J. Integr. Sci. Technol. 2025, 13(6), 1135

Journal of Integrated
SCIENCE & TECHNOLOGY

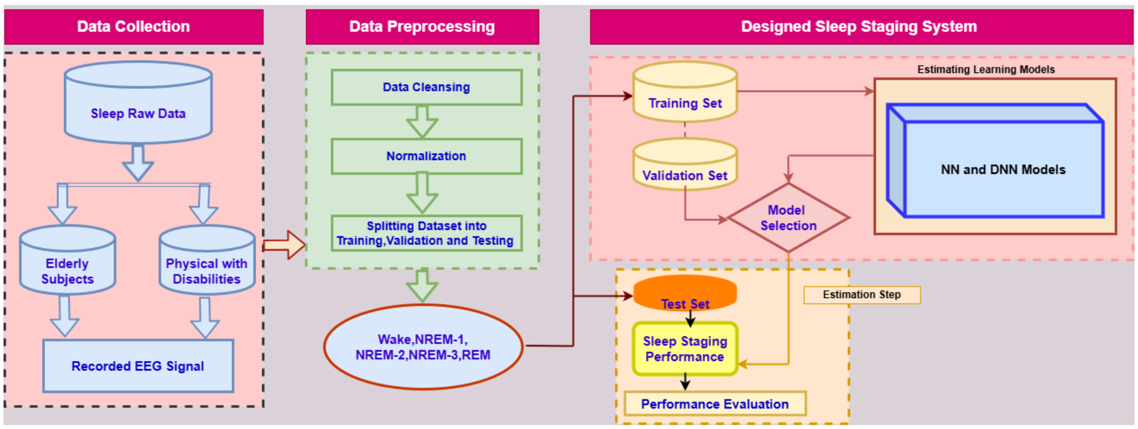
Sleep stages detection from EEG signal utilizing Backpropagation Neural Network and Deep Neural Network

Suren Kumar Sahu[1], Sudhir Kumar Mohapatra[1*] Santosh Kumar Satapathy[2]

[1]Faculty of Engineering & Technology, Sri Sri University, Cuttack, Odisha, India. [2]Department of Information & Communication Technology, Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Submitted on: 28-Oct-2024; Accepted and Published on: 07-May-2025

ABSTRACT



Prompt and accurate identification of sleep-related disorders is essential for preventing progression to serious neurodegenerative conditions. However, current diagnostic methods, such as biomarkers and cognitive screening tests, are expensive, time-consuming, or not user-friendly. This study evaluated a Neural Network (NN) and a Deep Neural Network (DNN) for classifying five sleep stages using EEG data. To ensure high-quality input, artifact correction, signal decomposition, and overlapping sliding window techniques were applied, followed by the extraction of time-domain, frequency-domain, and non-linear features. Model performance was assessed using precision, recall, and F1-score metrics. Overall, the DNN outperformed the NN, particularly in distinguishing wake (W) and rapid eye movement (R) stages, demonstrating a stronger ability to capture subtle EEG patterns. While the NN showed strength in classifying certain stages, it struggled with finer distinctions, such as between N1, N2, and R stages. This comparison highlights the advantage of deeper architectures like DNN for complex EEG analysis. However, the increased computational cost of DNNs suggests a need for future optimization to balance accuracy and efficiency

Keywords: Sleep stage classification, EEG signal analysis, Signal decomposition, Feature extraction, Machine Learning

INTRODUCTION

The study of sleep has existed for as long as there have been people since it is a complex process and a basic human need that directly affects a person's health and well-being. Regarding conscious activity, sleep is described as an active state of con-sciousness; sensory activity is considerably reduced; the brain is in a somewhat 'down' condition; it 'reacts' only to internal stimuli. It is generally divided into two main phases which include NREM, which can be referred to as slow wave sleep and second is REM sleep [1]. Every stage is characterized by certain EEG rhythms: Thus, non-REM sleep is indispensable for tissue repair, and REM sleep is associated with memory consolidation and regulation of emotions [2].

It also reveals them to be genuinely indispensable, not just for sleep, but for the optimization of physical and mental health, inclusive of aspects of habilitation being immunization, mental steadiness, and stress regulation[3]. Meanwhile, the conse-quences of sleep disorders can be very serious and extensive supported by research pointers pointing to detrimental effects such as lower productivity at work, getting sick easily as well as deadly accidents. For instance, based on the research conducted on Moroccan drivers it was estimated that drowsiness contributed to between 20–25% of the accident [4]; hence there is a need for increased focus on sleep issues and efforts.

Sleep disorders may be elicited by so many factors; it can be due to lifestyles, pa-thological conditions or infections. Out of all the

*Corresponding author: Prof. Sudhir Kumar Mohapatra, Sri Sri University, Cuttack, Odisha, India. Tel: +91- 9556878743; Email: sudhir.mohapatra@srisriuniversity.edu.in

Cite as: *J. Integr. Sci. Technol., 2025, 13(6), 1135.*
DOI: 10.62110/sciencein.jist.2025.v13.1135

diseases transmitted from insects, this paper shall concentrate on human African trypanosomiasis, better still known as sleeping sickness[5]. The disease that has ravaged more than 36 countries in Africa contributing to more than 50 million people and leading to the death of 100 people a day, according to the WHO affects people's sleep patterns[6]. The examination and analysis of the particular stage of sleep are therefore beneficial not only to improve the quality of sleep but also to diagnose sleep disorders such as apnea and insom-nia[7].

For this reason, several therapeutic as well as diagnostic techniques have been de-veloped to diagnose and subsequently treat sleep disorders particularly when linked to the brain [8]. Of these techniques, fMRI has been widely used in investigating the changes in the operational parameters of sleep phenotypes. fMRI on the other hand offers a good spatial resolution which allows for spatial localization of sources of signal. However, there is one more technique better suitable for the tasks mentioned above – EEG has a good temporal resolution, and real-time brain registration is possi-ble with less discomfort for the subject [9].

EEG functions in a way that it captures the brain's electrical output with electrodes placed on the scalp and such detail is recorded on a trace form[10]. When compared to other techniques, it is more suitable for sleep detection than for example, electro-cardiography in which the rate of the heartbeat is measured or electromyography, where muscle atonia in periods of REM sleep is determined[11]. Also, the extinction of electro-oculography, records movements of the eyes during rapid eye-moving sleep [12-15]. There are other global approaches to monitor sleep and breathing that are available but PSG involves several physiological signals and although PSG is less invasive it can be rather costly and cumbersome and patients may be uncomfortable having so many electrodes attached to their bodies.Fig.1 illustrates the behavior of the different sleep stages.

Therefore, up till now, EEG has been looked upon as being the most suitable for sleep identification because it is the method that occupies the first place based on efficacy and ked invasiveness as well as the ability to be applied in home and clinic classifications. Nonetheless, the ability of EEG to diagnose sleep highly depends on the right way of processing, analyzing, and categorizing the signals obtained from the tests. The pattern of studying the neural signals can therefore be categorized under temporal analysis, frequency analysis, and time-frequency analysis. Historical analy-sis involves searching for specific phenomena at a time, for example, electrical activi-ty in brains, and determining statistical characteristics such as the mean, variance, and standard deviation [16-17]. Whereas, the frequency analysis in EEG entails the separation of the signal into bands: Gamma, Beta, Alpha, Delta, and Theta, their power calculation as a function of the frequency and Fourier transforms of the fields for the signal have been developed [18]. Table 1 presents details of the behavior of the EEG waveforms.

The time-frequency analysis is based on the temporal and frequency one and can be used for identification of the dynamics of change in frequency with the help of computer techniques such as FFT or wavelet transform [19]. But these methods in-volve calculations and the formation of which new techniques are needed, which give important features in the time and frequency domain with faster and simpler calcula-tions. The method of non-linear characteristic extraction has been considered as a new approach in the analysis of a non-stationary process, especially on the EEG signals, which possess the characteristic methods based on temporal variability. It can cope with differences in frequency over time and thus helps to explain some of the phases of sleep[20].

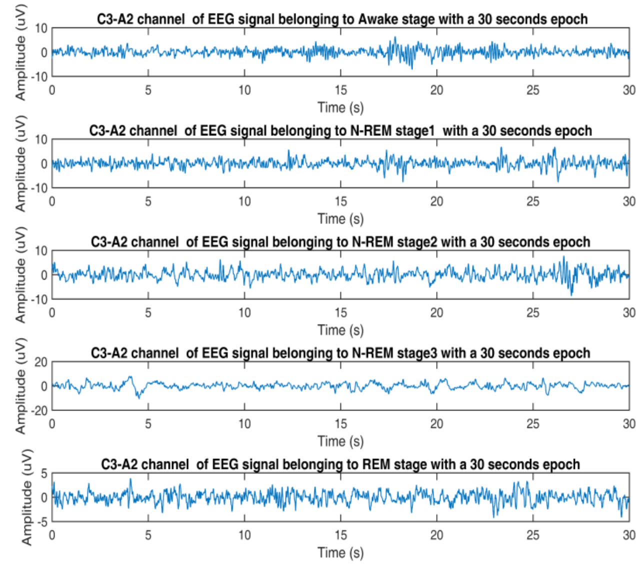


Figure 1. EEG signal segments under different sleep stages.

Table 1. Waves generated from the brain having dominant frequencies belong to the alpha, beta, delta & theta sub band

Sleep stages	Dominant Wave
W	β (12-30 Hz) and α (8-13 Hz)
N1	θ (4-8 Hz)
N2	K-Complex (1 Hz) and Spindle (12-14 Hz)
N3	δ (0.5-2 Hz)
R	Sawtooth wave (2-6 Hz), θ (4-8 Hz), α (8-13 Hz), and β (12-30 Hz)

Thus, the methods of analyzing signals and classifying EEG signals are instrumen-tal in the main process of the reduction of the dimensionality of the brain data that are raw and also in the task of mapping models of prediction and categorization of sleep stages. The future development in the study of sleep shall be the demand to advance to more sophisticated approaches of algorithm and analysis that shall warrant an enhanced understanding of sleep and the related disorders to enhance sleep health and disorders[21-22].

This study investigates the detection and classification of sleep stages through the analysis and optimization of EEG signals. By filtering and decomposing EEG signals into distinct brain waveforms (Alpha, Beta, Low Alpha, High Beta, Low Beta, High Alpha), we extract linear (Power Spectral Density) [23]. These features are then clas-sified into different sleep stages (W, N1, N2, N3, R) using machine learning models like Neural Networks (NN) and Deep Neural Networks (DNN)[24]. The study aims to improve classification accuracy and computational efficiency, focusing on optimal feature extraction and model architecture while minimizing patient discomfort by using non-invasive EEG techniques [25]. The outcomes contribute to advancing au-tomated sleep disorder detection and sleep health monitoring.

RELATED WORK

Various methods such as frequency and time distribution, graph theory, signal modeling, wavelet transform, and empirical mode decomposition are employed for signal processing in the separation of sleep stages. For the classification aspect, a range of models is applied, including support vector machines (SVM), neural networks (NN), and partial least squares. Ronzhina et al. [26] proposed a design utilizing the power spectral density of EEG signals combined with artificial neural networks, based on a single-channel EEG. Lajnef et al. [27] used multiple features such as entropy, variance, error, kurtosis, skewness, traversal entropy, and a multi-layer support vector machine on EOG, EMG, and EEG to automatically detect sleep stages. Hassan et al. [28] also employed a combination of wavelet transform and Taguchi-based neural networks to automatically identify sleep stages from EEG data. Berthomier et al. [29] extracted features from six EEG channels, three EOG channels, and one EMG channel, analyzing them with quadratic principles. Liang et al. [30] focused on entropy-based features to identify different sleep stages using EEG, while Liu et al. [31] generated a visual graph from a two-head EEG signal and applied nine features for classification via support vector machines. Kayikcioglu et al. [32] developed a feature extraction technique based on AR model technology and the partial least squares algorithm for sleep stage classification. Zhou et al. [33] introduced the empirical mode decomposition method, which is considered a key signal-processing technique in the time-frequency domain. Research on sleep stage detection utilizing EEG signals has been intensive in the past years, and many strategies have been suggested to enhance the accuracy and speed of the algorithms. The empirical mode decomposition method breaks signals down into several intrinsic mode functions, which are then used to process non-linear and non-stationary signals. To address some limitations of this method, Flandrin et al. [34] proposed adding white noise. Chang et al.[35] introduced the ensemble EEMD technique to solve the issue of mode mixing. Liu et al.[29] carried out a time-frequency analysis for feature extraction and used an accumulated auto-encoder algorithm for classification. Previous research has shown that most existing algorithms require more than one lead to automatically detect sleep stages, making it inconvenient for continuous use and limiting the effectiveness of sleep monitoring devices at home. Additionally, it was found that these methods provided less than 90% accuracy in classifying different sleep stages and were computationally intensive and time-consuming. In recent years, substantial research has been devoted to leveraging one-dimensional raw polysomnogram (PSG) data or PSG-derived features for sleep staging using deep learning models. Zhengling He et al. [36] introduced an end-to-end deep neural network based on single-channel EEG, incorporating a domain-adaptation module and a transfer attention mechanism. This network achieved accuracies of 83.9% and 78.8% on the Sleep-EDF-2018 dataset. Caihong Zhao et al. [37] developed a sleep staging model combining CNN and RNN, with average accuracies of 84.8% and 82.7% on single-channel EEG signals. In comparison to one-dimensional PSG data, two-dimensional time-frequency images offer more distinguishing information regarding sleep stages, facilitating richer feature extraction. Converting one-dimensional PSG data into two-dimensional time-frequency images allows for simultaneous analysis of signal dynamics in both the time and frequency domains [21-24]. Yang Dai et al. [38] proposed a transformer encoder-based automatic sleep stage classification network, tested on multi-channel time-frequency images derived from Short-Time Fourier Transform (STFT), achieving a peak accuracy of 85.0%. Huy Phan et al. [39] introduced a sequence-to-sequence sleep staging model that learned joint features from raw signals and FFT-based time-frequency images, reaching a maximum accuracy of 84.0% when using Fpz-Cz EEG and EOG signals. This article will focus on addressing these challenges.

Our contribution

This paper makes the following contributions:

- Data Preparation for Neural Networks: Prepared the dataset by splitting the data into features (X) and target (Y) for classification of sleep stages (W, N1, N2, N3, R). Further split the data into training and testing sets (80% training, 20% testing) for model evaluation.
- Implications: The DNN's deeper architecture allowed for more effective feature extraction, leading to improved classification accuracy in complex EEG signal patterns.
- Model Evaluation and Predictions: Predicted sleep stages using both NN and DNN models on the testing dataset. Compared the models based on accuracy, precision, recall, F1-score, and confusion matrices.
- Performance Metrics and Visualization: Calculated critical performance metrics such as accuracy, precision, recall, specificity, F1-score, and Cohen’s kappa score

METHODOLOGY

The method proposed in this section for sleep stage detection comprises four major steps. First, the acquisition of the EEG signal, then the processing and decomposition of the signal by using filters to eliminate noise that may disturb it and to obtain reliable and relevant information, employing different approaches to feature extraction and optimization. Algorithm 1 presents the details of the working procedures of sleep staging.

The final step is to classify the extracted data into six sleep categories (W, N1, N2, N3, R) by applying intelligent learning methods under the, which is a powerful open-source platform for data analysis and deployment of Deep Neural Network, and Neural Network through the creation of a workflow that allows visualization, transformations, and prediction operations based on feature extraction and statistics. The overall process of the approach chosen is depicted in Figure 2, which describes the various stages that will be detailed and justified later.

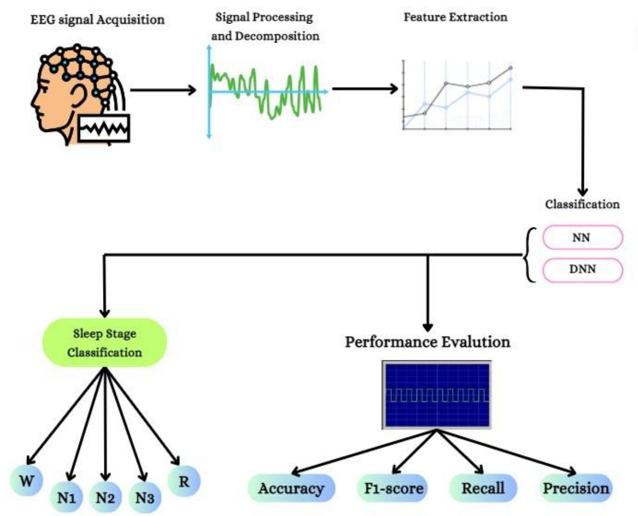


Figure 2. Flowchart of the proposed work

Dataset Description

In 1920, the neurologist Hans Berger discovered that the electroencephalogram (EEG) is a non-invasive device [21] based on the recording of cerebral electrical activities recorded by electrodes that are located on the head. Electrode positions and labels in the international 10–20 system, with the letters F, P, T, O, and C that represent the Frontal, Parietal, Temporal, Occipital, and Central zones respectively. The signals that are used in this paper are generated by taking the reference electrodes which are often used in sleep monitoring Fpz, Cz, Pz, where Z (Zero) represents the median section of the sagittal plane.Fig.3 illustrates the distributions of sleep epochs over the individual sleep stages used in this research work for experimental work[22].

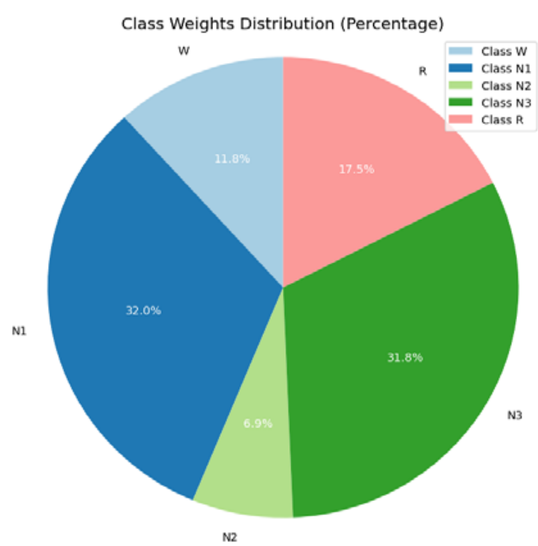


Figure 3. Distribution of epoch for data set utilized in this approach

Different brain waves such as Delta (0-4 Hz), Theta (4-8 Hz), Low Alpha (8-10 Hz), High Alpha (10-12 Hz), Low Beta (12-18 Hz) and High Beta(18-30 Hz) illustrated in Table 3 are produced by these electrical signals. The dataset is gathered by recording the sleep of two healthy women aged 33, for over two successive nights. The EEG recording results are saved in twenty casttes SC4101E0, SC4111E0, SC4121E0, SC4131E0, SC4141E0, SC4151E0, etc. where SC stands for 'Sleep Cassette'[23]. The sampling rate at which the recording is sampled is 30 Hz without overlapping following a well-defined standard procedure and recommendations of Rechtschaffen and Kales [24]. The selection of distinct and independent EEG signal segments is essential for simplifying the analysis of various sleep stages and enabling easier interpretation, lower computing complexity, and improving the outcomes, on the hand it will be advantageous to train the models using non-overlapping data to enhance generalization, prevent data loss, and prevent overlearning.

During this phase of the procedure, the following additional parameters were taken into account:

- Size of window: 30 seconds
- Rate of window overlapping: without overlapping
- Duration of recording: 24 hours
- Sampling rate: 100 Hz
- Labeling Method: The sleep stages were labeled according to the standard sleep staging criteria: Wakefulness (W), Stage 1 (N1), Stage 2 (N2), Stage 3 (N3), and REM sleep (R).
- Duration of Each Stage: Each sleep stage has a different duration for each recording, but average cycles last roughly 90 minutes.

Signal processing and analysis

The initial stage in processing physiological signals involves filtering to eliminate undesirable artifacts that could interfere with the accurate interpretation of the data. These artifacts may arise from various physiological sources, such as heartbeats, muscle movements, or external noise, all of which can distort the signal. In this context, Finite Impulse Response (FIR) filters are often employed due to their inherent stability and ability to maintain a linear phase [25].

The second step consists of splitting the signal into Theta, Delta, Alpha, Beta, sawtooth, and spindle waves by applying a Butterworth bandpass filter. The choice of this type of filter is justified by its facility of design and implementation based on the specific needs of the signal concerned, its adaptability compared with other types of filters such as elliptical and Chebyshev, and its ability to preserve the properties of the EEG signal thanks to its flat frequency response, which reduces distortion in the passband [26]. Table 2 provides a range of frequencies and EEG signal descriptions. Fig. 4 and Fig.5 depict raw signal and preprocessed signals, respectively.

- Wakefulness (W): Characterized by beta (12-30 Hz) and alpha (8- 13 Hz) waves. The brain is active and responsive to external stimuli.
- Stage 1 (N1): Dominated by theta waves (4-8 Hz). It is a light sleep stage where the transition from wakefulness to sleep occurs.
- Stage 2 (N2): Characterized by sleep spindles (12-14 Hz) and K Complexes (1 Hz). This stage is deeper than stage 1 and is crucial for maintaining sleep.
- Stage 3 (N3): Known as slow-wave sleep (SWS), dominated by delta waves (0.5-2 Hz). These stages are the deepest sleep phases, essential for physical and mental restoration.
- REM Sleep (R): Characterized by sawtooth waves (2-6 Hz), theta waves (4-8 Hz), alpha waves (8-13 Hz), and beta waves (12- 30 Hz). REM sleep is associated with vivid dreaming and brain activity similar to wakefulness [27].

Table 1. EEG frequencies range

EEG Rhythms	Frequency (Hz)	Psychological State
Delta	0-4	Deep Sleep, unconscious
Theta	4-8	Deep Relaxation, meditation
Low Alpha	8-10	Wakeful relaxation
High Alpha	10-12	Self-awareness
Low Beta	12-18	Making decisions and thinking critically
High Beta	18-30	Engagement in mental activity

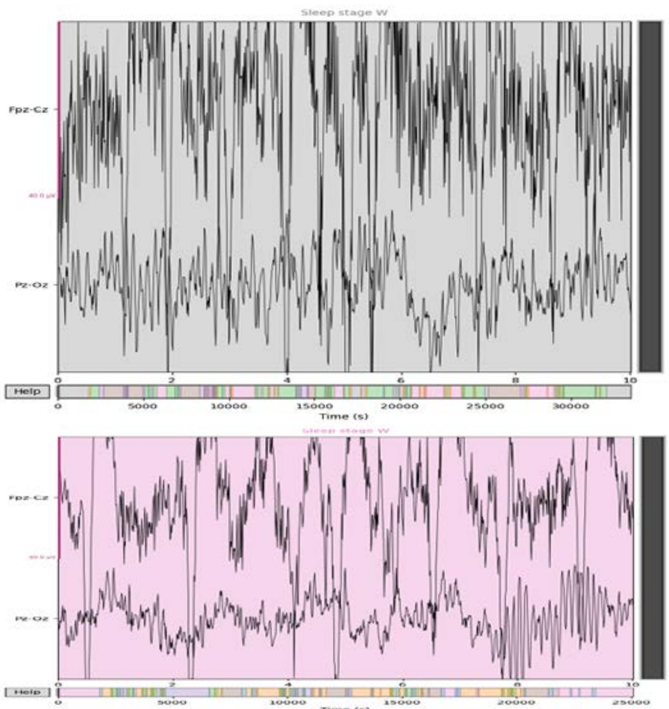


Figure 1. Raw EEG signal before preprocessing (Sleep-EDF database [22])

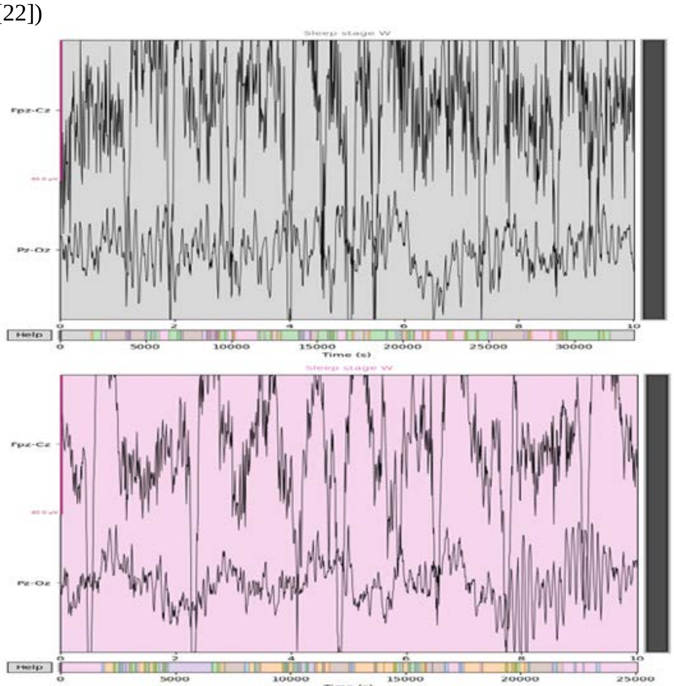


Figure 2. Filtered EEG signal after preprocessing

Feature Extraction

In the analysis of EEG signals at this stage, the emphasis lies on extracting and choosing important features from the processed EEG signal to decrease the data's dimensionality while retaining crucial information necessary for accurate classification or prediction. Power Spectral Density (PSD) is employed to derive significant attributes from various EEG frequency bands that correlate to different psychological and physiological states [27]. In contrast to non-linear feature extraction techniques like Hjorth parameters or fractal dimension analysis, PSD offers a linear method for examining the power distribution of the signal across frequency bands [28]. In Table 2, It is the frequency bands are fundamental to understanding EEG dynamics, and the Power Spectral Density (PSD) analysis quantifies the power in each of these bands.PSD quantifies how a signal's power is distributed across different frequency components, calculated using the Fourier Transform, often via Welch's method, which involves segmenting the EEG signal, applying windowing to each segment, and averaging the power spectra across all segments [29].

To compute PSD: The estimation of the Power Spectral Density (PSD) through built-in techniques, like those available in the SciPy library (e.g., scipy.signal.welch()), involves several crucial steps that ensure precise and stable assessment of the signal's frequency power. PSD is extensively utilized in signal processing to examine the power dispersion of a signal across different frequency components [30].

The following are the typical procedures involved in calculating PSD:

Divide the Signal into Segments (Optional):

The signal is segmented into overlapping or non-overlapping sections. This segmentation is performed to decrease noise and provide a more reliable PSD estimate by averaging across multiple segments. For example, Welch's method divides the signal into segments, often with 50% overlap, which enhances the robustness of the PSD estimate. When working with the entire signal at once, this step can be skipped, but most built-in PSD methods utilize segmentation for greater stability [31]. Fig.6 illustrates illustrations of the different power spectrum wave forms with the different frequency bands.

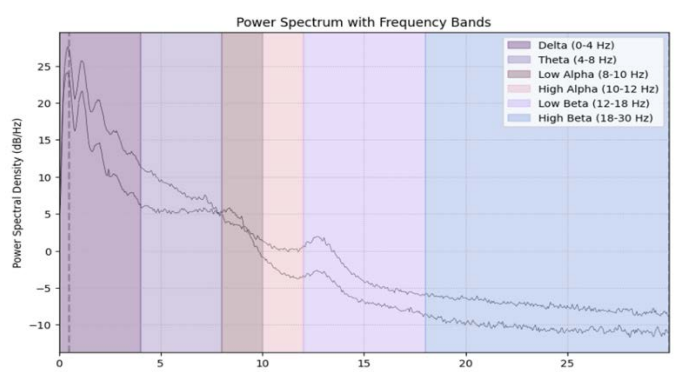


Figure 3. Illustration of Power spectral density features in different frequency bands

Application of Window Function:

A window function such as Hanning or Hamming is utilized on each segment. The window smoothens the segment edges and decreases spectral leakage, which can happen when the FFT assumes the signal to be periodic. Discontinuities at segment boundaries can distort the frequency analysis. Windowing aids in tapering the signal and reducing the impact of abrupt changes at the segment edges.

$$X(f) = FFT(x(t))$$

where x(t) is the time-domain signal, and X(f) is the Fourier-transformed signal in the frequency domain.

Perform Fourier Transform:

Following windowing, the Fast Fourier Transform (FFT) is applied to each segment. FFT converts the time-domain signal into the frequency domain, revealing the frequency components of the signal. Mathematically, this is represented as:

$$P(f) = |X(f)|^2$$

where P(f) represents the power at each frequency f.

Calculate Magnitude Squared:

The power for each frequency component is then computed by squaring the magnitude of the Fourier transform. The squared magnitude provides the power at each frequency:

$$PSD(f) = \frac{N.\Delta f}{|X(f)|^2}$$

Normalize the Power:

The power is normalized to compute the Power Spectral Density (PSD). The normalization takes into account the number of data points and the frequency resolution. For each segment, the PSD can be expressed as:

$$PSD(f) = \lim_{T\to\infty}(1/T)((|X(f,T)|)^2)$$

where f is the frequency, and $X(f,T)$ is the Fourier transform of x(t)

Average PSD Across Segments:

If the signal has been divided into multiple segments, the PSD for each segment is computed separately. The final PSD is then the average of all segment-wise PSDs, which helps reduce variance caused by noise:

$$PSD_{final}(f) = \frac{\sum_{i=1}^{N} PSD_i(f)}{M}$$

where M is the total number of segments and $PSD_i(f)$ is the PSD for the i^{th} segment.

Provide the PSD Results:

Finally, the method yields two outputs: the frequency bins and the corresponding PSD values. The frequency bins represent the frequencies present in the signal, while the PSD values indicate the power at those frequencies.

Classification Algorithms

Neural Networks (NNs) consist of an input layer, one or more hidden layers, and an output layer. Each neuron applies an activation function to its inputs, and the network learns by adjusting weights based on the error in predictions. Neural networks are adept at capturing nonlinear patterns in data, making them suitable for classification tasks with complex relationships [25].

Deep Neural Networks (DNNs) are a more advanced form of neural networks with multiple hidden layers. DNNs can model complex and abstract features from data, providing a deeper understanding of intricate patterns. This capability makes them especially effective for tasks where manual feature extraction is difficult, as they learn features directly from the data [26].

Algorithm 1: Feature Extraction and Classification Using NN and DNN

Step 1: Non-linear Feature Extraction
Initialize feature list
Extract features from each brain wave using Power Spectral Density (PSD)
Compute PSD for different frequency bands
Features = concatenate (PSD features)
Append extracted features to Fm
Step 2: Cross Validation
Split Features into 58% Training_Data, 13% Validation_Data, 29% Testing_Data
Validation_Results = crossValidate(Training_Data, Validation_Data, Testing_Data, k = 10)
Calculate average performance metrics
Step 3: Data Preparation for Neural Networks
Split data into Reduced_Data and Target
X = Reduced_Data // All columns except target
Y = Target // Sleep stages class column (W, N1, N2, N3, R)
Split data into training and testing sets
trainingData, testingData = split(X, Y, testSize=0.2)
Step 4: Classification and Performance Evaluation Using NN and DNN
1. Neural Network (NN) Architecture
NN_model = Sequential() // For shallow neural network
Add spatial convolutional layer: NN_model.add(Conv2d(1, 2, kernel_size=(2, 1), stride=(1, 1)))
- NN_model.add(Conv2d(1, 8, kernel_size=(1, 50), stride=(1, 1), padding=(0, 25)))
- NN_model.add(ReLU())
- NN_model.add(MaxPool2d(kernel_size=(1, 12), stride=(1, 12)))
- NN_model.add(Conv2d(8, 8, kernel_size=(1, 50), stride=(1, 1), padding=(0, 25)))
- NN_model.add(ReLU())
- NN_model.add(MaxPool2d(kernel_size=(1, 12), stride=(1, 12)))
- Fully connected layer:
- NN_model.add(Dropout(p=0.25))
- NN_model.add(Linear(in_features=320, out_features=5)) // 5 output classes: W, N1, N2, N3, R
2. Deep Neural Network(DNN) Architecture
DNN_model = Sequential()
- Add input layer: DNN_model.add(Dense(128, input_dim=number_of_features,activation='relu'))
- Add hidden layers:
- DNN_model.add(Dense(64, activation='relu'))
- DNN_model.add(Dense(32, activation='relu'))
- Add output layer: DNN_model.add(Dense(5, activation='softmax'))
3. Model Training
- For NN: NN_model.fit(trainingData.Reduced_Data, trainingData.Target, epochs=10, batch_size=32, validation_split)
- For DNN: DNN_model.fit(trainingData.Reduced_Data, trainingData.Target, epochs=10, batch_size=32, validation_split)
4. Model Evaluation

- NN_predictions = NN_model.predict(testingData.Reduced_Data)
- DNN_predictions = DNN_model.predict(testingData.Reduced_Data)

6. Performance Metrics Calculation: Confusion matrix, accuracy, precision, recall, specificity, F1 score, and Cohen's kappa score are calculated:
   - cm = confusion_matrix(y_true, y_pred, labels=[0, 1, 2, 3, 4])
   - Metrics like sensitivity , specificity , precision , accuracy, F1-score, and others are computed as described in the code above.
   - Visualize metrics like confusion matrix, accuracy, precision, recall, specificity, and F1 scores using plots.

End of Algorithm

Performance Evaluation

To evaluate the effectiveness of these classifiers, the following steps are performed:

- Data Preparation: This involves preprocessing the EEG data, which includes normalization, feature extraction, and splitting the data into training and testing sets.
- Training and Testing: The classifiers are trained on the training data and then assessed using the test data to gauge their performance.
- Performance Evaluation: Various metrics derived from the confusion matrix are used to measure classifier performance:

Sensitivity (*Sn*): Indicates the proportion of true positives among all actual positive cases[30].

$$Sn = \frac{TP}{TP + FN}$$

Specificity (*Sp*): Reflects the classifier's ability to correctly identify negative instances[31].

$$Sp = \frac{TN}{TN + FP}$$

Precision (*Pr*): This represents the ratio of true positives to the model's positive predictions[32].

$$Pr = \frac{TP}{TP + FP}$$

Accuracy (*Acc*): Shows the proportion of correct predictions (true positives and true negatives) relative to the total number of cases[33].

$$Acc = \frac{TP}{TP + FP + TN + FN}$$

F-measure (F1): Provides a single metric balancing precision and recall[34].

$$F1 = \frac{2 \times (Pr \times Sn)}{Pr + Sn}$$

Cohen's Kappa (k): Compares observed accuracy with the expected accuracy, accounting for random chance[35].

$$k = \frac{2 \times \big((TP \times TN) - (FN \times FP)\big)}{\big((FN + TN) \times (TP + FN)\big) + \big((TP + FP) \times (FP + TN)\big)}$$

Additional metrics include macro averages for precision, recall, and specificity, calculated as:

Macro Precision:

$$MPrecision = \frac{\sum_{i=1}^{N} Pr}{N}$$

Macro Recall:

$$MRecall = \frac{\sum_{i=1}^{N} Sn}{N}$$

Macro Specificity:

$$MSpecificity = \frac{\sum_{i=1}^{N} Sp}{N}$$

Result Analysis and Discussion

The study involved conducting simulations and tests to produce meaningful outcomes for identifying the best model in terms of both performance and architectural simplicity. As mentioned previously, the experimental results are derived from EEG signal data retrieved from the Physionet database [32]. This data was then processed and broken down using a series of brainwave filters such as FIR filters, in order to extract the key characteristics of each wave to classify sleep stages.Fig.7 and 8 presents the training samples used to train the obtained NN and DNN models.

The assessment of the examined models involved three main stages. To begin with, EEG signals were gathered as the initial data for the analysis. Next, the signals were processed and relevant features were extracted using filtering techniques to eliminate noise from the EEG signals. Lastly, sophisticated learning techniques such as Deep Neural Networks (DNN), Neural Networks (NN), or Multilayer Perceptron (MLP) were utilized to categorize the processed features into six distinct sleep stages (W, N1, N2, N3, R). Table 3 presents the configuration details of the proposed models.The proposed classification algorithm was developed using VSCode software version 1.89.1, running on a LAPTOP-UN1TDLNH with a Windows operating system. The system is equipped with an Intel(R) Core (TM) i5-10300H CPU @ 2.50 GHz processor and 8 GB of RAM (7.84 GB usable). It is a 64-bit operating system on an x64-based processor.

Table 2. Configuration of the proposed models

Classifiers	Configuration (Structure and training parameter)	Training Algorithm
NN	Layers: 7 (convolutional, max pooling, dropout, fully connected (linear) Number of convolutional layer: 3 Number of max pooling layer: 2 Number of dropout layer: 1 Number of fully connected layer: 1 Activation Functions: ReLU activation function for all layers.	Backpropagation
DNN	Layers: 5 (Input, 3 Hidden, Output) Hidden layer neurons: [128, 64, 32] Activation function: ReLU Output layer activation: Softmax Epochs: 10	Adam optimizer

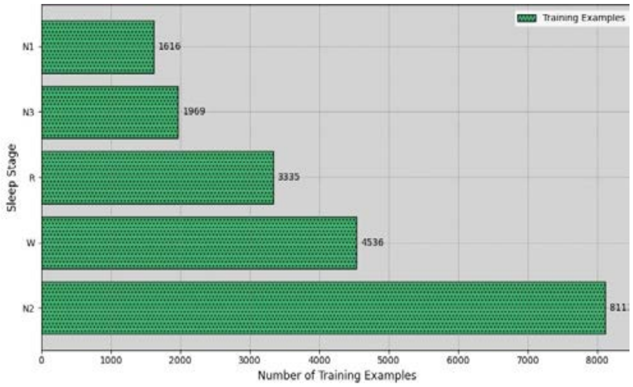


Figure 4. Training sample distributions over the individual sleep stages for training the NN model

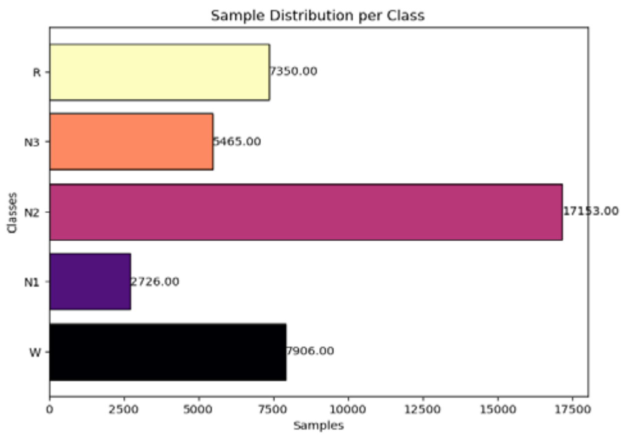
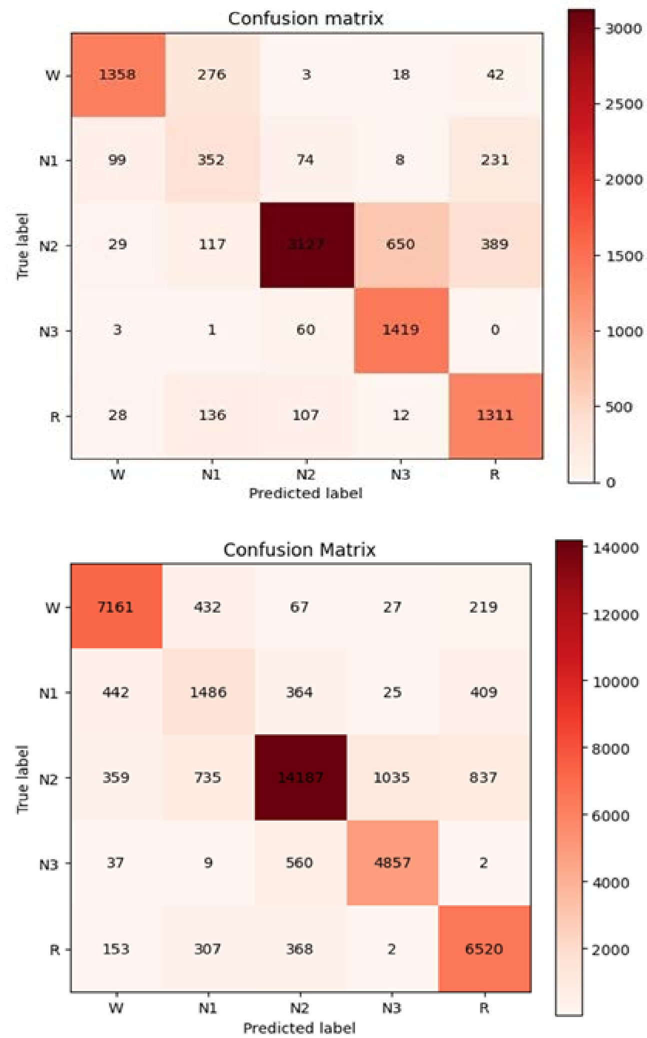


Figure 5. Training sample distributions over the individual sleep stages for training the DNN model

Figure 9(a-b) represents the confusion matrix for the obtained NN and DNN classifiers. The Neural Network (NN) model demonstrates robust performance in classifying stage N2, with a high number of correctly identified instances, highlighted by the deep red color in its confusion matrix, which is 3127. However, it exhibits significant misclassifications, especially confusing stage W with N1 and R with N2. This suggests a need for improved feature refinement or model adjustments. In contrast, the Deep Neural Network (DNN) model shows a more balanced performance across all sleep stages, with notable success in classifying N1. Despite this, it still struggles with misclassifications, particularly between stages W and R, and N3 and R, which may indicate areas for further optimization.

Table 3. Performance metrics results obtained from NN and DNN model

NN				DNN			
	Precision	**Recall**	**F1 Score**		**Precision**	**Recall**	**F1 Score**
W	0.85	0.8	0.82	**W**	0.89	0.96	0.93
N1	0.75	0.7	0.72	**N1**	0.48	0.49	0.48
N2	0.9	0.85	0.87	**N2**	0.89	0.87	0.88
N3	0.8	0.75	0.77	**N3**	0.86	0.83	0.85
R	0.7	0.65	0.68	**R**	0.82	0.83	0.83



a)

(b)

Figure 6. Confusion matrix results for classification of each sleep stage:(a) using the NN model, (b) using the DNN model

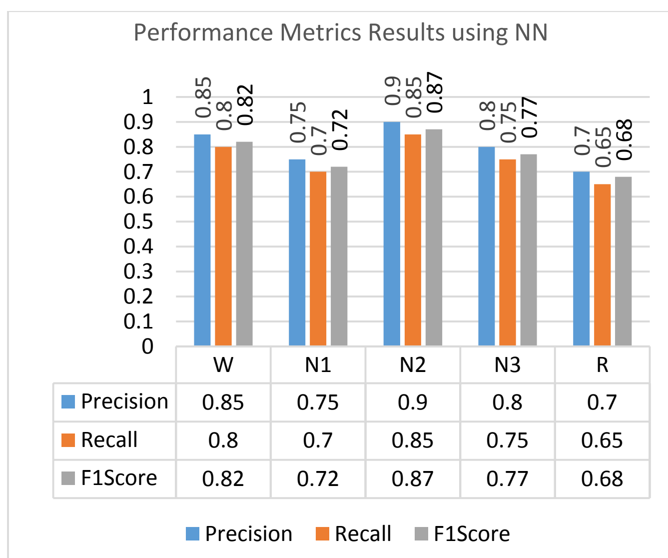


Figure 7. The performance of the different sleep stages about precision, Recall, and F1Score using the NN model

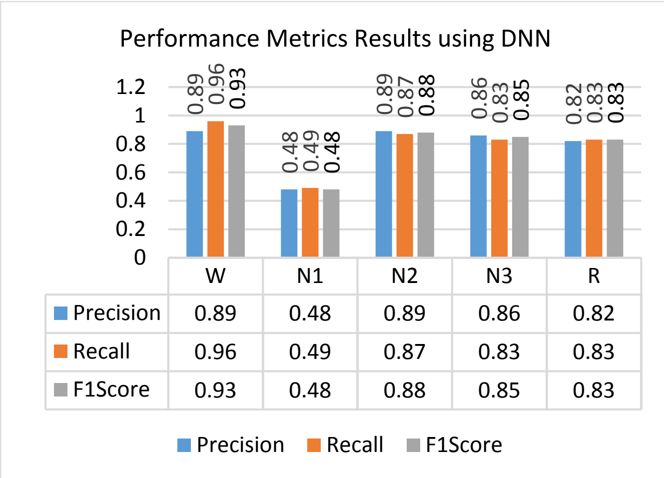


Figure 8. The performance of the different sleep stages about precision, Recall, and F1Score using the DNN model

Figures 11 and 12 reveal that the N2 stage is the best-performing stage for both the Neural Network (NN) and Deep Neural Network (DNN). In the W stage, the DNN surpasses the NN in all performance metrics. Conversely, in the N1 stage, the NN outperforms the DNN across all metrics. For the N2 stage, the performance of the DNN and NN is comparable, with the DNN achieving a slightly higher F1 Score. The DNN consistently outperforms the NN in the N3 and R stages across all metrics.Fig.12 and Fig.13 represent the performance of the macro and weighted precision, recall, and F1score.

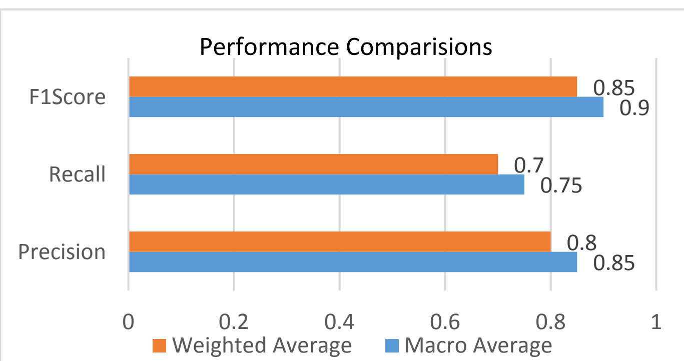


Figure 9. The performance of the different sleep stages about macro and weighted precision, Recall, and F1Score using the NN model

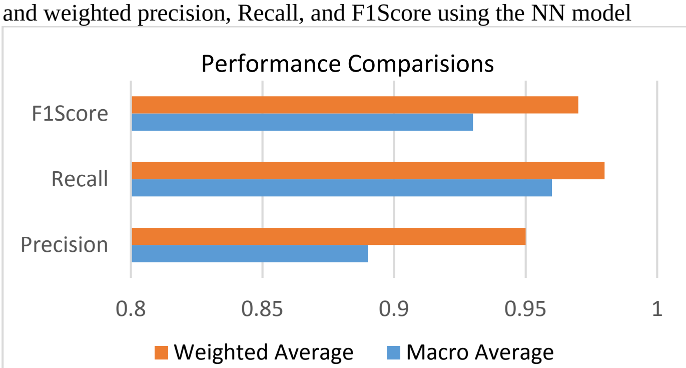


Figure 10. The performance of the different sleep stages about macro and weighted precision, Recall, and F1Score using the DNN model

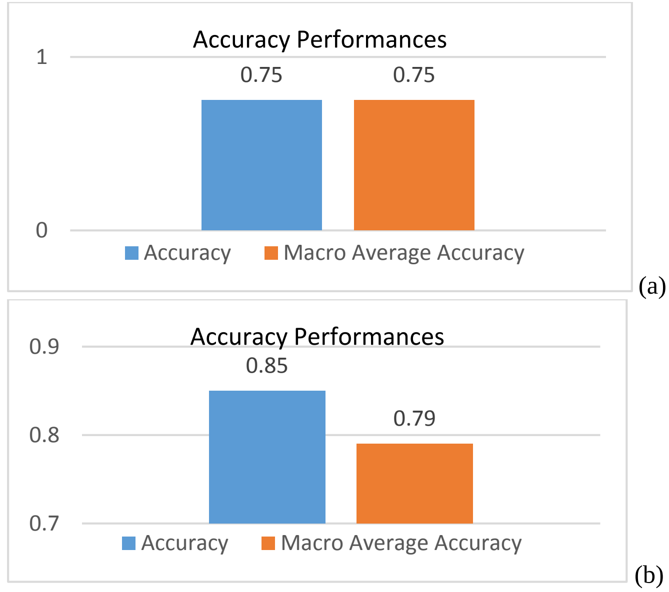
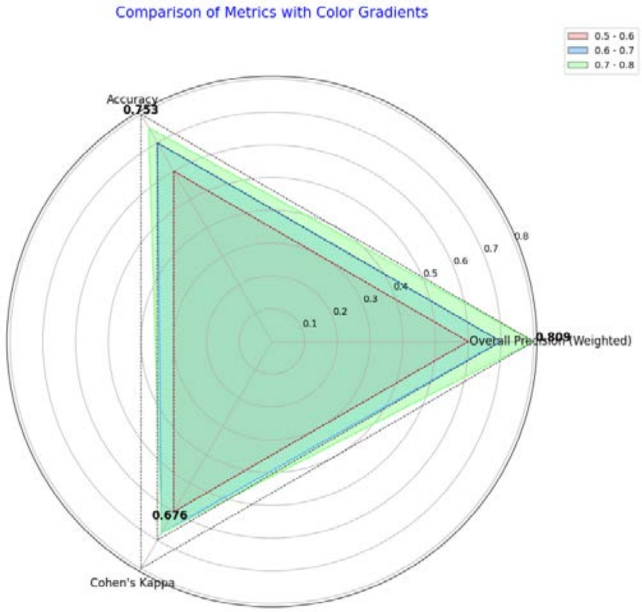


Figure 11. Weighted and Macro average accuracy performance of the obtained model:(a) using NN, (b) DNN model

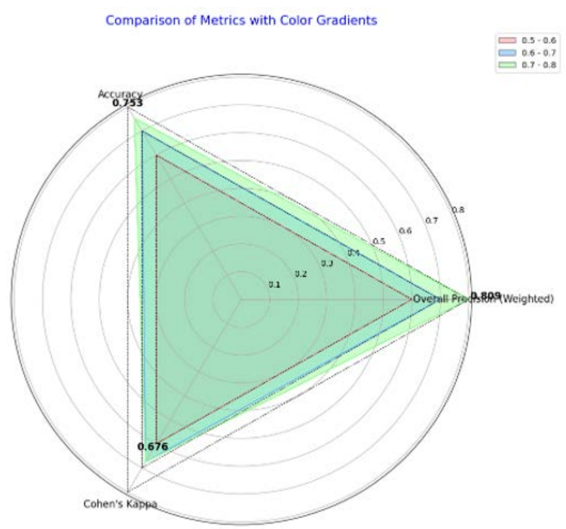
Figures 14(a) and (b) show the overall and macro accuracy performances for the proposed models. Similarly to better understand the behavior of the model, we used the radar graph representation. Figures 15 (a) and (b) illustrate represents radar graph presentation. Figure 16 (a) and (b) illustrate the hypnogram produced manually by a sleep expert and its corresponding hypnogram generated by our method for a subject for approximately 8 hours of sleep at night. It can be noted from the figure that around 85% manually scored a hypnogram and automatically scored correctly. The general sleep stage patterns are effectively represented by both models, but the DNN model exhibits a slightly more accurate match with the actual sleep stages than the NN model. The NN model encounters difficulty in differentiating between N2 and N3 stages, especially during rapid transitions. On the other hand, the DNN model displays enhanced precision in recognizing wakefulness (W) and REM sleep (R) stages, underscoring its superior performance in these specific areas.

Comparison with Prior Work

To further demonstrate the performance of the proposed NN and DNN model, a comparative analysis was conducted with state-of-the-art sleep staging techniques using the same publicly available dataset, Sleep-EDF Expanded. Many studies on automated sleep stage identification have been published in the literature. In this work, we proposed a new method for automatically detecting sleep disorders using two EEG channels (A1-C4 & C4-F4). We utilized both the combination of these channels and their components to classify the data [41]. The results showed that combining both EEG channels improved the classification accuracy. Our proposed model not only distinguished between participants with and without sleep disorder behavior but also identified the specific type of abnormalities in the sleep pattern. Notably, the N3 sleep stage was found to be more accurately classified compared to other stages. Table 5 summarizes the experimental results from earlier

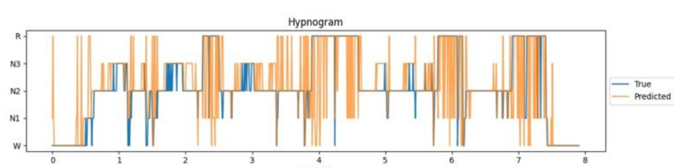


(a)

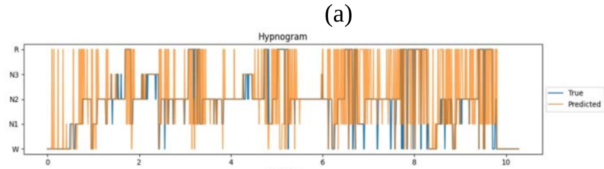


(b)

Figure 12. Radar graph comparing classification algorithms based on accuracy, precision and Cohen's Kappa:(a) using the NN model, (b) using the DNN model



(a)



(b)

Figure 13. The hypnogram scored by a human expert and the hypnogram scored by the proposed model: (a) using the NN model, (b) using the DNN model.

studies and the results obtained by applying the approach presented in this paper to Sleep-EDF Expanded. The overall performance of all sleep staging models was evaluated using two metrics, accuracy, and Cohen's kappa (κ). We successfully attained a maximum accuracy of classification of 85 percent by combining two EEG channels A1-C4 & C4-F4 using a DNN classifier.

Table 4. Comparison of our method with previous studies implemented on Sleep-EDF Expanded

Model	Channels	Input	Classifier	Accuracy	κ
Ref [38]	Fpz-Cz EEG Pz-Oz EEG ROC-LOC EOG	STFT	Multi-Channel CNN	85.0	0.79
Ref [47]	Fpz-Cz EEG	FRFT	Bi-LSTM	81.61	0.7468
Ref [44]	Fpz-Cz EEG ROC-LOC EOG	FFT	FCNN+RNN	84.0	0.778
Ref [45]	Fpz-Cz EEG	raw EEG and CWT	Double-branch CNN	81.0	0.73
	Pz-OZ EEG			77.4	0.68
Ref [39]	Fpz-Cz EEG Pz-Oz EEG ROC-LOC EOG	cross-epoch vector	LSTM	84.3	0.78
Ref [40]	Fpz-Cz EEG	raw EEG and the spectral sequence	Two-Stream Feature Extractor	82.21	0.7507
Ref [41]	Fpz-Cz EEG	noise-assisted bivariate empirical mode decomposition	1D-CNN+ BiLSTM	82.67	0.76
	Pz-OZ EEG			80.16	0.72
Ref [42]	Fpz-Cz EEG	raw EEG	CNN	83.9	0.78
	Pz-Oz EEG			78.8	0.71
Ref [43]	Fpz-Cz EEG	The long-term temporal context between consecutive EEG	CNN+RNN	82.7	0.76
Ours	**Fpz-Cz EEG**	**Raw EEG and PSD**	**NN, DNN**	**85%**	**0.84**

Conclusion

In this research, we conducted an assessment and comparison of the efficacy of a Neural Network (NN) and a Deep Neural Network (DNN) in categorizing five sleep stages using the Physionet EEG dataset. Our analysis aimed to evaluate both architectures in terms of precision, recall, and F1-score. The NN achieved an overall accuracy of 75.4%, while the DNN demonstrated a notably higher accuracy of 85%. The findings reveal that the DNN surpasses the NN in most metrics, particularly exhibiting improvement in the wake (W) and rapid eye movement (R) sleep stages. The DNN's higher precision and recall values for these stages indicate its strength in capturing subtle EEG patterns that differentiate various sleep stages. Conversely, the NN, while effective to some extent, displayed lower classification performance, especially for stages N1 and R, signifying limitations in discerning finer sleep stage disparities. This comparison underscores the benefit of utilizing deeper network architectures, such as DNN, for tasks involving intricate signal patterns like EEG data for sleep stage classification. The increased depth of the DNN facilitates a more thorough feature extraction process, resulting in enhanced overall classification accuracy. Nevertheless, employing DNN entails higher computational expenses and complexity. Although its superior performance mitigates this, future research should focus on optimizing the architecture to further balance accuracy and efficiency. It is also essential to conduct real-time testing to validate the practical applicability of these models in real-world sleep monitoring systems.

Despite the constraints associated with using a single-channel EEG, the current results suggest that DNNs offer substantial potential for advancing automated sleep stage detection with increased accuracy and dependability.

References

1. N. Baranwal, K. Y. Phoebe, N. S. Siegel. Sleep Physiology, Pathophysiology, and Sleep Hygiene. *Prog. Cardiovasc. Dis.* **2023**, *77*, 59–69.
2. S. K. Satapathy, D. Loganathan. SleepEEG Study: Automated Sleep Stage Scoring with Machine Learning Techniques from Dual-Channel Brain-EEG Signals. In *Soft Computing for Problem Solving*; Tiwari, A., Ahuja, K., Yadav, A., Bansal, J. C., Deep, K., Nagar, A. K., Eds.; Springer: Singapore, 2021; pp 239–249.
3. S. K. Satapathy, D. Loganathan. Automated Classification of Sleep Stages Using Single-Channel EEG: A Machine Learning-Based Method. *Int. J. Inf. Retr. Res.* **2022**, *12* (2), 1–19.
4. J. R. Franco, G. Cecchi, G. Priotto, M. Paone, A. Diarra, L. Grout, P. P. Simarro, W. Zhao, D. Argaw. Monitoring the Elimination of Human African Trypanosomiasis at Continental and Country Level: Update to 2018. *PLoS Negl. Trop. Dis.* **2020**, *14* (5), e0008261.
5. S. K. Satapathy, S. Pattnaik, B. Acharya, R. K. Rath. Automated Classification of Sleep Stages Using Single-Channel EEG Signal: A Machine Learning-Based Method. In *Advances in Computing and Data Sciences*; Singh, M., Tyagi, V., Gupta, P. K., Flusser, J., Ören, T., Eds.; Springer: Singapore, 2022; pp 268–278.
6. S. K. Satapathy, D. Loganathan. Automated Classification of Multi-Class Sleep Stages Using Polysomnography Signals: A Nine-Layer 1D-Convolution Neural Network Approach. *Multimed. Tools Appl.* **2022**, *81*, 12345–12360.
7. N. Sridhar, A. Shoeb, P. Stephens, A. Kharbouch, D. B. Shimol, J. Burkart, A. Ghoreyshi, L. Myers. Deep Learning for Automated Sleep Staging Using Instantaneous Heart Rate. *NPJ Digit. Med.* **2020**, *3*, 106.
8. S. K. Satapathy, D. Loganathan. Prognosis of Automated Sleep Staging Based on Two-Layer Ensemble Learning Stacking Model Using Single-Channel EEG Signal. *Soft Comput.* **2021**, *25*, 15445–15462.
9. S. K. Satapathy, A. K. Bhoi, D. Loganathan, B. Khandelwal, P. Barsocchi. Machine Learning with Ensemble Stacking Model for Automated Sleep Staging Using Dual-Channel EEG Signal. *Biomed. Signal Process. Control* **2021**, *69*, 102898.
10. A. El Hadiri, L. Bahatti, A. El Magri, R. Lajouad. Brain Signals Analysis for Sleep Stages Detection Using Virtual Instrumentation Platform. *Indones. J. Electr. Eng. Comput. Sci.* **2023**, *29*, 761–768.
11. F. Ebrahimi, I. Alizadeh. Automatic Sleep Staging by Cardiorespiratory Signals: A Systematic Review. *Sleep Breath.* **2021**, *25*, 1–17.
12. H. Zhu, C. Fu, F. Shu, H. Yu, C. Chen, W. Chen. The Effect of Coupled Electroencephalography Signals in Electrooculography Signals on Sleep Staging Based on Deep Learning Methods. *Bioengineering* **2023**, *10* (5), 573.
13. Z. Cui, X. Zheng, X. Shao, L. Cui. Automatic Sleep Stage Classification Based on Convolutional Neural Network and Fine-Grained Segments. *Complexity* **2018**, *2018* (1), 9248410.
14. P. Fonseca, M. Ross, A. Cerny, P. Anderer, F. van Meulen, H. Janssen, A. Pijpers, S. Dujardin, P. van Hirtum, M. van Gilst, et al. A Computationally Efficient Algorithm for Wearable Sleep Staging in Clinical Populations. *Sci. Rep.* **2023**, *13* (1), 9182.
15. Z. Khomsi, M. Elfezazi, L. Bellarbi. Deep Learning-Based Approach in Surface Thermography for Inverse Estimation of Breast Tumor Size. *Sci. Afr.* **2024**, *23*, e01987.
16. M. Diykh, Y. Li, P. Wen. EEG Sleep Stages Classification Based on Time Domain Features and Structural Graph Similarity. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *24* (11), 1159–1168.
17. K. Fu, J. Qu, Y. Chai, T. Zou. Hilbert Marginal Spectrum Analysis for Automatic Seizure Detection in EEG Signals. *Biomed. Signal Process. Control* **2015**, *18*, 179–185.
18. M. Sharma, D. Goyal, P. Achuth, U. R. Acharya. An Accurate Sleep Stages Classification System Using a New Class of Optimally Time-Frequency Localized Three-Band Wavelet Filter Bank. *Comput. Biol. Med.* **2018**, *98*, 58–75.
19. C. Li, H. Deng, S. Yin, C. Wang, Y. Zhu. sEMG Signal Filtering Study Using Synchrosqueezing Wavelet Transform with Differential Evolution Optimized Threshold. *Results Eng.* **2023**, *18*, 101150.
20. A. El Hadiri, L. Bahatti, A. El Magri, R. Lajouad. Sleep Stages Detection Based on Analysis and Optimisation of Non-Linear Brain Signal Parameters. *Results Eng.* **2024**, *23*, 102664.
21. S. Mousavi, F. Afghah, U. R. Acharya. SleepEEGNet: Automated Sleep Stage Scoring with Sequence to Sequence Deep Learning Approach. *PLoS One* **2019**, *14* (5), e0216456.
22. D. L. Frédéric, A. Rouijel, L.-C. Iyad, H. Elghazi. An Iterative Thresholding-Based Algorithm for Brain–Computer Interface (BCI) Application. *Sci. Afr.* **2023**, *21*, e01832.
23. S. K. Satapathy, B. Brahma, B. Panda, P. Barsocchi, A. K. Bhoi. Machine Learning-Empowered Sleep Staging Classification Using Multi-Modality Signals. BMC Med. Inform. Decis. Mak. 2024, 24 (1), 119.
24. E. A. Wolpert. A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. *Arch. Gen. Psychiatry* **1969**, *20* (2), 246–247.
25. M. Ronzhina, O. Janoušek, J. Kolářová, M. Nováková, P. Honzík, I. Provazník. Sleep Scoring Using Artificial Neural Networks. *Sleep Med. Rev.* **2012**, *16* (3), 251–263.

26. T. Lajnef, S. Chaibi, P. Ruby, P.-E. Aguera, J.-B. Eichenlaub, M. Samet, et al. Learning Machines and Sleeping Brains: Automatic Sleep Stage Classification Using Decision-Tree Multi-Class Support Vector Machines. *J. Neurosci. Methods* **2015**, *250*, 94–105.
27. A. R. Hassan, A. Subasi. A Decision Support System for Automated Identification of Sleep Stages from Single-Channel EEG Signals. *Knowl.-Based Syst.* **2017**, *128*, 115–124.
28. C. Berthomier, V. Muto, C. Schmidt, G. Vandewalle, M. Jaspar, J. Devillers, et al. Exploring Scoring Methods for Research Studies: Accuracy and Variability of Visual and Automated Sleep Scoring. *J. Sleep Res.* **2020**, *29* (5), e12994.
29. X. Li, Y. Gong, X. Jin, P. Shang. Sleep Posture Recognition Based on Machine Learning: A Systematic Review. *Pervasive Mob. Comput.* **2023**, 101752.
30. L. Liu, W. Zhang, Y. Song, Y. Jiang, X. Yu. Radar Emitter Structure Identification Based on Stacked Frequency Sparse Auto-Encoder Network. *IET Radar, Sonar Navig.* **2023**, *17* (5), 693–703.
31. T. Kayikcioglu, M. Maleki, K. Eroglu. Fast and Accurate PLS-Based Classification of EEG Sleep Using Single Channel Data. *Expert Syst. Appl.* **2015**, *42* (21), 7825–7830.
32. W. Zhou, Z. Feng, Y. F. Xu, X. Wang, H. Lv. Empirical Fourier Decomposition: An Accurate Signal Decomposition Method for Nonlinear and Non-Stationary Time Series Analysis. *Mech. Syst. Signal Process.* **2022**, *163*, 108155.
33. P. Flandrin, P. Gonçalves, G. Rilling. EMD Equivalent Filter Banks, from Interpretation to Applications. In *Hilbert-Huang Transform and Its Applications*; 2005; pp 57–74.
34. X. Chang, C. Shen, S. Liu, D. Zheng, S. Wang, C. Yang, et al. Robust Kramers–Kronig Holographic Imaging with Hilbert–Huang Transform. *Opt. Lett.* **2023**, *48* (15), 4161–4164.
35. Z. L. He, L. D. Du, P. Wang, P. Xia, Z. Liu, Y. L. Song, et al. Single-Channel EEG Sleep Staging Based on Data Augmentation and Cross-Subject Discrepancy Alleviation. *Comput. Biol. Med.* **2022**, *149*, 105996.
36. C. H. Zhao, J. B. Li, Y. H. Guo. SleepContextNet: A Temporal Context Network for Automatic Sleep Staging Based on Single-Channel EEG. *Comput. Methods Programs Biomed.* **2022**, *220*, 106806.
37. Y. Dai, X. L. Li, S. S. Liang, L. Wang, Q. T. Duan, H. Yang, et al. MultiChannelSleepNet: A Transformer-Based Model for Automatic Sleep Stage Classification with PSG. *IEEE J. Biomed. Health Inform.* **2023**, *27* (9), 4204–4215.
38. R. C. Li, B. Wang, T. Zhang, T. Sugi. A Developed LSTM-Ladder-Network-Based Model for Sleep Stage Classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 1418–1428.
39. G. D. Fu, Y. Y. Zhou, P. L. Gong, P. P. Wang, W. Shao, D. Q. Zhang. A Temporal-Spectral Fused and Attention-Based Deep Model for Automatic Sleep Staging. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 1008–1018.
40. J. Baek, C. Lee, H. Yu, S. Baek, S. Lee, S. Lee, et al. Automatic Sleep Scoring Using Intrinsic Mode Based on Interpretable Deep Neural Networks. *IEEE Access* **2022**, *10*, 36895–36906.
41. Z. L. He, L. D. Du, P. Wang, P. Xia, Z. Liu, Y. L. Song, et al. Single-Channel EEG Sleep Staging Based on Data Augmentation and Cross-Subject Discrepancy Alleviation. *Comput. Biol. Med.* **2022**, *149*, 105996.
42. C. H. Zhao, J. B. Li, Y. H. Guo. SleepContextNet: A Temporal Context Network for Automatic Sleep Staging Based on Single-Channel EEG. *Comput. Methods Programs Biomed.* **2022**, *220*, 106806.
43. H. Phan, O. Y. Chén, M. C. Tran, P. Koch, A. Mertins, M. De Vos. XSleepNet: Multi-View Sequential Model for Automatic Sleep Staging. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44* (9), 5903–5915.
44. Y. K. J. Fang, Y. Xia, P. Chen, J. Zhang, Y. L. Zhang. A Dual-Stream Deep Neural Network Integrated with Adaptive Boosting for Sleep Staging. *Biomed. Signal Process. Control* **2023**, *79*, 104150.
45. A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* **2000**, *101* (23), e215–e220.
46. Y. Y. You, X. Y. Zhong, G. Z. Liu, Z. H. Yang. Automatic Sleep Stage Classification: A Light and Efficient Deep Neural Network Model Based on Time, Frequency and Fractional Fourier Transform Domain Features. *Artif. Intell. Med.* **2022**, *127*, 102279.