-series for automatic sleep staging. IEEE Trans Artif Intell. 2021;1:248–57. https://doi.org/10.1109/tai.2021.3060350.

86. Jiang D, Ma Y, Wang Y. Sleep stage classification using covariance features of multi-channel physiological signals on Riemannian manifolds. Comput Methods Programs Biomed. 2019;178:19–30. https://doi.org/10.1016/j.cmpb.2019.06.008.

87. Jiang D, Lu Y, Ma Y, Wang Y. Robust sleep stage classification with single-channel EEG signals using multimodal decomposition and HMM-based refinement. Expert Syst Appl. 2019;121:188–203. https://doi.org/10.1016/j.eswa.2018.12.023.

88. Kang DY, DeYoung PN, Malhotra A, Owens RL, Coleman TP. A state space and density estimation framework for sleep staging in obstructive sleep apnea. IEEE Trans Biomed Eng. 2018;65:1201–12. https://doi.org/10.1109/TBME.2017.2702123.

89. Karimzadeh F, Boostani R, Seraj E, Sameni R. A distributed classification procedure for automatic sleep stage scoring based on instantaneous electroencephalogram phase and envelope features. IEEE Trans Neural Syst Rehabil Eng. 2018;26:362–70. https://doi.org/10.1109/TNSRE.2017.2775058.

90. Khalili E, Mohammadzadeh AB. Automatic sleep stage classification using temporal convolutional neural network and new data augmentation technique from raw single-channel EEG. Comput Methods Prog Biomed. 2021;204:106063. https://doi.org/10.1016/j.cmpb.2021.106063.

91. Khojandi A, Shylo O, Zokaeinikoo M. Automatic EEG classification: a path to smart and connected sleep interventions.

Ann Oper Res. 2019;276:169–90. https://doi.org/10.1007/s10479-018-2823-1.

92. Kim D, Lee J, Woo Y, Jeong J, Kim C, Kim DK. Deep learning application to clinical decision support system in sleep stage classification. J Pers Med. 2022. https://doi.org/10.3390/jpm12020136.

93. Kim H, Kim D, Oh J. Automation of classification of sleep stages and estimation of sleep efficiency using actigraphy. Front Public Health. 2022;10:1092222. https://doi.org/10.3389/fpubh.2022.1092222.

94. Kim H, Lee SM, Choi S. Automatic sleep stages classification using multi-level fusion. Biomed Eng Lett. 2022;12:413–20. https://doi.org/10.1007/s13534-022-00244-w.

95. Korkalainen H, Aakko J, Duce B, Kainulainen S, Leino A, Nikkonen S, et al. Deep learning enables sleep staging from photoplethysmogram for patients with suspected sleep apnea. Sleep. 2020. https://doi.org/10.1093/sleep/zsaa098.

96. Korkalainen H, Aakko J, Nikkonen S, Kainulainen S, Leino A, Duce B, et al. Accurate deep learning-based sleep staging in a clinical population with suspected obstructive sleep apnea. IEEE J Biomed Health Inform. 2020;24:2073–81. https://doi.org/10.1109/JBHI.2019.2951346.

97. Kwon K, Kwon S, Yeo WH. Automatic and accurate sleep stage classification via a convolutional deep neural network and nanomembrane electrodes. Biosensors. 2022. https://doi.org/10.3390/bios12030155.

98. Lee CH, Kim HJ, Kim YT, Kim H, Kim JB, Kim DJ. SleepExpertNet: high-performance and class-balanced deep learning approach inspired from the expert neurologists for sleep stage classification. J Ambient Intell Humaniz Comput. 2022. https://doi.org/10.1007/s12652-022-04443-2.

99. Li C, Qi Y, Ding X, Zhao J, Sang T, Lee M. A deep learning method approach for sleep stage classification with EEG spectrogram. Int J Environ Res Public Health. 2022. https://doi.org/10.3390/ijerph19106322.

100. Li M, Chen H, Cheng Z. An attention-guided spatiotemporal graph convolutional network for sleep stage classification. Life. 2022. https://doi.org/10.3390/life12050622.

101. Li QQ, Cakmak AS, Da Poian G, Bliwise DL, Vaccarino V, Shah AJ, Clifford GD. Transfer learning from ECG to PPG for improved sleep staging from wrist-worn wearables. Physiol Meas. 2021;42:44004. https://doi.org/10.1088/1361-6579/abf1b0.

102. Li T, Zhang B, Lv H, Hu S, Xu Z, Tuergong Y. CAttSleepNet: automatic end-to-end sleep staging using attention-based deep neural networks on single-channel EEG. Int J Environ Res Public Health. 2022. https://doi.org/10.3390/ijerph19095199.

103. Li X, Cui L, Tao S, Chen J, Zhang X, Zhang GQ. HyCLASSS: a hybrid classifier for automatic sleep stage scoring. IEEE J Biomed Health Inform. 2018;22:375–85. https://doi.org/10.1109/JBHI.2017.2668993.

104. Li Y, Xu Z, Zhang Y, Cao Z, Chen H. Automatic sleep stage classification based on a two-channel electrooculogram and one-channel electromyogram. Physiol Meas. 2022. https://doi.org/10.1088/1361-6579/ac6bdb.

105. Ling H, Luyuan Y, Xinxin L, Bingliang D. Staging study of single-channel sleep EEG signals based on data augmentation. Front Public Health. 2022;10:1038742. https://doi.org/10.3389/fpubh.2022.1038742.

106. Liu C, Yin Y, Sun Y, Ersoy OK. Multi-scale ResNet and BiGRU automatic sleep staging based on attention mechanism. PLoS ONE. 2022;17:e0269500–e0269500. https://doi.org/10.1371/journal.pone.0269500.

107. Liu C, Tan B, Fu M, Li J, Wang J, Hou F, Yang A. Automatic sleep staging with a single-channel EEG based on ensemble

108. empirical mode decomposition. Physica A. 2021. https://doi.org/10.1016/j.physa.2020.125685.

108. Liu GR, Lo YL, Malik J, Sheu YC, Wu HT. Diffuse to fuse EEG spectra ÔÇô Intrinsic geometry of sleep dynamics for classification. Biomed Signal Process Control. 2020. https://doi.org/10.1016/j.bspc.2019.101576.

109. Liu Z, Luo S, Lu Y, Zhang Y, Jiang L, Xiao H. Extracting multiscale and salient features by MSE based U-structure and CBAM for sleep staging. IEEE Trans Neural Syst Rehabil Eng. 2022. https://doi.org/10.1109/TNSRE.2022.3216111.

110. Lu C, Pathak S, Englebienne G, Seifert C. Channel contribution in deep learning based automatic sleep scoring: How many channels do we need? IEEE Trans Neural Syst Rehabil Eng. 2022. https://doi.org/10.1109/TNSRE.2022.3227040.

111. Luo Y, Li J, He K, Cheuk W. A hierarchical attention-based method for sleep staging using movement and cardiopulmonary signals. IEEE J Biomed Health Inform. 2022. https://doi.org/10.1109/JBHI.2022.3228341.

112. Lv X, Li J, Xu Q. A multilevel temporal context network for sleep stage classification. Comput Intell Neurosci. 2022;2022:6104736. https://doi.org/10.1155/2022/6104736.

113. Matsumori S, Teramoto K, Iyori H, Soda T, Yoshimoto S, Mizutani H. HARU sleep: a deep learning-based sleep scoring system with wearable sheet-type frontal EEG sensors. IEEE Access. 2022;10:13624–32. https://doi.org/10.1109/ACCESS.2022.3146337.

114. Mousavi S, Afghah F, Acharya UR. SleepEEGNet: automated sleep stage scoring with sequence to sequence deep learning approach. PLoS ONE. 2019;14:e0216456. https://doi.org/10.1371/journal.pone.0216456.

115. Neng W, Lu J, Xu L. CCRRSleepNet: a hybrid relational inductive biases network for automatic sleep stage classification on raw single-channel EEG. Brain Sci. 2021. https://doi.org/10.3390/brainsci11040456.

116. Paisarnsrisomsuk S, Ruiz C, Alvarez SA. Capturing the development of internal representations in a high-performing deep network for sleep stage classification. SN Comput Sci. 2021. https://doi.org/10.1007/s42979-021-00697-3.

117. Pan J, Zhang J, Wang F, Liu W, Huang H, Tang W, et al. Automatic sleep staging based on EEG-EOG signals for depression detection. Intell Autom Soft Comput. 2021;28:53–71. https://doi.org/10.32604/iasc.2021.015970.

118. Patanaik A, Ong JL, Gooley JJ, Ancoli-Israel S, Chee MWL. An end-to-end framework for real-time automatic sleep stage classification. Sleep. 2018. https://doi.org/10.1093/sleep/zsy041.

119. Pei W, Li Y, Siuly S, Wen P. A hybrid deep learning scheme for multi-channel sleep stage classification. Comput Mater Continua. 2022;71:889–905. https://doi.org/10.32604/cmc.2022.021830.

120. Perslev M, Darkner S, Kempfner L, Nikolic M, Jennum PJ, Igel C. U-Sleep: resilient high-frequency sleep staging. NPJ Digit Med. 2021. https://doi.org/10.1038/s41746-021-00440-5.

121. Phan H, Andreotti F, Cooray N, Chén OY, de Vos M. Joint classification and prediction CNN framework for automatic sleep stage classification. IEEE Trans Biomed Eng. 2019;66:1285–96. https://doi.org/10.1109/TBME.2018.2872652.

122. Phan H, Andreotti F, Cooray N, Chén OY, de Vos M. SeqSleepNet: end-to-end hierarchical recurrent neural network for sequence-to-sequence automatic sleep staging. IEEE Trans Neural Syst Rehabil Eng. 2019;27:400–10. https://doi.org/10.1109/TNSRE.2019.2896659.

123. Pini N, Ong JL, Yilmaz G, Nicholas IYN, Siting Z, Awasthi A, et al. An automated heart rate-based algorithm for sleep stage classification: Validation using conventional polysomnography and an innovative wearable electrocardiogram device. Front Neurosci. 2022;16:974192. https://doi.org/10.3389/fnins.2022.974192.

124. Qu W, Wang Z, Hong H, Chi Z, Feng DD, Grunstein R, Gordon C. A Residual based attention model for EEG based sleep staging. IEEE J Biomed Health Inform. 2020;24:2833–43. https://doi.org/10.1109/JBHI.2020.2978004.

125. Qureshi S, Karrila S, Vanichayobon S. GACNN SleepTunenet: a genetic algorithm designing the convolutional neural network architecture for optimal classification of sleep stages from a single EEG channel. Turk J Electr Eng Comput Sci. 2019;27:4203–19. https://doi.org/10.3906/ELK-1903-186.

126. Radha M, Fonseca P, Moreau A, Ross M, Cerny A, Anderer P, et al. Sleep stage classification from heart-rate variability using long short-term memory neural networks. Sci Rep. 2019;9:14149. https://doi.org/10.1038/s41598-019-49703-y.

127. Seo H, Back S, Lee S, Park D, Kim T, Lee K. Intra- and interepoch temporal context network (IITNet) using sub-epoch features for automatic sleep scoring on raw single-channel EEG. Biomed Signal Process Control. 2020. https://doi.org/10.1016/j.bspc.2020.102037.

128. Sharma M, Yadav A, Tiwari J, Karabatak M, Yildirim O, Acharya UR. An automated wavelet-based sleep scoring model using EEG, EMG, and EOG signals with more than 8000 subjects. Int J Environ Res Public Health. 2022. https://doi.org/10.3390/ijerph19127176.

129. Shen H, Ran F, Xu M, Guez A, Li A, Guo A. An automatic sleep stage classification algorithm using improved model based essence features. Sensors (Basel). 2020. https://doi.org/10.3390/s20174677.

130. Sokolovsky M, Guerrero F, Paisarnsrisomsuk S, Ruiz C, Alvarez SA. Deep learning for automated feature discovery and classification of sleep stages. IEEE/ACM Trans Comput Biol Bioinf. 2020;17:1835–45. https://doi.org/10.1109/TCBB.2019.2912955.

131. Sors A, Bonnet S, Mirek S, Vercueil L, Payen J-F. A convolutional neural network for sleep stage scoring from raw single-channel EEG. Biomed Signal Process Control. 2018;42:107–14. https://doi.org/10.1016/j.bspc.2017.12.001.

132. Sridhar N, Shoeb A, Stephens P, Kharbouch A, Shimol DB, Burkart J, et al. Deep learning for automated sleep staging using instantaneous heart rate. NPJ Digit Med. 2020;3:106. https://doi.org/10.1038/s41746-020-0291-x.

133. Sun C, Chen C, Fan J, Li W, Zhang Y, Chen W. A hierarchical sequential neural network with feature fusion for sleep staging based on EOG and RR signals. J Neural Eng. 2019. https://doi.org/10.1088/1741-2552/ab39ca.

134. Sun C, Chen C, Li W, Fan J, Chen W. A hierarchical neural network for sleep stage classification based on comprehensive feature learning and multi-flow sequence learning. IEEE J Biomed Health Inform. 2020;24:1351–66. https://doi.org/10.1109/JBHI.2019.2937558.

135. Sun S, Li C, Lv N, Zhang X, Yu Z, Wang H. Attention based convolutional network for automatic sleep stage classification. Biomed Tech (Berl). 2021;66:335–43. https://doi.org/10.1515/bmt-2020-0051.

136. Tao Y, Yang Y, Yang P, Nan F, Zhang Y, Rao Y, Du F. A novel feature relearning method for automatic sleep staging based on single-channel EEG. Complex Intell Syst. 2022. https://doi.org/10.1007/s40747-022-00779-6.

137. Urtnasan E, Park J-U, Joo EY, Lee K-J. Deep convolutional recurrent model for automatic scoring sleep stages based on single-lead ECG signal. Diagnostics (Basel). 2022. https://doi.org/10.3390/diagnostics12051235.

138. Vallat R, Walker MP. An open-source, high-performance tool for automated sleep staging. Elife. 2021. https://doi.org/10.7554/eLife.70092.

139. Vanbuis J, Feuilloy M, Baffet G, Meslier N, Gagnadoux F, Girault J-M. A new sleep staging system for type iii sleep studies

equipped with a tracheal sound sensor. IEEE Trans Biomed Eng. 2022;69:1225–36. https://doi.org/10.1109/TBME.2021.3120927.

140. Wang H, Guo H, Zhang K, Gao L, Zheng J. Automatic sleep staging method of EEG signal based on transfer learning and fusion network. Neurocomputing. 2022;488:183–93. https://doi.org/10.1016/j.neucom.2022.02.049.

141. Wang H, Lu C, Zhang Q, Hu Z, Yuan X, Zhang P, Liu W. A novel sleep staging network based on multi-scale dual attention. Biomed Signal Process Control. 2022. https://doi.org/10.1016/j.bspc.2022.103486.

142. Wei Y, Qi X, Wang H, Liu Z, Wang G, Yan X. A multi-class automatic sleep staging method based on long short-term memory network using single-lead electrocardiogram signals. IEEE Access. 2019;7:85959–70. https://doi.org/10.1109/ACCESS.2019.2924980.

143. Wongsirichot T, Elz N, Kajkamhaeng S, Nupinit W, Sridonthong N. An investigation of data mining based Automatic Sleep Stage Classification techniques. Int J Mach Learn Comput. 2019;9:520–6. https://doi.org/10.18178/ijmlc.2019.9.4.835.

144. Yan R, Li F, Zhou DD, Ristaniemi T, Cong F. Automatic sleep scoring: a deep learning architecture for multi-modality time series. J Neurosci Methods. 2021;348:108971. https://doi.org/10.1016/j.jneumeth.2020.108971.

145. You Y, Zhong X, Liu G, Yang Z. Automatic sleep stage classification: A light and efficient deep neural network model based on time, frequency and fractional Fourier transform domain features. Artif Intell Med. 2022;127:102279. https://doi.org/10.1016/j.artmed.2022.102279.

146. Yu R, Zhou Z, Wu S, Gao X, Bin G. MRASleepNet: a multi-resolution attention network for sleep stage classification using single-channel EEG. J Neural Eng. 2022. https://doi.org/10.1088/1741-2552/aca2de.

147. Yulita IN, Fanany MI, Arymurthy AM. Fast convolutional method for automatic sleep stage classification. Healthc Inform Res. 2018;24:170–8. https://doi.org/10.4258/hir.2018.24.3.170.

148. Zhang B, Yang Z, Cai H, Lian J, Chang W, Zhang Z. Ontology-based decision support tool for automatic sleep staging using dual-channel EEG data. Symmetry-Basel. 2020. https://doi.org/10.3390/sym12111921.

149. Zhang C, Yu W, Li Y, Sun H, Zhang Y, de Vos M. CMS2-Net: semi-supervised sleep staging for diverse obstructive sleep apnea severity. IEEE J Biomed Health Inform. 2022;26:3447–57. https://doi.org/10.1109/JBHI.2022.3156585.

150. Zhang J, Wu Y. Competition convolutional neural network for sleep stage classification. Biomed Signal Process Control. 2021. https://doi.org/10.1016/j.bspc.2020.102318.

151. Zhang X, Xu M, Li Y, Su M, Xu Z, Wang C, et al. Automated multi-model deep neural network for sleep stage scoring with unfiltered clinical data. Sleep Breath. 2020;24:581–90. https://doi.org/10.1007/s11325-019-02008-w.

152. Zhang X, Kou W, Chang EIC, Gao H, Fan Y, Xu Y. Sleep stage classification based on multi-level feature learning and recurrent neural networks via wearable device. Comput Biol Med. 2018;103:71–81. https://doi.org/10.1016/j.compbiomed.2018.10.010.

153. Zhong Q, Lei H, Chen Q, Zhou G. A sleep stage classification algorithm of wearable system based on multiscale residual convolutional neural network. J Sens. 2021. https://doi.org/10.1155/2021/8222721.

154. Zhou D, Wang J, Hu G, Zhang J, Li F, Yan R, et al. SingleChannelNet: a model for automatic sleep stage classification with raw single-channel EEG. Biomed Signal Process Control. 2022. https://doi.org/10.1016/j.bspc.2022.103592.

155. Zhou J, Wang G, Liu J, Wu D, Xu W, Wang Z, et al. Automatic sleep stage classification with single channel EEG signal based on two-layer stacked ensemble model. IEEE Access. 2020;8:57283–97. https://doi.org/10.1109/ACCESS.2020.2982434.