International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Traitement du Signal
Vol. 41, No. 5, October, 2024, pp. 2437-2448
Journal homepage: http://iieta.org/journals/ts

Analytical Features Assisted Hierarchical Classification for Automatic Sleep Stage Scoring

Voruchu Sai Babu[1,2], Avinash S Vaidya[1*]

[1] Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad 500075, India
[2] Department of Electronics and Communication Engineering, Guru Nanak Institutions Technical Campus, Hyderabad 501506, India

Corresponding Author Email: avinash.s@klh.edu.in

https://doi.org/10.18280/ts.410518

Received: 8 September 2023
Revised: 18 March 2024
Accepted: 10 July 2024
Available online: 31 October 2024

Keywords:
automatic sleep staging, Electroencephalogram (EEG), hierarchical classifier, instantaneous phase, envelope, entropy

ABSTRACT

In the area of sleep study, automatic sleep stage identification through Electroencephalogram (EEG) is an important step. However, the outcomes obtained with modern, cutting-edge techniques are not satisfactory; hence, they cannot be used in standard clinical procedures. In the following paper, we suggest a simple yet productive solution with automated sleep staging through analytical features and hierarchical classification. As an initial step, the proposed method segments the input EEG signal into different epochs and describes each epoch with 22 features extracted from multiple domains, including frequency-domain, time-frequency-domain, time-domain, and non-linear domain. The ability of analytical features to perfectly differentiate between sleep stages with similar characteristics demonstrates their efficacy. Further, this work introduces a new classification scheme called hierarchical classification, which solves the complex classification problem by breaking it into small problems. At classification, we employed the binary Support Vector Machine (SVM). In terms of performance, the proposed system is validated through a standard and publicly available Sleep-EDF dataset. In comparison with gold-standard manual scoring, we achieved 92.3500% accuracy and a 0.8646 kappa coefficient with our method. Further, the suggested method outperformed current cutting-edge automatic sleep stage classification techniques in terms of better results.

1. INTRODUCTION

One of life's most important daily activities, sleep is shown to follow the circadian rhythm of human psychological processes [1]. Sleep follows a stage-by-stage process that is indirectly linked to the functions of the autonomous nervous system [2]. The quality of sleep shows a significant impact on so many human activities like concentration, memorization, and learning [3]. Recently, some research has revealed that abnormal sleep leads to severe health issues like increased neurocognitive function disorders and the risk of cardiovascular strokes [4, 5]. So, a proper understanding and analysis of sleep disorders like sleep apnea and insomnia are essential. For this purpose, a precise sleep stage detection from the sleep cycles is necessary. Accurately identifying the various stages of sleep is crucial for the diagnosis and treatment of sleep disorders [6].

Currently, the mechanism for staging sleep is being carried out through a Polysomnograph (PSG) record acquired the whole night during sleep. Mainly, the PSG record involves the acquisition of six activities: respiratory rate, Blood Oxygen Level, Electromyogram (EMG), Electrooculogram (EOG), Electrocardiogram (ECG), and the Electroencephalogram (EEG). Based on these activities, sleep staging is performed by a well-trained professional based on common rules induced by Rechtschaffen and Kales (R&K) [7] and the American Academy of Sleep Medicine (AASM) [8]. Total sleep is divided into seven stages, as per R&K rules: movement time (MT), Rapid Eye Movement (REM), Stages from one to four (S1-S4), and Wake (W). However, the AASM amended the R&K guidelines and came up with new rules for classifying different stages of sleep, which are now used. According to the AASM, the stages of sleep are derived as REM, N1, N2, N3, and W [9]. AASM is considered to have three major attributes at the sleep staging; they are duration, pattern, and Spectral content of PSG signals. For instance, attributes like high EMG amplitudes combined with frequent eye movements and faster frequency rhythms in a 30s EEG window (called an epoch) or alpha band (8-12Hz) are used to characterize the wake stage. Next, the vertex waves and more than 50% of the epoch occupied by the alpha band are used to characterize the N1 stage [10]. Further, the K-complexes and the sleep spindles are major attributes to characterize the stage N2. Next, N3 Stage (called Slow Wave Sleep (SWS)) is distinguished by the presence of more than 20% of delta band (0-4Hz) activity in the epoch. At last, attributes like low EMG amplitudes, saccadic eye movements, and sawtooth waves are used to characterize the REM stage.

Manual sleep staging, however, takes a lot of time and is highly subjective in nature, and totally dependent on the expert.

Mostly, the agreement between two independent experts is unsatisfactory. For example, Danker-Hopfe et al. [11] discovered that the Inter-Rate Reliability (IRR) computed with the help of Kappa's Coefficient between two scorers is approximately 0.78 [11]. But the international score shows poor staging at the N1 stage and hence the IRR is approximated to be low in the range of 0.58 to 0.63 [12-15]. It has been proven that the approximated value of the N1 stage's agreement among the European laboratories is only 0.19, while it is approximated at 0.31 between international laboratories. Further, the manual sleep staging score gets worse if the patient is experiencing any medical conditions like Obstructive Sleep Apnea (OSA) [16]. Compared with automatic scoring, manual scoring enhances sleep staging consistency between healthcare centers and hospitals. Furthermore, automatic sleep stage scoring ensures minimal computational complexity by considering only a few signals. Therefore, we are driven to develop an automated, reliable sleep staging system based on signal processing and machine learning methods.

Over the past few years, several automatic staging methods have been developed by researchers. In general, automated approaches depend on predefined rules, features extracted from signals, and classification methods [17]. However, in most past methods, the classifier was trained with all stage features which in turn rises the computational burden over the system. Moreover, the Pats methods employed only a limited set of features to describe the sleep stages through EEG. Hence, this paper proposes a new sleep staging framework that considers multiple and diversified features and employs a hierarchical classification. The following are the main contributions of this paper:

- To give a clear differentiation of the various stages of sleep, a new method that elucidates EEG signal-based sleep stages is proposed. Totally, we use 22 different features to represent each EEG signal epoch. The features include non-linear, time-domain, frequency-domain, and time-frequency domain features.
- To achieve a reduced computational burden during training and testing, this work employs a hierarchal classification based on brute-force methodology. Support Vector Machine (SVM) is employed at multiple phases to classify each and every sleep stage.

The rest sections of the study are organized as follows: specifics of the literature survey are explored in the second section. Following that, the third section explores the specifics of the suggested sleep stage classification mechanism. The details of experimental investigations are examined in the fourth section, and the study is concluded in the fifth.

2. LITERATURE SURVEY

Earlier, the automatic methods of sleep staging were broadly classified into two learning-based methods: deep learning and machine learning-based. The former methods extracted handcrafted features and fed them to machine learning algorithms to get class labels, while the next methods employed deep learning algorithms for both component extraction and classification. Since the withdrawal of features that are handcrafted increases time complexity, recent deep learning algorithms have been replaced with traditional machine learning methods. However, deep learning algorithms employ a set of convolutional operations as they increase computational complexity. In this section, we outline the details of both methods briefly.

2.1 Machine learning based staging

In the method proposed by Memar and Faradji [18], initially, every EEG epoch is partitioned into eight sub-bands based on the rhythm of EEG (i.e., gamma 1, gamma 2, beta 1, beta 2, sigma, alpha, theta, and delta). Totally, each epoch is represented with 104 features, and they employed a feature significance examination by employing the Kruskal-Wallis test, which helped in eliminating the less significant features. Then they removed redundant features with the help of Minimal Redundancy Maximal Relevance(MRMR) and finally fed them to random forest classifier. Sharma et al. [19] proposed a new Least Squares design with a Bi-orthogonal Wavelet Filter Bank (BWFB) for feature extraction and employed different supervised classifiers for automatic sleep stage scoring. They used only EEG channels that are unipolar, such asC3-M2 andO1-M2, and extracted Hjorth Parameters (HP) from the wavelet sub-bands.

Alickovic and Subasi [20] considered single-channel EEG and proposed executing the automatic sleep staging system in three modules. In the first module, the input EEG signal acquired from the Pz-Oz electrode is subjected to denoising through Multiscale Principal Component Analysis (MSPCA) [21]. The 2nd module applies the discrete wavelet transform (DWT) to withdraw informative features through the computation of statistical values from DWT bands. At last, they employed an ensemble classifier called Rotational SVM for classifying the signal. Widasari et al. [22] proposed a five-stage methodology for an automatic sleep staging. They are: (1) Pre-processing through R-peak detection, filtering, and interpolation of R-R intervals; (2) Extraction of spectral features through the Hanning window with the Welch method; (3) Decision tree-based SVM for sleep stage detection; and (4) Sleep quality features assessment and classification of sleep disorders through the ensemble of bagged tree classifiers.

Zapata et al. [23] employed the Multitaper with Convolution (MT&C) methodology for the feature withdraws from EEG signals. They employed two methods for classification of sleep stage. In the 1ststage, EEG waves are used directly, and the sleep stages classification is predefined and rule-based. The next method uses SVM with a quadratic equation (SVM-Q) as a classifier to classify the stages of sleep based on the scores of the experts. Ye et al. [24] employed self-supervised learning (SSL) [25, 26] mechanisms that make the system learn roust and generalizable features from physiological signals. They employed a novel co-training mechanism that exploits complementary information from multiple orientations (like time and frequency) of EEG signals to dig out the more positive samples.

Karimzadeh et al. [27] proposed to extract two informative features from the EEG signal: the Shannon entropy of the instantaneous envelope and an instantaneous frequency. Then, the obtained set of features is used to build a decision tree-based classifier in a distributed manner. The design is done according to brute-force methodology, and the K-Nearest Neighbor classifier is used at each decision node in the distributed decision tree model. Li et al. [28] proposed a scoring method, hybrid automatic sleep stage called Hy CLASS, which considers the single-channel EEG as input. Hy

CLASS employed two types of features, namely signal features and transition features [29], to describe each epoch of the EEG signal. Under signal features, totally they extracted 30 features, which include both time and frequency features. At classification, they employed a random forest algorithm along with correction rules. The stage transition characteristics, which import the sleep property and describe the transition of the sleep stage, are the foundation upon which the correction rules are created.

2.2 Deep learning based staging

Under this category, the sleep staging [30] methods for sleep stages classification employed deep learning methods. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are the two approaches that have been widely used.

In order to score sleep stages, Zhou et al. [31] suggested a new deep learning architecture called a segmented Attention Network (SAN). The entire architecture is two-fold: one is the Feature Extraction (FE) and the other is the Time Sequence Encoder (TSE). The reinforced features are taken from numerous, identical-length EEG epochs using FE, which consists of Multiple Multiscale CNN (MMCNN), Residual Squeeze, and Excitation Block. The features extracted by the FE module are used by TSE to derive temporal information.

Abdollahpour et al. [32] considered single-channel EEG and extracted only one feature (i.e., standard deviation) from the frequency sub-bands of EEG. For classification, they employed a two-stream CNN and a two-stage learning model. Similarly, Sun et al. [33] proposed classification methods consisting of two stages that include a learning network of two stages for automatic sleep staging. The first stage consisted of the handcrafted feature extraction, and 2nd stage involved the RNN for full utilization of learning the temporal information between EEG epochs. Further, they proposed a data augmentation scheme to solve the data imbalance problems.

Leino et al. [34] combined the CNN and RNN models to diagnose several sleep disorders from the single-channel (i.e., F4-M1) EEG signals recorded through an ambulatory electrode set (AES). Liu et al. [35] built a new deep learning network by combining Multi-Scale Extraction (MSE) based Convolutional Block Attention Module (CBAM) and U-Structure for the extraction of multi-scale salient features from the single-channel EEG signals. First, the U-structured CNN model with MSE extracts multi-scale features from EEG, and then CBAM is employed to leverage the salient variations and then learn transition rule among different stages.

Phan et al. [36] proposed a joint prediction and classification architecture for automatic sleep staging depends on CNNs. For a given EEG epoch, their architecture jointly identifies its label and also the labels of its neighboring epochs. Due to this kind of joint classification, the method can leverage the dependency between successive epochs while surpassing the traditional Machine learning-based method’s problems. An energy-based optimization technique for the improvisation of hypnograms produced by automatic sleep staging methods was developed by Aristimunha et al. [37]. For each epoch, they compute energy followed by conditional probabilities and then apply an energy minimization process for sleep stage prediction.

Recently, transfer learning-based deep learning architectures have been put forth to enhance categorization efficiency. Based on this fact, Abdollahpour et al. [38] introduced a new method that uses EEG and EOG signals for performing automatic sleep staging. At feature extraction, they extracted two sets of features: features of EEG and features of a composite of EEG and EOG. Then the obtained features are transformed into an HV (Horizontal Visibility) Graph and the HV Graph image is classified through the transfer learning CNN for Fusion (TLCNN-DF). Table 1 shows the comparison between the state of the art methods.

Table 1. Comparison between state of the art methods

Ref. No	Feature Extraction	Classification	Demerit
Memar and Faradji [18]	1. Epochs extraction from rhythm of EEG 2. 104 features from each epoch 3. Kruskal-Wallis test for feature significance	Random forest classifier	Huge computational complexity due to more number of features
Sharma et al. [19]	Least Squares design with a Bi-orthogonal Wavelet Filter Bank (BWFB)	Multiple supervised learning algorithms for classification	Scale Invariance occurrence due to wavelet filter
Alickovic and Subasi [20]	Multiscale Principal Component Analysis (MSPCA) and DWT	Rotational SVM	DWT is a multiscale transform and PCA over such features results in information loss
Widasari et al. [22]	Spectral Features and Welch features	Ensemble of bagged tree classifiers	The width of main lobe is more in Hanning window
Karimzadeh et al. [27]	Shannon entropy of the instantaneous envelope and an instantaneous frequency	brute-force methodology and the K-Nearest Neighbor	Very less features
Zhou et al. [31]	Reinforced features are taken from numerous, identical-length EEG epochs	Time Sequence Encoder (TSE) for classification	Huge complexity due to TSE
Sun et al. [33]	Temporal features	Recurrent neural Networks	Not involved the statistical features
Liu et al. [35]	U-structure for the extraction of multi-scale salient features	Multi-Scale Extraction based Convolutional Block Attention Module	More errors due to MSE

3. METHODOLOGY

3.1 Overview

Under the methodology, we explore the complete details of the proposed automatic sleep staging system. Figure 1 shows the overall working schematic of the proposed system, which consists of two stages: feature extraction and classification. As we noticed a huge computational burden through deep learning approaches, we used simple and effective machine learning methods for sleep stage scoring. To ensure better differentiation between various stages of sleep, we employed a robust and composite set of feature extraction in multiple domains: the non-linear domain, frequency domain, time domain, and time-frequency domain. Unlike the conventional ML-based method, which employs direct classification, we employed hierarchical classification to lessen the computational burden on the system. The major novelty of this system is a hierarchical classification based on an exhaustive search strategy over different feature set combinations. Further, the consideration of multiple features makes the scoring system fully aware and helps in the accurate classification of sleep stages.

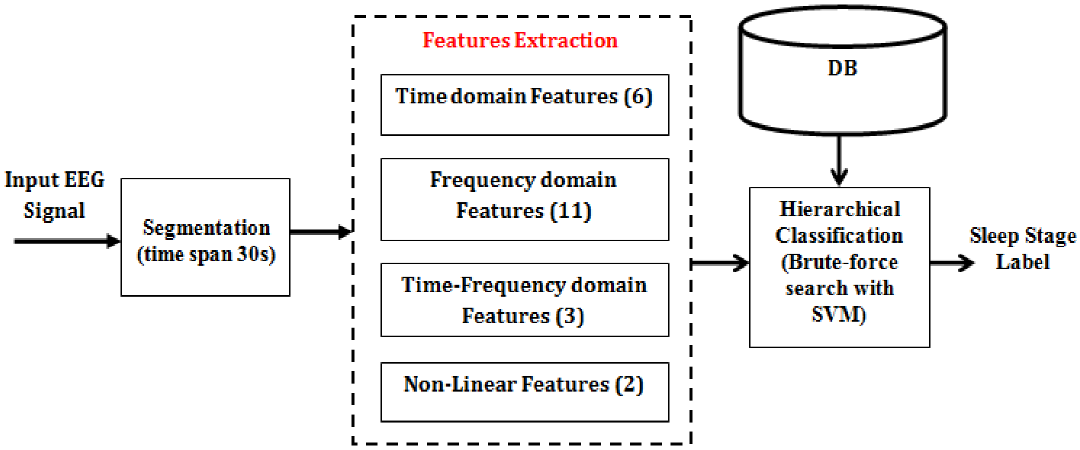

Figure 1. Overall working schematic of proposed sleep staging system

3.2 Pre-processing

At this stage, the EEG signal is subjected to quality enhancement. Towards such an aspect, we apply an 8th order Butterworth bandpass filter with a pass band ranging from 0.5 Hz to 35 Hz. Let's consider $x(t)$ being an input EEG signal, after filtering, it is denoted as $x_f(t)$. Then $x_f(t)$ is processed for segmentation where it is divided into different epochs each having a time span of 30 seconds, let the i^{th} epoch is denoted as $x_f^i(t)$. Instead of processing the complete EEG signal for scoring, epochs are processed where each epoch is associated with one sleep stage.

3.3 Features extraction

For feature extraction, the earlier features that are reported are considered for the study. From each epoch signal, a total of 22 features are extracted from four different perspectives: eleven features in the frequency domain, six features in the time domain, two non-linear features and three time-frequency domain features. The complete details of the features are explored here.

A. Time domain features

Time domain features play a fundamental role in EEG classification by capturing the temporal dynamics and morphological characteristics of EEG signals, facilitating the discrimination of different brain states and conditions for various clinical and research applications. EEG signals are inherently temporal in nature, representing the electrical activity of the brain over time. Time domain features capture various temporal characteristics of EEG signals, such as amplitude, frequency, duration, and temporal dynamics. Time domain features allow capturing the shape, amplitude, and timing of various EEG waveform patterns, such as spikes, sharp waves, and rhythmic oscillations, which can provide valuable insights into brain function and dysfunction. Under the time domain, each epoch is characterized by six different features; they are namely Mean, Minimum, maximum, Mean Absolute Deviation (MAD), Standard Deviation (SD), and Root Sum of Squared Level (RSS). They are expressed as follows.

Mean: The mean explores the average amplitude variations in the epoch. For a given epoch, the mean is computed as a sum of the amplitudes of all samples, followed by the division of the sum by the total number of the samples present in the epoch. Mathematically, expressed as:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} x_f^i(j) \tag{1}$$

where, $x_f^i(j)$ denotes the j^{th} sample's amplitude in i^{th} epoch of filtered EEG signal and N express the total number of samples present in the epoch.

Minimum and Maximum: Minimum and Maximum values explore the least and most values among the given input samples. These features help in the discrimination between sag and swell signals, as well as between sag with harmonics and swell with harmonics. For a given window *w*, the minimum and maximum are computed as follows:

$$Mx_i = Maximum\left(x_f^i(j)\right), j \in 1, \ldots, N \tag{2}$$

$$Mn_i = Minimum\left(x_f^i(j)\right), j \in 1, \ldots, N \tag{3}$$

SD: SD explores the EEG signal's statistical distribution with respect to the mean. For a given epoch, the initial mean is computed, and then each sample's amplitude is subtracted from the mean, followed by summation, normalization, and square root computation. Mathematically, SD is computed as:

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\mu_i - x_f^i(j)\right)^2} \quad (4)$$

MAD: It reveals the signal's variability. For a given epoch, the mean is calculated first, and then each sample's amplitude is subtracted from the mean, followed by summation and normalization. Since the means of epochs are different in nature, the MAD ensures better discrimination between sleep stages. Mathematically, MAD is computed as:

$$M_i = \frac{1}{N}\sum_{j=1}^{N}\left(\mu_i - x_f^i(j)\right) \quad (5)$$

RSS: It is measured as the square root of the mean of summation of the squared amplitudes of each sample in the epoch. For a given epoch, initially, each sample is squared, and then all the values are subjected to summation, followed by normalization and square root computation. RSS alleviates the difference between EEG amplitudes and noises perfectly since the squared amplitude clears the ambiguity. For a given epoch, the RSS is computed as follows:

$$R_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(x_f^i(j)\right)^2} \quad (6)$$

Table 2. Different frequency bands of EEG signal

Notation	Name	Frequency Range
Beta 1	x_{β_1}	[15.0, 20.0]
Beta 1	x_{β_2}	[20.0, 30.0]
Spindle	x_S	[10.5, 14.5]
Alpha	x_α	[8.0, 10.5]
Theta	x_θ	[4.0, 8.0]
Delta	x_δ	[1.0, 4.0]
SWA	x_{SWA}	[0.5, 5.5]
Sigma	x_Σ	[12.0, 15.0]
Fast_Sigma	$x_{\Sigma\mathrm{F}}$	[13.5, 15.0]
Slow_Sigma	$x_{\Sigma\mathrm{S}}$	[12.0, 13.5]
Osscilate_Slow	x_{OS}	[0.5, 2.0]

B. Features of frequency domain

Frequency domain features are essential for EEG classification by capturing the spectral characteristics, functional connectivity patterns, and neurophysiological correlates of brain activity, which are critical for discriminating between different cognitive states, tasks, and neurological conditions. EEG signals contain information about the rhythmic electrical activity of the brain, which can be decomposed into different frequency bands, such as delta (0.5-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz), and gamma (30-100Hz). Frequency domain analysis allows capturing the spectral characteristics of EEG signals, revealing the distribution of power across different frequency bands. The diverse sleep stage EEG signals have distinctive frequency-domain characteristics, and many past methods have analyzed such characteristics [39]. The importance of spectral power is reported by most of the previous methods, which showed better performance at stage scoring [40] of automatic sleep. For example, the rise in the depth of sleep makes the power of EEG signals stronger at low frequencies. Next, the amplitudes of stage 1 sleep are larger in the frequency range between 2Hz and 7Hz. Next, stage 2 is easily determined with the help of sleep spindles at a frequency range between 12Hz and 14Hz. SWS is scored if the frequency is observed to be less than 2Hz. The epoch is first sent through the Fast Fourier Transform (FFT) at this stage to extract features in the frequency domains, and then the spectral power is calculated as features in the frequency domain. Table 2 shows all the frequency bands & their ranges.

C. Time-Frequency domain features

Under this category, we extract two features from each epoch; they are namely instantaneous phase and envelope. Generally, the conventional methods employed a three-phase strategy for the determination of envelope and phase, which is (1) Narrow-band filtering, (2) Computation of the narrowband signal analytical form, and (3) estimation of envelope and phase from analytical expression [41, 42]. However, the inaccurate measurement of these two features results in a false diagnosis in the presence of background EEG activity (cerebral activity is treated as noise). Moreover, the phase and amplitudes of the input signal must be least affected by the process of narrow-band filtering.

To sort out these problems, we propose robust instantaneous envelope (IE), instantaneous frequency (IF), and instantaneous phase (IP) estimation mechanisms for each epoch. These two features explore the presence of instantaneous variations in the temporal attributes of EECG which directly associated with brain functioning. Especially, these three features help in determination of early stage sleep stage disorders. The proposed mechanism is inspired by the Mote-Carlo simulation developed for IP and IE prediction from the EEG's analytical form, with the help of infinitesimal perturbations over the transfer function of a narrow band filter [43]. This mechanism employs a zero-phase backward and forward strategy through very narrow-band FIR or IIR filters. Then they estimate the phase with the help of EEG signals analytical expression. Here, we utilized the new toolbox introduced by Seraj and Sameni [44] for the extraction of robust IE and IP features. We use a narrow-band IIR filter of frequency 1 Hz as a primary filter bank for IE and IP computation. From the filter, center frequency is passed through the entire range of frequencies that lie within the pass band range.

In each frequency bin, the analytical expression of the EEG epoch is computed as:

$$y_f^i(t,f) = x_f^i(t,f) + j\mathcal{H}\{x_f^i(t,f)\} \quad (7)$$

where, $x_f^i(t,f)$ is the filtered i^{th} epoch at frequency bin f and $\mathcal{H}\{x_f^i(t,f)\}$ is the Hilbert transform of $x_f^i(t,f)$. With the help of above analytical expression, we compute the IE and IP and let they are denoted as $IP_f^i(t,f)$ and $IE_f^i(t,f)$. Mathematically they are expressed as follows:

$$IP_f^i(t,f) = arctan\left(\frac{\mathcal{H}\{x_f^i(t,f)\}}{x_f^i(t,f)}\right) \quad (8)$$

And

$$IE_f^i(t,f) = \sqrt{x_f^i(t,f)^2 + \mathcal{H}\{x_f^i(t,f)\}^2} = \left|y_f^i(t,f)\right| \quad (9)$$

Based on the obtained instantaneous phase, the instantaneous frequency is derived as:

$$IF_f^i(t,f) = \frac{1}{2\pi}\frac{d}{dt}IP_f^i(t,f) \tag{10}$$

At the computation of time-frequency features, the filter specifications are altered for multiple times. The final values of IE, IP and F are obtained through the average of the values obtained at past iterations.

D. Non-linear features

Under this category, we compute the non-linear features from each epoch. For this purpose, we consider computing two entropy features from each epoch: Shannon entropy [45] and spectral entropy [46, 47]. The main reason behind the computation of entropies is that they are directly linked with the information present within the Epoch. Mainly, the irregular variations present in the epoch are characterized by entropy. Even though entropy and variance explore the same information, the entropy is non-sensitive while the variance is feature amplitude sensitive. Hence, we can state that larger entropy values signify more information, which is required for getting better classification results.

Here, the Shannon entropy measures the irregularity of the signal in time domain and directly associated with the irregular temporal patterns of the epoch. For a given i^{th} epoch x_f^i the Shannon entropy is mathematically is expressed as:

$$H_s = -\sum p_k \log p_k \tag{11}$$

where, k denotes the entire discrete amplitude bin's range of an epoch and p_k denotes the probability of a sample in the epoch having the k^{th} amplitude. Here we consider each amplitude as a bin, and hence the total number of bins present in each epoch is the total number of unique amplitudes of the epoch.

The spectral entropy is measured with the help of Shannon entropy. Spectral entropy explores the digital signal's irregularity in the frequency domains. For the computation of Spectral entropy, initially the signal is subjected to Fourier transform followed by power spectrum computation. Consider x_f^i and $p(x_f^i)$ be the corresponding probability, the Shannon entropy is calculated with Eq. (11) while the Spectral entropy is calculated as Shannon entropy of the Probability distribution of a signal in the frequency domain. Hence the Shannon entropy is considered as spectral entropy if we assume $p(x_f^i)$ as the probability distribution of a power spectrum. The power spectrum is defined as $psd(p(x_f^i))$ which signifies the absolute value of the digital signal in the Discrete Fourier Transform (DFT).

3.4 Hierarchical classification

Once the feature extraction is completed, then they are trained to the system through machine learning algorithms. Here, we built a hierarchical classifier based on brute-force search (BFS) methodology. The major reason behind this model is the idea of decision tree structures and ensemble classification [48] where the typical classification issues are solved by breaking them into small problems and solved by the simple traditional classifiers. The proposed classifier architecture is a unique architecture that can classify one class against the remaining classes. In such a way, due to the presence of only a few classes, we employed a general exhaustive search technique using which a class or a group of classes can be well separated from each other.

Here, to find the better combination through the BFS mechanism, initially one class is separated against all the two classes; then the separation of a set of two-classes compared to a single-class and a set of two-classes compared to two-classes. Further, the set of three classes is separated in comparison with a single-class, against two classes, and then against three-classes. This process is repeated until the investigation of all combinations of classes is completed. The classifier is formulated into a hierarchical tree structure by such kind of separable set of classes as: The full classes are attended for classification in the 1st level of the tree, and optimal combination of one or two classes is determined against the remaining classes. As a result, the number of classes will be reduced to the next level. The remaining classes are attended for classification and determined against one or two classes in the 2nd level of the tree. This process is repeated until the tree structure reaches the leaves.

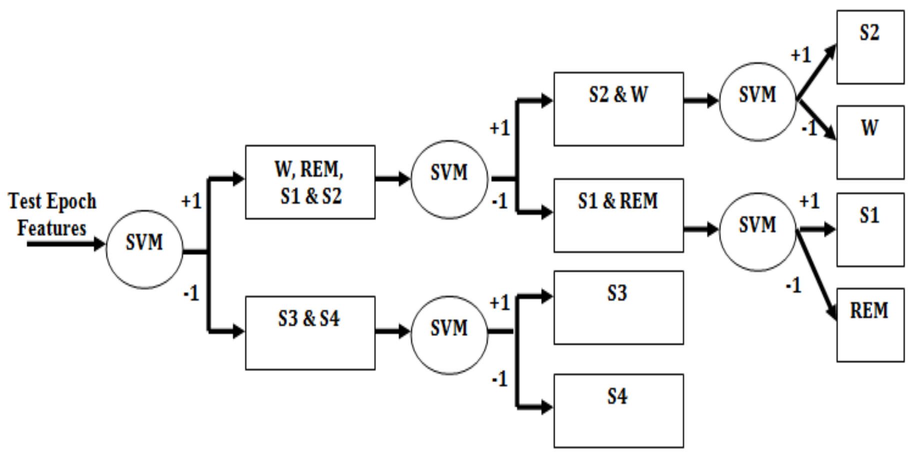

Figure 2. Hierarchical classification over Sleep-EDF dataset

Here, we employed a simple SVM classifier at each node of the proposed hierarchical classification tree. The main reason is that it didn't require any complex training process, and it also won't produce any bias for the classes with more samples in the database. Further, the SVM is superior in local decision making which is a more suitable technique for class distributions that are multi-modal in nature. Though the SVM's decision-making is completely dependent on the majority voting, it ensures better performance even for classes that have scattered feature distributions in multi-dimensional space. In such situations, SVM suits more than several state-of-the-art classifiers including K-NN, Random Forest, and Linear Discriminant Analysis (LDA) [49]. Figure 2 shows the hierarchical classification of six classes in the Sleep-EDF dataset. Algorithm 1 shows the step by step process of sleep stage disorders classification.

Algorithm: Hierarchical Classification

Input: Test Epoch Features (Ftr_{test}), Trained Epoch features and Labels ($Ftr^{j}_{Train}, Lableset$)
Output: Class Label (CL)

Step 1: Stage 1 classification, $Label_1 = Svaclassify(Ftr_{test}, Ftr^{j}_{Train})$
If $Label_1 = +1$,
    CL=W, REM, S1 and S2
Else if $Label_1 = -1$,
    CL=S3 and S4
End if
Step 2a: Stage 2 classification, $Label_{2a} = Svaclassify(Ftr_{test}, Ftr^{j}_{Train})$
If $Label_{2a} = +1$,
    CL=W and S2
Else if $Label_{2a} = -1$,
    CL=S1 and REM
End if
Step 2b: Stage 2 classification, $Label_{2b} = Svaclassify(Ftr_{test}, Ftr^{j}_{Train})$
If $Label_{2b} = +1$,
    CL=S3
Else if $Label_{2b} = -1$,
    CL=S4
End if
Step 3a: Stage 3 classification, $Label_{3a} = Svaclassify(Ftr_{test}, Ftr^{j}_{Train})$
If $Label_{3a} = +1$,
    CL=S2
Else if $Label_{3a} = -1$,
    CL = W
End if
Step 3b: Stage 3 classification, $Label_{3b} = Svaclassify(Ftr_{test}, Ftr^{j}_{Train})$
If $Label_{3b} = +1$,
    CL =S1
Else if $Label_{3b} = -1$,
CL =REM
End if

4. EXPERIMENTAL INVESTIGATIONS

This section explores the details of experimental investigations carried out on the proposed sleep staging mechanism. For experimental investigation, we used MATLAB tools with signal processing and statistics toolboxes. In the current section, we initially explore the details of the dataset and then the results derived from various simulation studies.

A. Dataset

A standard and publicly available dataset called Sleep-EDF [50, 51] is used here for experimental investigations. Sleep-EDF is composed of two distinct sets of files: Sleep Cassette (SC) files and Sleep Telemetry (ST) files. The Sleep Cassette files are related to twenty healthy subjects (10 male and 10 female, ages ranging from 21-35) without any sleep associated issues or medications. All the SC files were acquired over a period of two consecutive day and nights, approximately 20 hours. During the acquisition, the subjects continued their normal activities but wore a modified Walkman like cassette tape recorder. The files are named in the form SC4ssNEO-PSG.edf, where ss are the subject number and N is the night. During the time of recording, the 2nd night of subject 13 and the 1st nights of subjects 36 and 52 were lost because their laserdisc or cassette has fallen down. The sampling frequency is maintained at 100 Hz for both EOG and EEG signals. A sample EEG waveform and its staging details are shown in Figure 3.

On the other hand, the ST files are associated with subjects who have mild sleep issues, like difficulty falling asleep. The acquisition of ST files was done after Temazepam and Placebo intakes. The PSGs were acquired for about nine hours from the subjects who stayed in the hospital for two nights. The individuals were equipped with a small, high-quality telemetry equipment at the time of capture. ST7ssNEJ0-PSG.edf is the format used for the file names, where ss stand for the topic number and N for the night. Similar to SC files, the sampling frequency is maintained at 100 Hz for both EOG and EEG signals.

With the aid of R&K rules, each EEG wave is manually segmented into 30s epochs. W, R, S1, S2, S3, S4, M (movement Time), and U (Uncensored) were the annotations. In this research, we use a single-channel EEG acquired at the Pz-Oz electrodes because past approaches [52] suggested that it has deeper sleep stages and can be detected effectively. Further, the last two annotations, namely M and U, are eliminated since there is a very little chance that they will occur; consequently, there are six classes overall. An illustration of hierarchical classification on the Sleep-EDF is shown in Figure 2 where we employed the binary SM classifier, which classifies only two classes at a time. In the first level, the binary SVM Classified the set of S3&S4 classes with other classes such as S1, S2, REM, and W. This means that the first level of classification identified whether or not the epoch belongs to S3 or S4. Next, the 2nd level classification executes the classification of three subset of classes. They are S3 & S4, S2 & W, and R & S1. At the second level, the input features are fed to a binary SVM to determine the input epoch into one of three subsets. Finally, at the third, two SVMs are applied to classify the two subsets into individual classes. One SVM separates S1 from R, and another classifier separates W from S2. At every level, the SVM produces two labels for two classes: +1 and -1. For example, in the third level, S2 is assigned for +1 while W is assigned for -1. Since SVM is a binary classifier, it can classify only two classes at a time. Hence, the complexity of SVM is evaluated as $O(N^2)$.

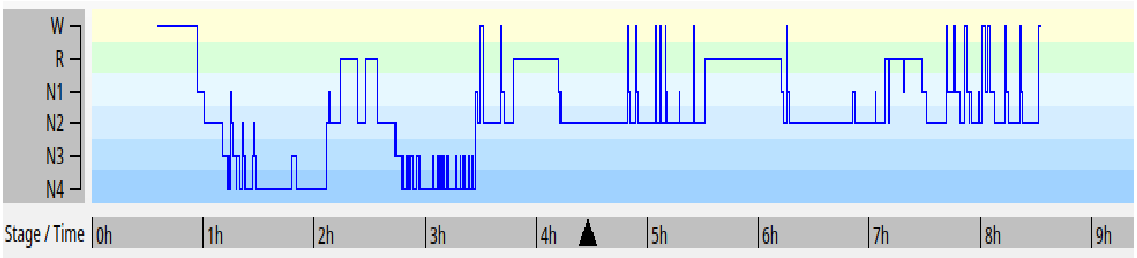

Figure 3. Example EEG wave form with sleep staging

B. Performance evaluation

Here we assess the proposed approach's performance with the help of four metrics of performance that include the Accuracy, Specificity, the Sensitivity, and the Kappa Coefficient. Mathematically, they are described as:

$$Accuracy = \frac{Estimated\ Results \cap Groundtruth\ Results}{length(Groundtrugth\ Results)} \quad (12)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (13)$$

$$Specificity = \frac{TP}{TP + FP} \quad (14)$$

where, FP denotes False Positives, TP denotes the True positives, and FN denotes False Negatives. Next, kappa coefficient (k) is defined as:

$$k = \frac{P_a - P_c}{1 - P_c} \quad (15)$$

where, P_a denotes the portion of correctly classified epochs and P_c denotes the portion of epochs expected to get classified correctly. They are mathematically described as:

$$P_a = \sum_{i=1}^{N} C_{ii} \Big/ \sum_{i=1}^{N}\sum_{j=1}^{N} C_{ij} \quad (16)$$

And

$$P_c = \sum_{i=1}^{N}\left(\sum_{j=1}^{N} C_{ij} \sum_{j=1}^{N} C_{ji}\right) \Big/ \left(\sum_{j=1}^{N}\sum_{j=1}^{N} C_{ij}\right)^2 \quad (17)$$

where,C_{ii} denotes i^{th} predicted class and i^{th} ground truth class. The range of kappa coefficient is defined as 0-1 where 0 denotes poor and 1 denotes excellent. Further the kappa coefficient is arranged in six levels as 0.00 to 0.2 (slight), 0.21 to 0.40 (Fair), 0.41 to 0.60 (Moderate), 0.61 to 0.80 (Substantial) and >0.80 (excellent).

Table 3 shows the proposed approach's performance in accuracy and sensitivity terms and is specific to different classes of the Sleep-EDF dataset. As per the results, the mean accuracy, mean sensitivity, and mean specificity are observed as 92.3500%, 82.3652%, and 92.9614%, respectively. Further, we can see that the specificity of all sleep stages is greater than 88%, which infers that the suggested strategy is effective in terms of accurate classification. Next, Table 4 shows the correct classification and misclassification rates for individual classes. Here, we evaluate the performance of the suggested strategy in the terms of excellent and bad classifications by measuring TPR and FPR. From the results, we observed that the maximum TPR is at Wake and the minimum TPR is at S3. Next, the maximum misclassification of wake is observed at S1 and S2, while for the remaining stages, such as REM, S1, S2, S3, and S4, it is observed at S1, REM, S4, S4, and S3, respectively. Since S3 and S4 resemble similar characteristics, their misclassification rates are observed to be high.

Table 3. Performance of proposed approach in terms of sensitivity, accuracy, and specificity

Class	Accuracy	Sensitivity	Specificity
Wake	97.4120	97.3562	97.4420
REM	94.3756	92.7689	94.5560
S1	91.6320	81.8472	93.5560
S2	88.2300	74.7520	90.8140
S3	79.7780	56.2380	93.1856
S4	82.8850	63.2300	88.2150

Table 4. False Positive Rates (FPR) and True Positive Rate (TPR) assessment on Sleep-EDF dataset

	Wake	REM	S1	S2	S3	S4
Wake	**97.3562**	0	1.3219	1.3219	0	0
REM	2.0751	**92.7689**	5.1560	0	0	0
S1	2.2564	8.2580	**81.8472**	4.6630	1.4554	1.5200
S2	2.9162	0	6.7750	**74.7520**	3.3368	12.2200
S3	2.5950	5.3210	3.2250	10.2560	**56.2380**	22.3650
S4	1.1550	7.5620	11.2230	2.5680	14.2620	**63.2300**

Next, to check the effectiveness of the proposed hierarchical classification mechanism, we used different classifiers (namely KNN, RF, and LDA), and obtained results are reported in Figure 4 and Figure 5 with sensitivity and specificity, respectively. The outcomes demonstrate that, when compared to traditional classifiers, the proposed method performed better. The average sensitivity of HSVM is observed as 82.3652%, while for remaining methods such as KNN, RF, and LDA, it is observed as 47.3300, 64.2033, and 72.3520%, respectively. Further, the maximum sensitivity is observed by the proposed HSVM at the wake-sleep stage, while the minimum sensitivity is observed at S3. Next, the average specificity of HSVM is observed as 92.9583%, while for remaining methods such as KNN, RF, and LDA, it is observed as 86.1200, 87.5420, and 73.6350%, respectively. Further, the maximum specificity is observed by the proposed HSVM at the wake-sleep stage, while the minimum sensitivity is observed at S4.

Figure 6 and Figure 7 show the details of FNR and FPR at different classifiers respectively. The average FNR of HSVM, KNN, RF and LDA classifiers is observed as 22.8333%, 52.6667%, 46.0000% and 47.8333% respectively. Similarly, the average FPR of HSVM, KNN, RF and LDA classifiers is observed as 7.0417%, 14%%, 12.3333% and 27.00% respectively. From the results, it can be noticed that the proposed HSCM has less misclassification than the other classifiers. The main reason is that the HSVM ensures stage-wise discrimination capability to the scoring system through SVM algorithm.

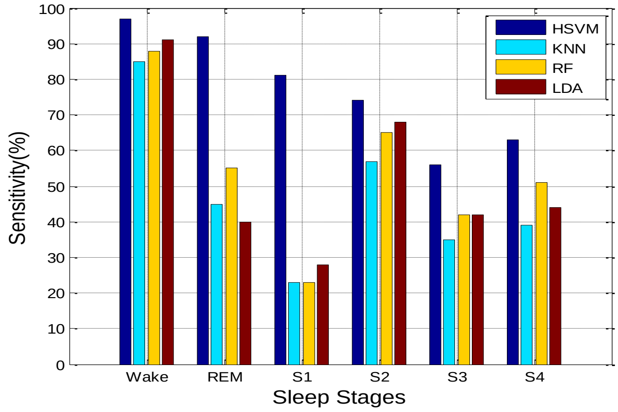

Figure 4. Sensitivity comparison between different classifiers

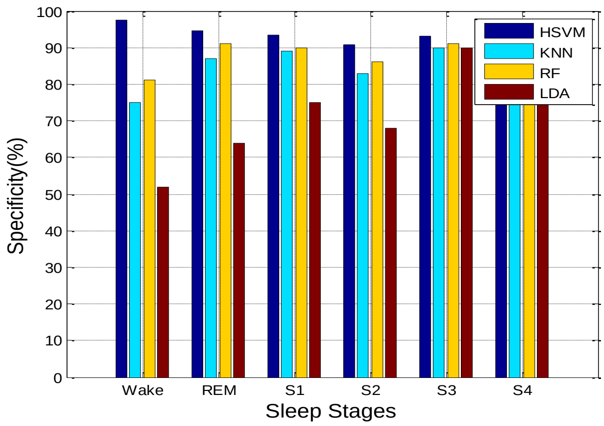

Figure 5. Specificity comparison between different classifiers

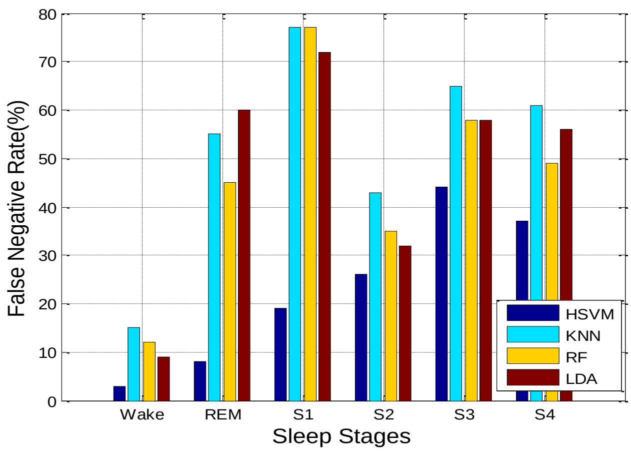

Figure 6. FNR comparison between different classifiers

The comparison between the suggested method and various other ways that are currently in use is demonstrated in Table 5. Here we refer to recent methods (both machine learning-based and deep learning-based) for comparison purposes. From the results, we can see that the suggested method has shown superior performance to the traditional approaches. Even though some authors, like Zhou et al. [31] and Phan et al. [36], applied deep learning approaches, they didn't identify the discriminative features between sleep stages. They directly applied deep learning methods to the Raw EEG signal, which cannot provide discriminative knowledge to the system. On the other hand, the machine learning methods employed different features to describe the epoch and processed them through traditional classifiers. But they didn't focus on the local level discrimination, i.e., the two different sleep stages may resemble at some epochs if the global features like Wavelet features and frequency features are considered. In such instances, time domain features such as mean, SD, MAD, and RSS give better discrimination.

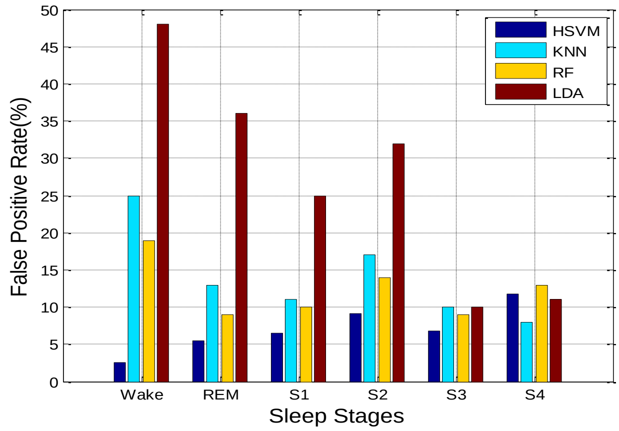

Figure 7. FPR comparison between different classifiers

Table 5. Overall system performance at different classifiers

Method/Metric		**Accuracy (%)**	**Kappa Coefficient**	**Testing Time (sec)**
Proposed	SVM	92.3500	0.8646	20.2
	RF	79.6360	0.7715	15.6
	LDA	65.8900	0.6235	12.2
Karimzadeh et al. [27]	KNN	88.9700	-	-
	RF	77.6600	-	-
	LDA	61.0100	-	-
Alickovic and Subasi [20]	Ensemble SVM	91.1000	-	-
Zapata et al. [23]	MT & C	87.6000		-
	SVM-Q	90.0000		-
Zhou et al. [31]	TSE+SAN	85.8000	0.8540	42.1
Phan et al. [36]	SeqSleepNet	85.000	0.8470	-

The recent existing method proposed by Karimzadeh et al. [27] considered Shannon entropy for different frequency domain images constructed based on IP and IE. However, such image formation induces more redundant information, which adds additional processing burden to the system. Moreover, the redundant information results in higher false positive rates between similar sages like S3 and S4. In such a situation, there is a need for additional features that can overcome the confusion and ensure perfect discrimination between sleep stages, and the proposed method can provide such flexibility. Hence, our method gained better accuracy and kappa coefficient.

5. CONCLUSION

This study suggests a straightforward and accurate system for classifying sleep stages that is based on composite features and hierarchical classification. The composite features concept is an alternative solution for sleep analysis that ensures perfect discrimination between sleep stages, especially those with a shorter transition period. The composite features include a total of 22 features in multiple domains, including time, frequency, tie-frequency, and non-linear. In addition, we suggested a hierarchical classifier, which solves the complex classification problem by breaking it into several small problems. At the classification, we employed the Binary SVM, which ensures a perfect decision between two different classes. An extensive set of simulation experiments on the developed system through the standard Sleep-EDF database proves its effectiveness. The average improvement in accuracy gained by the proposed system is observed at 3.38% compared to the recent existing methods.

As the manual diagnosis process for sleep stage disorders identification is a lengthy process, the proposed automatic system is more beneficial. The proposed system can be integrated into current clinical workflows at diagnosis stage because the current diagnosis methods are time consuming. This system reduces manual errors along with manual burden and no limitations are found. In perspective of ethical considerations, the periodic EEG signal recordings are requiring which makes the patients discomfort.

As we considered only machine learning algorithms for automatic sleep staging, in the future we can apply deep learning algorithms for better identification of sleep stages through EEG. Deep learning can provide more distinctive features which can ensure a better classification performance at sleep staging.

REFERENCES

[1] Panda, S., Hogenesch, J.B., Kay, S.A. (2002). Circadian rhythms from flies to human. Nature, 417(6886): 329-335. https://doi.org/10.1038/417329a

[2] Chung, M.H., Kuo, T.B., Hsu, N., Chu, H., Chou, K.R., Yang, C.C. (2009). Sleep and autonomic nervous system changes-enhanced cardiac sympathetic modulations during sleep in permanent night shift nurses. Scandinavian Journal of Work, Environment & Health, 35(3): 180-187. http://www.jstor.org/stable/40967775

[3] Ellenbogen, J.M., Payne, J.D., Stickgold, R. (2006). The role of sleep in declarative memory consolidation: Passive, permissive, active or none? Current Opinion in Neurobiology, 16(6): 716-722. https://doi.org/10.1016/j.conb.2006.10.006

[4] Torabi-Nami, M., Mehrabi, S., Borhani-Haghighi, A., Derman, S. (2015). Withstanding the obstructive sleep apnea syndrome at the expense of arousal instability, altered cerebral autoregulation and neurocognitive decline. Journal of Integrative Neuroscience, 14(02): 169-193. https://doi.org/10.1142/S0219635215500144

[5] Karimzadeh, F., Nami, M., Boostani, R. (2017). Sleep microstructure dynamics and neurocognitive performance in obstructive sleep apnea syndrome patients. Journal of Integrative Neuroscience, 16(2): 127-142.

[6] Jo, H.G., Park, J.Y., Lee, C.K., An, S.K., Yoo, S.K. (2010). Genetic fuzzy classifier for sleep stage identification. Computers in Biology and Medicine, 40(7): 629-634. https://doi.org/10.1016/j.compbiomed.2010.04.007

[7] Wolpert, E.A. (1969). A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Arch. Arch Gen Psychiatry, 20(2): 246-247. https://doi.org/10.1001/archpsyc.1969.01740140118016

[8] Berry, R.B., Budhiraja, R., Gottlieb, D.J., Gozal, D., Iber, C., Kapur, V.K., Marcus, C.L., Mehra, R., Parthasarathy, S., Quan, S.F., Redline, S., Strohl, K.P., Sally, L., Ward, D., Tangredi, M.M. (2012). Rules for scoring respiratory events in sleep: Update of the 2007 AASM manual for the scoring of sleep and associated events: Deliberations of the sleep apnea definitions task force of the American Academy of Sleep Medicine. Journal of Clinical Sleep Medicine, 8(5): 597-619. https://doi.org/10.5664/jcsm.2172

[9] Moser, D., Anderer, P., Gruber, G., Parapatics, S., Loretz, E., Boeck, M., Kloesch, G., Heller, E., Schmidt, A., Danker-Hopfe, H., Saletu, B., Zeitlhofer, J., Dorffner, G. (2009). Sleep classification according to AASM and Rechtschaffen& Kales: effects on sleep scoring parameters. Sleep, 32(2): 139-149. https://doi.org/10.5665/sleep/32.2.139

[10] Güneş, S., Polat, K., Yosunkaya, Ş. (2010). Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting. Expert Systems with Applications, 37(12): 7922-7928. https://doi.org/10.1016/j.eswa.2010.04.043

[11] Danker-Hopfe, H., Anderer, P., Zeitlhofer, J., Boeck, M., Dorn, H., Gruber, G., Heller, E., Loretz, E., Moser, D., Parapatics, S., Saletu, B., Schmidt, A., Dorffner, G. (2009). Interrater reliability for sleep scoring according to the Rechtschaffen& Kales and the new AASM standard. Journal of Sleep Research, 18(1): 74-84. https://doi.org/10.1111/j.1365-2869.2008.00700.x

[12] Magalang, U.J., Chen, N.H., Cistulli, P.A., Fedson, A.C., Gíslason, T., Hillman, D., Penzel, T., Tamisier, R., Tufik, S., Phillips, G., Pack, A.I., SAGIC Investigators. (2013). Agreement in the scoring of respiratory events and sleep among international sleep centers. Sleep, 36(4): 591-596. https://doi.org/10.5665/sleep.2552

[13] Zhang, X., Dong, X., Kantelhardt, J.W., Li, J., Zhao, L., Garcia, C., Glos, M., Penze, T., Han, F. (2015). Process and outcome for international reliability in sleep scoring. Sleep and Breathing, 19: 191-195.

[14] Rosenberg, R.S., Van Hout, S. (2014). The American academy of sleep medicine inter-scorer reliability program: Respiratory events. Journal of Clinical Sleep Medicine, 10(4): 447-454. https://doi.org/10.5664/jcsm.3630

[15] Penzel, T., Zhang, X., Fietze, I. (2013). Inter-scorer reliability between sleep centers can teach us what to improve in the scoring rules. Journal of Clinical Sleep Medicine, 9(1): 89-91. https://doi.org/10.5664/jcsm.2352

[16] Norman, R.G., Pal, I., Stewart, C., Walsleben, J.A., Rapoport, D.M. (2000). Interobserver agreement among sleep scorers from different centers in a large dataset. Sleep, 23(7): 901-908.

[17] Aboalayon, K.A.I., Faezipour, M., Almuhammadi, W.S., Moslehpour, S. (2016). Sleep stage classification using

EEG signal analysis: A comprehensive survey and new investigation. Entropy, 18(9): 272. https://doi.org/10.3390/e18090272
[18] Memar, P., Faradji, F. (2017). A novel multi-class EEG-based sleep stage classification system. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(1): 84-95. https://doi.org/10.1109/TNSRE.2017.2776149
[19] Sharma, M., Makwana, P., Chad, R.S., Acharya, U.R. (2023). A novel automated robust dual-channel EEG-based sleep scoring system using optimal half-band pair linear-phase biorthogonal wavelet filter bank. Applied Intelligence, 53(15): 18681-18699. https://doi.org/10.1007/s10489-022-04432-0
[20] Alickovic, E., Subasi, A. (2018). Ensemble SVM method for automatic sleep stage classification. IEEE Transactions on Instrumentation and Measurement, 67(6): 1258-1265. https://doi.org/10.1109/TIM.2018.2799059
[21] Akinduko, A.A., Gorban, A.N. (2014). Multiscale principal component analysis. Journal of Physics: Conference Series, 490(1): 012081 https://doi.org/10.1088/1742-6596/490/1/012081
[22] Widasari, E.R., Tanno, K., Tamura, H. (2020). Automatic sleep disorders classification using ensemble of bagged tree based on sleep quality features. Electronics, 9(3): 512. https://doi.org/10.3390/electronics9030512
[23] Zapata, I.A., Li, Y., Wen, P. (2022). Rules-Based and SVM-Q methods with multitapers and convolution for sleep EEG stages classification. IEEE Access, 10: 71299-71310. https://doi.org/10.1109/ACCESS.2022.3188286
[24] Ye, J., Xiao, Q., Wang, J., Zhang, H., Deng, J., Lin, Y. (2021). Cosleep: A multi-view representation learning framework for self-supervised learning of sleep stage classification. IEEE Signal Processing Letters, 29: 189-193. https://doi.org/10.1109/LSP.2021.3130826
[25] Sarkar, P., Etemad, A. (2020). Self-supervised ECG representation learning for emotion recognition. IEEE Transactions on Affective Computing, 13(3): 1541-1554. https://doi.org/10.1109/TAFFC.2020.3014842
[26] Eldele, E., Ragab, M., Chen, Z., Wu, M., Kwoh, C.K., Li, X., Guan, C. (2023). Self-supervised contrastive representation learning for semi-supervised time-series classification. IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/TPAMI.2023.3308189
[27] Karimzadeh, F., Boostani, R., Seraj, E., Sameni, R. (2017). A distributed classification procedure for automatic sleep stage scoring based on instantaneous electroencephalogram phase and envelope features. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(2): 362-370. https://doi.org/10.1109/TNSRE.2017.2775058
[28] Li, X., Cui, L., Tao, S., Chen, J., Zhang, X., Zhang, G.Q. (2017). Hyclasss: A hybrid classifier for automatic sleep stage scoring. IEEE Journal of Biomedical and Health Informatics, 22(2): 375-385. https://doi.org/10.1109/JBHI.2017.2668993
[29] Laffan, A., Caffo, B., Swihart, B.J., Punjabi, N.M. (2010). Utility of sleep stage transitions in assessing sleep continuity. Sleep, 33(12): 1681-1686. https://doi.org/10.1093/sleep/33.12.1681
[30] Loh, H.W., Ooi, C.P., Vicnesh, J., Oh, S.L., Faust, O., Gertych, A., Acharya, U.R. (2020). Automated detection of sleep stages using deep learning techniques: A systematic review of the last decade (2010-2020). Applied Sciences, 10(24): 8963. https://doi.org/10.3390/app10248963
[31] Zhou, W., Zhu, H., Shen, N., Chen, H., Fu, C., Yu, H., Shu, F., Chen, C., Chen, W. (2022). A lightweight segmented attention network for sleep staging by fusing local characteristics and adjacent information. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31: 238-247. https://doi.org/10.1109/TNSRE.2022.3220372
[32] Abdollahpour, M., Rezaii, T.Y., Farzamnia, A., Saad, I. (2022). A two-stage learning convolutional neural network for sleep stage classification using a filterbank and single feature. IEEE Access, 10: 60597-60609. https://doi.org/10.1109/ACCESS.2022.3180730
[33] Sun, C., Fan, J., Chen, C., Li, W., Chen, W. (2019). A two-stage neural network for sleep stage classification based on feature learning, sequence learning, and data augmentation. IEEE Access, 7: 109386-109397. https://doi.org/10.1109/ACCESS.2019.2933814
[34] Leino, A., Korkalainen, H., Kalevo, L., Nikkonen, S., Kainulainen, S., Ryan, A., Duce, B., Sipilä, K., Sahlman, J., Miettinen, T., Töyräs, J., Myllymaa, S., Leppänen, T., Myllymaa, K. (2022). Deep learning enables accurate automatic sleep staging based on ambulatory forehead EEG. IEEE Access, 10: 26554-26566. https://doi.org/10.1109/ACCESS.2022.3154899
[35] Liu, Z., Luo, S., Lu, Y., Zhang, Y., Jiang, L., Xiao, H. (2022). Extracting multi-scale and salient features by MSE based U-structure and CBAM for sleep staging. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31: 31-38. https://doi.org/10.1109/TNSRE.2022.3216111
[36] Phan, H., Andreotti, F., Cooray, N., Chén, O.Y., De Vos, M. (2018). Joint classification and prediction CNN framework for automatic sleep stage classification. IEEE Transactions on Biomedical Engineering, 66(5): 1285-1296. https://doi.org/10.1109/TBME.2018.2872652
[37] Aristimunha, B., Bayerlein, A.J., Cardoso, M.J., Pinaya, W.H.L., De Camargo, R.Y. (2023). Sleep-Energy: An energy optimization method to sleep stage scoring. IEEE Access, 11: 34595-34602. https://doi.org/10.1109/ACCESS.2023.3263477
[38] Abdollahpour, M., Rezaii, T.Y., Farzamnia, A., Saad, I. (2020). Transfer learning convolutional neural network for sleep stage classification using two-stage data fusion framework. IEEE Access, 8: 180618-180632. https://doi.org/10.1109/ACCESS.2020.3027289
[39] Acharya, R., Faust, O., Kannathal, N., Chua, T., Laxminarayan, S. (2005). Non-linear analysis of EEG signals at various sleep stages. Computer Methods and Programs in Biomedicine, 80(1): 37-45. https://doi.org/10.1016/j.cmpb.2005.06.011
[40] Fraiwan, L., Lweesy, K., Khasawneh, N., Wenz, H., Dickhaus, H. (2012). Automated sleep stage identification system based on time-frequency analysis of a single EEG channel and random forest classifier. Computer Methods and Programs in Biomedicine, 108(1): 10-19. https://doi.org/10.1016/j.cmpb.2011.11.005
[41] Boashash, B. (1992). Estimating and interpreting the

instantaneous frequency of a signal. I. Fundamentals. Proceedings of the IEEE, 80(4): 520-538. https://doi.org/10.1109/5.135376
[42] Picinbono, B. (1997). On instantaneous amplitude and phase of signals. IEEE Transactions on Signal Processing, 45(3): 552-560. https://doi.org/10.1109/78.558469
[43] Sameni, R., Seraj, E. (2017). A robust statistical framework for instantaneous electroencephalogram phase and frequency estimation and analysis. Physiological Measurement, 38(12): 2141. https://doi.org/10.1088/1361-6579/aa93a1
[44] Seraj, E., Sameni, R. (2017). Robust electroencephalogram phase estimation with applications in brain-computer interface systems. Physiological Measurement, 38(3): 501. https://doi.org/10.1088/1361-6579/aa5bba
[45] Bruhn, J., Lehmann, L.E., Röpcke, H., Bouillon, T.W., Hoeft, A. (2001). Shannon entropy applied to the measurement of the electroencephalographic effects of desflurane. The Journal of the American Society of Anesthesiologists, 95(1): 30-35. https://doi.org/10.1097/00000542-200107000-00010
[46] Rezek, I.A., Roberts, S.J. (1998). Stochastic complexity measures for physiological signal analysis. IEEE Transactions on Biomedical Engineering, 45(9): 1186-1191. https://doi.org/10.1109/10.709563
[47] Inouye, T., Shinosaki, K., Sakamoto, H., Toi, S., Ukai, S., Iyama, A.,Katsuda, Y., Hirano, M. (1991). Quantification of EEG irregularity by use of the entropy of the power spectrum. Electroencephalography and Clinical Neurophysiology, 79(3): 204-210. https://doi.org/10.1016/0013-4694(91)90138-T
[48] Che, D., Liu, Q., Rasheed, K., Tao, X. (2011). Decision tree and ensemble learning algorithms with their applications in bioinformatics. Software Tools and Algorithms for Biological Systems, Springer, New York, NY, USA, 191-199. https://doi.org/10.1007/978-1-4419-7046-6_19
[49] Grčar, M., Fortuna, B., Mladenič, D., Grobelnik, M. (2006). kNN versus SVM in the collaborative filtering framework. In Data Science and Classification. Springer Berlin Heidelberg, pp. 251-260. https://doi.org/10.1007/3-540-34416-0_27
[50] Kemp, B., Zwinderman, A.H., Tuk, B., Kamphuisen, H.A., Oberye, J.J. (2000). Analysis of a sleep-dependent neuronal feedback loop: The slow-wave microcontinuity of the EEG. IEEE Transactions on Biomedical Engineering, 47(9): 1185-1194. https://doi.org/10.1109/10.867928
[51] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23): e215-e220. https://doi.org/10.1161/01.CIR.101.23.e215
[52] Bajaj, V., Pachori, R.B. (2013). Automatic classification of sleep stages based on the time-frequency image of EEG signals. Computer Methods and Programs in Biomedicine, 112(3): 320-328. https://doi.org/10.1016/j.cmpb.2013.07.006