Sleeping Disorder Diagnosis Methods – A Systematic Review

M. Nirubama[1], N. Arivazhagan[1,*]

[1]CINTEL Department, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu, India

Abstract

Significant advancements in the identification of sleep disorders, such as obstructive sleep apnea and insomnia, cardiovascular illnesses, diabetes, and mental health issues, have resulted from the combination of machine learning and deep learning approaches. These techniques leverage physiological signals and patient data to provide automated, accurate, and efficient diagnostic tools. Comparing these sophisticated computational techniques to conventional diagnostic procedures, there is potential for significant gains in efficiency and accuracy. This review examines current research (2021–2024) on applying ML and DL to diagnose distinct sleep disorders. It highlights approaches, datasets for important comparisons, performance measures, outcomes, potential future directions, and gaps in the field. With the incorporation of new technologies, the diagnosis of mental health illnesses, cardiovascular diseases, diabetes, and sleep disorders including obstructive sleep apnea and insomnia has changed dramatically.

Keywords

Deep learning, Explainable AI, Sleeping disorder

1. Introduction

Sleep disorders such obstructive sleep apnea , insomnia, restless leg syndrome, narcolepsy, and comorbid insomnia and sleep apnea , which can cause major health problems like neurological, metabolic, and cardiovascular problems. Early detection and accurate diagnosis are essential for managing and treating patients well. Polysomnography and clinical as- sessments, the conventional technique of identifying these diseases, are labor-intensive, costly, time-consuming, and prone to human error. By using massive datasets and complex algorithms to find patterns suggestive of different sleep disorders, recent developments in machine learning and deep learning provide potential options for the effective, scalable solutions and accurate diagnosis of sleep problem.Obstructive sleep apnea , insomnia, narcolepsy, restless legs syndrome, and concomitant PhysioNet ECG Sleep Apnea v1.0.0 dataset for sleep apnea detection. Achieved performance with the highest accuracy of 88.13%.SVM, logistic regression, Gaussian naïve Bayes, discriminate analyses, nearest neighbor, decision tree, random forest, Ada- Boost, gradient boosting, MLP, recurrent networks , and hybrid convolutional-recurrent networks are used with majority voting. This method is more complex implementation [1].

For the detection of insomnia, the heart rate variability of ECG Signal Power spectral density. LDA classifier achieves the finest insomnia detection accuracy with 99.0%. Fine-tuned and evaluated by the free public PhysioNet dataset over fivefold trails cross-validation. This method is not generalized for different data set [2].

2. Diagnosis Methods

This section discusses various methods of sleeping disorders.

BISEC'2024: 15th International Conference on Business Information Security, November 28-29, 2024, Niš, Serbia
*Corresponding author.
nm1203@srmist.edu.in (M. Nirubama); arivazhn@srmist.edu.in (N. Arivazhagan)
0009-0004-3850-2564 (M. Nirubama); 0000 0001 9515 4237 (N. Arivazhagan)

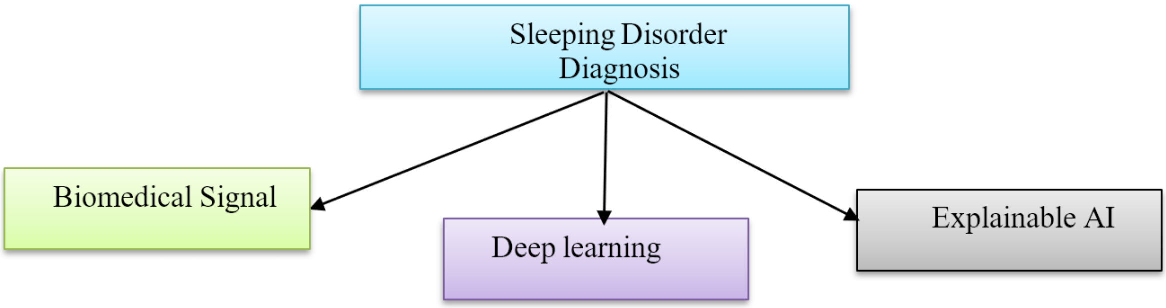

Figure 1: Types of Sleeping Disorder diagnosis methods.

2.1. Biomedical Signal Processing Methods

Detecting bio-signals based-sleep stages .on Model Agnostic Meta-Learning achieved 5.4% to 17.7% range upgrading with statistical difference in the mean. This method iis Computational Overhead. Sensitivity to Hyperparameters. This method is limited when tasks are too dissimilar or when the task distribution changes drastically [3].

Detection of Insomnia presented with electroencephalogram (EEG) Demonstrated Gaussian mixturehidden Markov model (GMM-HMM) achieves 86 accuracy. This method is having less accuracy [4].

Detection of Respiratory disturbances during sleep.A bed-integrated radio-frequency sensor through near-field coherent sensing was applied. Apneic event detection attained a sensitivity up to 88.6% and 89.0% for k-fold validation. For this method more controlled environment is required [5].

Hybrid neural network with semi-supervised learning at the same time sleep arousal and sleep stage finding with features of single-channel electroencephalography. On the Physio2018 dataset achieves an overall accuracy of 0.78. This method is having less accuracy [6].

1,111 characteristics were produced after several criteria were suggested in the literature. The 112 worthies tones for automated sleep grading were given by the actometer, respiratory inductance plethysmography belts, pulse oximeter, PneaVoX sensor (which records tracheal sounds), nasal cannula, and respiratory inductance plethysmography belts. The system gets substantial agreement with manual scoring for classifications into two stages (wake vs. sleep: mean Cohen's Kappa κ of 0.63 and accuracy rate Acc of 87.8%) and three stages (wake vs. R stage vs. NREM stage: mean κ of 0.60 and Acc of 78.5%) [7]. Finally, implemented a three-step model, consisting of category using a multi-layer perceptron, sleep transition rules correction, and sequence corrections using a Viterbi hidden Markov model.

Improved cardiovascular OSA phenotyping is required to rank treatment of high-risk individuals. Methods: SpO 2 records from 1987 overnight polysomnography are included in the study. Of these, 974 come from patients who may have OSA, 931 from the Sleep Heart Health Study, which is based on data from the general community, and 83 from healthy controls. For every oxygen desaturation, the amplitude ratio of desaturation over resaturation, the minimum SpO 2 value, and the SpO 2 upslope are retrieved and averaged per patient. Findings: The mICS performs 2.7% better when the SpO 2 parameters are included together with age and BMI. This results in a test area under the curve of 69.5% for the identification of any cardiovascular comorbidity [8]. Although wearable sensor technology has advanced dramatically over the past ten years, the absence of large and representative datasets concurrently obtained with polysomnography (PSG) limits its clinical utility for the evaluation of obstructive sleep apnea. Methods: respiratory effort and electrocardiogram data were used to create an artificial neural network that would identify instances of sleep disturbed breathing. Findings: four-class sleep staging distinguished between waking, combined N1-N2, N3, and REM with a κ of 0.69 compared to PSG. AHI estimate performed well in terms of diagnosis for various OSA severity thresholds, with an intraclass correlation value of 0.91 [9].

Although research on respiratory and metabolic issues has been the focus of central sleep apnea (CSA), the neuronal dysfunction that underlies central sleep apnea is still largely understood. Here, using hypnograms to analyze the sleep-wake dynamics, we explore the underlying neural mechanism of central sleep apnea. Techniques: We reviewed the sleep records of seven subject groups: adults without CSA ($n = 25$), adults with CSA ($n = 29$), adults with obstructive sleep apnea (OSA) ($n = 28$), strong children ($n = 40$), children with OSA ($n = 18$), children with CSA ($n = 73$), and children with CSA treated with CPAP ($n = 10$). We have discovered that, in differ to the scale-invariant (i.e., power-law) distribution that has been documented for stimulations in healthy sleep, the sleep arousals of CSA patients exhibit a distinctive temporal scale (i.e., exponential distribution) [10].

The paper analyzes consumer sleep technology such as wearable sensors, bed sensors, smartphone applications, ambient room sensors, and artificial intelligence, as well as sleep lab technologies like polysomnography. The study also classifies the various learning approaches and gives an overview of many clinical datasets for sleep staging. In conclusion, the paper provides our perspectives and suggestions on the utilization of the examined sleep technologies [11].

2.2. Deep Learning Methods

Photoplethysmography (PPG) time series data used with residual convolutional network achieved median Cohen's Kappa (κ) score of 0.75 compare to 0.69 for existing method. This method is having less accuracy [12, 13]. Single-channel EEG recording with long short-term memory along with convolutional neural network for healthy-unhealthy, and disease grouping with an accuracy of 91.45% and 90.55%.This method is having less accuracy [14].

Pressure-sensor-based smart mattress to realize sleep status finding and quality evaluation. CNN model for four various sleep postures archives accuracy of up to 96.987%.In this method Cross-validation using medical data is not evaluated [15].

Deep learning-based sleep staging was used to detect sleep phases by assessing the hypothesis, overlap 30-second epochs with 15-, 5-, 1-, or 0.5-second epoch-to-epoch duration. With a period of one second between epochs, the hazard ratio, which indicates the risk of fragmented sleep, was 1.14 ($p = 0.39$) for mild OSA, 1.59 ($p < 0.01$) for moderate OSA, and 4.13 ($p < 0.01$) for severe OSA. The findings show that, in order to properly diagnose sleep problems, a more thorough examination of sleep architecture is required [16].

Using wrist-worn consumer sleep technology (CST), categorization and detection of sleep apnea (SA) is a deep transfer learning strategy for sleep stage. Methods: The model is based on a deep convolutional neural network (DNN) that has been trained with information from accelerometers and photoplethysmography recordings made at night. Using a hold-out test dataset containing raw data from a wrist-worn CST, an external validation was performed. Using internal datasets that include raw data from clinical and wrist-worn sensors, the DNN was trained and assessed. Findings: Training on clinical data leads to a large improvement in performance, while feature enrichment using a sleep stage stream only slightly improves performance. In CST datasets, raw data input performs better than feature-based input. When comparing wearable device data to clinical data, the system performs marginally worse, although it still generalizes well [17, 18].

A new network called SwSleepNet is suggested that is capable of accurately offline sleep staging as well as online sleep stage prediction and calibration. In order to balance the network's operational efficiency and comprehensive feature extraction, For offline analysis, the sequence consolidating module (SCM), squeeze and excitation (SE) block, sequential CNN (SCNN), and sequence broadening module (SBM) are coordinated by the suggested network. In the context of online analysis, the only models used to predict the sleep state within a brief video clip are SCNN and SE.Two publicly accessible datasets, The Sleep-EDF and the Montreal Archive of Sleep Studies Huashan Hospital Fudan University (HSFU), as well as one clinical dataset, have been expanded. are used to validate SwSleepNet's performance. The result shows that SwSleepNet outperforms state-of-the-art ways with offline accuracy of 84.5%, 86.7%, and 81.8%, respectively [19].

To develop an accurate deep learning technique for the automatic classification of sleep phases and to look into how the severity of OSA affects classification accuracy. Two distinct datasets' worth of nightly polysomnographic recordings were used to build a mixed convolutional and long short-term memory neural network: one from a clinical dataset ($n = 891$) of patients with suspected OSA, and the other from a public dataset of healthy persons (Sleep-EDF, $n = 153$). The model obtained an accuracy of 83.7% ($\kappa = 0.77$) in sleep staging on the public dataset using a single frontal EEG channel, and 83.9% ($\kappa = 0.78$) when augmented with EOG. The model's accuracy for the clinical dataset was 82.9% ($\kappa = 0.77$) for a single EEG channel and 83.8% ($\kappa = 0.78$) for two channels (EEG+EOG) [20]. A deep learning model was contructed to score respiratory events and sleep phases at the same time. Pulse oximetry data alone should be sufficient to accomplish the scoring and subsequent AHI computation, according to the hypothesis. Methods: The deep learning models were trained using 877 polysomnography recordings of people who may have had OSA. Three distinct input signal combinations were used to train the same architecture: Photoplethysmogram and oxygen saturation (SpO2) were included in model 1; PPG, SpO2, and nasal pressure were included in model 2; and respiratory belts, electroencephalogram, nasal pressure, SpO2, and oronasal thermocouple were included in model 3. Results: Model 1 performed comparably to models 2 and 3 in terms of REM- AHI and AHI estimation as well as REM-AHI [21, 22].

2.3. Explainable AI in Sleep Diagnosis

Using optical, differential air pressure, and acceleration readings from a chest-worn sensor, five somnographic-like signals are generated and fed into a deep network. To predict three patterns related to breathing (normal, apnea, irregular), three patterns related to sleep (normal, snoring, loud), and the overall signal quality (normal, corrupted),this solves a three-fold classification issue. Saliency maps and confidence indices are two examples of qualitative and quantitative information that the created architecture provides to enhance explainability and aid in prediction interpretation. The accuracy of breathing rhythms was higher (0.93) than that of sleep patterns (0.76). Compared to apnea (0.97), Using optical, differential air pressure, and acceleration readings from a chest-worn sensor, five somnographic-like signals are generated and fed into a deep network. To predict three patterns related to breathing (normal, apnea, irregular), three patterns related to sleep (normal, snoring, loud), and the overall signal quality (normal, corrupted), Consequently, this is a step in the direction of gradually closer clinical translation of the usage of AI-based techniques for sleep problem detection [23].

3. Challenges and Future Directions

Although ML and DL provide important improvements in the diagnosis of sleep disor- ders, there are still a number of difficulties. These include the difficulty of integrating various physiological signals, the requirement for sizable, annotated datasets for effi- cient model training, and the assurance of model interpretability and clinical accepta- bility. In order to enable continuous monitoring and early action, future research is probably going to concentrate on creating more reliable, understandable AI models and incorporating these systems into wearable technology.

1. Data Availability: To ensure the resilience of these models across various populations, larger and more diverse datasets are required for both training and validation.
2. Integration into Clinical Practice: Further investigation is necessary to optimize the incorporation of these sophisticated models into standard clinical procedures, tackling concerns pertaining to interpretability of the models and their real-time implementation.
3. Cross-Disorder Applicability: The majority of current research focuses on certain illnesses, such as sleep apnea. Increasing the application's scope to cover more sleep disorders might improve the overall effect.
4. Generalization Across Diverse Populations: The generalizability of the models across various demographics and settings is impacted by dataset variety, which limits the majority of investigations. Integration of Multimodal Data: While combining multichannel physiological signals has increased accuracy, further study is required to successfully integrate various data sources, including patient health records and wearable technology.

5. Generalization Across Diverse Populations: The majority of research is constrained by dataset variety, which impacts the models' applicability to various situations and demography . Integration of Multimodal Data: Although combining multichannel physiological signals has increased accuracy, further study is required to success- fully integrate various data sources, including wearable technology and medical records.
6. Real-World Implementation: There are several obstacles to overcome when moving from research to clinical practice, such as the requirement for thorough validation in real-world situations, patient privacy issues, and model interpretability.
7. Dataset Diversity: A lot of research uses narrow datasets, which restricts how broadly applicable models may be. To increase the robustness of the model, different, multi-ethnic datasets are required.
8. Real-time Monitoring: Many of the current models do not have real-time diagnostic features.Future studies ought to concentrate on creating wearable, real-time diagnostic instruments.
9. Interpretability: The black-box character of many ML and DL models makes clinical interpretation difficult. Improving the interpretability of the model will be essential to its clinical acceptance.
10. Longitudinal Studies: Studies evaluating the long-term efficacy of ML and DL models in clinical contexts are few. Longitudinal research is required to verify these models in the long run.

4. Conclusion

A potential area of medical technology is the combination of deep learning and machine learning for the identification of sleep problems. An important development in sleep medicine is the use of ML and DL in the diagnosis of sleep disorders. These cutting-edge techniques offer effective, precise, and less invasive substitutes for conventional diagnostic techniques, possibly revolutionizing the identification, treatment, and successful integration of these technologies into clinical practice. These technologies are anticipated to advance in sophistication, accessibility, and widespread clinical adoption as research continues. These technologies will be further improved by ongoing research that focuses on various datasets, multimodal integration, and useful deployment, ultimately leading to better patient outcomes and more effective healthcare delivery. However, in order to integrate these technical advances into clinical practice, it is imperative that the identified research gaps be addressed.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] M. Bahrami, M. Forouzanfar, Sleep apnea detection from single-lead ecg: A comprehensive analysis of machine learning and deep learning algorithms, IEEE Transactions on Instrumentation and Measurement 71 (2022) 1–11.
[2] P. Tripathi, M. Ansari, T. K. Gandhi, R. Mehrotra, M. B. B. Heyat, F. Akhtar, C. C. Ukwuoma, A. Y. Muaad, Y. M. Kadah, M. A. Al-Antari, et al., Ensemble computational intelligent for insomnia sleep stage detection via the sleep ecg signal, IEEE Access 10 (2022) 108710–108721.
[3] N. Banluesombatkul, P. Ouppaphan, P. Leelaarporn, P. Lakhan, B. Chaitusaney, N. Jaimchariyatam, E. Chuangsuwanich, W. Chen, H. Phan, N. Dilokthanakul, et al., Metasleeplearner: A pilot study on fast adaptation of bio-signals-based sleep stage classifier to new individual subject using meta-learning, IEEE Journal of Biomedical and Health Informatics 25 (2020) 1949–1963.
[4] J. Huang, L. Ren, L. Feng, F. Yang, L. Yang, K. Yan, Ai empowered virtual reality integrated systems for sleep stage classification and quality enhancement, IEEE Transactions on Neural Systems and Rehabilitation Engineering 30 (2022) 1494–1503.

[5] Z. Zhang, T. B. Conroy, A. C. Krieger, E. C. Kan, Detection and prediction of sleep disorders by covert bed-integrated rf sensors, IEEE Transactions on Biomedical Engineering 70 (2022) 1208–1218.

[6] H. Liu, H. Zhang, B. Li, X. Yu, Y. Zhang, T. Penzel, Msleepnet: A semi-supervision based multi-view hybrid neural network for simultaneous sleep arousal and sleep stage detection, IEEE Transactions on Instrumentation and Measurement (2024).

[7] J. Vanbuis, M. Feuilloy, G. Baffet, N. Meslier, F. Gagnadoux, J.-M. Girault, A new sleep staging system for type iii sleep studies equipped with a tracheal sound sensor, IEEE Transactions on Biomedical Engineering 69 (2021) 1225–1236.

[8] M. Deviaene, P. Borzée, M. Van Gilst, J. van Dijk, S. Overeem, B. Buyse, D. Testelmans, S. Van Huffel, C. Varon, Multilevel interval coded scoring to assess the cardiovascular status of sleep apnea patients using oxygen saturation markers, IEEE Transactions on Biomedical Engineering 67 (2020) 2839–2848.

[9] P. Fonseca, M. Ross, A. Cerny, P. Anderer, F. Schipper, A. Grassi, M. van Gilst, S. Overeem, Estimating the severity of obstructive sleep apnea using ecg, respiratory effort and neural networks, IEEE Journal of Biomedical and Health Informatics (2024).

[10] H. Dvir, S. Guo, S. Havlin, N. Xin, T. Jun, D. Li, X. Zhifei, R. Kang, R. P. Bartsch, Central sleep apnea alters neuronal excitability and increases the randomness in sleep-wake transitions, IEEE Transactions on Biomedical Engineering 67 (2020) 3185–3194.

[11] G. Cay, V. Ravichandran, S. Sadhu, A. H. Zisk, A. L. Salisbury, D. Solanki, K. Mankodiya, Recent advancement in sleep technologies: A literature review on clinical standards, sensors, apps, and ai methods, IEEE Access 10 (2022) 104737–104756.

[12] K. Kotzen, P. H. Charlton, S. Salabi, L. Amar, A. Landesberg, J. A. Behar, Sleepppg-net: A deep learning algorithm for robust sleep staging from continuous photoplethysmography, IEEE Journal of Biomedical and Health Informatics 27 (2022) 924–932.

[13] D. Marzorati, A. Dorizza, D. Bovio, C. Salito, L. Mainardi, P. Cerveri, Hybrid convolutional networks for end-to-end event detection in concurrent ppg and pcg signals affected by motion artifacts, IEEE Transactions on Biomedical Engineering 69 (2022) 2512–2523.

[14] A. Wadichar, S. Murarka, D. Shah, A. Bhurane, M. Sharma, H. S. Mir, U. R. Acharya, A hierarchical approach for the diagnosis of sleep disorders using convolutional recurrent neural network, IEEE Access (2023).

[15] L.-J. Kau, M.-Y. Wang, H. Zhou, Pressure-sensor-based sleep status and quality evaluation system, IEEE Sensors Journal 23 (2023) 9739–9754.

[16] H. Korkalainen, T. Leppänen, B. Duce, S. Kainulainen, J. Aakko, A. Leino, L. Kalevo, I. O. Afara, S. Myllymaa, J. Töyräs, Detailed assessment of sleep architecture with deep learning and shorter epoch-to-epoch duration reveals sleep fragmentation of patients with obstructive sleep apnea, IEEE journal of biomedical and health informatics 25 (2020) 2567–2574.

[17] M. Olsen, J. M. Zeitzer, R. N. Richardson, V. H. Musgrave, H. B. Sørensen, E. Mignot, P. J. Jennum, A deep transfer learning approach for sleep stage classification and sleep apnea detection using wrist-worn consumer sleep technologies, IEEE Transactions on Biomedical Engineering (2024).

[18] K. McClure, B. Erdreich, J. H. Bates, R. S. McGinnis, A. Masquelin, S. Wshah, Classification and detection of breathing patterns with wearable sensors and deep learning, Sensors 20 (2020) 6481.

[19] H. Zhu, Y. Wu, Y. Guo, C. Fu, F. Shu, H. Yu, W. Chen, C. Chen, Towards real-time sleep stage prediction and online calibration based on architecturally switchable deep learning models, IEEE Journal of Biomedical and Health Informatics 28 (2023) 470–481.

[20] H. Korkalainen, J. Aakko, S. Nikkonen, S. Kainulainen, A. Leino, B. Duce, I. O. Afara, S. Myllymaa, J. Töyräs, T. Leppänen, Accurate deep learning-based sleep staging in a clinical population with suspected obstructive sleep apnea, IEEE journal of biomedical and health informatics 24 (2019) 2073–2081.

[21] R. Huttunen, T. Leppänen, B. Duce, E. S. Arnardottir, S. Nikkonen, S. Myllymaa, J. Töyräs, H. Korkalainen, A comparison of signal combinations for deep learning-based simultaneous sleep staging and respiratory event detection, IEEE Transactions on Biomedical Engineering 70 (2022)

1704–1714.

[22] R. Lazazzera, M. Deviaene, C. Varon, B. Buyse, D. Testelmans, P. Laguna, E. Gil, G. Carrault, Detection and classification of sleep apnea and hypopnea using ppg and spo _2 signals, IEEE Transactions on Biomedical Engineering 68 (2020) 1496–1506.

[23] M. Rossi, D. Sala, D. Bovio, C. Salito, G. Alessandrelli, C. Lombardi, L. Mainardi, P. Cerveri, Sleep-see-through: Explainable deep learning for sleep event detection and quantification from wearable somnography, IEEE journal of biomedical and health informatics 27 (2023) 3129–3140.