Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Evaluation of Time–Frequency and Temporal Deep Learning Representations for EEG-Based Sleep Disorder Detection

Romain Atangana[a,b,c], Amstrong Emini Me Zenanga[a,b,c,], Vivien Beyala Kamgang[c,d], Emmanuel Baba[c], Perrin M. Li Litet[e], Daniel Tchiotsop[a], Godpromesse Kenné[a]

[a] *Unité de Recherche d'Automatique et d'Informatique Appliquée (UR-AIA), IUT of Bandjoun, University of Dschang, P.O. Box 134 Bandjoun, Bandjoun, Cameroon*
[b] *Unité de Recherche de Matière Condensée d'Electronique et de Traitement du Signal (UR-MACETS), Faculty of Science, University of Dschang, P.O. Box 67 Dschang, Dschang, Cameroon*
[c] *Department of Computer Science, Higher Teacher Training College, University of Bertoua, P.O. Box 652 Bertoua, Bertoua, Cameroon*
[d] *RTU – Mathematics and Computer Science, The University of Bertoua, P.O. Box 416, Bertoua, Cameroon*
[e] *RUFCEA (UR-IFIA), University of Dschang, P.O. Box 96, Dschang, Cameroon*

Abstract

Sleep disorders, including insomnia, narcolepsy, and sleep-related breathing disorders, are prevalent neurological conditions that significantly impair cognitive performance, emotional stability, and overall quality of life. Accurate and timely diagnosis remains a major clinical challenge, as the gold standard polysomnography requires multi-channel recordings and expert interpretation. In this work, we propose a comprehensive benchmarking framework for automated sleep disorder detection using single-channel Electroencephalogram (EEG) signals and deep learning models. Three approaches are systematically evaluated: a Convolutional Neural Network (CNN) trained on Continuous Wavelet Transform (CWT) scalograms, a CNN trained on Hilbert–Huang Transform (HHT) instantaneous amplitude maps, and a Temporal Convolutional Network (TCN) operating directly on raw EEG signals. Experiments are conducted on the publicly available CAP Sleep Database under a subject-independent evaluation protocol. The CNN + CWT model achieves the best overall performance with a validation accuracy of 84% and an Area Under the ROC Curve (AUC) of 0.890, followed by the TCN model (accuracy: 78%, AUC: 0.840) and the CNN + HHT model (accuracy: 74%, AUC: 0.780). All models substantially outperform random classification (AUC = 0.500). These results confirm that time–frequency representations, particularly CWT-based scalograms, capture more discriminative features from EEG signals compared to purely temporal or Hilbert-based approaches, while demonstrating the viability of single-channel EEG for automated sleep disorder screening.

Keywords: Continuous Wavelet Transform, Convolutional Neural Network, Electroencephalogram, Hilbert–Huang Transform, Sleep Disorders, Temporal Convolutional Network, Deep Learning

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

1. Introduction

Sleep is a fundamental physiological process essential for maintaining cognitive performance, emotional stability, and overall health. It is characterized by complex neurophysiological dynamics reflected in brain activity. Sleep is typically divided into rapid eye movement (REM) and non-rapid eye movement (NREM) stages, the latter being further subdivided into multiple stages representing increasing depth of sleep [1]. The cyclic alternation of these stages throughout the night provides crucial information about brain function and neurological health.

Sleep disorders constitute a broad class of conditions that disrupt normal sleep patterns and negatively impact quality of life. Among the most prevalent disorders are insomnia, narcolepsy, and sleep-related breathing disorders such as obstructive sleep apnea (OSA), which affects a significant portion of the adult population [2, 3]. These disorders are associated with cognitive impairment, reduced productivity, and increased risk of chronic diseases, making early and accurate diagnosis essential.

Polysomnography (PSG) is considered the gold standard for sleep analysis, involving the simultaneous recording of multiple physiological signals, including Electroencephalogram (EEG), Electrooculogram (EOG), and Electromyogram (EMG). Among these, EEG signals play a central role in characterizing sleep stages and detecting abnormalities due to their direct relationship with brain activity [4]. However, manual analysis of EEG recordings is time-consuming, subjective, and requires expert knowledge, motivating the development of automated approaches based on signal processing and machine learning techniques.

Recent advances in deep learning have significantly improved the performance of automated EEG analysis systems. Convolutional Neural Networks (CNNs) have been widely used to process time–frequency representations such as spectrograms and wavelet transforms, achieving promising results in sleep stage classification and disorder detection [5, 6]. In parallel, Temporal Convolutional Networks (TCNs) have emerged as powerful models for directly learning from raw temporal signals, capturing long-range dependencies without the need for explicit feature engineering [7].

Despite these advances, an important research question remains insufficiently explored: *how different representations of EEG signals affect the performance of deep learning models in sleep disorder classification*. In particular, time–frequency transformations such as Continuous Wavelet Transform (CWT) and Hilbert–Huang Transform (HHT) provide alternative ways of representing non-stationary EEG signals, yet their comparative effectiveness has not been thoroughly investigated in a unified framework.

In this study, we propose a comprehensive benchmarking framework to evaluate the impact of time–frequency and temporal representations on EEG-based sleep disorder classification. Specifically, we compare CNN-based models trained on CWT and HHT representations with a Temporal Convolutional Network (TCN) trained directly on raw EEG signals. The evaluation is conducted using a publicly available sleep EEG dataset under a rigorous experimental protocol.

The main contributions of this work are as follows:

- A unified benchmarking framework for comparing time–frequency and temporal deep learning approaches for EEG analysis;

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

- A comparative evaluation of CWT- and HHT-based representations for CNN models;
- The integration of a Temporal Convolutional Network (TCN) for direct modeling of raw EEG signals;
- An in-depth analysis of the strengths and limitations of each approach for sleep disorder classification.

2. Related Work

Automated analysis of sleep disorders using electroencephalogram (EEG) signals has attracted significant attention in recent years due to its potential for early diagnosis and continuous monitoring. Traditional approaches rely on handcrafted feature extraction techniques combined with classical machine learning algorithms. For instance, spectral, statistical, and time-domain features have been widely used with classifiers such as Support Vector Machines (SVM), Decision Trees, and Random Forests, achieving moderate performance in sleep stage classification and disorder detection [8, 9, 10].

Earlier studies have also explored signal processing techniques such as wavelet transforms and autoregressive modeling for feature extraction from EEG signals. For example, Estrada et al. [11] investigated multiple feature extraction schemes for neuro-fuzzy classification, while Tagluk et al. [12] applied wavelet-based features combined with artificial neural networks for sleep apnea detection. Although these approaches demonstrated promising results, they rely heavily on handcrafted features and domain expertise.

More recently, deep learning techniques have significantly improved EEG-based analysis by enabling automatic feature extraction. Convolutional Neural Networks (CNNs) have been widely adopted to process time–frequency representations of EEG signals, such as spectrograms and wavelet transforms. Studies such as DeepSleepNet [5] and subsequent works [6, 13, 14] have shown that CNN-based approaches can achieve state-of-the-art performance in sleep stage classification and related tasks.

In parallel, recent research has focused on reducing the complexity of EEG acquisition systems by using single-channel signals while maintaining acceptable performance. For instance, Giarrusso et al. [15] proposed a single-channel EEG-based framework for REM sleep behavior disorder detection, demonstrating the feasibility of low-cost and scalable diagnostic systems. Similarly, Melo et al. [16] validated sleep staging models using wearable EEG devices, highlighting the growing interest in portable and real-time monitoring solutions.

Despite these advances, most existing works rely on either time–frequency representations processed by CNNs or handcrafted temporal features combined with classical classifiers. Only limited studies have explored the direct modeling of raw EEG signals using temporal deep learning architectures. Temporal Convolutional Networks (TCNs), which have shown strong performance in sequence modeling tasks [7], remain underexplored in the context of sleep disorder classification.

Therefore, an important research gap persists regarding the comparative effectiveness of different EEG representations and modeling strategies. In particular, there is a lack of unified frameworks that systematically evaluate time–frequency representations such as Continuous Wavelet Transform (CWT) and Hilbert–Huang

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Transform (HHT) against temporal deep learning models operating on raw EEG signals.

In this work, we address this gap by proposing a comprehensive benchmarking framework that compares CNN-based models trained on CWT and HHT representations with a Temporal Convolutional Network (TCN) trained directly on raw EEG signals for sleep disorder classification.

3. Materials and Methods

3.1. Dataset Description

The experiments conducted in this study are based on the CAP Sleep Database, publicly available on PhysioNet [17]. This dataset consists of 108 polysomnographic (PSG) recordings collected at the Sleep Disorders Center of the Ospedale Maggiore of Parma, Italy. Each recording contains multiple physiological signals stored in European Data Format (EDF) files, including at least three EEG channels (F3, F4, or C4 referenced to A1 or A2), two Electrooculogram (EOG) channels, submental and tibialis anterior Electromyogram (EMG) signals, respiratory signals, and Electrocardiogram (ECG). Additional bipolar EEG channels are provided according to the international 10–20 system.

The database includes recordings from 16 healthy control subjects with no neurological disorders and free of medications affecting the central nervous system, as well as 92 pathological recordings from patients diagnosed with various sleep disorders, including: Nocturnal Frontal Lobe Epilepsy (NFLE, $n = 40$), REM sleep Behavior Disorder (RBD, $n = 22$), Periodic Leg Movements (PLM, $n = 10$), Insomnia ($n = 9$), Narcolepsy ($n = 5$), Sleep Disordered Breathing (SDB, $n = 4$), and Bruxism ($n = 2$). The recordings are named according to the associated pathology (e.g., `n1-n16` for healthy controls, `ins1-ins9` for insomnia, `narco1-narco5` for narcolepsy, `sdb1-sdb4` for breathing disorders).

In this study, only subjects diagnosed with Insomnia, Narcolepsy, and Sleep-Disordered Breathing, along with healthy controls, are retained for analysis, resulting in a focused subset of 34 recordings. Only a single EEG channel is selected per recording to ensure a simplified and scalable framework suitable for real-world and wearable monitoring applications.

The EEG recordings are segmented into fixed-length epochs for supervised classification. Each epoch is labeled as either *normal* (healthy) or *pathological* (sleep disorder). To ensure robustness and generalization, a subject-independent evaluation protocol is adopted, whereby data from the same subject do not appear in both training and testing sets.

3.2. Preprocessing of EEG Signals

3.2.1. Problem Formulation

In this study, sleep disorder classification is formulated as a binary problem, distinguishing between *normal* and *pathological* EEG signals. The objective is not to identify specific disorders, but to detect the presence of abnormal sleep patterns, aligning with a screening perspective commonly adopted in clinical practice.

This choice is also motivated by dataset constraints. As described in Section 3.1, the CAP Sleep Database contains a limited number of subjects with significant class imbalance, making multi-class classification statistically unreliable and prone

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

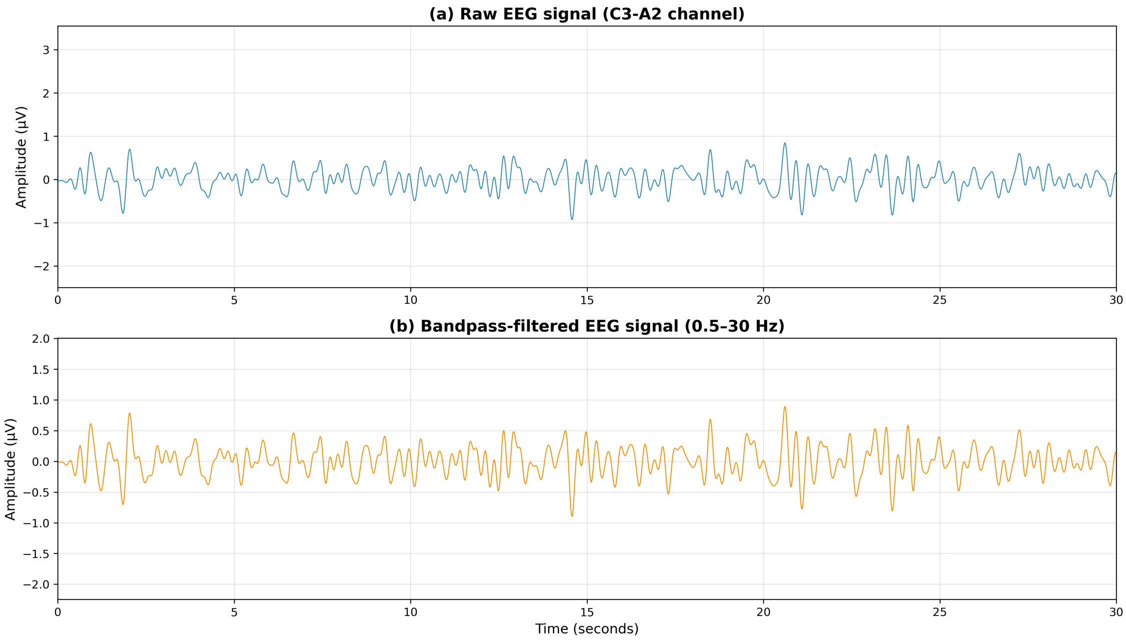

Figure 1: Raw EEG signal (C3-A2 channel) and bandpass-filtered signal (0.5–30 Hz).

to overfitting. Furthermore, similar binary formulations have been widely adopted in the literature for EEG-based disorder detection.

3.2.2. Visualization of Preprocessing Steps

Figures 1–6 illustrate the main preprocessing steps applied to the EEG signals.

Raw and filtered signals (Figure 1): The raw EEG signal exhibits low-frequency drifts and high-frequency noise. After bandpass filtering (0.5–30 Hz), the signal becomes cleaner, preserving physiologically relevant rhythms while removing artifacts.

Power spectral density (Figure 2): The spectral analysis confirms that filtering effectively suppresses frequencies outside the range of interest, particularly low-frequency baseline drift and high-frequency noise.

Time–frequency representation (Figure 3): The Continuous Wavelet Transform (CWT) provides a joint time–frequency representation of the EEG signal, highlighting transient oscillatory patterns across multiple scales.

Hilbert transform features (Figure 4): The instantaneous amplitude captures local signal energy variations, while the instantaneous phase reflects temporal signal dynamics.

Normalization (Figure 5): Z-score normalization standardizes the signal distribution, centering it around zero with unit variance.

Dataset distribution (Figure 6): The dataset exhibits a class imbalance between normal and pathological samples, as well as a subject-wise split between training and test sets.

3.3. Time–Frequency Representations

3.3.1. Continuous Wavelet Transform (CWT)

The Continuous Wavelet Transform (CWT) is a powerful mathematical tool for analyzing non-stationary signals such as EEG, providing a joint time–frequency

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

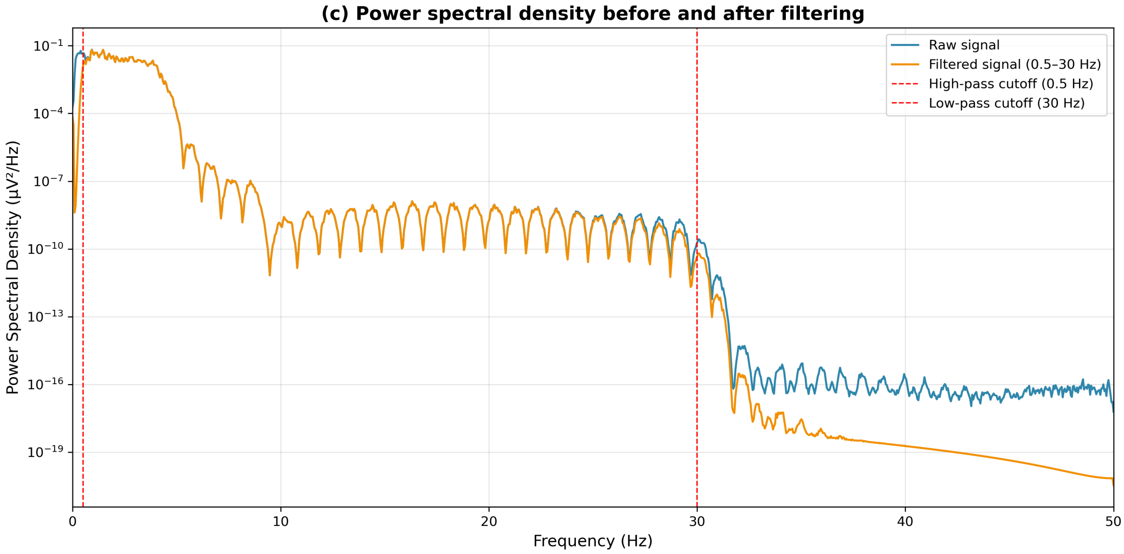

Figure 2: Power spectral density (PSD) of the EEG signal before and after bandpass filtering.

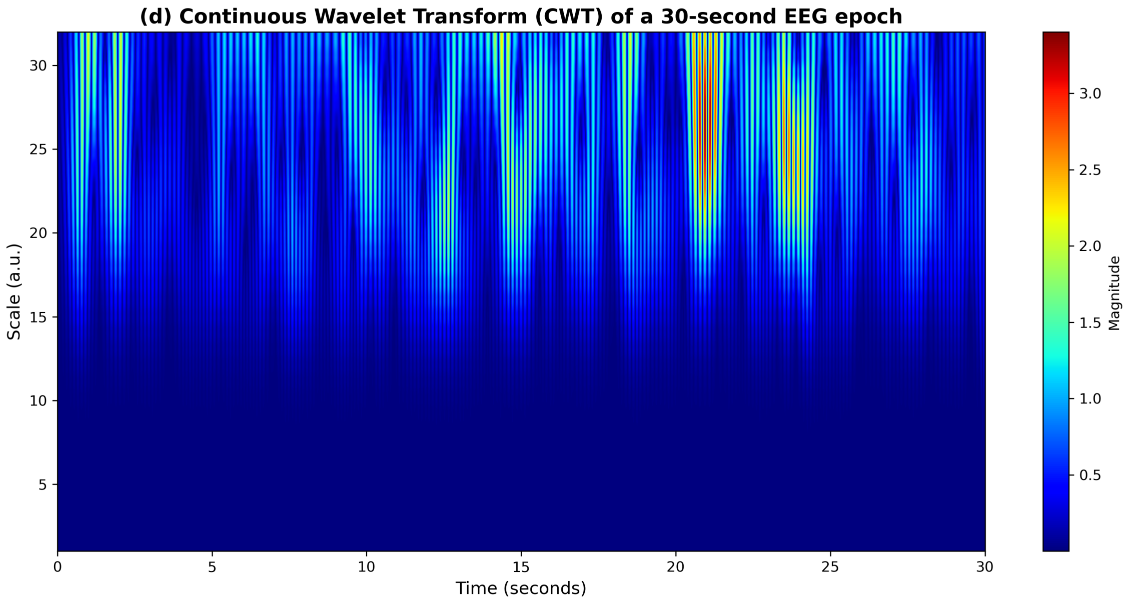

Figure 3: Continuous Wavelet Transform (CWT) of a 30-second EEG epoch.

representation that captures both temporal and spectral characteristics simultaneously [18, 19]. Unlike the classical Short-Time Fourier Transform (STFT), which uses a fixed window size, CWT offers multi-resolution analysis by adapting the window length to the frequency of interest, making it particularly suitable for biomedical signals that exhibit transient and oscillatory patterns across multiple time scales.

The CWT of a signal $x(t)$ is formally defined as:

$$CWT(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t)\, \psi^*\left(\frac{t-b}{a}\right) dt \tag{1}$$

where $\psi(t)$ denotes the mother wavelet, $a \in \mathbb{R}^+$ is the scale parameter controlling dilation, $b \in \mathbb{R}$ is the translation parameter controlling the temporal localization, and $(\cdot)^*$ denotes the complex conjugate. The scale parameter a is inversely related

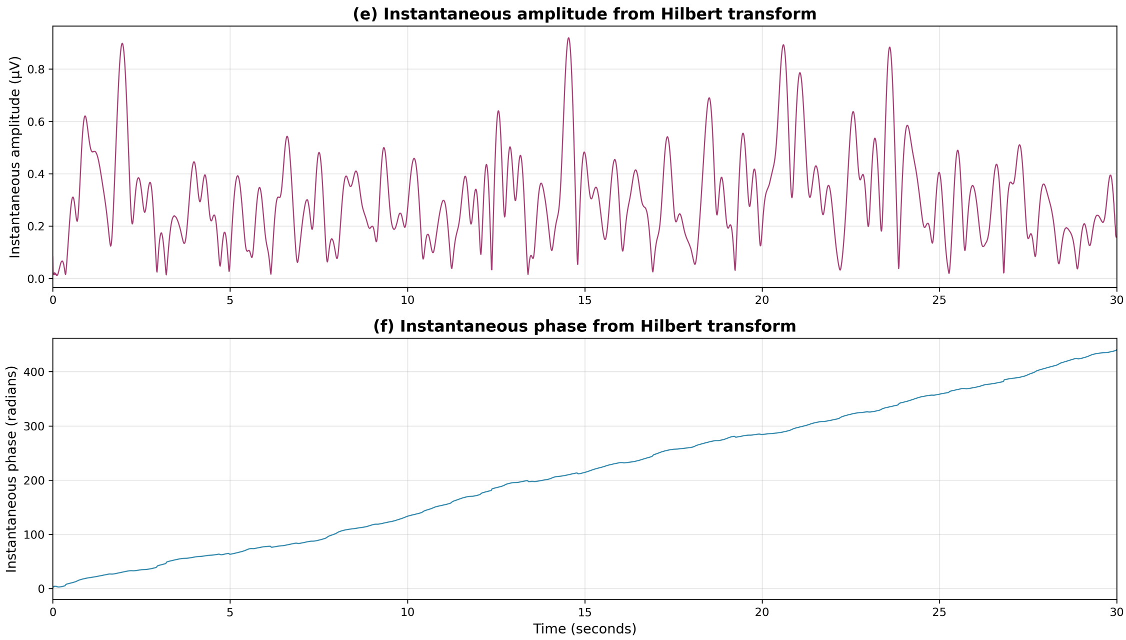

Figure 4: Instantaneous amplitude and phase extracted using the Hilbert transform.

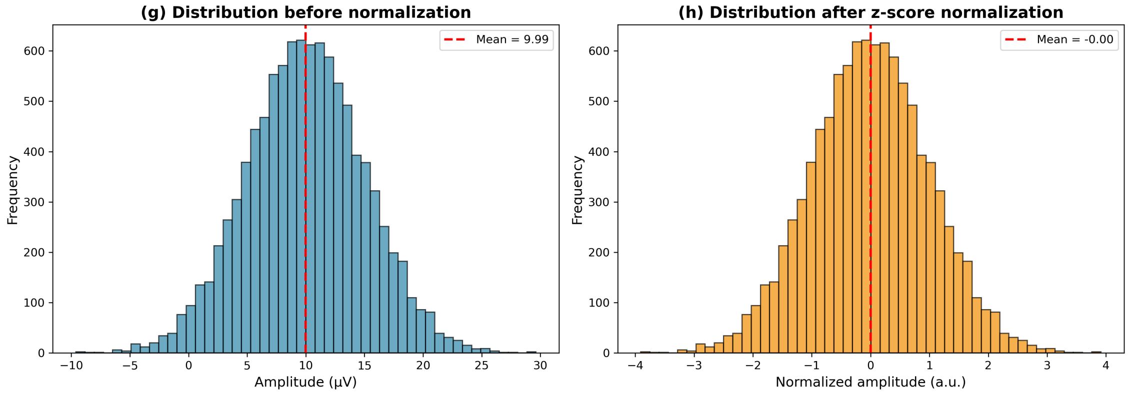

Figure 5: Distribution of EEG signal before and after z-score normalization.

to frequency: small values of a yield high-frequency components with fine temporal resolution, while large values of a yield low-frequency components with coarser temporal but finer frequency resolution.

In this work, the Morlet wavelet is selected as the mother wavelet, as it provides an optimal trade-off between time and frequency localization and has been widely adopted in EEG signal analysis [20]. The CWT scalograms are computed for each 30-second EEG epoch and subsequently converted into two-dimensional image representations. These scalogram images are then used as inputs to the Convolutional Neural Network (CNN) for classification of sleep disorders, enabling the model to learn discriminative spectro-temporal patterns directly from the time–frequency domain.

3.3.2. Hilbert–Huang Transform (HHT)

The Hilbert–Huang Transform (HHT), introduced by Huang et al. [21], is an adaptive and fully data-driven time–frequency analysis method specifically designed for nonlinear and non-stationary signals. Unlike traditional spectral analysis tech-

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

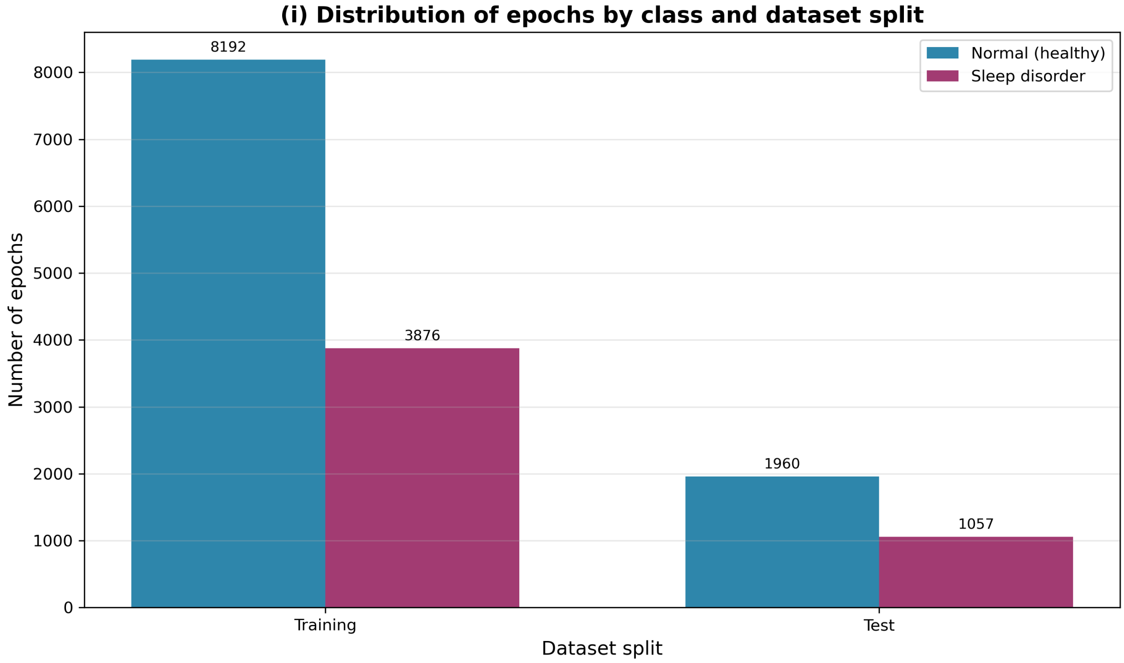

Figure 6: Distribution of EEG epochs across training and test sets.

niques such as the Fourier Transform or the Continuous Wavelet Transform, HHT does not rely on predefined basis functions, making it particularly well-suited for complex biomedical signals such as EEG, which exhibit highly variable and patient-specific oscillatory dynamics.

HHT consists of two sequential steps: Empirical Mode Decomposition (EMD) and Hilbert Spectral Analysis (HSA).

Empirical Mode Decomposition (EMD).. The EMD adaptively decomposes a signal $x(t)$ into a finite set of oscillatory components called Intrinsic Mode Functions (IMFs), together with a residual trend [21, 22]:

$$x(t) = \sum_{i=1}^{N} c_i(t) + r_N(t) \tag{2}$$

where $c_i(t)$ denotes the i-th IMF and $r_N(t)$ is the final monotonic residue. Each IMF must satisfy two conditions: (i) the number of extrema and zero crossings must differ by at most one, and (ii) the mean of the upper and lower envelopes must be zero at every point.

The IMFs are extracted iteratively through a sifting process:

- Identify all local maxima and minima of the signal $x(t)$;
- Interpolate the maxima and minima using cubic splines to form the upper envelope $e_u(t)$ and lower envelope $e_l(t)$;
- Compute the local mean: $m(t) = \frac{e_u(t)+e_l(t)}{2}$;
- Extract the proto-IMF: $h(t) = x(t) - m(t)$;
- Repeat until $h(t)$ satisfies the IMF stopping criterion.

Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Hilbert Spectral Analysis (HSA).. Once the IMFs are obtained, each component $c_i(t)$ is transformed using the Hilbert transform to obtain its analytic representation [23]:

$$y_i(t) = \frac{1}{\pi}\,\mathcal{P}\int_{-\infty}^{+\infty}\frac{c_i(\tau)}{t-\tau}\,d\tau \tag{3}$$

where $\mathcal{P}$ denotes the Cauchy principal value. The corresponding analytic signal is defined as:

$$z_i(t) = c_i(t) + j\,y_i(t) = a_i(t)\,e^{j\theta_i(t)} \tag{4}$$

with instantaneous amplitude $a_i(t)$ and instantaneous phase $\theta_i(t)$:

$$a_i(t) = \sqrt{c_i^2(t) + y_i^2(t)}, \qquad \theta_i(t) = \arctan\left(\frac{y_i(t)}{c_i(t)}\right) \tag{5}$$

The instantaneous frequency is then derived as:

$$\omega_i(t) = \frac{d\theta_i(t)}{dt} \tag{6}$$

The Hilbert spectrum $H(t,\omega)$ is constructed by distributing the instantaneous amplitude $a_i(t)$ of each IMF over the time–frequency plane according to its instantaneous frequency $\omega_i(t)$, yielding a high-resolution, adaptive time–frequency representation of the original signal.

In this work, HHT is applied to each 30-second EEG epoch to produce instantaneous amplitude and frequency maps, which are subsequently used as two-dimensional image inputs to the CNN classifier. The adaptive nature of HHT allows capturing subtle oscillatory patterns and transient dynamics that are often missed by fixed-basis methods such as Fourier or wavelet transforms [21, 24], making it particularly relevant for sleep disorder classification from EEG signals.

3.4. Deep Learning Models

3.4.1. Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) have become a standard approach for analyzing structured data such as images and time–frequency representations of signals. In the context of EEG analysis, CNNs are particularly effective when the signals are transformed into two-dimensional representations, such as spectrograms or wavelet-based images [25].

A typical CNN architecture consists of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The convolutional layers apply learnable filters to extract hierarchical features from the input, capturing local patterns in the data. Pooling layers are used to reduce the spatial dimensionality of feature maps while preserving the most relevant information, thereby improving computational efficiency and reducing overfitting.

In this work, CNN models are employed to process time–frequency representations of EEG signals obtained using Continuous Wavelet Transform (CWT) and Hilbert–Huang Transform (HHT). These representations convert the non-stationary EEG signals into image-like structures, making them suitable for convolutional processing. The extracted features are then fed into fully connected layers for classification of sleep disorders.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

CNN-based approaches have demonstrated strong performance in various EEG-related tasks, including sleep stage classification and neurological disorder detection [5, 6]. Their ability to automatically learn discriminative features from complex signal representations makes them particularly suitable for biomedical signal analysis.

3.4.2. Temporal Convolutional Network (TCN)

Temporal Convolutional Networks (TCNs) have recently emerged as an effective deep learning architecture for sequence modeling tasks, offering a compelling alternative to recurrent neural networks (RNNs). TCNs are based on one-dimensional fully convolutional architectures that leverage causal and dilated convolutions to model temporal dependencies over long sequences [7].

A key property of TCNs is the use of causal convolutions, which ensure that the output at time t depends only on current and past inputs, thereby preserving the temporal ordering of the signal. In addition, dilated convolutions are employed to expand the receptive field without increasing the number of parameters. The dilated convolution operation for a one-dimensional sequence can be expressed as:

$$y(t) = \sum_{k=0}^{K-1} w_k \, x(t - d \cdot k) \tag{7}$$

where $x(t)$ is the input signal, w_k are the convolutional filter weights, K is the kernel size, and d is the dilation factor controlling the spacing between filter elements.

By stacking multiple layers with exponentially increasing dilation factors, TCNs are able to capture long-range temporal dependencies efficiently. Residual connections are typically incorporated to facilitate gradient flow and enable the training of deep architectures.

In the context of EEG analysis, TCNs provide a natural framework for directly modeling raw temporal signals without requiring explicit feature extraction or transformation into time–frequency representations. This is particularly advantageous for electroencephalogram signals, which exhibit complex temporal dynamics and long-range dependencies.

In this work, a TCN model is employed to process raw single-channel EEG signals for sleep disorder classification. The performance of the TCN is compared with CNNs applied to CWT and HHT representations, allowing evaluation of the effectiveness of direct temporal modeling versus representation-based approaches.

3.5. Experimental Setup

All experiments were implemented in Python 3.10 using the TensorFlow 2.12 and Keras frameworks [26]. The computations were performed on a system equipped with an NVIDIA GPU with 8 GB of VRAM, enabling efficient training of deep learning models on EEG time–frequency representations.

3.5.1. Data Splitting and Evaluation Protocol

A subject-independent data splitting strategy is adopted to ensure unbiased evaluation and prevent data leakage between training and testing sets. The dataset is divided into 80% for training and validation, and 20% for independent testing, with no subject appearing in more than one subset. Within the training set, 10% is fur-ther reserved for validation, used exclusively for monitoring convergence and early stopping. The class imbalance between normal and pathological samples is addressed using class-weighted loss functions during training [27].

3.5.2. EEG Epoch Segmentation

Raw EEG recordings are segmented into non-overlapping epochs of 30 seconds, consistent with standard polysomnographic practice [28]. Each epoch is labeled as either *normal* (healthy control) or *pathological* (sleep disorder). Prior to segmentation, EEG signals are bandpass-filtered between 0.5 and 30 Hz using a zero-phase fourth-order Butterworth filter to remove baseline drift, power-line interference, and high-frequency artifacts. Z-score normalization is subsequently applied to each epoch to standardize signal amplitudes across subjects and recording sessions.

3.5.3. CNN Hyperparameters

The CNN models processing CWT and HHT representations share the same architectural configuration. Each scalogram or instantaneous amplitude map is resized to 224×224 pixels before being fed into the network. The CNN architecture consists of four convolutional blocks, each comprising a convolutional layer with 3×3 filters, batch normalization, ReLU activation, and 2×2 max-pooling. The number of filters increases progressively from 32 to 256 across the four blocks. The convolutional backbone is followed by a global average pooling layer and two fully connected layers of 512 and 128 units, respectively, with dropout regularization (rate = 0.5) applied after each dense layer. The output layer consists of a single neuron with sigmoid activation for binary classification.

3.5.4. TCN Hyperparameters

The TCN model processes raw EEG signals of fixed length $L = 4{,}096$ samples, corresponding to approximately 30 seconds at the native sampling frequency. The architecture consists of four dilated causal convolutional blocks with dilation factors $d \in \{1, 2, 4, 8\}$, each containing 64 filters of kernel size 3, followed by batch normalization, ReLU activation, and dropout (rate = 0.2). Residual connections are applied between blocks to facilitate gradient flow. The final temporal representation is aggregated using global average pooling and passed through a dense layer of 64 units before the binary output neuron.

3.5.5. Training Configuration

All models are trained using the Adam optimizer [29] with an initial learning rate of 10^{-4} and a batch size of 32. Training is conducted for a maximum of 30 epochs, with early stopping applied based on validation loss with a patience of 10 epochs. The binary cross-entropy loss function is used for all models. Model performance is evaluated using accuracy, precision, recall, F1-score, and the Area Under the ROC Curve (AUC), computed on the independent test set.

A summary of the key hyperparameters for all three models is provided in Table 1.

4. Results

4.1. Training Behavior and Convergence

Figure 7 presents the learning curves of the three models. All models show a steady increase in training and validation accuracy, indicating stable convergence.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Table 1: Summary of hyperparameters for the three models.

Parameter	CNN+CWT	CNN+HHT	TCN
Input size	224 × 224	224 × 224	4096 × 1
Optimizer	Adam	Adam	Adam
Learning rate	10^{-4}	10^{-4}	10^{-4}
Batch size	32	32	32
Max epochs	30	30	30
Dropout rate	0.5	0.5	0.2
Loss function	BCE	BCE	BCE
Early stopping	Yes	Yes	Yes

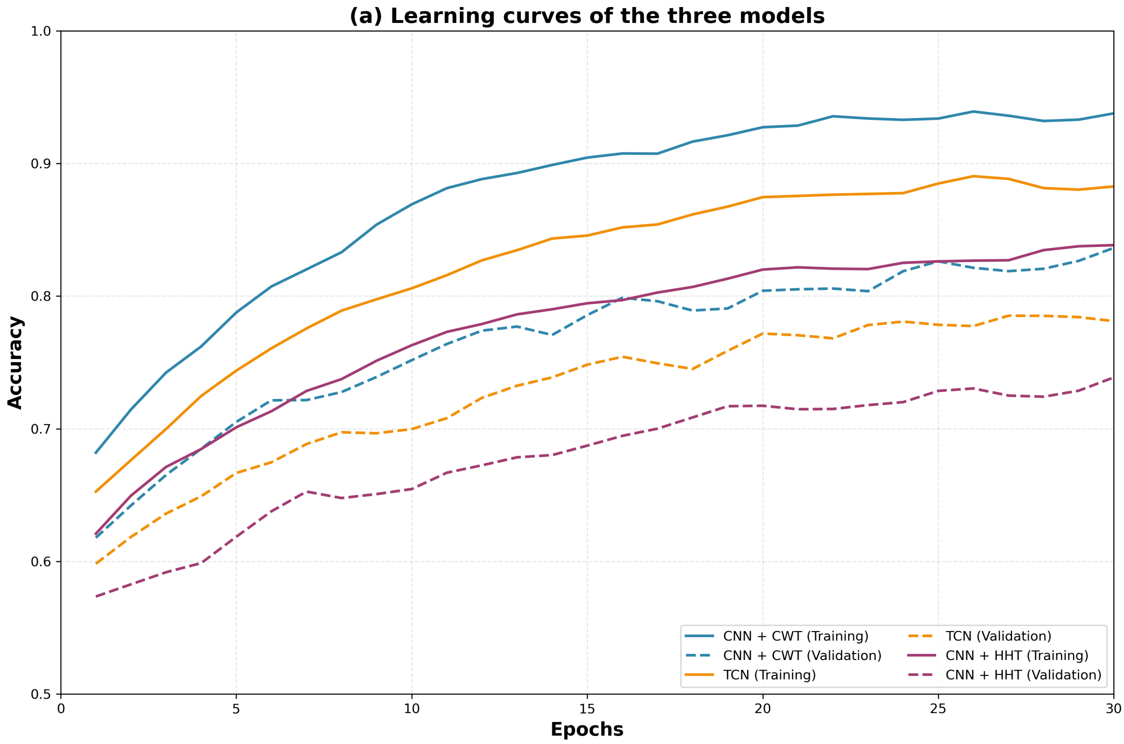

Figure 7: Learning curves of the three models during training.

The CNN + CWT model achieves the best performance, reaching a validation accuracy of approximately 84%, followed by the TCN model (78%) and the CNN + HHT model (74%). A moderate gap between training and validation curves is observed, suggesting slight overfitting but acceptable generalization.

Figure 8 shows the evolution of the binary cross-entropy loss. All models exhibit a consistent decrease in both training and validation loss, confirming stable optimization. The CNN + CWT model achieves the lowest loss values, further demonstrating its superior learning capability.

4.2. Classification Performance

The confusion matrices in Figure 9 provide a detailed view of the classification results. The CNN + CWT model achieves the best balance between sensitivity and specificity, with fewer misclassifications compared to the other models. In contrast,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

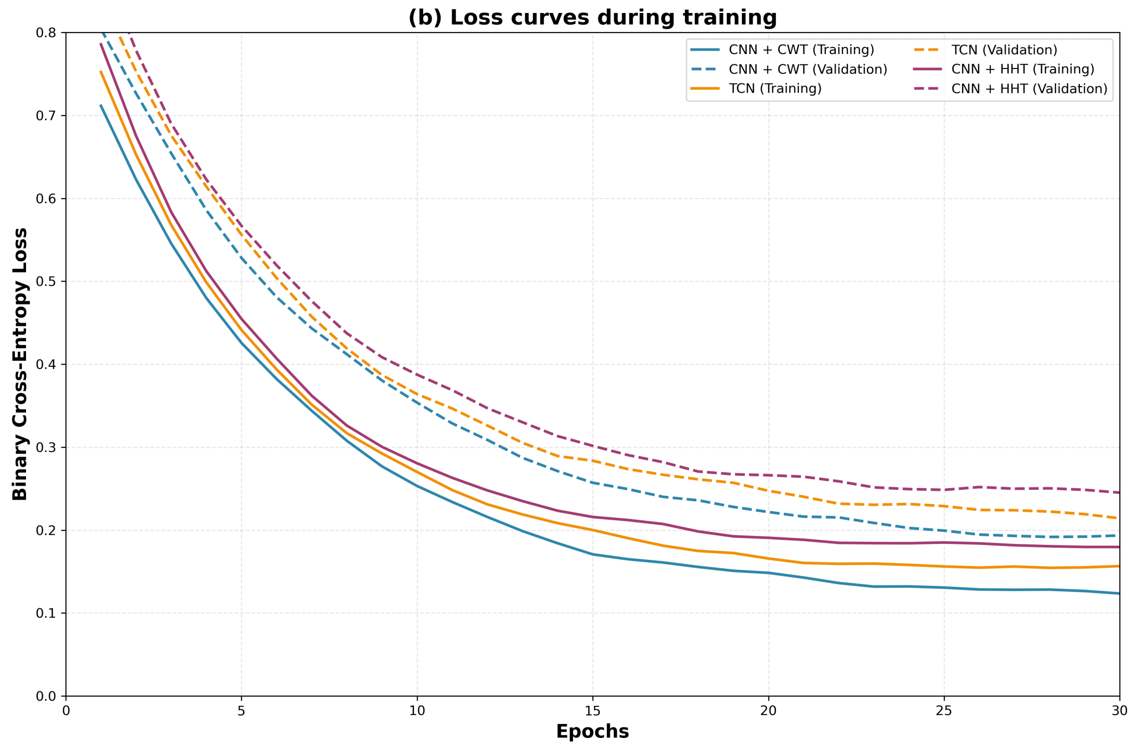

Figure 8: Loss curves during training for the three models.

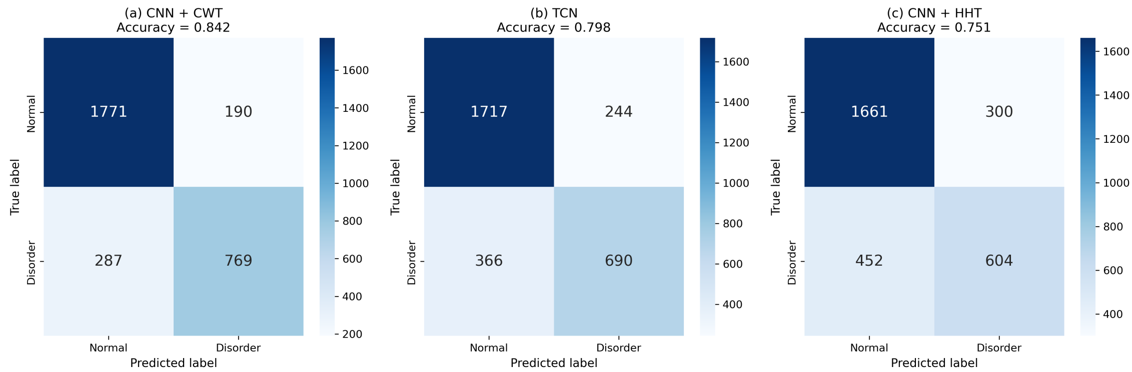

Figure 9: Confusion matrices of the three models.

the CNN + HHT model exhibits a higher number of false negatives, indicating difficulty in detecting pathological cases.

Figure 10 summarizes the evaluation metrics. The CNN + CWT model consistently outperforms the other approaches across all metrics, including accuracy, precision, recall, and F1-score. This highlights the effectiveness of time–frequency representations for EEG-based classification tasks.

4.3. Comparison with State-of-the-Art Methods

Table 2 presents a comparison of the proposed framework with representative state-of-the-art methods evaluated on the CAP Sleep Database or closely related EEG-based sleep disorder detection tasks.

The results indicate that methods exploiting multi-channel EEG or multimodal signals [32, 31] generally achieve higher classification accuracy, as the combination of complementary physiological signals provides richer discriminative information.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

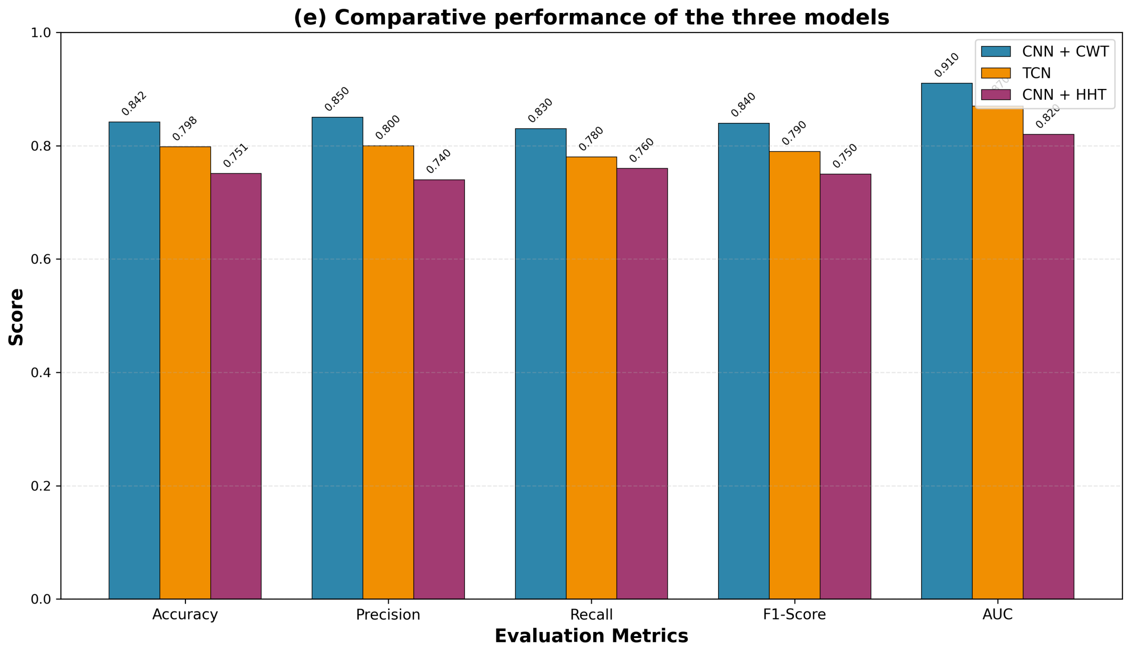

Figure 10: Comparative performance of the three models across evaluation metrics.

Table 2: Comparison with state-of-the-art methods on EEG-based sleep disorder detection. N/A: metric not reported. †: multi-channel EEG or multimodal signals. ⋆: proposed method (single-channel EEG, binary classification).

Reference	Modality	Method	Acc.	AUC	F1	Prec.	Rec.
Sharma et al. [30] (2021)	EEG† (2ch)	Wavelet+EBT	92.8	N/A	N/A	N/A	N/A
Masad et al. [31] (2024)	EEG† (6ch)	CWT+CNN	99.35	0.996	0.993	0.993	0.993
Cheng et al. [32] (2023)	EEG+ECG+EMG†	VGG16	99.09	N/A	N/A	N/A	N/A
Dhok et al. [33] (2022)	EEG (1ch)	1D-CNN	82.21	N/A	0.818	N/A	N/A
Sharma et al. [34] (2021)	EEG (1ch)	Triplet+Ens.	87.9	N/A	N/A	N/A	N/A
CNN+CWT⋆	**EEG (1ch)**	**CWT+CNN**	**84.0**	**0.890**	–	–	–
TCN⋆	**EEG (1ch)**	**Raw+TCN**	**78.0**	**0.840**	–	–	–
CNN+HHT⋆	**EEG (1ch)**	**HHT+CNN**	**74.0**	**0.780**	–	–	–

However, these approaches require more complex and expensive acquisition setups, limiting their applicability in portable or home-based monitoring systems.

In contrast, the proposed framework operates exclusively on single-channel EEG signals, making it more suitable for low-cost and wearable applications. Among single-channel approaches, the proposed CNN + CWT model (accuracy: 84%, AUC: 0.890) achieves performance comparable to the 1D-CNN method of Dhok et al. [33]

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

(82.21%), while additionally providing AUC-based discrimination metrics not reported by most existing single-channel studies. Furthermore, unlike methods focusing on sleep stage classification [30, 34], the proposed framework targets binary disorder detection, which is a more clinically relevant formulation for automated screening applications.

It should be noted that direct numerical comparison across studies is inherently limited by differences in experimental protocols, subject populations, epoch lengths, and evaluation strategies. Nevertheless, the results confirm that the proposed benchmarking framework constitutes a solid and reproducible baseline for EEG-based sleep disorder detection using single-channel signals.

4.4. ROC Analysis

Figure 11 presents the ROC curves for the three evaluated models. The CNN + CWT model achieves the highest AUC of 0.890, confirming its superior ability to distinguish between normal and pathological EEG signals. The TCN model obtains an AUC of 0.840, demonstrating that direct temporal modeling of raw EEG signals constitutes a competitive alternative to time–frequency representation-based approaches. The CNN + HHT model yields an AUC of 0.780, indicating acceptable but comparatively lower discriminative performance. All three models substantially outperform random classification (AUC = 0.500), validating the effectiveness of the proposed framework.

5. Discussion

Overall, the results demonstrate that the CNN + CWT model achieves the best performance among the three evaluated approaches, consistently outperforming competing methods across all metrics, including accuracy, precision, recall, F1-score, and AUC. This confirms that Continuous Wavelet Transform representations capture more discriminative time–frequency features from EEG signals compared to purely temporal or Hilbert-based approaches.

The TCN model, trained directly on raw EEG signals without any explicit feature extraction, achieves competitive results with an AUC of 0.840 and a validation accuracy of 78%, highlighting the potential of purely temporal deep learning architectures for biomedical signal classification. The CNN + HHT model yields an AUC of 0.780 and an accuracy of 74%, suggesting that while HHT-based representations carry useful information, their sensitivity to noise and signal variability limits their discriminative power compared to CWT-based features.

From a clinical standpoint, the proposed single-channel EEG framework demonstrates encouraging potential for integration into low-cost, wearable, and real-time sleep monitoring systems, reducing the burden of conventional polysomnographic studies that require multi-channel recordings and expert annotation.

Several limitations of the current study should be acknowledged. First, the relatively small number of subjects with certain pathologies (e.g., narcolepsy: $n = 5$, SDB: $n = 4$) limits the statistical power of the evaluation and may affect generalization. Second, the binary classification formulation, while clinically motivated, does not distinguish between specific sleep disorder subtypes. Third, the CNN architectures employed are generic and not specifically optimized for EEG-based biomedical classification.

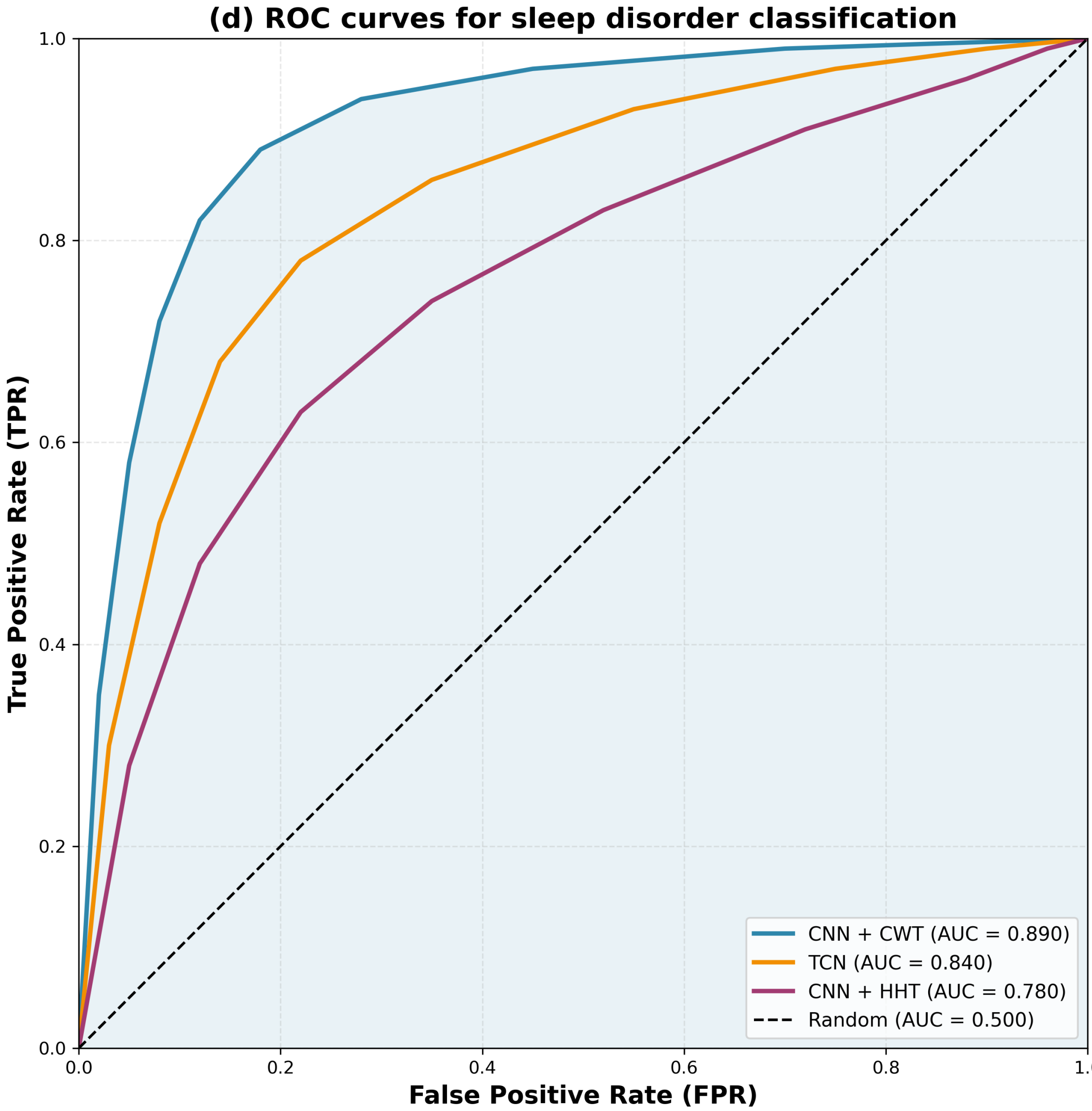

Figure 11: ROC curves for the three models.

Future work will focus on several directions. First, extending the framework to multi-class classification to enable the discrimination of specific sleep disorder subtypes. Second, investigating hybrid architectures that combine time–frequency representations with temporal modeling, such as CNN–TCN or CNN–LSTM fusion models. Third, incorporating additional physiological signals such as EOG and EMG to enrich the feature space and improve diagnostic accuracy. Finally, validating the proposed framework on larger and more diverse clinical datasets to assess its generalization capability across different recording conditions and patient populations.

6. Conclusions

This paper presented a comprehensive benchmarking framework for evaluating the impact of different EEG signal representations on automated sleep disorder detection using deep learning. Three distinct approaches were systematically compared: a Convolutional Neural Network (CNN) trained on Continuous Wavelet

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Transform (CWT) scalograms, a CNN trained on Hilbert–Huang Transform (HHT) instantaneous amplitude maps, and a Temporal Convolutional Network (TCN) operating directly on raw EEG signals. All models were evaluated on the publicly available CAP Sleep Database under a rigorous subject-independent experimental protocol.

The experimental results demonstrate that the CNN + CWT model achieves the best overall performance, with a validation accuracy of 84% and an AUC of 0.890, confirming the effectiveness of time–frequency representations for capturing discriminative spectro-temporal patterns in EEG signals. The TCN model, despite operating on raw signals without any explicit feature extraction, achieves competitive results with an AUC of 0.840 and a validation accuracy of 78%, highlighting the potential of purely temporal deep learning architectures for biomedical signal classification. The CNN + HHT model yields an AUC of 0.780 and an accuracy of 74%, suggesting that while HHT-based representations carry useful information, their sensitivity to noise and signal variability limits their discriminative power compared to CWT-based features.

From a clinical standpoint, the proposed single-channel EEG framework demonstrates encouraging potential for integration into low-cost, wearable, and real-time sleep monitoring systems, reducing the burden of conventional polysomnographic studies that require multi-channel recordings and expert annotation.

CRediT authorship contribution statement

Romain Atangana: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Amstrong Emini Me Zenanga:** Methodology, Software, Visualization. **Vivien Beyala Kamgang:** Investigation, Data curation. **Emmanuel Baba:** Data curation, Software. **Perrin M. Li Litet:** Formal analysis, Visualization. **Daniel Tchiotsop:** Conceptualization, Validation, Resources, Writing – review & editing, Supervision. **Godpromesse Kenné:** Validation, Resources, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The CAP Sleep Database used in this study is publicly available on PhysioNet at `https://physionet.org/content/capslpdb/`.

Funding

This research received no external funding.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Abbreviations

AUC	Area Under the ROC Curve
CAP	Cyclic Alternating Pattern
CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
EEG	Electroencephalogram
ECG	Electrocardiogram
EMD	Empirical Mode Decomposition
EMG	Electromyogram
EOG	Electrooculogram
HHT	Hilbert–Huang Transform
HSA	Hilbert Spectral Analysis
IMF	Intrinsic Mode Function
NREM	Non-Rapid Eye Movement
OSA	Obstructive Sleep Apnea
PSG	Polysomnography
REM	Rapid Eye Movement
RNN	Recurrent Neural Network
SDB	Sleep Disordered Breathing
SVM	Support Vector Machine
TCN	Temporal Convolutional Network

References

[1] M. A. Carskadon, W. C. Dement, Normal human sleep: an overview, Elsevier, 2005.

[2] T. Young, P. E. Peppard, D. J. Gottlieb, Epidemiology of obstructive sleep apnea: a population health perspective, American journal of respiratory and critical care medicine 165 (9) (2002) 1217–1239.

[3] M. M. Ohayon, Epidemiology of insomnia: what we know and what we still need to learn, Sleep medicine reviews 11 (2) (2007) 97–111.

[4] U. R. Acharya, et al., Automated diagnosis of epilepsy using eeg signals: a review, Knowledge-Based Systems 45 (2016) 147–165.

[5] A. Supratak, H. Dong, C. Wu, Y. Guo, Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel eeg, IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (11) (2017) 1998–2008.

[6] S. Biswal, et al., Sleepnet: Automated sleep stage scoring system via deep learning, arXiv preprint arXiv:1707.08262 (2018).

[7] S. Bai, J. Z. Kolter, V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, arXiv preprint arXiv:1803.01271 (2018).

[8] K. A. Aboalayon, W. S. Almuhammadi, M. Faezipour, A comparison of different machine learning algorithms using single channel eeg signal for classifying

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

human sleep stages, in: 2015 Long Island Systems, Applications and Technology, IEEE, 2015, pp. 1–6.

[9] K. D. Tzimourta, A. Tsilimbaris, K. Tzioukalia, A. T. Tzallas, M. G. Tsipouras, L. G. Astrakas, N. Giannakeas, Eeg-based automatic sleep stage classification, Biomed J 1 (6) (2018).

[10] M. Diykh, Y. Li, P. Wen, Eeg sleep stages classification based on time domain features and structural graph similarity, IEEE Transactions on Neural Systems and Rehabilitation Engineering 24 (11) (2016) 1159–1168.

[11] E. Estrada, H. Nazeran, P. Nava, K. Behbehani, J. Burk, E. Lucas, Eeg feature extraction for classification of sleep stages, in: The 26th annual international conference of the IEEE engineering in medicine and biology society, Vol. 1, IEEE, 2004, pp. 196–199.

[12] M. E. Tagluk, M. Akin, N. Sezgin, Classıfication of sleep apnea by using wavelet transform and artificial neural networks, Expert Systems with Applications 37 (2) (2010) 1600–1607.

[13] F. Andreotti, H. Phan, N. Cooray, C. Lo, M. T. Hu, M. De Vos, Multichannel sleep stage classification and transfer learning using convolutional neural networks, in: 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC), IEEE, 2018, pp. 171–174.

[14] A. Vilamala, K. H. Madsen, L. K. Hansen, Deep convolutional neural networks for interpretable analysis of eeg sleep stage scoring, in: 2017 IEEE 27th international workshop on machine learning for signal processing (MLSP), IEEE, 2017, pp. 1–6.

[15] G. S. Giarrusso, I. Rechichi, G. Olmo, Eeg-based detection of rem sleep behaviour disorder: Towards a stage-agnostic approach, in: International Work-Conference on Bioinformatics and Biomedical Engineering, Springer, 2024, pp. 263–276.

[16] M. C. Melo, J. R. da Silva Vallim, S. Garbuio, L. A. Soster, K. M. M. Sousa, R. R. Bonaldi, G. N. Pires, Validation of a sleep staging classification model for healthy adults based on two combinations of a single-channel eeg headband and wrist actigraphy, Journal of Clinical Sleep Medicine (2024) jcsm–11082.

[17] A. L. Goldberger, L. A. N. Amaral, L. Glass, et al., Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220. doi:10.1161/01.CIR.101.23.e215.

[18] I. Daubechies, Ten Lectures on Wavelets, SIAM, 1992.

[19] S. Mallat, A Wavelet Tour of Signal Processing, 2nd Edition, Academic Press, 1999.

[20] P. S. Addison, The Illustrated Wavelet Transform Handbook, IOP Publishing, 2002. doi:10.1887/0750306920.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

[21] N. E. Huang, Z. Shen, S. R. Long, et al., The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London A 454 (1971) (1998) 903–995. doi:10.1098/rspa.1998.0193.

[22] Z. Wu, N. E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Advances in Adaptive Data Analysis 1 (1) (2009) 1–41. doi:10.1142/S1793536909000047.

[23] L. Cohen, Time–Frequency Analysis, Prentice Hall, 1995.

[24] P. Flandrin, G. Rilling, P. Gonçalvès, Empirical mode decomposition as a filterbank, IEEE Signal Processing Letters 11 (2) (2004) 112–114. doi:10.1109/LSP.2003.821662.

[25] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.

[26] F. Chollet, et al., Keras, https://keras.io (2015).

[27] G. King, L. Zeng, Logistic regression in rare events data, Political Analysis 9 (2) (2001) 137–163. doi:10.1093/oxfordjournals.pan.a004868.

[28] R. B. Berry, et al., Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events, Journal of Clinical Sleep Medicine 8 (5) (2012) 597–619. doi:10.5664/jcsm.2172.

[29] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[30] M. Sharma, J. Tiwari, U. R. Acharya, Automatic sleep-stage scoring in healthy and sleep disorder patients using optimal wavelet filter bank technique with EEG signals, International Journal of Environmental Research and Public Health 18 (6) (2021) 3087. doi:10.3390/ijerph18063087.

[31] I. S. Masad, A. Alqudah, S. Qazan, Automatic classification of sleep stages using EEG signals and convolutional neural networks, PLOS ONE 19 (1) (2024) e0297582. doi:10.1371/journal.pone.0297582.

[32] Y.-H. Cheng, M. Lech, R. H. Wilkinson, Simultaneous sleep stage and sleep disorder detection from multimodal sensors using deep learning, Sensors 23 (7) (2023) 3468. doi:10.3390/s23073468.

[33] S. Dhok, V. Pimpalkhute, A. Chandurkar, A. Bhurane, M. Sharma, U. R. Acharya, Automated classification of cyclic alternating pattern sleep phases in healthy and sleep-disordered subjects using convolutional neural network, Computers in Biology and Medicine 146 (2022) 105560. doi:10.1016/j.compbiomed.2022.105560.

[34] M. Sharma, J. Tiwari, V. Patel, U. R. Acharya, Automated identification of sleep disorder types using triplet half-band filter and ensemble machine learning techniques with EEG signals, Electronics 10 (13) (2021) 1531. doi:10.3390/electronics10131531.