



A COMPREHENSIVE REVIEW ON SLEEP APNEA DETECTION USING HYBRID MACHINE LEARNING AND DEEP LEARNING MODELS

Anmol Kaur

Department of Computer Science and Engineering
Punjab University
Patiala, Punjab, India

Nirvair Neeru

Department of Computer Science and Engineering
Punjab University
Patiala, Punjab, India

Abstract: Sleep apnea is a prevalent and serious sleep disorder characterized by repeated interruptions in breathing during sleep, increasing risks of cardiovascular disease, cognitive impairment, and reduced quality of life. The disorder occurs in three main forms: obstructive, central, and complex sleep apnea, with obstructive sleep apnea (OSA) being the most prevalent. Common risk factors include obesity, aging, smoking, alcohol use, and genetic predisposition. Traditionally, diagnosis relies on polysomnography, which is accurate but costly and time-intensive. To address these limitations, recent studies have applied machine learning (ML) and deep learning (DL) techniques for automated detection and classification of sleep apnea. Hybrid and ensemble models, such as Gradient Boosting and CatBoost, have shown promising results with classification accuracies exceeding 97%. These advancements suggest that AI-driven approaches can offer scalable, cost-effective alternatives to conventional methods, supporting early detection and improved management of sleep apnea in both clinical and home settings.

Keywords: Sleep Disorders, Obstructive Sleep Apnea (OSA), Machine Learning (ML), Deep Learning (DL), Random Forest, Automated Diagnosis

I. INTRODUCTION

Sleep Apnea is a major sleep disorder which can be characterized due to repetitive breaks in breathing while sleeping. Basically, it is a sleep-related problem where the airway becomes partially or completely blocked during sleep. The common symptoms of this disorder are that if a person snores loudly and feels tired even after sufficient sleep. A person not only feels groggy during the day, but it also increases the chance the chances of Cardiac Arrest too [1].

The three main categories of Sleep Apnea are as under:

- 1) **Obstructive Sleep Apnea (OSA):** OSA is the predominant type of sleep apnea seen in patients. During OSA, relaxation of the throat muscles can block the airway, and affected individuals frequently experience loud snoring.
- 2) **Central Sleep Apnea (CSA):** In CSA, the brain tends to fail to send the actual signals that are helpful in controlling breathing.
- 3) **Complex (Mixed) Sleep Apnea:** It is a coalition of both OSA and CSA.

Obesity is the biggest risk factor for sleep apnea because excess weight can increase fat deposits around the upper airway, making it more likely to collapse during sleep. Additionally, it is seen that apnea tends to be observed more in older adults and men, though it can affect people of all ages and genders. Other factors which are responsible for increasing risk factors include genetic history, tobacco, consumption of alcohol, and certain anatomical features like a thick neck or narrow airway [2].

Severity levels of Sleep Apnea

The severity level of Sleep Apnea is determined with the help of apnea-hypopnea index (AHI), which is conducted by sleep

medical professionals. The AHI evaluates the average number of apnea events (complete pauses in breathing) and hypopnea events (partial reductions in airflow) that occur per hour of sleep. Basically, sleep apnea severity level is divided into three groups, which are [3][4][5]:

- Mild sleep apnea: 5 to 14 apnea or hypopnea events per hour.
- Moderate sleep apnea: 15 to 29 apnea or hypopnea events per hour.
- Severe sleep apnea: 30 or more apnea or hypopnea events per hour.

It has been observed that in the United States, out of five women, one woman is and nearly one third of men have sleep apnea. Sometimes it is very difficult to predict whether the patient is suffering from OSA or CA because of the overlap of apnea symptoms [6].

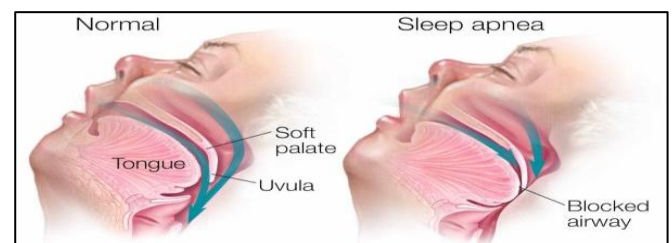


Fig. 1: Illustrates the difference between normal breathing and sleep apnea

The following are the most common symptoms shown in patients who are affected by sleep apnea [7]:

- Dryness in the mouth during morning awakening
- Night sweats
- Excessive daytime fatigue
- Waking up with a headache

- Heartburn
- Loud and uninterrupted snoring
- Weak concentration
- Gasping for air during sleep

II. WHY SLEEP APNEA IS A SERIOUS DISEASE ACCORDING TO THE WORLD HEALTH ORGANIZATION (WHO)?

The WHO acknowledges that Sleep Apnea is one of the notable problems rising these days. It is considered a serious disorder because of the discontinuous and non-regenerative sleep, which leads to severe health issues like anxiety, diabetes insipidus, and heart ailments. The poor quality of sleep increases various kinds of risks, such as impaired vigilance, poor concentration at work, feeling lethargic during office hours, and road accidents [8].

The elaboration of these risk factors is explained as under:

1. **Poor Sleep Quality:** Frequent awakenings prevent restful, restorative sleep, leading to chronic fatigue and daytime sleepiness.
2. **Cardiovascular Risks:** Sleep apnea patients face a higher risk of hypertension, heart attacks, and irregular heart rhythms.
3. **Metabolic Issues:** It is linked to insulin resistance and type 2 diabetes.
4. **Cognitive and Mood Impairments:** Sleep deprivation affects concentration, memory and can cause irritability or depression.
5. **Increased Accident Risk:** Daytime drowsiness increases the risk of motor vehicle and workplace accidents.
6. **Complications in Surgery and Medication:** Sleep apnea can complicate anesthesia and recovery from surgery.

To get rid of these challenges WHO encourages refinements in life lifestyle of the people, awareness among the public, and routine checkups to detect and treat the ailments [9].

WHO's Perspective on Addressing the Problem

- **Public Awareness and Screening:** WHO draws attention to individuals to detect the initial indications. People are encouraged to undergo well-timed screening so that suitable treatment can be taken to prevent the bigger risks of life.
- **Avoidable peril factors:** There is a special focus given on enhancing the healthy lifestyle of people by the World Health Organization. To curtail the possibility of the disorder, masses are inspired to do physical activity, lessen the consumption of alcohol and smoking [9] [10].

III. LITERATURE WORK

With the integration of machine learning (ML) and deep learning (DL) techniques, significant advancements have been made in the detection of sleep disorders. This section

reviews the literature and various studies that have aimed to automate the detection of sleep disorders using these approaches. By examining prior research in this field, we hope to identify the most advanced methodologies, emphasise key findings, and lay the groundwork for our own investigation.

Padovano *et al.* (2025) [11] proposed an automatic detection framework for Obstructive Sleep Apnea (OSA) that leverages deep learning with recurrence information from heart rate variability (HRV). Unlike traditional approaches relying on hand-crafted recurrence quantification analysis (RQA) features, the authors transformed HRV into distance matrix images and applied Convolutional Neural Networks (CNNs), including a pre-trained AlexNet model. Evaluations on three public databases (Apnea-ECG, MIT-BIH, and UCD Sleep Apnea) demonstrated that CNNs trained on distance matrix representations achieved superior performance ($\approx 75\%$ accuracy, sensitivity, and specificity) compared to classical machine learning classifiers and wavelet-based scalograms. The study highlights the potential of recurrence-based image representations combined with CNNs to improve OSA detection, while also emphasizing the need for further optimization to enhance clinical applicability.

A. Taher *et al.* (2024) [12] used the Sleep Health and Lifestyle Dataset and evaluated multiple Machine Learning algorithms to detect sleep disorders. In this research, data processing is performed using Label Encoding, and Normalization is applied for min-max scaling, followed by Parameter tuning to enhance accuracy. It is concluded that Gradient Boosting performs best, with an accuracy of 93.80%, compared to Random Forest (90.26%), Logistic Regression (91.15%), and AdaBoost (91.15%).

M. Kolhar *et al.* (2025) [13] focused on utilizing ECG signals to frame a dual-branch convolutional neural network (CNN) for the early detection of obstructive sleep apnea (OSA). The major drawbacks of the traditional polysomnography (PSG) are eliminated using the above technique. Moreover, it is observed that both CNN and dual-branch CNN models are able to find OSA. Both models attain test and validation accuracy of approximately 93%. ROC curves, confusion matrices, and class-wise precision/recall are calculated to evaluate the performance. Research Findings show that the dual-branch CNNs are more efficacious in classifying apnea.

A. Javeed *et al.* (2023) [14] conducted a study that helps to identify the major causes, such as obesity, heart attacks, and arrhythmias, that cause Sleep Apnea. This research comprises two modules, which are XGBoost and Bidirectional Long Short-Term Memory Networks (BiLSTM). The proposed methodology achieves 97% accuracy. In addition, the performance of the model is evaluated against that of traditional long-short-term memory networks (LSTM). It is concluded that by identifying risk factors performance of the model is enhanced.

M. A. Rahman *et al.* (2025) [15] proposed two methodologies using the Sleep Health and Lifestyle Dataset. In first method, the Mean Decrease Impurity (MDI) technique is used, in which five prominent features are selected using a Gradient Boosting Regressor. On the other hand, in the second procedure, by keeping the same methodology, only two core components are chosen. Researchers used 15 machine

learning classifiers, in which the maximum accuracy of 97.33% is achieved by Gradient Boosting, Voting, CatBoost, and Stacking. Out of all classifiers, Gradient Boosting reached to highest AUC of 0.9953. As compared to the other mentioned classifiers, Gradient Boosting is 20.16 times faster. On the other hand, the Decision Tree achieved 96% accuracy and demonstrated notable computational efficiency. Therefore, it has been concluded that Gradient Boosting is the fastest and most capable amongst other classifiers.

Fakhim Babaei *et al.* (2025) [16] designed a multimodal framework utilizing physiological signals to examine sleep apnea using deep learning. Spatial-temporal patterns are being captured by merging both the framework named convolutional neural networks (CNN) and gated recurrent units (GRU). To strengthen the features, integration of a multi-domain extractor (MDFE) with a squeeze-and-excitation (SE) block is done. In this study, the Sleep Heart Health Study (SHHS) dataset is used, and the proposed methodology is able to gain Accuracy and an AUC of 83.87% and 91.94, respectively.

B. Halder *et al.* (2023) [17] developed a multi-resolution deep neural network model with temporal and channel attention was presented for detecting the cyclic alternating pattern (CAP) of electroencephalogram (EEG) activity during sleep. The goal was to identify and track sleep disorders and inadequate sleep. The model utilised a multi-branch one-dimensional convolutional neural network (1D-CNN) with varying kernel sizes to automatically extract features of varied frequency resolutions. A network of attention-based transformers was utilised to capture the dynamic and temporal relationship between CAP event characteristics. In A-phase detection, the model outperformed recent approaches with an accuracy of 90.31%, a specificity of 95.30%, and an F1-Score of 65.73 percent. The model detected A-phase subtypes with an accuracy of 86.72%, a specificity of 89.53%, and an F1-Score of 59.59%.

Saini *et al.* (2024) [18] developed a machine learning framework for the identification of sleep disorders, aiming to

improve diagnostic accuracy and reduce reliance on traditional, time-consuming methods such as polysomnography. The study applied multiple classifiers to clinical and behavioral datasets, evaluating their performance in distinguishing sleep disorder cases from normal sleep patterns. Experimental results demonstrated that the proposed approach achieved high accuracy and reliability, highlighting the potential of machine learning as a practical tool for early screening and diagnosis of sleep disorders. The work emphasizes computational efficiency and the ability of ML models to capture complex patterns in sleep-related data, thereby supporting faster and more accessible diagnostic solutions.

Shah *et al.* (2024) [19] evaluated various machine learning models by employing EEG signals. To fetch the relevant features, the data signals are segmented into frames and subframes. Various classifiers such as Naive Bayes, Random Forest, K-Nearest Neighbors, and Decision Tree are assessed. In this work, it has been concluded that amongst all the classifiers, Naive Bayes has the highest performance. Their results show that Naive Bayes performs best, particularly in the Alpha band and C3–A2 EEG channel, achieving the highest accuracy among the tested models.

Tareq (2024) [20] employed the Sleep Health and Lifestyle Dataset, which contains demographic, lifestyle, and sleep-related features. Researchers have implemented multiple models, which include Random Forests, Support Vector Machines, and Decision Trees. Among these, Random Forest shows the better performance by giving its results with an accuracy of 88%.

Analysis

Based on the literature work, a comparison has been done to analyse the work of the machine and deep learning models for different sleep disorders and datasets, along with their outcomes and limitations in Table 1.

Table 1: Literature Review

Study	Method / Model	Dataset	Best Performance	Key Contributions	Limitations
Padovano et al. (2025) [11]	CNNs (custom & AlexNet) on recurrence images (distance matrices of HRV) vs. ML (SVM, RF, etc.)	Apnea-ECG, MIT-BIH, UCD Sleep Apnea DB	Accuracy≈75%	Used distance matrices instead of handcrafted RQA features; external validation improved generalization	Accuracy moderate; still not clinically deployable
Taher & Ayon (2024) [12]	Gradient Boosting, RF, Logistic Regression, AdaBoost	Sleep Health & Lifestyle dataset (~375 samples, 13 features)	Gradient Boosting: 93.8% Acc	Showed lifestyle/health data can effectively classify sleep disorders; identified BMI, BP, stress as key predictors	Small dataset; limited generalizability
Kolhar et al. (2025) [13]	Dual-branch CNN on ECG; compared with ML classifiers	ECG dataset (not specified in	≈93% Acc (dual-branch CNN)	Novel CNN architecture capturing local +	Generalization to noisy / diverse

		detail; OSA annotated)		global ECG features; used SMOTE to balance classes	ECG signals remains open
Javeed et al. (2023) [14]	Hybrid XGBoost + BiLSTM using electronic health records	EHR dataset (clinical health data)	Reported high predictive accuracy (noted in study; >90%)	Combined temporal sequence learning (BiLSTM) with feature boosting (XGBoost)	Needs validation across larger & diverse EHR datasets
Rahman et al. (2025) [15]	Optimized ML approaches (various)	Clinical sleep disorder datasets	Gradient Boosting, Voting, CatBoost, and Stacking approaches all reached the same classification accuracy level of 97.33%	Systematic optimization of ML for sleep disorder diagnosis.	Limited physiological feature depth; may lack generalizability.
Fakhim Babaei et al. (2025) [16]	Multimodal DL (CNN + GRU + Attention, custom loss function)	SHHS dataset	AUC 91.94, Accuracy 83.87	Introduced MDFE + SE blocks with novel loss function for robust apnea detection.	Computationally complex; requires high resources.
Halder et al. (2023) [17]	One dimensional convolutional neural network	Real time dataset	Accuracy = 90.31% Specificity = 95.30% F1 score= 65.73%	Novel architecture with attention for cyclic alternating pattern detection.	Focused on CAP (not general apnea); limited to single EEG channel.
Saini et al. (2024) [18]	Machine Learning classifiers	Sleep Health & Lifestyle dataset	.SVM, RF, and LR models have an accuracy of 90.99%, 90.29, and 87.61% respectively.	Applied ML to general sleep disorder classification.	No detailed metrics; lacks deep model comparison.
Shah et al. (2024) [19]	ML models on EEG	ISRUC-Sleep dataset	Accuracy ~87%	Demonstrated EEG-only detection using ML.	Small dataset; subject variability issues.
Tareq et al. (2024) [20]	Random forest classifier	Sleep Health & Lifestyle dataset	Accuracy = 88%	pre-assembled data is used	Small dataset

IV. RESEARCH FINDINGS

Recent studies have applied a wide spectrum of machine learning (ML) and deep learning (DL) methods that aimed at increasing the accuracy and efficiency of sleep disorder detection, particularly sleep apnea.

Padovano et al. [11] use a combination of deep learning with recurrence information analysis to detect obstructive sleep apnea automatically, demonstrating the effectiveness of recurrence-based features in improving classification accuracy. Similarly, Kolhar et al. [13] proposed an AI-driven dual-branch CNN combined with ML models, which captured both temporal and spatial features from physiological signals, enhancing apnea event detection. Javeed et al. [14] introduced a hybrid XGBoost-BiLSTM model using electronic health data, which outperformed conventional LSTM approaches in predictive power and computational efficiency.

In a broader context, Taher and Ayon [12] assessed numerous of machine learning classifiers using sleep health and lifestyle datasets, identifying Gradient Boosting as the most efficient method, achieving high classification accuracy across apnea, insomnia, and normal sleep classes. Rahman et al. [15] further optimized ML pipelines for general sleep disorder diagnosis, highlighting improved diagnostic accuracy by selecting critical features and fine-tuning models.

Beyond conventional ML, several works leveraged novel deep learning strategies. Babaei et al. [16] designed a multimodal deep learning framework with a custom loss function, significantly boosting feature diversity and apnea detection robustness. Halder et al. [17] focused on cyclic alternating patterns in EEG. Shah et al. [19] also utilized EEG signals for apnea detection via ML models, validating EEG's role as a strong biomarker.

Other studies explored alternative signal modalities. Saini et al. [18] employed ML models for sleep disorder

classification, emphasizing the importance of accessible and cost-effective tools.

Similarly, Tareq [10] employed the same dataset, which included demographic, lifestyle, and sleep-related attributes, to compare models such as Random Forests, Support Vector Machines, and Decision Trees. Among these, the Random Forest classifier outperformed other approaches, achieving an accuracy of 88% due to its ability to capture non-linear relationships and reduce overfitting. Together, these studies underscore the potential of machine learning techniques in leveraging lifestyle and health-related data for accurate, scalable, and cost-effective sleep disorder detection.

Overall, the collective findings emphasize three major trends:

1. Hybrid and ensemble methods (e.g., XGBoost-BiLSTM, Gradient Boosting) consistently outperform single models.
2. Multimodal and recurrence-based feature extraction enhances robustness and generalization across diverse datasets.
3. Accessible physiological and behavioral data (sleep sound, SpO₂, lifestyle attributes) are emerging as alternatives to traditional PSG/EEG, paving the way for more scalable and cost-efficient diagnostic solutions.

V. CONCLUSION

Sleep apnea remains a critical health concern due to its strong association with cardiovascular and metabolic risks. While polysomnography is the diagnostic gold standard, its limitations underscore the need for scalable alternatives. Recent studies show that ML and DL methods, particularly ensemble models such as Gradient Boosting and CatBoost, achieve high classification accuracy and offer practical solutions for automated detection. These advances highlight the potential of AI-driven approaches to improve early diagnosis, reduce healthcare costs, and enhance patient outcomes.

VI. FUTURE WORK

Future research should focus on explainable AI models to enhance clinical interpretability and trust. Integrating multimodal data, such as wearable sensors and lifestyle information, could improve diagnostic accuracy. Additionally, large-scale validation studies are needed to ensure generalizability across diverse populations. Real-time monitoring through portable devices may further enable cost-effective and continuous assessment of sleep apnea.

References

- [1] J. Yayan and K. Rasche, "A Systematic Review of Risk factors for Sleep Apnea," *Preventive Medicine Reports*, vol. 42, p. 102750, 2024
- [2] R. J. Henning and W. M. Anderson, "Sleep apnea is a common and dangerous cardiovascular risk factor," *Current Problems in Cardiology*, vol. 50, Issue no. 1, p. 100214, Jan. 2025.
- [3] A. E. Bluher, T. Kearney, and T. Vazifedan, "Vitamin D deficiency and pediatric obstructive sleep apnea severity," *JAMA Otolaryngology–Head & Neck Surgery*, vol. 151, no. 1, pp. 72–77, Jan. 2025.
- [4] Y. Wang, W. Yue, B. Zhou, J. Zhang, Y. He, M. Wang, and K. Hu, "The Hourly Apnea-Hypopnea Duration Better Correlates with OSA-Related Nocturnal Hypoxemia and Excessive Daytime Sleepiness Rather Than AHI," *Nat. Sci. Sleep*, vol. 17, pp. 1101–1112, May 2025.
- [5] P. Sharma, S. Thakur, D. K. Rai, S. Karmakar, and A. Harishkumar, "Connecting the dots: analysing the relationship between AHI and ODI in obstructive sleep apnea," *Sleep Sci. Pract.*, vol. 8, no. 1, p. 9, 2024.
- [6] A. C. Quan et al., "The Public Health Burden of Obstructive Sleep Apnea," *Sleep Medicine Reviews*, vol. 63, pp. 101643, 2023.
- [7] D. M. Chervin et al., "Projecting the 30-Year Burden of Obstructive Sleep Apnea," *Sleep Health*, vol. 11, no. 4, pp. 545–555, 2025.
- [8] American Academy of Sleep Medicine, "Rising Prevalence of Sleep Apnea in the U.S. Threatens Public Health," 2024.
- [9] M. Benjafield et al., "The Global Burden of Obstructive Sleep Apnea: Executive Summary," *Diagnostics*, vol. 15, no. 9, p. 1088, 2025.
- [10] National Sleep Foundation, "Public Health Impact of Sleep Apnea is Severe," 2024. D. Padovano, A. Martinez-Rodrigo, J. M. Pastor, J. J. Rieta, and R. Alcaraz, "Deep learning and recurrence information analysis for the automatic detection of obstructive sleep apnea," *Applied Sciences*, vol. 15, no. 1, p. 123, Jan. 2025.
- [11] A. Taher and W. I. Z. Ayon, "Exploring sleep disorders: A comparative analysis of machine learning algorithms on sleep health and lifestyle data," in *Proc. IEEE Int. Conf. Power, Electrical, Electronics and Industrial Applications (PEEIACON)*, Rajshahi, Bangladesh, Sept. 2024, pp. 71–76.
- [12] M. Kolhar, M. M. Alfridan, and R. A. Siraj, "AI-driven detection of obstructive sleep apnea using dual-branch CNN and machine learning models," *Biomedicines*, vol. 13, no. 5, p. 1090, May 2025.
- [13] A. Javeed, J. S. Berglund, A. L. Dallora, M. A. Saleem, and P. Anderberg, "Predictive power of XGBoost_BiLSTM model: A machine-learning approach for accurate sleep apnea detection using electronic health data," *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, art. no. 188, 2023.
- [14] M. A. Rahman, I. Jahan, M. Islam, T. Javid, M. S. Ali, M. R. A. Rashid, M. Manzurul Islam, and M. H. Ferdous, "Improving Sleep Disorder Diagnosis Through Optimized Machine Learning Approaches," *IEEE Access*, vol. 13, pp. 20 989–21 004, 2025.
- [15] A. Fakhim Babaei, J. Tanha, M. A. Balafar, and S. Roshan, "A Novel Multimodal Deep Learning Approach with Loss Function for Detection of Sleep Apnea Events," *IEEE Access*, vol. 13, pp. 52 085–52 099, 2025.
- [16] B. Halder, T. Anjum, and M. I. H. Bhuiyan, "An attention-based multi-resolution deep learning model for automatic A-phase detection of cyclic alternating pattern

- in sleep using single-channel EEG,” *Biomed. Signal Process. Control*, vol. 85, p. 104934, Jan. 2023.
- [17] S. K. Saini, O. T. Jemima, A. Shayamano, and G. Chandel, “Identification of Sleep Disorder Using Machine Learning,” in *Proc. IEEE Int. Conf. on Futuristic Computing and Research (ICFCR)*, 2024.
- [18] R. Shah, D. Gaur, and G. Premalatha, “Machine Learning Based Sleep Apnea Detection Using EEG Signals,” in *Proc. Int. Conf. on Electronics and Renewable Systems (ICEARS)*, Mar. 2024, pp. 1593–1599.
- [19] Tareq, Wadhah Zeyad Tareq. "Sleep Disorders Detection and Classification Using Random Forests Algorithm." *Decision Making in Healthcare Systems*. Cham: Springer International Publishing, 2024. 257-266