Deepnap: Explainable Cnn Framework For Eeg-Based Sleep Disorder Detection With Web Visualization

[1]Priyanka P D'souza, [2]Nisarga S G, [3]Prajwalkumar S Halagi, [4]Chandan S B, [5]Vinutha H M

[1-4]Student, Dept. of CSE, PES Institute of Technology & Management, Shivamogga,Karnataka, India

[5]Assistant Professor, Dept. of CSE, PES Institute of Technology & Management,Shivamogga,Karnataka, India

Abstract: Sleep disorders like Insomnia, Narcolepsy, Noctur nal Frontal Lobe Epilepsy (NFLE), Rapid Eye Movement (REM) Sleep Behavior Disorder (RBD), and Periodic Limb Movement (PLM) are becoming more common. These neurological and physiological conditions harm human health and well-being. Detecting and classifying these disorders is crucial for effective treatment and preventing long-term issues. In this study, we suggest a deep learning approach that uses a hybrid model combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to predict multiple sleep disorders. The CNN part automatically extracts spatial and local features from input data, including physiological signals, body movements, and eye activity. Meanwhile, the LSTM part captures the time dependencies and sequential dynamics during sleep cycles. This combination helps the model learn complex patterns linked to abnormal sleep behaviors. Our deep learning framework offers a non-invasive, efficient, and automated way to identify sleep disorders with high accuracy. This reduces reliance on manual diagnosis and traditional polysomnography (PSG) tests. Experimental results indicate that the CNN-LSTM model surpasses traditional machine learning methods, making it a promising tool for clinical support and real-time sleep health monitoring.

***IndexTerms* -** Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Polysomnography (PSG),Sleep Disorder Classification, Non-invasive Diagnosis.

I. INTRODUCTION

Sleep is a basic physiological process that is crucial for physical health, mental stability, and cognitive performance. Disruptions in normal sleep can lead to various sleep disorders, such as insomnia, narcolepsy, nocturnal frontal lobe epilepsy (NFLE), REM sleep behavior disorder (RBD), and periodic limb movement (PLM). These disorders can greatly impact a person's quality of life, causing fatigue, lower concentration, and a higher risk of chronic conditions like depression, heart problems, and neurodegenerative diseases like Parkinson's and dementia.

Typically, diagnosing sleep disorders involves polysomnog raphy (PSG), which monitors several physiological signals, including EEG, EMG, and EOG during sleep. While PSG is considered the clinical gold standard, it is often costly, time-consuming, and needs specialized medical oversight. This makes it impractical for widespread or ongoing screening. These challenges have driven researchers to look into artificial intelligence (AI) and deep learning methods for automatic and non-invasive detection of sleep disorders.

Deep learning models, especially convolutional neural net works (CNNs) and recurrent neural networks (RNNs) like long short-term memory (LSTM) networks, have proven very effective in pattern recognition and time-series analysis. CNNs excel at recognizing spatial hierarchies and pulling relevant features from complex physiological or behavioral data. On the other hand, LSTMs are good at capturing long-term dependencies in sequential data. By combining these two models, a CNN-LSTM hybrid can analyze both spatial and temporal patterns found in sleep-related signals.

Recent progress in computational neuroscience and biomed ical signal processing has improved the potential of deep learning models in healthcare. By combining different types of physiological data with advanced neural networks, we can now detect subtle biomarkers and changes linked to various sleep stages and health issues. Using CNN-LSTM frameworks allows for strong classification of complex sleep patterns and enables real-time monitoring and early diagnosis of disorders through wearable sensors or mobile apps. These automated systems mark an important move toward personalized sleep medicine. Continuous data-driven insights can help clinicians create targeted treatment plans, ultimately benefiting patient health and lessening the global impact of sleep-related diseases.

In this project, we are creating a deep learning-based CNN-LSTM model to automatically predict five major sleep disorders: insomnia, narcolepsy, NFLE, RBD, and PLM. The model will use various data types, including speech patterns, eye movements, and body movement signals, to improve diagnostic accuracy. Our goal is to develop a non-invasive, intelligent, and scalable diagnostic system that can detect and monitor sleep disorders early. This method will not only lower the need for manual analysis but will also help clinicians enhance diagnostic efficiency and improve patient outcomes.

II. NEED OF THE STUDY

Sleep disorders are becoming increasingly prevalent in modern society, with millions of people from different age brackets affected. Stress, changes in lifestyle, and increased neurological ailments stand out as factors that continue to increase cases of disrupted or poor-quality sleep among an increased number of people. These, if not diagnosed, might result in serious detrimental effects, including poorer cognitive performance and emotional instability, as well as cardiovascular risks and decreased overall well-being. Yet, even with the severity of such conditions, most people remain undiagnosed simply because early symptoms are minimal and hardly detectable without the right tools.

While PSG is a highly reliable gold standard in the diagnosis of sleep disorders, it is equally resource-intensive. It has to involve patients staying overnight in special sleep laboratories while trained clinicians interpret complex EEG recordings by hand. This makes PSG time-consuming, expensive, and inaccessible for people in most rural areas and/or with poor medical infrastructure. Manual interpretation of EEG signals is labor-intensive and prone to missing some fine patterns due to fatigue or individual variability in expertise. All these represent just some of the main issues pointing to a gap between needs and traditional diagnostics in performing sleep disorder assessments on a wide scale.

With this growth in the availability of EEG datasets and advancements in artificial intelligence, there is a strong need for automated solutions that can support clinicians and provide faster, more accessible diagnostics. Deep learning models, especially hybrid architectures like CNN-LSTM, are capable of eliciting both spatial and temporal patterns in brain signals, patterns which might not be readily visible by a human observer. Such a system will go a long way in reducing the burden on sleep specialists, enabling early diagnosis, and making sleep health evaluation inclusive and scalable. Eventually, an automated deep learning–based framework can help bridge the gap between rising sleep-related health concerns and the limited availability of conventional diagnostic resources, improving both clinical outcomes and patient quality of life.

III.RESEARCH METHODOLOGY

The proposed system provides an end-to-end framework for automatic sleep disorder classification using single-channel EEG signals. The main steps include dataset preparation, model design, training and optimization, and deployment for real-time testing.

3.1 Dataset Preparation and Input Handling

The EEG data used in this study are taken from the CAP Sleep Database, using the C4–A1 channel. The recordings are originally in EDF format and are converted to CSV for training.

The continuous EEG signals are divided into 2-second segments (about 1024 samples each). Overlapping segments are used to increase the number of training samples and ensure smoother transitions.

Each segment is normalized using Min–Max normalization to maintain consistency:

xnorm=xmaxxmin/xmaxxmin

Every segment is labeled as either healthy or with a specific sleep disorder. The data are split into training, validation, and testing sets in a 70:15:15 ratio, ensuring that the same subject's data does not appear in multiple sets.

3.2 Model Architecture and Network Design

The classification model is built using a 1D Convolutional Neural Network (1D CNN) combined with Inception modules, skip connections, and temporal modeling layers. Two types of models are developed:

Binary model– classifies subjects as healthy or unhealthy. Multi-class model– identifies different types of sleep dis orders.

The CNN layers capture short-term signal patterns such as spikes and bursts. Inception modules use multiple kernel sizes (3, 7, and 15) to learn features at different time scales, while skip connections help retain low-level features and improve training stability.

For temporal learning, both LSTM and GRU were tested. The final model uses GRU since it provides similar accuracy to LSTM but trains faster. In some model variants, Global Max Pooling (GMP) replaces the GRU layer to make training more efficient. GMP selects the most important activation in each feature map:

zi= maxyi(t) This reduces the number of parameters and speeds up the process, making it ideal for real-time applications.

The output from the GRU or GMP layer is passed to fully connected layers with ReLU activation, and finally to a Softmax (multi-class) or Sigmoid (binary) output layer.

3.3 Training and Optimization

Training is performed on the CSV-formatted EEG segments, while inference uses EDF files that are converted internally. The model uses the Adam optimizer (learning rate 0.001) with categorical cross-entropy loss.

Batch normalization and dropout (0.3–0.5) are applied to prevent overfitting. K-fold cross-validation is used to fine-tune hyperparameters such as kernel size, dropout rate, and GRU hidden units.

The model's performance is measured using accuracy, preci sion, recall, F1-score, and confusion matrix. The training time and inference speed are compared between GRU and GMP models to evaluate efficiency.

3.4 EDF Conversion and Front-End Integration

The deployed system accepts EDF files directly. When a f ile is uploaded, the backend automatically extracts the C4–A1 channel, segments and normalizes the data, and converts it to CSV format before making predictions.

A Gradio web interface is developed for user interaction. It allows users to upload EEG recordings, visualize the EEG waveform, and view the predicted class (healthy or disorder). The interface can switch between GRU and GMP models, demonstrating the balance between accuracy and speed.

3.5 Implementation and Deployment

The models are implemented in TensorFlow/Keras, and pyEDFlib is used for EDF file processing. Training is carried out on GPU-enabled platforms such as Jupyter Notebook for faster computation.

For deployment:

The GMP-based model is suitable for fast, low-resource environments.
The GRU-based model is recommended for high-accuracy, detailed analysis.

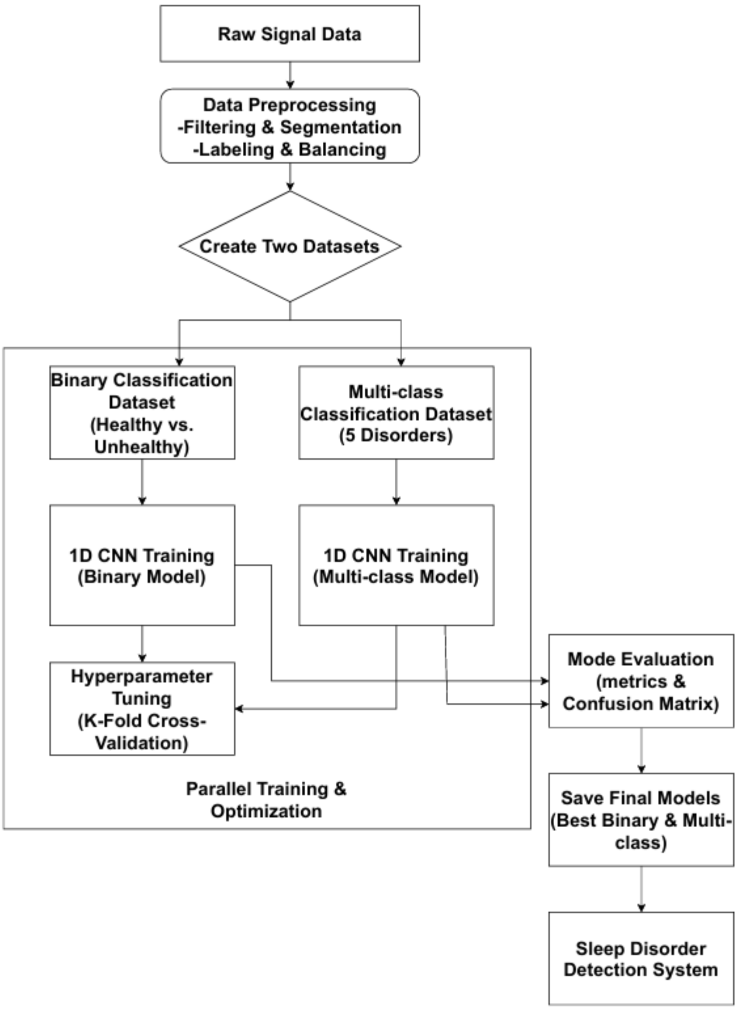

Fig.3.1 Model workflow

IV. RESULTS AND DISCUSSION

The performance of the proposed DeepNap framework was carefully evaluated across three interconnected stages:

- distinguishing between healthy and disordered EEG pat terns,
- identifying the specific type of disorder, and
- detecting Cyclic Alternating Pattern (CAP) phases that reflect variations in sleep stability.

Each model was trained and validated using EEG signals stored in .edf and .csv formats, with a 70:30 train–validation split. The evaluation focused on commonly accepted metrics such as accuracy, precision, recall, F1-score, specificity, and ROC–AUC to ensure a balanced assessment of the model's performance.

4.1 Healthy vs Disorder Classification

Table 4.1: Performance metrics for Healthy vs Disorder Classification

Metric	Value
Accuracy	91.45%
Specificity	92.87%
Precision	92.72%
Sensitivity (Recall)	90.04%
F1-Score	91.36%
Cohen’s Kappa	0.829
ROC-AUC	0.968

The first stage of the DeepNap framework focuses on distinguishing normal (healthy) sleep patterns from signals that may show a sleep disorder. This binary classification is a key clinical checkpoint. It helps avoid unnecessary evaluations for people who have normal sleep behavior.

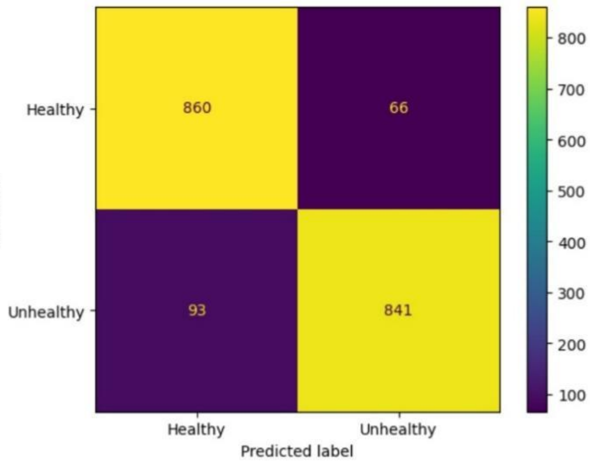

Fig. 4.1 Confusion matrix for healthy vs unhealthy classification

To do this, a strong CNN-LSTM architecture was created to capture both spatial EEG features and their changing patterns over time during sleep. This model was trained and validated using standardized EEG recordings. This approach allows it to effectively interpret changes in neural signals that are often linked to sleep issues.

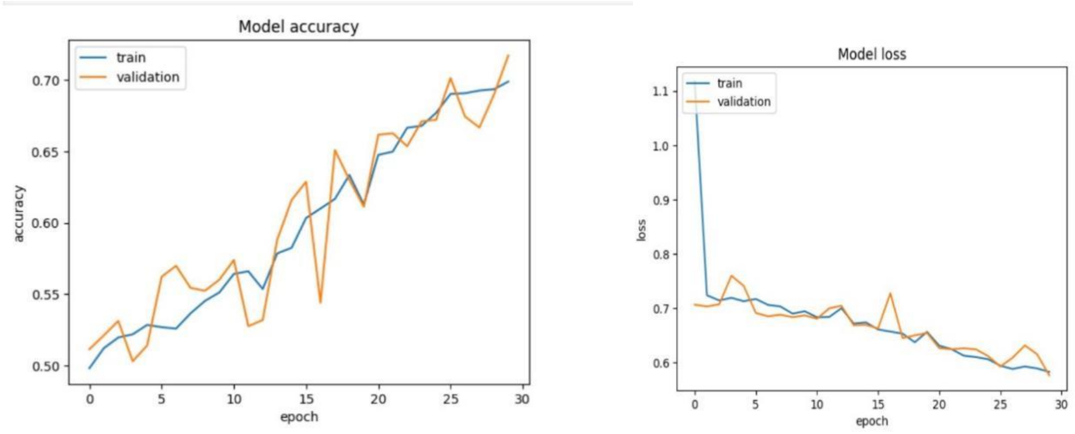

Fig. 4.2 Accuracy and Loss graph for healthy vs unhealthy classification

The high ROC-AUC score and Kappa coefficient confirm that the model performs significantly better than chance and shows strong agreement between predicted and actual out comes. The confusion matrix (Figure 3) indicates very few false negatives and false positives, reinforcing the model's reliability. The performance curves (Figure 4) show smooth convergence with no signs of overfitting, which points to an effective training strategy.

Beyond its quantitative performance, the model shows stable results across different subjects and EEG formats (.edf and .csv). This reliability is especially important for real-world use, where consistent outcomes across various equipment, hospitals, and datasets are needed. Additionally, the classifier identifies subtle neural disruptions and micro-architectural changes that often serve as early signs of sleep disorders.

Clinically, this stage acts as a quick and reliable screening tool. It allows DeepNap to send only high-risk EEG record ings to the next classification layer, improving efficiency and speeding up clinical decision-making. Ultimately, this dual purpose—early detection and reduced workload—highlights the importance of this binary classification module within the larger DeepNap system.

4.2 Multi-Disorder Classification

Once a sample is flagged as disordered, it passes through the second model, which classifies the signal into one of five categories: Insomnia (INS), Nocturnal Frontal Lobe Epilepsy (NFLE), Narcolepsy (NARCO), REM Behavior Disorder (RBD), or Periodic Limb Movement (PLM).

The multi-class CNN model achieved an overall accuracy of 87.32 percentage, a strong indication of its effectiveness in distinguishing among multiple complex disorders.

These results show that the model performs especially well in identifying NFLE and RBD, which exhibit clear and distinctive EEG patterns. A few overlapping signals between Insomnia and PLM led to occasional misclassifications — anexpected outcome given the similar low-frequency wave char acteristics of both conditions. The confusion matrix (Figure 5) confirms that most samples are correctly mapped to their respective classes, while the training accuracies reveals a stable learning pattern with no signs of overfitting. Overall, this stage demonstrates that DeepNap can handle the nuanced variability among different sleep disorders with high reliability.

Table 4.2: Performance Metrics for Multi-Class Classification

Metric	Value
Accuracy	87.32%
Loss	0.3918
Precision (per class)	[0.857, 0.957, 0.746, 0.941, 0.890]
Recall (per class)	[0.886, 0.956, 0.891, 0.887, 0.748]

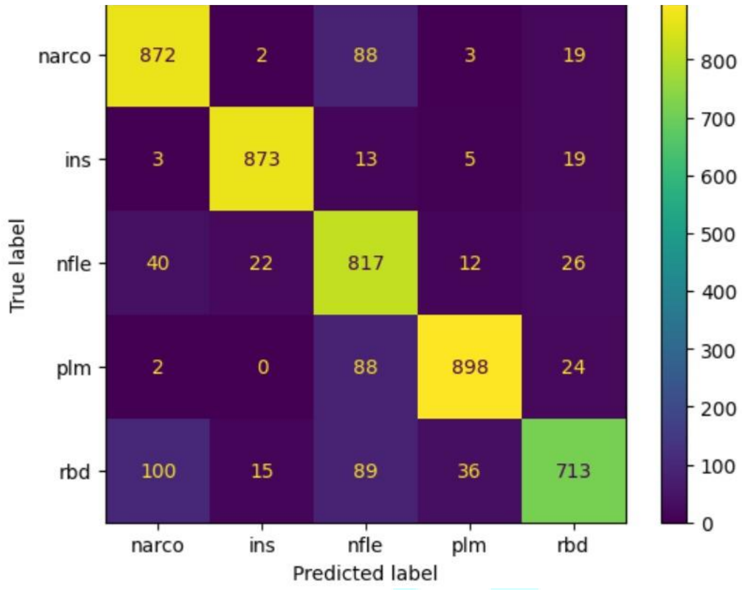

Fig. 4.3 Confusion matrix for disorder classification

4.3 CAP Phase Detection

The final stage of the framework examines the Cyclic Alternating Pattern (CAP) — a microstructural element of sleep that reflects the brain's stability and reactivity during rest. For each disorder category, a dedicated binary CNN was trained to distinguish between Phase A (activation) and Phase B (quiet state).

Acknowledgment

We would like to extend our most sincere gratitude to PES Institute of Technology & Management, Shivamogga, for providing us support during the period and for successful accomplishment of this project. We are extremely thankful to our guide, Ms. Vinutha H M, for her priceless guidance and encouragement right from inception to completion of our research work.

References

[1] R. Ferri et al., "A Hierarchical Approach for the Diagnosis of Sleep Disorders Using Convolutional Recurrent Neural Network,"IEEE Trans. Neural Syst. Rehabil. Eng., vol. 29, pp. 1721–1732, Nov. 2021.
[2] K. Wang et al., "Wearable Sleep Monitoring Devices and AI-Based Healthcare Applications: A Review," IEEE Sensors Journal, vol. 23, no. 8, pp. 8900–8914, 2023.
[3] S. Sankari et al., "EEG-Based Biomarkers of Sleep Quality and Neurodegenerative Disorders," Nature and Science of Sleep, vol. 15, pp. 455–471, 2023.
[4] G. D. Mitsis, "CAP Analysis and Sleep Disorders: An Updated Review," Sleep Medicine, vol. 77, pp. 340–353, 2021.
[5] S. Z. Li et al., "Sleep Disorder Severity Estimation Using Machine Learning Techniques on EEG Data," Expert Systems with Applications, vol. 167, 2021.
[6] R. Das, "A Review of EEG-Based Automated Sleep Stage Scoring Using Deep Learning," IEEE Rev. Biomed. Eng., vol. 14, pp. 364–376, 2021.
[7] T. Perslev et al., "U-Sleep: Resilient High-Frequency Sleep Staging," npj Digital Medicine, vol. 4, pp. 1–10, 2021.
[8] U. R. Acharya et al., "Automated Detection of Sleep Disorders from EEG Signals Using Feature Learning," Computers in Biology and Medicine, vol. 111, 2019.
[9] A. Phan et al., "Joint Classification and Prediction CNN Framework for Automatic Sleep Stage Classification," IEEE Trans. Biomed. Eng., vol. 66, no. 5, pp. 1285–1296, 2019.
[10] M. Aboalayon, H. Faezipour, W. Almuhammadi and S. Moslehpour, "Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey," IEEE Access, vol. 6, pp. 14412–14431, 2018.
[11] S. Chambon, V. Thorey, N. Arnaldi, P. Mignot and M. Gramfort, "A Deep Learning Architecture for Temporal Sleep Stage Classification Using Multivariate and Multimodal Time Series," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 26, no. 4, pp. 758–769, 2018.

[12] J. L. Supratak, W. Dong, C. Wu and Y. Guo, ”DeepSleepNet: End-to End CNN-RNN Based Automatic Sleep Stage Scoring,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 25, no. 11, pp. 1998–2008, 2017.
[13] A. Tsinalis, P. Matthews and Y. Guo, ”Automatic Sleep Stage Scoring Using Time-Frequency Analysis and Stacked Sparse Autoencoders,” Ann. Biomed. Eng., vol. 44, pp. 1587–1597, 2016.
[14] M. D. Prerau et al., ”Sleep EEG Analysis Using Multitaper Spectral Methods,” Proceedings of the National Academy of Sciences, vol. 111, no. 38, pp. E4277–E4286, 2014