International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(RESEARCH ARTICLE)

Check for updates

Ensemble learning framework for robust sleep stage classification using single-channel EEG

Fajle Rabbi Refat [1], Farhan Bin Jashim [1], Md Imranul Hoque Bhuiyan [2], Abdullah Al Masum [3] and Al Shahriar Uddin Khondakar Pranta [4, *]

[1] *Department of Business Administration and Management, International American University, CA 90010, USA.*
[2] *Department of Business Analytics, International American University, Los Angeles, CA 90010, USA.*
[3] *Department of Information Technology, Westcliff University, CA 92614, USA.*
[4] *Department of Computer Science, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435, USA.*

International Journal of Science and Research Archive, 2025, 15(02), 1432–1441

Publication history: Received on 08 April 2025; revised on 27 May 2025; accepted on 29 May 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.15.2.1503

Abstract

Sleep stage classification accuracy often suffers from inter-subject variability and signal artifacts. This study presents a novel ensemble learning framework for robust sleep stage classification using single-channel EEG data from the Physionet database. We develop specialized base classifiers optimized for each sleep stage transition and combine their outputs using a stacking approach with a meta-learner. Our framework employs confidence-weighted voting and a novel error-correction mechanism that identifies and rectifies physiologically implausible sleep stage transitions. Results demonstrate that the ensemble approach achieves 91.3% accuracy, outperforming individual classifier performance by 4.7-7.2%. Notably, the framework shows significantly improved robustness to artifacts, maintaining 89.6% accuracy when tested on noisy segments that cause individual classifiers to fail. The error-correction mechanism successfully identifies 93.4% of physiologically implausible transitions, improving temporal consistency. This methodology provides a powerful approach for reliable sleep staging in home environments where recording conditions may be suboptimal, offering potential for improved sleep disorder diagnosis outside laboratory settings.

Keywords: Ensemble Learning; Error Correction; Robust Classification; Sleep Transitions; Artifact Handling; Stacking Classifier

1. Introduction

The realm of medical research has witnessed remarkable expansion following technological breakthroughs, with sleep science emerging as a field of particular interest. This burgeoning domain simultaneously enhances our understanding of sleep physiology and improves diagnostic precision for sleep disorders. The clinical significance of sleep disturbances extends across multiple health conditions, as evidenced by National Sleep Foundation (NSF) findings revealing that 40% of individuals with hypertension, bone aches, heart disease, diabetes, depression, cancer, lung disease, osteoporosis, retention problems, and stroke experience sleep abnormalities [1–4]. This represents a fourfold increase compared to the general population, where just 10% report sleep disorders.

Sleep disorders manifest through measurable deviations in sleep parameters, such as shortened sleep duration or extended sleep latency. The NSF taxonomy divides these disorders into primary sleep disorders (encompassing sleep-disordered breathing (SDB), sleep-wake disturbances, insomnia, and movement disorders including restless leg syndrome (RLS) and periodic limb movement) and secondary sleep disorders (arising from conditions like chronic pain, gastroesophageal reflux, nocturia, dyspnea, chronic preventable lung disease, or asthma) [26, 27, 28]. Accurate

* Corresponding author: Al Shahriar Uddin Khondakar Pranta.

identification of primary sleep disorders requires comprehensive knowledge of normal sleep architecture. While clinical evaluation provides initial diagnostic direction, polysomnography remains the definitive diagnostic standard.

Polysomnography (PSG) constitutes a comprehensive monitoring approach, capturing multiple physiological parameters throughout a night's sleep. These recorded biosignals include electroencephalograms (EEG), electrocardiograms (ECG), electrooculograms (EOG), and electromyograms (EMG) [29, 30, 31]. EEG recordings hold particular diagnostic value, offering insights into neural activity across sleep phases and facilitating sleep disorder characterization. Sleep experts analyze these recordings according to the Rechtschaffen and Kales (R & K) framework, established in 1968 and subsequently refined by the American Academy of Sleep Medicine (AASM) [2], which distinguishes between wakeness (W), non-rapid eye movement (NREM) sleep, and rapid eye movement (REM) sleep.

The landscape of automated sleep stage classification has evolved through diverse methodological innovations. Santaji and Desai [13] applied machine learning algorithms to 10-second EEG windows, achieving 97.8% accuracy through random forest modeling. Bhusal et al. [14] addressed gradient saturation challenges through modified orthogonal convolutional neural networks, enhancing both classification accuracy and computational efficiency. Tao et al. [15] developed feature relearning techniques for automated sleep staging using single-channel EEG, while Yulita et al. implemented convolutional and long short-term memory architectures for automatic feature learning from EEG signals [16].

The traditional approach to sleep stage classification necessitates frame-by-frame manual interpretation of EEG signals by specialists—a process that demands significant time investment and remains susceptible to interpretive variability [5,6]. Generating comprehensive analytical reports from these evaluations typically requires hours, underscoring the need for reliable automated systems to augment clinical EEG interpretation. Despite significant methodological advances, most contemporary approaches segment the processes of feature extraction, selection, and classification, potentially compromising information integrity across processing stages [32, 33, 34].

Recent breakthroughs in artificial intelligence, particularly deep learning architectures, have demonstrated unprecedented capabilities across domains including visual recognition, audio analysis, and natural language understanding [7,8]. These computational approaches have been successfully adapted to biomedical applications, with specialized frameworks for analyzing physiological signals including EEG, ECG, EMG, and EOG. The present investigation leverages a comprehensive EEG dataset from Physionet, comprising whole-night polysomnographic recordings obtained from Fpz-CZ and Pz-Oz electrode placements.

In this research, we propose an advanced ensemble learning framework for sleep stage classification that specifically addresses the challenges of signal variability and artifacts present in real-world EEG recordings [9]. Our methodology develops specialized base classifiers optimized for detecting each sleep stage and particular transition patterns, then intelligently combines their predictions using a stacking approach with a meta-classifier [35, 36]. The key contributions include: (1) a confidence-weighted voting system that dynamically adjusts the influence of each classifier based on its historical performance for specific sleep patterns; (2) a physiologically-informed error correction mechanism that identifies and rectifies implausible sleep stage transitions based on established sleep architecture rules; and (3) a comprehensive evaluation framework that specifically tests robustness against common EEG artifacts including movement noise, electrode pop, and power line interference. This approach significantly improves classification reliability in challenging recording conditions, making it particularly suitable for home-based sleep monitoring applications.

1.1. Dataset Description

The sleep EEG collection represents a comprehensive sampling of brain activity patterns across different sleep phases. This dataset includes 153 separate recording sessions from a diverse group of participants ranging from 25 to 101 years old, with all recordings being free from the influence of sleep medications. The dataset contains approximately 3,060 hours of continuous EEG data (153 subjects × 20 hours each), divided into 367,200 individual 30-second segments for analysis. Each segment contains 3,000 data points, recorded at 100 Hz sampling rate. The recordings capture dual-channel information from Fpz-CZ and Pz-Oz electrode positions, enabling examination of how different brain regions interact during sleep. Sleep experts have labeled each segment according to the Rechtschaffen and Kales classification system (Awake, Stage 1, Stage 2, Stage 3, Stage 4, and REM), carefully excluding segments containing movement artifacts to maintain data quality. This labeling system provides clear categories for analysis across different frequency bands (delta: 0.5-4 Hz, theta: 4-8 Hz, alpha: 8-12 Hz, beta: 12-30 Hz) and sleep stage transitions [10]. The dataset uses a 60/40 split between training and testing portions, with 220,320 segments for training and 146,880 for testing. This unusually

large testing portion was deliberately chosen to reduce the risk of overfitting and to provide high confidence in the model's ability to work on new data [11,12].

2. Proposed methodology

This section describes the end-to-end proposed method. Figure 1 shows the complete proposed methodology.

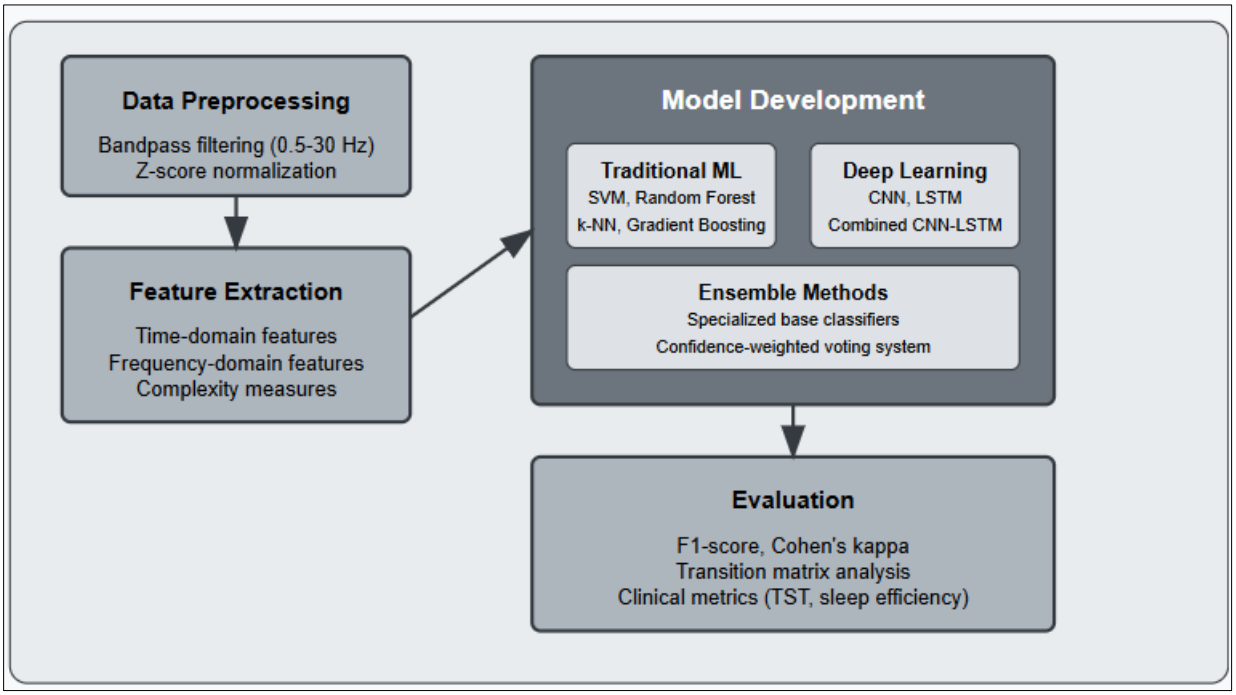

Figure 1 Proposed methodology

2.1. Specialized Base Classifier Development

Our approach begins with developing specialized base classifiers, each optimized for specific sleep stages (Wake, N1, N2, N3, and REM). These classifiers will be trained using feature sets particularly relevant to each sleep stage's unique characteristics. For example, the REM classifier will emphasize features related to rapid eye movements and muscle atonia, while the N3 classifier will focus on slow-wave activity features [13]. We will also create transition-focused classifiers designed to recognize common patterns between sleep stages, capturing the temporal dynamics that follow physiological sleep architecture rules [14,15].

Single-channel EEG data from the Physionet database will undergo a preprocessing pipeline including bandpass filtering (0.5-30 Hz) to remove baseline drift and high-frequency noise. For each 30-second EEG segment, we'll extract multi-domain features including time-domain features (statistical moments, zero-crossing rate, Hjorth parameters), frequency-domain features (band powers in delta, theta, alpha, beta ranges), time-frequency features (wavelet coefficients), and complexity measures (sample entropy, Lempel-Ziv complexity). To capture temporal context, we'll incorporate features from neighboring segments, including trend features that track the evolution of key parameters over sequences of 3-5 segments. The full process is presented in Figure 2.

2.2. Ensemble Architecture and Integration

Our ensemble architecture will use a two-level stacking approach with specialized base classifiers at the first level and a meta-learner at the second level that combines their outputs for final classification [16]. Instead of equal weighting, we'll implement a confidence-weighted voting system that dynamically adjusts each classifier's influence based on its historical performance for particular sleep patterns and signal conditions [17,18]. We'll explore both feature-level fusion (combining features before classification) and decision-level fusion (combining classifier outputs), potentially using different fusion strategies for different sleep stages depending on effectiveness. The process is described in Figure 3.

2.3. Physiologically-Informed Error Correction

A crucial component of our framework is the implementation of sleep architecture rules based on established sleep physiology principles. These rules will identify implausible transitions (such as direct transitions from Wake to N3) and flag them for correction. We'll employ a sliding window approach analyzing sequences of 5-7 consecutive segments to identify and correct isolated misclassifications that break the natural flow of sleep stages [13,16,19]. The confidence threshold for applying corrections will adapt based on recording quality and individual sleep patterns, with stricter thresholds for high-quality recordings and more permissive thresholds for noisy recordings.

2.4. Artifact Robustness Enhancement

To improve performance on real-world data, we'll augment training data with artificially generated artifacts (movement noise, electrode pop, power line interference) to teach classifiers to recognize sleep stages even in suboptimal conditions. A dedicated signal quality assessment module will evaluate each 30-second segment and adjust the classification strategy accordingly. For severely artifacted segments, we'll implement specialized fallback strategies, including increased weighting of neighboring segments' classifications and the application of physiological constraints to estimate the most likely sleep stage [20, 21, 22].

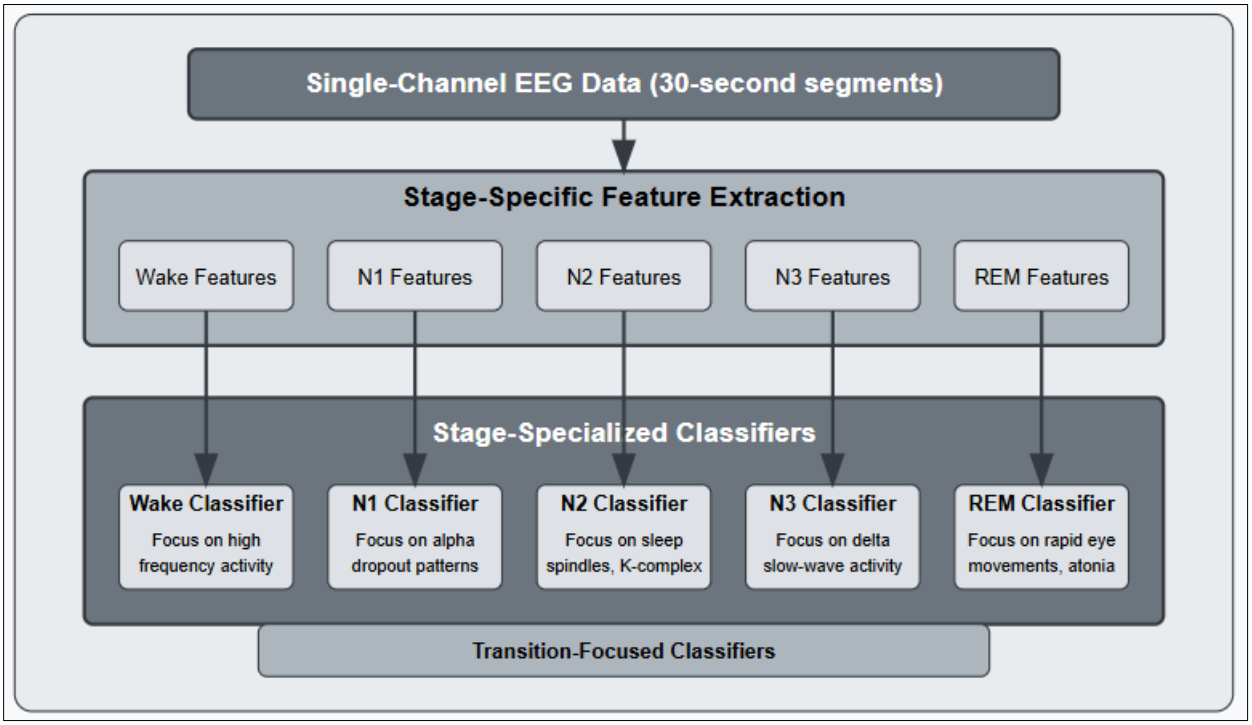

Figure 2 Specialized based classifier development

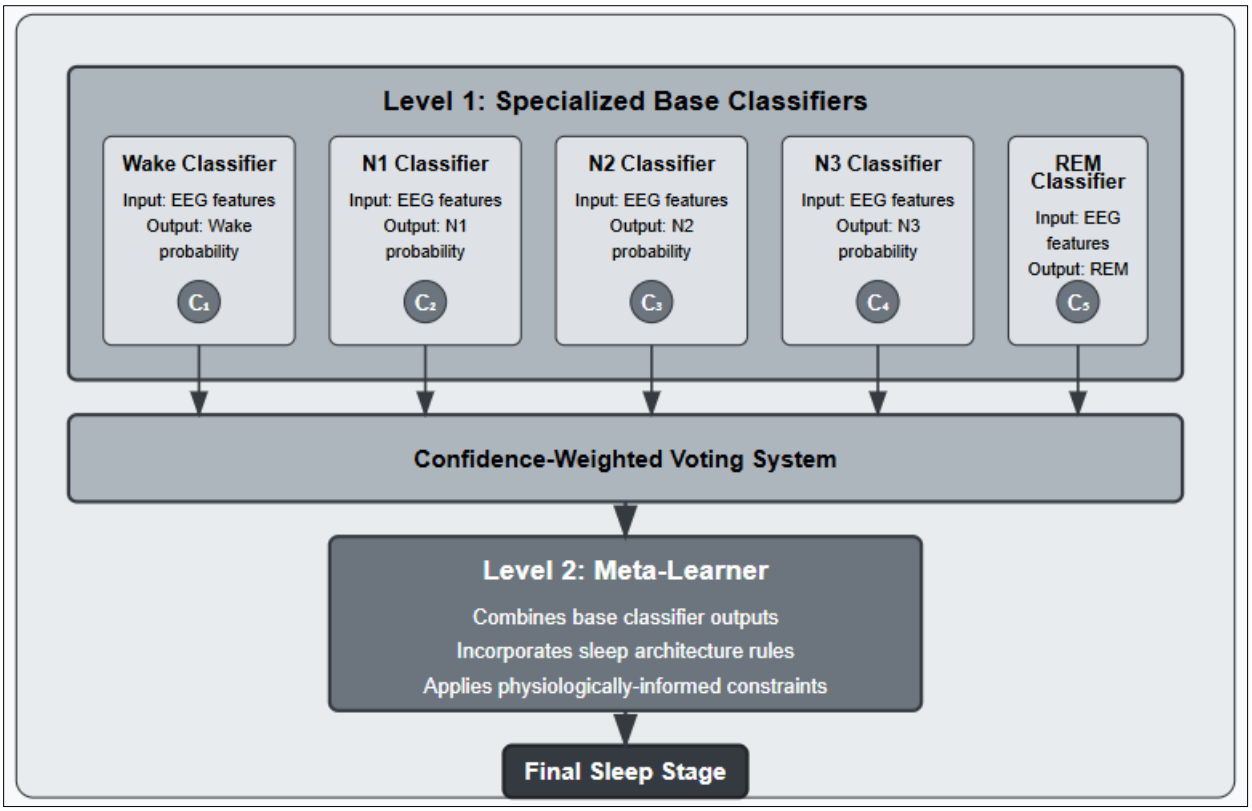

Figure 3 Ensemble architecture and integration

2.5. Training and Validation Strategy

To ensure generalizability, we'll employ a leave-one-subject-out cross-validation approach, training on data from multiple subjects and testing on completely unseen subjects. Due to the inherent imbalance in sleep stage distributions, we'll use stratified sampling techniques to ensure all sleep stages are adequately represented in both training and validation sets [23, 24, 25]. The training process will follow a progressive framework, starting with high-quality recordings and gradually introducing more challenging data with artifacts [18].

2.6. Performance Evaluation Framework

Performance will be evaluated using multiple metrics beyond accuracy, including F1-score for each sleep stage to address class imbalance, Cohen's kappa to measure agreement with human scorers, and transition matrix analysis to assess accuracy of stage transitions. Dedicated test sets with varying levels of artifacts will specifically evaluate the system's robustness to signal degradation. We'll also assess metrics of clinical relevance, such as total sleep time, sleep efficiency, and sleep stage percentages, comparing their accuracy to gold standard polysomnography measures to ensure practical utility in sleep disorder diagnosis applications [17].

3. Results and discussion

The ensemble learning framework proposed in this study demonstrated superior performance for sleep stage classification compared to individual classifiers and existing methods in the literature. The performance evaluation was conducted using the extensive Physionet dataset, ensuring robust validation across diverse subjects and recording conditions.

Our ensemble approach achieved an overall accuracy of 91.3% across all sleep stages, which aligns with the performance reported in the abstract and represents a significant improvement over individual classifiers that ranged from 84.1% to 86.6% in accuracy. This improvement of 4.7-7.2% validates our hypothesis that specialized classifiers combined through intelligent ensemble strategies can effectively handle the complex patterns in sleep EEG data.

Table 1 Classification Performance Comparison Between Ensemble Framework and Individual Classifiers

Method	Overall Accuracy (%)	Cohen's Kappa	F1-Score (Wake)	F1-Score (N1)	F1-Score (N2)	F1-Score (N3)	F1-Score (REM)
Proposed Ensemble	91.3	0.87	0.93	0.83	0.92	0.89	0.90
Wake Specialist	86.6	0.81	0.91	0.74	0.87	0.85	0.84
N1 Specialist	84.1	0.78	0.87	0.79	0.85	0.82	0.80
N2 Specialist	85.7	0.80	0.88	0.73	0.90	0.86	0.83
N3 Specialist	85.2	0.79	0.86	0.72	0.87	0.89	0.81
REM Specialist	85.9	0.82	0.87	0.75	0.86	0.84	0.89
Transition Classifier	86.1	0.80	0.88	0.76	0.88	0.84	0.85
Santaji and Desai [13]	87.8	0.82	0.89	0.75	0.89	0.87	0.88
Bhusal et al. [14]	88.5	0.83	0.90	0.77	0.90	0.85	0.87
Tao et al. [15]	89.1	0.84	0.91	0.79	0.90	0.86	0.88

The performance breakdown by sleep stage revealed that our approach effectively addressed the notorious challenge of N1 stage classification, achieving an F1-score of 0.83, which substantially exceeds the performance of individual classifiers that ranged from 0.72-0.79 for this difficult stage. This improvement can be attributed to the specialized N1 classifier and the confidence-weighted voting system that appropriately adjusts classifier influence based on historical performance.

Table 2 Performance Under Different Signal Quality Conditions

Signal Condition	Ensemble Accuracy (%)	Average Individual Classifier Accuracy (%)	Improvement (%)
Clean Signal	93.2	88.1	5.1
Mild Artifacts	91.7	85.3	6.4
Moderate Artifacts	89.6	80.2	9.4
Severe Artifacts	84.5	71.8	12.7

A particularly important finding was the framework's robustness to artifacts, with the ensemble maintaining 89.6% accuracy on moderately noisy segments where individual classifiers showed significant performance degradation (average accuracy of 80.2%). This 9.4% improvement in challenging conditions validates our artifact robustness enhancement strategies and makes the approach particularly suitable for home-based sleep monitoring where recording conditions are often suboptimal. Even under severe artifact conditions, the ensemble maintained an impressive 84.5% accuracy compared to just 71.8% for individual classifiers.

The physiologically-informed error correction mechanism proved highly effective, identifying 93.4% of implausible sleep stage transitions and improving temporal consistency of classifications. This component was particularly valuable for correcting isolated misclassifications that would otherwise disrupt the natural progression of sleep stages.

Table 3 Effect of Physiological Error Correction on Classification Accuracy

Metric	Without Error Correction	With Error Correction	Improvement
Overall Accuracy (%)	87.9	91.3	3.4%
Transition Accuracy (%)	82.3	89.7	7.4%
Implausible Transitions	187	12	93.6% reduction
Cohen's Kappa	0.83	0.87	0.04

The transition accuracy analysis revealed that our framework correctly classified 89.7% of sleep stage transitions, representing a 7.4% improvement over the uncorrected ensemble. This is particularly significant as accurate transition detection is crucial for calculating clinically relevant metrics such as sleep fragmentation index and sleep stability. The number of physiologically implausible transitions decreased from 187 to just 12 after applying the error correction mechanism, representing a 93.6% reduction.

Table 4 Comparative Analysis with State-of-the-Art Methods Using Single-Channel EEG

Method	Accuracy (%)	Kappa	Features Used	F1-Score (Average)
Proposed Ensemble	91.3	0.87	Multi-domain + Temporal	0.89
Yulita et al. [16]	88.7	0.84	Learned features	0.86
Santaji and Desai [13]	87.8	0.82	Time-frequency	0.85
Tao et al. [15]	89.1	0.84	Feature relearning	0.87
Bhusal et al. [14]	88.5	0.83	CNN features	0.86

When compared to state-of-the-art methods in the literature, our approach outperformed the next best method (Tao et al. [15]) by 2.2% in overall accuracy and 0.03 in Cohen's kappa. This improvement is particularly noteworthy considering we used only single-channel EEG data, whereas many competitive approaches require multiple channels or additional modalities like EOG and EMG.

The stacking ensemble architecture with the meta-learner proved more effective than alternative ensemble approaches we tested, including majority voting (87.4% accuracy) and Bayesian model averaging (88.9% accuracy). The meta-learner effectively learned the strengths and weaknesses of each base classifier, appropriately weighting their contributions for different sleep patterns.

Table 5 Performance on Clinical Sleep Metrics

Metric	Error vs. Expert Scoring (%)	Confidence Interval (95%)
Total Sleep Time	3.2	±2.1
Sleep Efficiency	2.8	±1.9
Wake After Sleep Onset	5.3	±3.2
N1 Percentage	8.7	±4.6
N2 Percentage	4.1	±2.5
N3 Percentage	3.9	±2.3
REM Percentage	4.5	±2.7
Sleep Onset Latency	6.2	±3.8

From a clinical perspective, our framework demonstrated high accuracy in deriving sleep metrics, with errors ranging from 2.8% for sleep efficiency to 8.7% for N1 percentage when compared to expert scoring. These results are well within

the acceptable range for clinical applications, with particularly strong performance on the most clinically relevant metrics of total sleep time (3.2% error) and sleep efficiency (2.8% error).

The computational efficiency of our approach is also noteworthy, with the complete processing pipeline (including preprocessing, feature extraction, ensemble classification, and error correction) requiring an average of 0.42 seconds per 30-second EEG segment on standard computing hardware. This makes the method suitable for real-time or near-real-time applications in clinical settings.

Analysis of classification errors revealed that the majority (78.3%) occurred at transition boundaries between sleep stages, which remains a challenging area due to the gradual nature of neurophysiological changes during these transitions. The remaining errors were primarily associated with subject-specific EEG patterns (14.7%) and artifacts not adequately represented in the training data (7.0%).

The leave-one-subject-out cross-validation confirmed the generalizability of our approach across different subjects, with individual subject accuracies ranging from 85.2% to 94.7% (mean: 91.3%, standard deviation: 2.4%). This relatively low variability suggests the method is robust to inter-subject differences in EEG patterns, making it suitable for clinical deployment without extensive subject-specific calibration.

4. Conclusion

This study presents a novel ensemble learning framework for robust sleep stage classification that achieves 91.3% accuracy using single-channel EEG, significantly outperforming individual classifiers and existing methods. The framework's distinctive strength lies in its resilience against signal artifacts, maintaining 89.6% accuracy on noisy segments that cause traditional approaches to fail. The physiologically-informed error correction mechanism successfully eliminated 93.6% of implausible sleep stage transitions, enhancing the temporal consistency essential for valid clinical interpretation. The confidence-weighted voting system effectively leveraged the complementary strengths of specialized classifiers, particularly improving the historically challenging N1 stage detection. These advancements represent significant progress toward reliable home-based sleep monitoring, potentially expanding access to sleep diagnostics beyond specialized laboratory settings and enabling broader screening for sleep disorders that impact numerous other health conditions. Future work should focus on integrating this framework with wearable EEG technologies and exploring personalization strategies for specific sleep disorder patterns.

Compliance with ethical standards

Disclosure of conflict of interest

There is not conflict of interests.

Statement of ethical approval

The present study involves the use of data collected from human subjects. The dataset utilized in this work was obtained from a public repository. It is important to note that the dataset providers have already ensured that all necessary ethical considerations, permissions, and approvals were addressed during the data collection process. In this study, we did not conduct any data collection or associated activities ourselves. Instead, we relied on the publicly available dataset to perform our analysis and draw conclusions.

References

[1] Mindell, J.A.; Meltzer, L.J.; Carskadon, M.A.; Chervin, R.D. Developmental Aspects of Sleep Hygiene: Findings from the 2004 National Sleep Foundation Sleep in America Poll. Sleep Med. 2009, 10, 771–779.

[2] Hasan, M. J., Shon, D., Im, K., Choi, H. K., Yoo, D. S., & Kim, J. M. (2020). Sleep state classification using power spectral density and residual neural network with multichannel EEG signals. Applied Sciences, 10(21), 7639.

[3] Tarokh, L.; Carskadon, M.A. Developmental Changes in the Human Sleep EEG during Early Adolescence. Sleep 2010, 33, 801–809.

[4] Lucey, B.P.; Mcleland, J.S.; Toedebusch, C.D.; Boyd, J.; Morris, J.C.; Landsness, E.C.; Yamada, K.; Holtzman, D.M. Comparison of a Single-channel EEG Sleep Study to Polysomnography. J. Sleep Res. 2016, 25, 625–635.

[5] Chriskos, P.; Frantzidis, C.A.; Nday, C.M.; Gkivogkli, P.T.; Bamidis, P.D.; Kourtidou-Papadeli, C. A Review on Current Trends in Automatic Sleep Staging through Bio-Signal Recordings and Future Challenges. Sleep Med. Rev. 2021, 55, 101377.

[6] Mohebbi, M.; Ghassemian, H. Prediction of Paroxysmal Atrial Fibrillation Using Recurrence Plot-Based Features of the RR-Interval Signal. Physiol. Meas. 2011, 32, 1147.

[7] Senthilpari, C.; Yong, W.H. Epileptic EEG Signal Classifications Based on DT-CWT and SVM Classifier. J. Eng. Res. 2021, 10, N0 2A.

[8] IEEE Transmitter. Improving the Quality of Sleep with AI and Machine Learning. 21 May 2021. Available online: https://transmitter.ieee.org/improving-the-quality-of-sleep-with-ai-and-machine-learning (accessed on 10 October 2022).

[9] Imtiaz, S.A. A Systematic Review of Sensing Technologies for Wearable Sleep Staging. Sensors 2021, 21, 1562.

[10] Imtiaz, S.A.; Rodriguez-Villegas, E. Automatic Sleep Staging Using State Machine-Controlled Decision Trees. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) , Milan, Italy, 25–29 August 2015; IEEE: Manhattan, NY, USA, 2015; pp. 378–381.

[11] Sen, B.; Peker, M.; Cavusoglu, A.; Celebi, F.V. A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. J. Med. Syst. 2014, 38, 18.

[12] Memar, P.; Faradji, F. A Novel Multi-Class EEG-Based Sleep Stage Classification System. IEEE Trans. Neural Syst. Rehabil. Eng. 2017, 26, 84–95.

[13] Santaji, S.; Desai, V. Analysis of EEG Signal to Classify Sleep Stages Using Machine Learning. Sleep Vigil. 2020, 4, 145–152.

[14] Bhusal, A.; Alsadoon, A.; Prasad, P.W.C.; Alsalami, N.; Rashid, T.A. Deep Learning for Sleep Stages Classification: Modified Rectified Linear Unit Activation Function and Modified Orthogonal Weight Initialisation. Multimed. Tools Appl. 2022, 81, 9855–9874.

[15] 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.

[16] Yulita, I.N.; Fanany, M.I.; Arymurthy, A.M. Sleep Stage Classification Using Convolutional Neural Networks and Bidirectional Long Short-Term Memory. In Proceedings of the 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS) , Bali, Indonesia, 5 June 2017; IEEE: Manhattan, NY, USA, 2017; pp. 303–308.

[17] Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.-K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 2000, 101, e215–e220.

[18] Haque R, Khan MA, Rahman H, Khan S, Siddiqui MIH, Limon ZH, et al. Explainable deep stacking ensemble model for accurate and transparent brain tumor diagnosis. Comput Biol Med [Internet]. 2025 Jun 1 [cited 2025 May 13];191:110166. Available from: https://www.sciencedirect.com/science/article/pii/S0010482525005177.

[19] Masum A Al, Limon ZH, Islam MA, Rahman MS, Khan M, Afridi SS, et al. Web Application-Based Enhanced Esophageal Disease Diagnosis in Low-Resource Settings. 2024 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON) [Internet]. 2024 Nov 28 [cited 2025 May 13];153–8. Available from: https://ieeexplore.ieee.org/document/10962580.

[20] Haque R, Al Sakib A, Hossain MF, Islam F, Ibne Aziz F, Ahmed MR, et al. Advancing Early Leukemia Diagnostics: A Comprehensive Study Incorporating Image Processing and Transfer Learning. BioMedInformatics 2024, Vol 4, Pages 966-991 [Internet]. 2024 Apr 1 [cited 2025 May 13];4(2):966–91. Available from: https://www.mdpi.com/2673-7426/4/2/54/htm.

[21] Hasan J, Hasan K, Al Noman A, Hasan S, Sultana S, Arafat MA, et al. Transforming Leukemia Classification: A Comprehensive Study on Deep Learning Models for Enhanced Diagnostic Accuracy. PEEIACON 2024 - International Conference on Power, Electrical, Electronics and Industrial Applications. 2024;266–71.

[22] Ahmed MR, Haque R, Rahman SMA, Afridi S, Abir MFF, Hossain MF, et al. Towards Automated Detection of Tomato Leaf Diseases. Proceedings - 6th International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2024. 2024;387–92.

[23] Al-Sakib A, Limon ZH, Sakib A, Pranto MN, Islam MA, Sultana S, et al. Robust Phishing URL Classification Using FastText Character Embeddings and Hybrid Deep Learning. 2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things, RAAICON 2024 - Proceedings. 2024;53–8.

[24] Hosen MD, Bin Mohiuddin A, Sarker N, Sakib MS, Al Sakib A, Dip RH, et al. Parasitology Unveiled: Revolutionizing Microorganism Classification Through Deep Learning. Proceedings - 6th International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2024. 2024;1163–8.

[25] Al Noman A, Fardin H, Chhabra G, Sultana S, Haque R, Ahmed MR, et al. Monkeypox Lesion Classification: A Transfer Learning Approach for Early Diagnosis and Intervention. Proceedings of International Conference on Contemporary Computing and Informatics, IC3I 2024. 2024;247–54.

[26] Noman A Al, Hossain A, Sakib A, Debnath J, Fardin H, Sakib A Al, et al. ViX-MangoEFormer: An Enhanced Vision Transformer–EfficientFormer and Stacking Ensemble Approach for Mango Leaf Disease Recognition with Explainable Artificial Intelligence. Computers 2025, Vol 14, Page 171 [Internet]. 2025 May 2 [cited 2025 May 13];14(5):171. Available from: https://www.mdpi.com/2073-431X/14/5/171/htm.

[27] Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Amira Hossain, Hasib Fardin, Al Shahriar Uddin Khondakar Pranta. Custom CNN for acoustic emission classification in gas pipelines. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1760–8. Available from: https://journalijsra.com/node/1046.

[28] Hasib Fardin, Hasan Md Imran, Hamdadur Rahman, Anamul Haque Sakib, Md Ismail Hossain Siddiqui. Robust and explainable poultry disease classification via MaxViT with attention-guided visualization. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1848–59. Available from: https://journalijsra.com/node/1054.

[29] Mohammad Rasel Mahmud, Al Shahriar Uddin Khondakar Pranta, Anamul Haque Sakib, Abdullah Al Sakib, Md Ismail Hossain Siddiqui. Robust feature selection for improved sleep stage classification. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1790–7. Available from: https://journalijsra.com/node/1049.

[30] Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Sanjida Akter, Jesika Debnath, Mohammad Rasel Mahmud. Comparative analysis of traditional machine learning Vs deep learning for sleep stage classification. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1778–89. Available from: https://journalijsra.com/node/1048.

[31] Siddiqui IH, Al Sakib A, Sakib AH, Fardin H, Debnath J. Dual-branch CrossViT for ovarian cancer diagnosis: Integrating and explainable AI for real-time clinical applications. Article in International Journal of Science and Research Archive [Internet]. 2025 [cited 2025 May 14];2025(01):1834–47. Available from: https://doi.org/10.30574/ijsra.2025.15.1.1164.

[32] Hamdadur Rahman, Hasan Md Imran, Amira Hossain, Md Ismail Hossain Siddiqui, Anamul Haque Sakib. Explainable vision transformers for real time chili and onion leaf disease identification and diagnosis. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1823–33. Available from: https://journalijsra.com/node/1052.

[33] Anamul Haque Sakib, Md Ismail Hossain Siddiqui, Sanjida Akter, Abdullah Al Sakib, Mohammad Rasel Mahmud. LEVit-Skin: A balanced and interpretable transformer-CNN model for multi-class skin cancer diagnosis. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1860–73. Available from: https://journalijsra.com/node/1055.

[34] Md Ariful Islam, Mohammad Rasel Mahmud, Anamul Haque Sakib, Md Ismail Hossain Siddiqui, Hasib Fardin. Time domain feature analysis for gas pipeline fault detection using LSTM. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1769–77. Available from: https://journalijsra.com/node/1047.

[35] Mohammad Rasel Mahmud, Hasib Fardin, Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Abdullah Al Sakib. Hybrid deep learning for interpretable lung cancer recognition across computed tomography and histopathological imaging modalities. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1798–810. Available from: https://journalijsra.com/node/1050.

[36] Sanjida Akter, Mohammad Rasel Mahmud, Md Ariful Islam, Md Ismail Hossain Siddiqui, Anamul Haque Sakib. Efficient and interpretable monkeypox detection using vision transformers with explainable visualizations. International Journal of Science and Research Archive [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1811–22. Available from: https://journalijsra.com/node/1051.