



# A novel automated robust dual-channel EEG-based sleep scoring system using optimal half-band pair linear-phase biorthogonal wavelet filter bank

Manish Sharma<sup>1</sup> · Paresh Makwana<sup>1</sup> · Rajesh Singh Chad<sup>1</sup> · U Rajendra Acharya<sup>2,3,4</sup>

Accepted: 21 December 2022 / Published online: 8 February 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

Nowadays, the hectic work life of people has led to sleep deprivation. This may further result in sleep-related disorders and adverse physiological conditions. Therefore, sleep study has become an active research area. Sleep scoring is crucial for detecting sleep-related disorders like sleep apnea, insomnia, narcolepsy, periodic leg movement (PLM), and restless leg syndrome (RLS). Sleep is conventionally monitored in a sleep laboratory using polysomnography (PSG) which is the recording of various physiological signals. The traditional sleep stage scoring (SSG) done by professional sleep scorers is a tedious, strenuous, and time-consuming process as it is manual. Hence, developing a machine-learning model for automatic SSG is essential. In this study, we propose an automated SSG approach based on the biorthogonal wavelet filter bank's (BWFB) novel least squares (LS) design. We have utilized a huge Wisconsin sleep cohort (WSC) database in this study. The proposed study is a pioneering work on automatic sleep stage classification using the WSC database, which includes good sleepers and patients suffering from various sleep-related disorders, including apnea, insomnia, hypertension, diabetes, and asthma. To investigate the generalization of the proposed system, we evaluated the proposed model with the following publicly available databases: cyclic alternating pattern (CAP), sleep EDF, ISRUC, MIT-BIH, and the sleep apnea database from St. Vincent's University. This study uses only two unipolar EEG channels, namely O1-M2 and C3-M2, for the scoring. The Hjorth parameters (HP) are extracted from the wavelet subbands (SBS) that are obtained from the optimal BWFB. To classify sleep stages, the HP features are fed to several supervised machine learning classifiers. 12 different datasets have been created to develop a robust model. A total of 12 classification tasks (CT) have been conducted employing various classification algorithms. Our developed model achieved the best accuracy of 83.2% and Cohen's Kappa of 0.7345 to reliably distinguish five sleep stages, using an ensemble bagged tree classifier with 10-fold cross-validation using WSC data. We also observed that our system is either better or competitive with existing state-of-art systems when we tested with the above-mentioned five databases other than WSC. This method yielded promising results using only two EEG channels using a huge WSC database. Our approach is simple and hence, the developed model can be installed in home-based clinical systems and wearable devices for sleep scoring.

**Keywords** EEG (Electroencephalography) · PSG (Polysomnographic) · Classification · Ensemble Bagged Tree (EBT) · Sleep stages · Wavelets · Filter bank · Wisconsin Sleep Cohort (WSC)

## 1 Introduction

Sleep is one of the most natural and essential requirements for the human body to function properly, both physically and mentally [1]. While we sleep, our muscles are relaxed

while the brain retains new informations [2], eliminates toxic waste, and releases growth hormones. The body repairs cells and replenishes energy [3]. Normal healthy sleep is defined by adequate length, sleep quality, suitable timing, regularity, and the absence of sleep disruptions or disorders. Insomnia, excessive drowsiness and obstructive sleep apnea (OSA) are among sleep disorders that can negatively affect our mental, physical, social, and emotional well-being. This explains why neuro-anatomists and neuro-physiologists have been researching sleep for over a century [4, 5].

---

✉ Manish Sharma  
manishsharma.iitb@gmail.com

Extended author information available on the last page of the article.

Sleep scoring is the process of extracting sleep cycle information which can be used for detecting and treating sleep-related disorders. Traditionally Sleep scoring is performed using polysomnogram (PSG) recordings obtained from patients during their overnight sleep in a sleep laboratory. Further, sleep scoring professionals analyze PSG recordings and classify the sleep into stages using either Rechtschaffen and Kales (R&K) guidelines, or the more recent recommendations published by the American Academy of Sleep Medicine (AASM). According to R&K's guidelines PSG recordings are first divided in 20 or 30-second epochs, which are then classified as wakefulness (W), rapid eye movement (REM) sleep, or Non-REM sleep (NREM). Non-REM sleep is further subdivided into the first, second, third, and fourth stages (also called S1, S2, S3, and S4). In contrast, AASM standards merge the NREM stages S3 and S4 to a single deep sleep stage known as N3, often known as slow-wave sleep (SWS) [6].

PSG recordings comprise electrooculograms (EOG), electroencephalograms (EEG), electrocardiograms (ECG), and electromyograms (EMG). Among them, EEG is amongst the most important physiological signals in terms of clinical relevance and practical utility. The EEG spectrum is categorized into five rhythms (frequency bands), namely  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  and  $\theta$ . The frequency and the amplitude of these different rhythms can reveal important insight for identifying, observing, and treating neurological abnormalities and diseases [7]. When a subject is awake, the EEG shows  $\beta$  waves, high in frequency and low in amplitude. N1 is the dozing off stage; during this stage, the frequency of the brain waves begin to slow, and this stage generally lasts 5-10 minutes/cycle. The body undergoes a

more subdued condition in the second stage of sleep, which includes a decrease in body temperature and pulse rate. The brain activity in this stage is slower than in earlier stages, and there is a new pattern of acute and quick fall and rise in amplitude of brain waves termed the K-complex [8] followed by short bursts of activity called sleep spindle [9]. The second stage lasts for 20-25 minutes/cycle in a healthy patient. The third stage of sleep is also known as deep sleep. This is very crucial since it is when the body is most relaxed and the brain produces growth hormones, allowing muscular tissues to mend and grow. Inadequate deep sleep may cause a person to be heavy and sluggish in the morning. This stage lasts about 20-40 minutes/cycle, depending on the person's health and age. As humans grow older, growth hormones decrease, resulting in less deep sleep [10]. During REM stage of sleep, the eyes move swiftly from one side to the other beneath closed eyelids, breathing becomes rapid and heart rate accelerates, giving the appearance of wakefulness. A person is more likely to experience a dream when in the REM stage. Their arms and legs are briefly paralysed to prevent them from acting out the dream [11].

The sleep stage begins with N1 and advances to N2, N3, further, the body returns to N2 and then eventually reaches REM. After this, the body returns to N2, and the complete sleep cycle repeats. A healthy individual experiences this sleep cycle four to five times every night, with each cycle lasting longer than the one before it [12]. Figure 1 shows the EEG sample epochs of a subject taken from the database for each sleep stage.

Traditional sleep scoring is a time-consuming, labor-intensive procedure that is vulnerable to human errors due to many hours of continuous examination. Aside from the

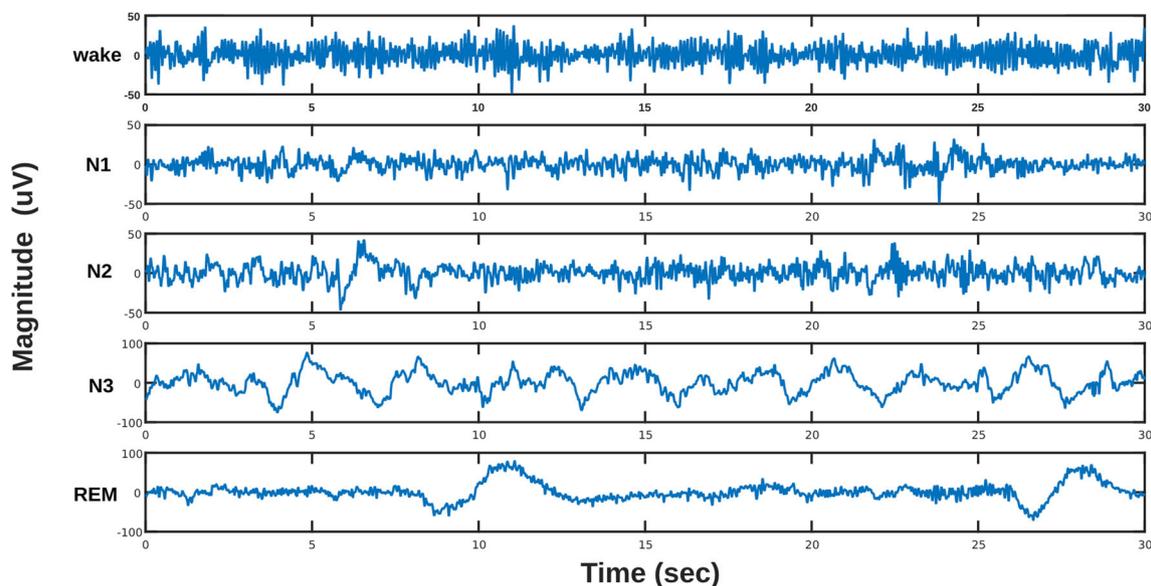


Fig. 1 Signal representing all the stages of a subject (subject ID : 10198)

expense and difficulty of assessing sleep patterns, patients must spend the night inside a fully equipped sleep laboratory with sticky electrodes and cables attached to their heads. Although sleeping in this unfamiliar environment is uncomfortable, these treatments may impact the patient's sleep efficiency. Due to these challenges, several automated methods have been proposed for sleep-stage scoring [13–16]. In this paper, we propose an automated sleep scoring system that uses subbands (SBs) obtained from optimal stopband energy (SBE) from a biorthogonal wavelet filter bank (BWFB) involving only two channels of EEG signals. For each SB, Hjorth parameters (HP) are computed for classification purposes. Since this suggested approach uses only two EEG channels, it may be easier to install compared to prior cutting-edge systems that used PSG or numerous EEG channels and other physiological inputs for sleep scoring [17]. As a result, the patient's comfort may be enhanced.

In literature, many studies [18–30] have been proposed for automatic sleep stage classification. Sharma et al. [13] in their recent study, have presented a sleep scoring method for healthy (good sleepers) and unhealthy subjects using cyclic alternating pattern (CAP) sleep database [31, 32] consisting of 108 PSG files [31]. They employed norm-based features and attained an accuracy of 85.3%. Langkvist et al. [33] proposed a five-stage automated sleep stage classification system based on data from St. Vincent's University Hospital and University College Dublin (SVUH) [31]. They used one EEG, one EMG, two EOG signal for their study. They developed deep belief network for the purpose and achieved the best result of 72.2% using the database containing only 25 subjects. Tzimourta et al. [34] utilised the Institute of System and Robotics—University of Coimbra (ISRUC) Sleep dataset [35] for EEG-Based autonomous sleep stage classification [35]. They employed a total of 118 subjects, healthy and unhealthy, out of which they used 100 subjects and 87187 epochs (30 sec) in total. This data was used for five-stage classifications using six EEG channels. Using time and frequency-based features coupled with random forest classifier, they got the best accuracy of 75.3%. Michielli et al. [5] utilised Physiobank's sleep-EDF database [36, 37] for automated sleep stage classification based on a long short-term memory (LSTM) approach [31]. They achieved 83.6% accuracy for two-class classification and 86.74% for five-class classifications using ten healthy subjects. Willemen et al. [38] in their work, achieved 69% accuracy for five-class classification (WAKE-REM-N1-N2-N3) and 81% for three-class classification using 85 PSG files obtained from 36 healthy subjects acquired from Vrije Universiteit Brussel, with SVM classifier. Fonseca et al. [39] conducted sleep stage classification using cardio-respiratory signals obtained from PSG recording of 48 subjects acquired from

the SIESTA project [40]. They achieved an accuracy of 69% for four-class classification and 80% for three-class classification using multi-class Bayesian linear discriminant with time-varying probabilities. Recently Phan et al. [41] proposed a convolutional neural network (CNN) framework for classification and prediction of automatic sleep stage scoring, using two different data-sets (i) sleep cassette (SC) from Sleep EDF [31] containing 20 subjects and (ii) Montreal Archive of Sleep Studies (MASS) [42] containing 200 subjects. They achieved an accuracy of 82.3% for SC and 83.6% for MASS for five-class classification. Shi et al. [43] achieved an accuracy of  $81.1 \pm 0.15\%$  for a two-class classification employing multiple kernel learning and using data from SVUH Dublin [31] containing 25 PSG from 25 subjects with the suspected sleep disorder. Yuan et al. [44] in their recent study on sleep stage classification, used the dataset from SVUH Dublin [31] containing a total of 25 PSG. They employed hybrid CNN as a classifier and obtained an accuracy of  $0.7424 \pm 0.594$ . Radha et al. [45] in their recent work on sleep stage classification used LSTM model with dataset obtained from EU SIESTA project [40] containing 292 subjects. Thus, all above mentioned studies have used various publicly available databases such as CAP, Sleep EDF, ISURC sleep data, SVUH Dublin, Massachusetts Institute of Technology at Boston's Beth Israel Hospital (MIT-BIH), Vrije University Brussel, SIESTA project and MASS. In all the above studies, The sleep model's classification accuracy ranged from 75% to 85%, while the number of subjects employed in the study ranged from 10 to 292. Hence, it is essential to use a large and diverse database that contains a large number of diverse subjects to generalize the results. The Wisconsin Sleep Cohort (WSC) seems to be a good candidate that satisfies the above requirement, which consists of two databases containing 2570 subjects, including good sleepers and patients suffering from various sleep disorders. Hence, we employed the WSC sleep database in this study.

The salient features of the proposed study are mentioned below:

- This study designed a new class of linear phase optimum wavelet bi-orthogonal filter bank using the least squares (LS) method wherein the filter bank is halfband pair filter bank (HPFB).
- We have used a huge Wisconsin Sleep Cohort (WSC) which includes WSC\_dataset.1 (1715 PSG) and WSC\_dataset.2 (716 PSG) compared to other state-of-the-art studies.
- From the total of 2431 subjects, 21,65,205 (30 sec) epochs have been extracted and used in this study.
- Our proposed model used only two unipolar EEG signals (O1\_M2 and C3\_M2). Hence, it can easily be incorporated into user-friendly hardware and devices.

- Both single and dual-channel combined are considered in this study.
- Feature extraction is done using HP, a good discriminator for EEG signals.
- For classification, we employed supervised machine learning classifiers with a 10-fold cross-validation technique.
- The existing proposed systems employed a single database to develop models, so generalization to new databases is unknown. To investigate the generalization of the proposed system, we evaluated the model with the following publicly available databases: CAP, sleep EDF, ISRUC, MIT-BIH, and the sleep apnea database from St. Vincent's University.
- Our system is found to be either better or competitive with existing state-of-art systems when we tested our model with the above-mentioned five databases other than WSC.
- The proposed method yielded good accuracy and Kappa score using two unipolar EEG signals. Thus, the proposed model may be installed in real-time, low-cost IoT configurations and hardware, without compromising the comfort of the subjects.

## 2 Dataset used

The dataset used for this study was obtained from the National Sleep Research Resource (NSRR) and provided by the Wisconsin Sleep Cohort (WSC). The WSC cohort contains two different databases collected using different systems and at different periods. The WSC\_dataset\_1 was collected during 2000-2009 using Grass Heritage System, this dataset has 15 signals collected from 1715 recordings, while the second dataset WSC\_dataset\_2 (from 2009 - to present) was collected using Grass Comet Lab-Based system, and has 18 signals obtained from 716 recordings. The detailed descriptions of both datasets are mentioned in Table 1.

The majority of the PSG files contain at least two EEG, two EOG, two EMG, one ECG signal, and other signals such as snore, nasal airflow, oral airflow, thoracic, abdominal, position, SpO2. We employed only two EEG channels in the proposed study, namely O1-M2 and C3-M2 (O1 and C3 with reference to M2).

The WSC cohort contains PSG recordings of 2570 subjects with 24,62,228 epochs of 30 seconds duration. For EEG recording either O1-M2 or C3-M2 channels are used.

**Table 1** Detailed explanation of the two datasets used in this work

Channel	Label	WSC_dataset_1		WSC_dataset_2	
		SF (Hz)	$f_c$ (Hz)	SF (Hz)	$f_c$ (Hz)
Left EOG	E1	100	30	200	35
Right EOG	E2	100	30	200	35
Left Central EEG	C3_M2	100	30	200	35
Left Occipital EEG	O1_M2	100	30	200	35
Left Frontal EEG	F3_M2	NA	NA	200	35
Frontal EEG	Fz_M2	NA	NA	200	35
Central EEG	Cz_M2	NA	NA	200	35
Parietal EEG	Pz_M2	NA	NA	200	35
Linked Center & Left Chin EMG	cchin_1	NA	NA	200	70
Chin EMG	chin	100	30	NA	NA
Linked Left & Right Leg EMG	lleg_r	100	30	200	70
Snore	snore	100	30	200	70
ECG	ECG	100	30	200	35
Airflow	flow	NA	NA	200	15
Nasal Airflow	nasalflow	100	30	NA	NA
Oral Airflow	oralflow	100	30	NA	NA
Nasal Pressure	nas_pres	100	30	200	15
Thoracic	thorax	100	NA	200	15
Abdominal	abdomen	100	NA	200	15
Position	position	100	NA	200	NA
SpO2	spo2	100	NA	200	NA

Cut-off frequency ( $*f_c$ ), Sampling Frequency (SF), NA(Not Available)

**Table 2** Details of PSG and channel used in the study

Channels	Dataset	No. of PSG
O1_M2 and C3_M2	WSC_dataset.1	1715
O1_M2 and C3_M2	WSC_dataset.2	716
Total	WSC	2431

The EEG recordings from both EEG channels (O1-M2 and C3-M2) have been collected in most of the subjects except for a few subjects. In this study, we have recruited only those subjects for which EEG recordings of both channels are available as indicated in Table 2. Thus, out of 2570 subjects, we used 2431 subjects. Hence, 21,65,205 epochs of each 30s are generated from 2431 PSG recordings. Table 3 contains information about WSC utilized in this study. Further, to evaluate the generalization of the proposed system, we have tested the model with five more publicly available databases: CAP, Slep-EDF, MIT-BIH, ISURC and SVUH can be found in [31, 32, 35–37, 37] and [46] respectively.

### 3 Methodology

The Fig. 2 depicts the proposed method. In the following sections, we have explained the data preparation and the technique we employed.

#### 3.1 Segmentation of the dataset

We used two unipolar EEG channels O1\_M2 and C3\_M2 in this work, and classified them both individually and collectively. As indicated in Table 4, the dataset was segmented into nine data subsets.

**Up-sampling datasets** The EEG channels of WSC\_dataset.1 were sampled at 100 Hz, whereas the EEG signals of WSC\_dataset.2 were sampled at 200 Hz. We have upsampled EEG signals of WSC\_dataset.1 to 200 Hz to have all EEG signals at the same sampling frequency. The

upsampled dataset was segmented into three data subsets, as indicated in Table 5. The samples were zero-padded to increase the sampling rate from 100 to 200 Hz.

Furthermore, the EEG signal of every subject is segmented into 30 seconds epochs. The epochs are labeled according to the annotation files provided with the dataset.

##### 3.1.1 Filtering and wavelet decomposition

The wavelet transform has shown to be an effective technique for evaluating non-linear, non-stationary EEG data. The wavelets provide both the frequency and temporal domain information. Hence, the feature extraction using wavelets can be an ideal approach for sleep scoring using EEG signals [47, 48]. Many signal processing applications have used two-channel orthogonal and wavelet biorthogonal filter banks [49]. Linear phase filters are desired in applications such as image and communication coding [50–56]. Except for Haar filter banks, orthogonal filters cannot attain a linear phase. In few situations, biorthogonal filter banks are favored over orthogonal filter banks. This study's FB was created utilizing a least-squares (LS) technique. The FB is designed by minimizing a quadratic objective function for the given linear constraints. The quadratic function of pass-band and stop-band errors is used as the objective function. With a half band analysis filter, we designed a linear phase optimum biorthogonal wavelet filter bank.

The initial stage in designing the filter bank was to create a half band analysis lowpass filter (HALF) by creating a linearly restricted convex quadratic optimization problem with a convex mixture of passband and stopband errors. Following the design of the HALF, the synthesis lowpass filter (SLPF) is constructed similarly to the HALF, but with certain modifications such as (i) avoiding SLPF being confined as halfband filters, and (ii) along with the regularity criteria (vanishing moment) such that the perfect reconstruction conditions are fulfilled.

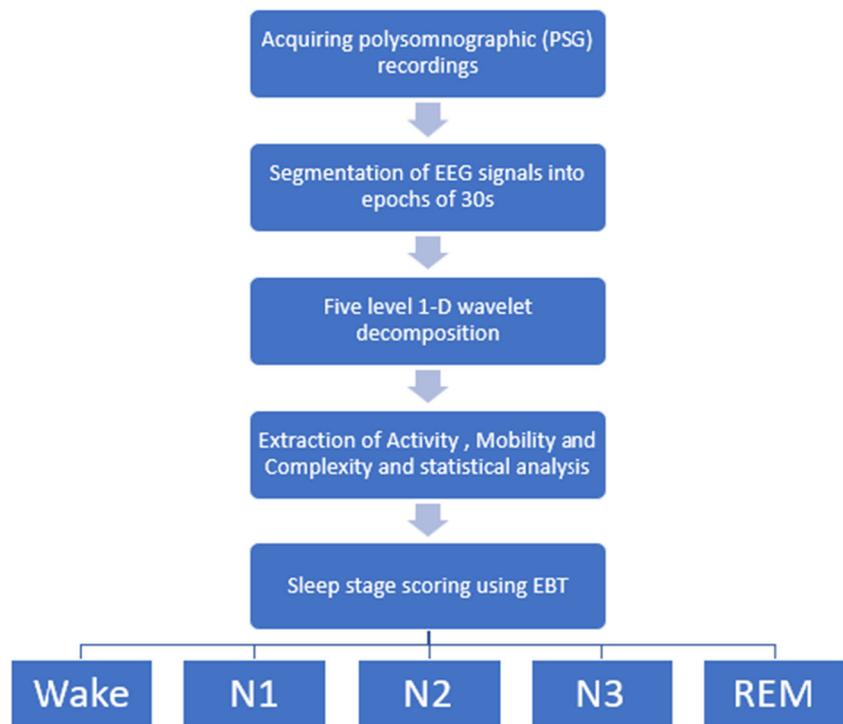
##### 3.1.2 Design of wavelet filter bank

We used a two-stage design process to create a novel class of linear-phase biorthogonal filter banks (FBs). The FB

**Table 3** Distribution of epochs across sleep stages in two datasets of WSC

Stages	WSC_dataset.1	WSC_dataset.2	WSC
Wake	2,85,395 (18.5%)	1,34,267 (21.6%)	4,19,662 (19.3%)
N1	1,28,988 (8.36%)	44,091 (7.08%)	1,73,079 (7.9%)
N2	8,34,612 (54.1%)	3,26,516 (52.4%)	11,61,128 (53.6%)
N3	81,425 (5.27%)	40,539 (6.51%)	1,21,964 (5.6%)
REM	2,12,106 (13.7%)	77,266 (12.4%)	2,89,372 (13.3%)
Total	15,42,526 (71.2%)	6,22,679 (27.4%)	21,65,205

**Fig. 2** Proposed method used for the automated classification of sleep stages



designed is a halfband pair filter bank (HPFB) wherein the lowpass filter of the two-channel FB is a halfband filter [57, 58]. The FB's design has been characterized as a two-stage optimization issue. In the initial stage, HALF is designed, and in the second stage, the SLPF filter has been designed. Both HALF and SLPF have been obtained as solutions of least squares (LS) optimization problem. The design process involves solving linear equations for the given linear constraints with the two-stage approach. Using the technique, one can design optimal filters with different lengths and vanishing moments [13–15, 59–61] designed easily. The filters are designed to minimize stop-band energy (SBE). The SBE can be expressed in a quadratic form. The vanishing moment (VM) restrictions are stated as

a set of linear equations. Hence, the optimization problem is a LS problem whose solution can be obtained easily and effectively. We have used the CVX toolbox to solve optimization problems [62, 63]. In this study, we have used 29/15 filter bank in which the orders of HALF and SLFP are 28 and 14, respectively. The HALF and SLFP both contain 4 VMs each. The frequency responses and pole-zero plots of the filters used are shown in Figs. 3 and 4.

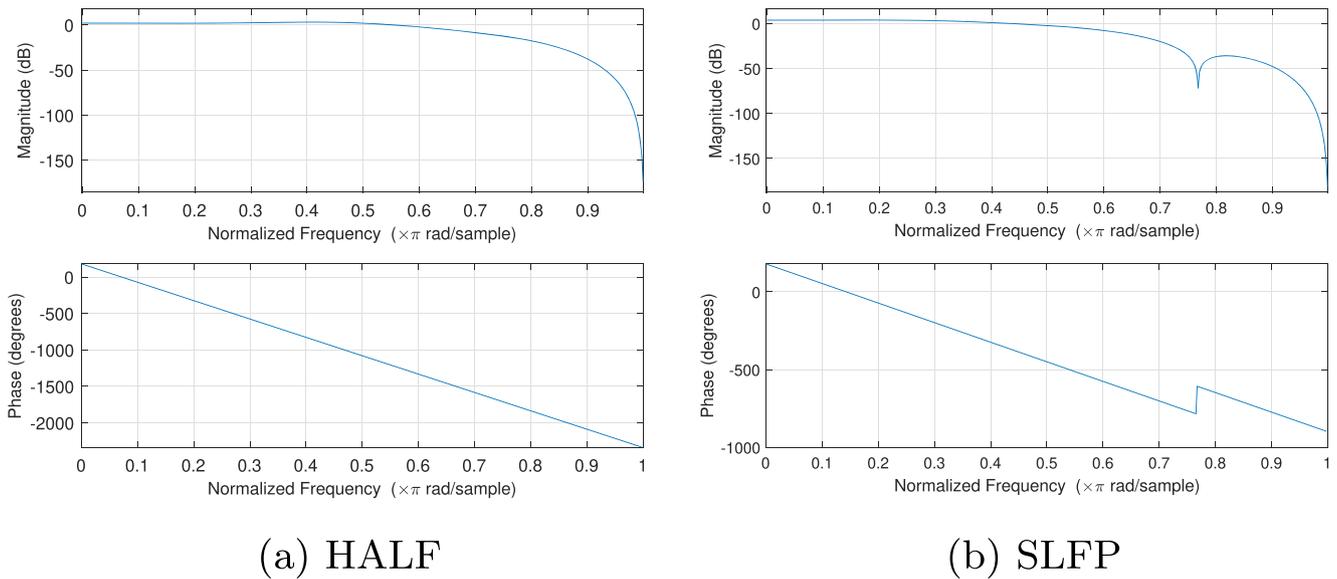
Determining the level of decomposition is critical. We have tested 3-level, 5-level and 7-level wavelet decomposition and got best results from 5-level decomposition, thus 5-level wavelet decomposition is chosen. The frequency components that dominate the signal and the most significant sub-bands for categorization are retained. Based on this, one-dimensional five-level wavelet decomposition is conducted using the designed 29/15 FB. The wavelet decomposition yielded the six sub-bands with frequency ranges of 0–3.125 Hz (A), 3.125–6.25 Hz (D1), 6.25–12.5 Hz (D2), 12.5–25 Hz (D3), 25–50 Hz (D4), 50–100 Hz (D5). In this study, the lower frequency band 0–3.125 Hz (A) is

**Table 4** Details of classification tasks (CT) and corresponding data subsets

CT	Data subset used	Channel employed
CT1	WSC_dataset_1	O1-M2
CT2	WSC_dataset_1	C3-M2
CT3	WSC_dataset_1	O1-M2 + C3-M2
CT4	WSC_dataset_2	O1-M2
CT5	WSC_dataset_2	C3-M2
CT6	WSC_dataset_2	O1-M2 + C3-M2
CT7	WSC_dataset_1 + WSC_dataset_2	O1-M2
CT8	WSC_dataset_1 + WSC_dataset_2	C3-M2
CT9	WSC_dataset_1 + WSC_dataset_2	O1-M2 + C3-M2

**Table 5** Details of classification tasks and data subset post upsampling

CT	Dataset	Channel
CT10	WSC_dataset_1	O1-M2
CT11	WSC_dataset_1	C3-M2
CT12	WSC_dataset_1	O1-M2 + C3-M2



**Fig. 3** Frequency response obtained using the biorthogonal filter

referred as approximate coefficients [64], while the remaining bands with higher frequencies are referred as detailed coefficients.

### 3.1.3 Feature extraction

In our study, we extracted Hjorth parameters(HP) namely activity, mobility, and complexity developed by Bo Hjorth [65], from six sub-bands of each epoch’s five level one dimensional wavelet decomposition. These parameters primarily indicate the time domain properties of EEG signals. Table 6 provides a brief explanation for these parameters. After obtaining six sub-bands from each epoch’s five-level one-dimensional wavelet decomposition, the Activity, Complexity and Mobility of the subbands are extracted.

Here  $y(t)$  represents the signal in the time domain and  $\frac{dy(t)}{dt}$  denotes the first derivative of the signal.

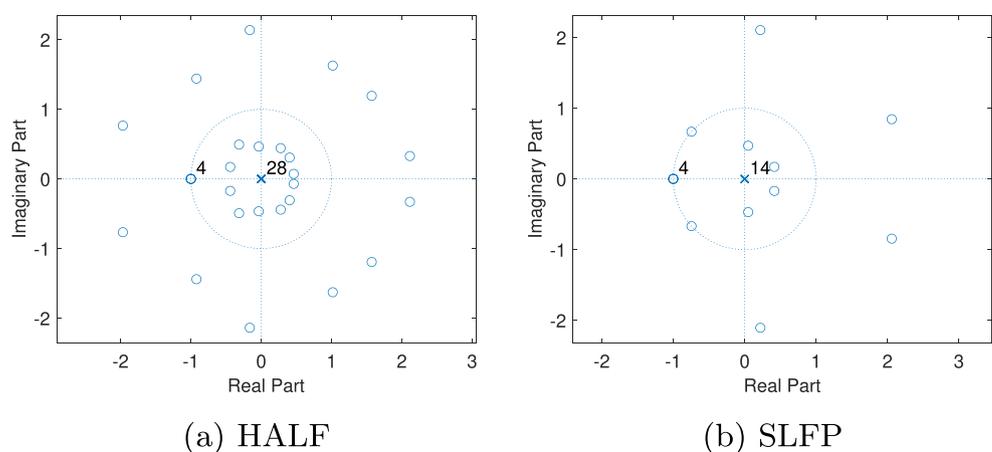
### 3.1.4 Classification

The HP features that have been chosen are fed into supervised machine learning classifiers. In order to classify, Support vector machines (SVM) [66, 67], ensemble bagged trees (EBT) [68, 69], decision trees [68, 70], K-nearest neighbours (KNN) [71] and naive Bayes [72] are used. Among the above-mentioned classifiers, EBT delivered the best classification as indicated in Table 7. In this work, we have developed the classifiers using ten-fold cross-validation (CV) strategy.

## 4 Results

The final classification results obtained using the extracted features are obtained using the classification learner app

**Fig. 4** Visualization of filter pair in Z-plane



**Table 6** Details of HP used in this study

HP	Description	Formula
Activity	Activity represents the variance of the signal $y(t)$ , Also indicates the energy of the signal $y(t)$	$\text{var}(y(t))$
Mobility	It provides an estimation of the mean frequency, Also represents the proportion of standard deviation of the power spectrum.	$\sqrt{\frac{\text{var}(\frac{dy(t)}{dt})}{\text{var}(y(t))}}$
Complexity	It provides an estimate of the bandwidth frequency, Also represents change in frequency.	$\frac{\text{Mobility}(\frac{dy(t)}{dt})}{\text{Mobility}(y(t))}$

of MATLAB. It may be noted from Table 3 that the observations and epochs between different sleep stages are not the same, thus making the entire dataset very unbalanced. Therefore, to evaluate such unbalanced systems we have used Cohen's Kappa coefficient ( $\kappa$ ). Classification accuracy and  $\kappa$  value for respective data subset is mentioned in Table 8.

We have performed ANOVA test to evaluate the statistical significance of the features, from the test performed we obtained p-values of all the features as zero which indicates null hypothesis is rejected and all features are statistically significant. Therefore, we have used all features during the classification. We have also ranked the features. The summary of the features extracted and the rankings are mentioned in Tables 9 and 10 respectively.

The WSC database used for this study contains two main datasets the WSC\_dataset\_1 and WSC\_dataset\_2, the method of collecting the PSG signals from subjects, make these two datasets different from each other in terms of sampling frequencies and channels present. As previously mentioned, we chose these O1\_M2 and C3\_M2 unipolar EEG channels since they are present in the majority of subject files for this study. The selected channels (O1\_M2 and C3\_M2) from each dataset had different sampling frequencies (100 Hz for WSC\_dataset\_1 and 200 Hz for WSC\_dataset\_2), thus taking into account the diversity in sampling frequency. Hence we decided to classify the entire dataset in 9 different classification tasks. Ranging from CT1-CT9, which is described in Table 4. The accuracies obtained by these different classification tasks using EBT

classifier are presented in Table 8. The best overall accuracy of 83.2% was obtained with Kappa value of 0.7345.

Below is a detailed overview of the aforementioned classification tasks. For CT-1, we had an overall accuracy of 73.4% and Kappa Value of 0.5584, the confusion matrix for which is Table 11(c). Individual accuracies for different classes, namely wake, N1, N2, N3, and REM, are 82.6%, 9.9%, 90%, 27.9% and 52.2%, respectively. In CT-2, an overall accuracy of 77% and Kappa value of 0.6275 is achieved with wake, N1, N2, N3 and REM having 83.7%, 16.4%, 91%, 34.3% and 67.3% respectively, the confusion matrix of which is shown in Table 11(d). CT-3 yielded an overall accuracy of 78.5% and Kappa value of 0.6499 with individual accuracies for wake, N1, N2, N3 and REM as 85.8%, 17.4%, 92%, 36% and 35.9% respectively the confusion matrix for which is Table 11(h). For CT-4, we had an overall accuracy of 78.4% and a Kappa Value of 0.6524 the confusion matrix for which is Table 11(b). Individual accuracies for different classes namely wake, N1, N2, N3 and REM are 89.2%, 12.4%, 90.7%, 56.5% and 56.5% respectively. In CT-5, an overall accuracy of 81.1% and Kappa value of 0.7017 is achieved with the wake, N1, N2, N3, and REM, having 90.1%, 18.2%, 91.1%, 64.3% and 67.5%, respectively, confusion matrix of which is shown

**Table 7** Performance of sleep stage classification achieved using different classifiers

Classifier	Classification accuracy(%)
Linear discriminant	68.4
Subspace KNN	70.3
Linear SVM	71.0
Fine gaussian SVM	74.7
Ensamble bagged trees	83.2
Subspace discriminant	67.1

**Table 8** Performance of sleep stage classification achieved using EBT classifier and 10-fold Cross Validation

CT	Accuracy	Cohens Kappa
CT-1	73.4%	0.5584
CT-2	77.0%	0.6275
CT-3	78.5%	0.6499
CT-4	78.4%	0.6524
CT-5	81.1%	0.7017
CT-6	83.2%	0.7345
CT-7	74.0%	0.4164
CT-8	77.5%	0.6363
CT-9	79.2%	0.5151
CT-10	75.6%	0.6524
CT-11	79.7%	0.6743
CT-12	81.5%	0.7020

**Table 9** Summary of features extracted from O1\_M2 EEG channel

Feature	Sub-band	Rank	wake(mean ± std)	N1(mean ± std)	N2(mean ± std)	N3(mean ± std)	N5(mean ± std)
Activity	1	16	6.9037e4 ± 6.2871e5	1.2468e4 ± 2.6774e5	8.7065e3 ± 1.4342e5	1.2790e4 ± 1.2504e5	6.7375e3 ± 1.8402e5
	2	4	437.0687 ± 7.1043e3	339.7297 ± 6.3899e3	264.5039 ± 5.5367e3	347.0319 ± 6.7100e3	374.3967 ± 6.5828e3
	3	11	189.7352 ± 1.8063e3	74.8133 ± 757.7816	68.1989 ± 700.0650	72.3573 ± 765.3363	68.0964 ± 805.0741
	4	14	493.3151 ± 7.8830e3	163.8336 ± 817.5680	193.0237 ± 1.0324e3	220.9435 ± 608.7185	118.2067 ± 1.0921e3
	5	2	1.2348e3 ± 2.8391e4	360.8364 ± 3.8548e3	507.2561 ± 4.4276e3	738.8226 ± 3.1324e3	267.9632 ± 4.2247e3
	6	17	3.3895e3 ± 7.8151e4	685.2076 ± 1.4498e4	1.0084e3 ± 1.7647e4	1.7201e3 ± 1.0102e4	531.0209 ± 1.2740e4
Mobility	1	1	0.8820 ± 0.2437	0.9777 ± 0.2214	1.0406 ± 0.2165	1.0353 ± 0.2116	0.9918 ± 0.2049
	2	13	1.8022 ± 0.0648	1.7879 ± 0.0623	1.7538 ± 0.0685	1.7457 ± 0.0695	1.7818 ± 0.0724
	3	5	1.8003 ± 0.0788	1.8147 ± 0.0676	1.8269 ± 0.0733	1.8131 ± 0.0789	1.8193 ± 0.0702
	4	10	1.6777 ± 0.1369	1.7642 ± 0.0713	1.7944 ± 0.0621	1.8159 ± 0.0467	1.7649 ± 0.0639
	5	12	1.6795 ± 0.1073	1.7141 ± 0.0735	1.7419 ± 0.0640	1.7719 ± 0.0579	1.7218 ± 0.0641
Complexity	6	3	1.6684 ± 0.1056	1.6544 ± 0.0841	1.6843 ± 0.0708	1.7205 ± 0.0554	1.6697 ± 0.0750
	1	9	1.0959 ± 0.1750	1.1287 ± 0.1503	1.0731 ± 0.1540	0.9943 ± 0.1396	1.1237 ± 0.1409
	2	7	0.4588 ± 0.1071	0.4774 ± 0.1052	0.5099 ± 0.1136	0.5258 ± 0.1256	0.4825 ± 0.1183
	3	18	0.4862 ± 0.1152	0.4623 ± 0.0823	0.4328 ± 0.0803	0.4534 ± 0.0801	0.4513 ± 0.0739
	4	15	0.6467 ± 0.1661	0.5819 ± 0.1088	0.5451 ± 0.1094	0.5046 ± 0.0804	0.5867 ± 0.0998
	5	6	0.7175 ± 0.1398	0.7003 ± 0.1138	0.6630 ± 0.1042	0.6097 ± 0.1021	0.6889 ± 0.0979
	6	8	0.7466 ± 0.1463	0.7811 ± 0.1115	0.7411 ± 0.0986	0.6827 ± 0.0831	0.7534 ± 0.1008

**Table 10** Summary of features extracted from C3\_M2 EEG channel

Feature	Sub-band	Rank	wake(mean + std)	N1(mean + std)	N2(mean + std)	N3(mean + std)	REM(mean + std)
	1	16	63428.48 ± 734470.44	10533.01 ± 286369.62	9754.33 ± 192770.99	23543.15 ± 359730.30	7082.80 ± 212060.26
	2	4	876.62 ± 11025.78	824.46 ± 10927.85	699.71 ± 9992.90	602.80 ± 8990.26	700.88 ± 9957.39
	3	14	295.60 ± 2816.43	147.58 ± 1379.64	136.02 ± 1234.05	117.29 ± 1013.98	115.58 ± 1277.34
	4	2	663.29 ± 10937.64	214.98 ± 1614.92	255.51 ± 1538.86	307.82 ± 637.72	157.63 ± 1432.63
	5	11	1731.37 ± 40327.51	504.10 ± 6203.37	736.77 ± 6212.60	1165.77 ± 3284.11	400.31 ± 4513.21
Activity	6	1	4567.52 ± 108866.64	983.59 ± 17309.70	1799.19 ± 21056.05	3684.05 ± 10858.84	823.63 ± 12978.18
	1	17	0.90 ± 0.23	0.98 ± 0.21	1.10 ± 0.22	1.10 ± 0.21	0.98 ± 0.19
	2	5	1.82 ± 0.05	1.81 ± 0.05	1.76 ± 0.07	1.75 ± 0.09	1.82 ± 0.05
	3	13	1.78 ± 0.08	1.80 ± 0.07	1.82 ± 0.09	1.81 ± 0.09	1.79 ± 0.07
	4	3	1.70 ± 0.10	1.75 ± 0.07	1.76 ± 0.08	1.78 ± 0.07	1.76 ± 0.05
	5	8	1.69 ± 0.09	1.74 ± 0.06	1.76 ± 0.05	1.78 ± 0.04	1.75 ± 0.05
Mobility	6	10	1.67 ± 0.10	1.66 ± 0.08	1.71 ± 0.07	1.75 ± 0.05	1.67 ± 0.08
	1	18	1.09 ± 0.16	1.14 ± 0.14	1.06 ± 0.16	0.98 ± 0.15	1.14 ± 0.13
	2	15	0.41 ± 0.10	0.41 ± 0.10	0.46 ± 0.12	0.48 ± 0.15	0.40 ± 0.10
	3	6	0.51 ± 0.11	0.48 ± 0.09	0.42 ± 0.09	0.44 ± 0.10	0.49 ± 0.09
	4	7	0.64 ± 0.14	0.60 ± 0.10	0.58 ± 0.13	0.54 ± 0.12	0.57 ± 0.08
	5	9	0.70 ± 0.14	0.65 ± 0.10	0.61 ± 0.09	0.57 ± 0.08	0.63 ± 0.08
Complexity	6	12	0.73 ± 0.14	0.76 ± 0.11	0.68 ± 0.10	0.61 ± 0.08	0.74 ± 0.10

**Table 11** Confusion matrix obtained corresponding to data subsets using EBT classifier with 10-fold CV

True class	Predicted class				
	Wake	N1	N2	N3	REM
<b>(a) Confusion matrix for CT-1</b>					
Wake	82.6%	2.5%	10.4%	<1%	4.3%
N1	19.2%	9.9%	48.5%	<1%	22.2%
N2	2.9%	1.2%	90.0%	1.8%	4.1%
N3	<1%	<1%	71.0%	27.9%	<1%
REM	8.3%	4.0%	35.1%	<1%	52.2%
Accuracy= 73.4% & $\kappa = 0.5584$ Sensitivity=75.8% & Specificity = 93.4%					
<b>(b) Confusion matrix for CT-4</b>					
Wake	89.2%	2.0%	6.6%	<1%	2.1%
N1	26.3%	12.4%	41.7%	<1%	19.5%
N2	3.0%	<1%	90.7%	2.1%	3.2%
N3	<1%	<1%	42.8%	56.5%	<1%
REM	6.5%	3.3%	33.7%	<1%	56.5%
Accuracy = 78.4% & $\kappa = 0.6524$ Sensitivity=79.0% & Specificity = 94.6%					
<b>(c) Confusion matrix for CT-7</b>					
Wake	83.0%	2.0%	11.0%	<1%	4.0%
N1	21.0%	9.0%	49.0%	<1%	21.0%
N2	3.0%	<1%	90%	2.0%	4.0%
N3	<1%	<1%	62.0%	37.0%	<1%
REM	9.0%	4.0%	37.0%	<1%	50.0%
Accuracy= 74.0% & $\kappa = 0.4164$ Sensitivity=76.2% & Specificity = 93.5%					
<b>(d) Confusion matrix for CT-2</b>					
Wake	83.7%	3.4%	9.1%	<1%	3.5%
N1	20.9%	16.4%	35.3%	<1%	27.1%
N2	2.8%	1.4%	91%	2.0%	3.1%
N3	<1%	<1%	64.2%	34.4%	<1%
REM	6.9%	6.4%	18.8%	<1%	67.3%
Accuracy = 77.0% & $\kappa = 0.6275$ Sensitivity=78.1% & Specificity = 94.3%					
<b>(e) Confusion matrix for CT5</b>					
Wake	90.1%	2.5%	5.5%	<1%	1.9%
N1	28.8%	18.2%	29.6%	<1%	23.3%
N2	2.8%	1.1%	91.1%	2.1%	2.9%
N3	<1%	<1%	34.9%	64.3%	<1%
REM	6.3%	5.1%	21.1%	<1%	67.5%
Accuracy = 81.1% & $\kappa = 0.7017$ Sensitivity=80.9% & Specificity = 95.3%					
<b>(f) Confusion matrix for CT8</b>					
Wake	85.0%	3.0%	8.0%	<1%	3.0%
N1	23.0%	15.0%	35.0%	<1%	26.0%
N2	3.0%	<1%	91%	2.0%	3.0%
N3	<1%	<1%	55.0%	44.0%	<1%
REM	8.0%	6.0%	22.0%	<1%	64.0%
Accuracy = 77.5% & $\kappa = 0.6363$ Sensitivity=78.5% & Specificity = 94.3%					

**Table 11** (continued)

True class	Predicted class				
	Wake	N1	N2	N3	REM
<b>(g) Confusion matrix for CT-6</b>					
Wake	91.5%	2.5%	4.5%	<1%	1.5%
N1	25.5%	20.8%	33.1%	<1%	20.5%
N2	2.4%	<1%	92.7%	1.7%	2.1%
N3	<1%	<1%	34.2%	65.0%	<1%
REM	5.4%	3.7%	17.3%	<1%	73.6%
Accuracy = 83.2% & $\kappa = 0.7345$ Sensitivity=82.9% & Specificity = 95.8%					
<b>(h) Confusion matrix for CT-3</b>					
Wake	84.8%	3.1%	8.2%	<1%	3.6%
N1	19.4%	17.5%	38.3%	<1%	24.7%
N2	2.4%	1.4%	92%	1.8%	2.5%
N3	<1%	<1%	62.9%	36.0%	<1%
REM	7.1%	5.1%	17.0%	<1%	35.9%
Accuracy = 78.5% & $\kappa = 0.6499$ Sensitivity=79.4% & Specificity = 94.6%					
<b>(i) Confusion matrix for CT-9</b>					
Wake	86.0%	3.0%	8.0%	<1%	3.0%
N1	22.0%	17.0%	37.0%	<1%	23.0%
N2	3.0%	<1%	92%	2.0%	2.0%
N3	<1%	<1%	54.0%	45.0%	<1%
REM	8.0%	5.0%	19.0%	<1%	69.0%
Accuracy = 79.2% $\kappa = 0.5151$ Sensitivity=80.0% & Specificity = 94.8%					
<b>(j) Confusion matrix for CT-10</b>					
Wake	86.7%	2.6%	6.7%	0.0%	4.0%
N1	18.4%	11.8%	46.1%	0.0%	23.7%
N2	2.3%	1.3%	91%	1.8%	4.1%
N3	0.5%	0.0%	62.7%	36.7%	0.1%
REM	8.4%	4.6%	31.4%	0.0%	55.5%
Accuracy = 75.6% & $\kappa = 0.6524$ Sensitivity=76.9% & Specificity = 93.8%					
<b>(k) Confusion matrix for CT-11</b>					
Wake	87.9%	3.4%	5.9%	0.0%	2.8%
N1	20.2%	19.1%	32.1%	0.0%	28.6%
N2	2.1%	1.4%	92%	1.9%	2.8%
N3	0.6%	0.0%	55.1%	44.2%	0.0%
REM	6.9%	7.2%	14.6%	0.0%	71.3%
Accuracy = 79.7% & $\kappa = 0.6743$ Sensitivity=80.3% & Specificity = 95%					

**Table 11** (continued)

True class	Predicted class				
	Wake	N1	N2	N3	REM
(l) Confusion matrix for CT-12					
Wake	89.2%	3.2%	4.9%	0.0%	2.8%
N1	18.5%	21.8%	34.2%	0.0%	25.4%
N2	1.7%	1.5%	93%	1.6%	2.2%
N3	0.4%	0.0%	52.1%	47.4%	0.0%
REM	6.9%	5.8%	12.4%	0.0%	74.9%
Accuracy = 81.5% & $\kappa$ = 0.7020					
Sensitivity=81.3% & Specificity = 95.3%					

in Table 11(e). CT-6 yielded an overall accuracy of 83.2% and Kappa value of 0.7345 with individual accuracies for the wake, N1, N2, N3, and REM, as 91.5%, 20.8%, 92.7%, 65% and 73.6%, respectively the confusion matrix for which is given in Table 11(g). For CT-7, we had an overall accuracy of 74% and Kappa value of 0.4164 the confusion matrix for which is Table 11(c). Individual accuracies for different classes namely wake, N1, N2, N3 and REM are 83%, 9%, 90%, 37% and 50% respectively. In CT-8 an overall accuracy of 77.5% and Kappa value of 0.6363 is achieved with wake, N1, N2, N3 and REM having 85.0%, 15.0%, 91%, 44.0% and 64.0% respectively, confusion matrix of which is shown in Table 11(f). CT-9 yielded an overall accuracy of 79.2% and Kappa value of 0.5151 with individual accuracies for wake, N1, N2, N3 and REM as 86.0%, 17.0%, 92%, 45% and 69% respectively the confusion matrix for which is given in Table 11(i).

From analysing the results of the above classification task CT1-CT9, it seems that C3-M2 gives better classification accuracy than O1-M2. We can observe that CT-2, CT-5, and CT-8 (CT with C3-M2 channels) provide better accuracy than CT1, CT4 and CT7 (CT with O1-M2 channel). For each WSC\_dataset, the results get better after combining both the channels. By comparing accuracies of CT-1,2,3 (CT from WSC\_dataset\_1) and CT-4,5,6 (CT from WSC\_dataset\_2), we can conclude that WSC\_dataset\_2 provides better accuracy than WSC\_dataset\_1.

Due to the inherent difference in the sampling frequency of the dataset available, we experimented by upsampling the WSC\_dataset\_1 from 100 Hz to 200 Hz and then performing classification task to check whether it would improve the accuracy. The description of these classification task is given in Table 5. The accuracies obtained by these different classification tasks using EBT classifier are presented in Table 8. After upsampling the WSC\_dataset\_1, it showed an improvement in both individual and overall accuracies. CT-10 showed an overall accuracy of 75.6% and Kappa value of 0.6524, with individual accuracies of wake, N1,

N2, N3, and REM as 86.7%, 11.8%, 91%, 36.7%, and 55.5% respectively the confusion matrix of which is given in Table 11(j). In CT-11, an overall accuracy of 79.7%, and Kappa value of 0.6743 is achieved with wake, N1, N2, N3, and REM have 87.9%, 19.1%, 92%, 44.2%, and 71.3%, respectively, confusion matrix of which is shown in Table 11(k). CT-12 yielded an overall accuracy of 81.5% and Kappa value of 0.7020 with individual accuracies for wake, N1, N2, N3 and REM as 89.2%, 21.8%, 93%, 47.4%, and 74.9%, respectively. The confusion matrix for which is given in Table 11(l).

## 5 Discussion

Our proposed method is the first study on sleep scoring developed using 2431 subjects belonging to the entire WSC sleep database. We employed two unipolar EEG channels (O1-M2 or/and C3-M2 channels). The proposed model is developed using a large number of 21,65,205 EEG epochs of 30s duration. The model is trained and tested using optimal wavelet-based, highly discriminating HP features. To ensure the robustness of the model, it is tested with 12 diverse data subsets (refer to Tables 4 and 5). To ensure all EEG signals have the same sampling frequency, the channels with an original sampling frequency of 100 Hz was upsampled to 200 Hz. The proposed model has yielded the best classification accuracy of 83.2% and kappa value of 0.7345. The high kappa value indicates the good discriminating ability of the proposed model.

The results obtained indicate that HP extracted from the C3-M2 channel presents better classification performance than the O1-M2 channels. Also, the model presents better classification performance for WSC\_dataset\_2 than WSC\_dataset\_1. The best overall result was achieved for CT-6 channel, with an accuracy of 83.2% and  $\kappa$  value of 0.7340. It is observed that the best individual accuracy is

obtained for N2 followed by wake, REM, N3, then N1. This trend can be observed in all classification tasks (CT-1 to CT-12). The lesser percentage of epochs of the N1 and N3 stage in the total epoch (Table 3) might contribute to this trend. The CT-10,11,12 (with upsampling) showed better accuracy and  $\kappa$  values than CT-1,2,3 (without upsampling). Thus, the up-sampling has been proved to enhance accuracy. The  $\kappa$  value gives a better insight into the robustness of the model in addition to the accuracy while using the unbalanced datasets. Therefore, we have also used  $\kappa$  value to evaluate our model, and our model yielded a high  $\kappa$  value of 0.74. We used a combination of optimal bi-orthogonal filter bank and HP for our datasets, which yielded better accuracies and Kappa values than other wavelet decomposition and feature extraction methods [56, 73]. As compared to the state-of-the-art methods [74–76], it may be noticed that our proposed method gave better accuracy for the N2 stage. Sharma et al. [15] did a six-class SSC using the sleep-EDF database (80 subjects). They used six sub-bands using optimal wavelet filter bank, norm features, and EBT classifier to get a best model with accuracy 85.3% (unbalanced) and 92.8% (unbalanced) and  $\kappa$  of 0.786 and 0.915. However, their model reported the best accuracy of 89% for the S2 class. Phan et al. [41] proposed a 5 class SSC using a joint classification-and-prediction framework based on convolutional neural networks (CNNs) using sleep EDF and MASS datasets. They obtained an individual accuracy of 88% (sleep EDF) and 86.9% (MASS). In the proposed work, the best accuracy of 92.7% is obtained for the N2 stage, as shown in Table 11. Sharma et al. [77] did SSC using sleep-EDF database with 76 subjects and reported the best accuracy of 98.3% and achieved only 17.3% accuracy for the N1 class, while our model obtained an accuracy of 21.1% for N1 stage detection. Thus, the proposed model also performed well in detecting the N1 class.

To assess the robustness of the proposed method, the model was tested with other databases along with the WSC database. We evaluated the model with the following publicly available data, namely CAP [36, 37], Sleep EDF [31, 32], ISRUC [35], MIT-BIH [37] and sleep apnea database from SVUH [46]. We observed that our model is either better or more competitive with other mentioned models when we tested with these five databases. While employing the same number of subjects and EEG derivations, from the SLEEP-EDF, CAP, ISRUC, SVUH, and MIT-BIH databases, we achieved an accuracy of 90.6%, 83.6%, 77.4%, 77.4%, 75.4%, and 73.6%, respectively. From Table 12, it is clear that our method attained an accuracy of 83.6% for the CAP database, which is the highest among all other studies mentioned for the CAP database. For the Sleep-EDF database, our model achieved an accuracy of 90.6%, which is almost equal to the highest accuracy of 90.8% achieved by Yildirim et al. [28]. Similarly, for the

ISRUC database, the proposed model got an accuracy of 77.4%, which is the second highest among the other studies mentioned in the Table 12. Besides, for the MIT-BIH data, our model's accuracy is almost close to the highest accuracy obtained by Tripathy et al. [78]. For the SVUH database, our model surpassed all other methods mentioned in the table. We attained an accuracy of 83.6% with the WSC database, which is the first work to the best of our knowledge. The comparison result indicates that our model not only performed well on the WSC database but its performance is also encouraging with all five popularly used public databases, indicating that the model is robust and accurate. It also indicates the generalizability of the model for any unseen new database.

The advantages of the proposed method are as follows:

- A two-stage technique is used to design a new class of linear phase WFB, wherein the analysis lowpass filter is a half-band filter. The LS criterion has been used in both the stages of the design stages, which resulted in a simple optimization problem involving a set of linear equations.
- This is the first study to incorporate the entire WSC-dataset for SSG.
- We have extracted and used 21,65,205 epochs of 30 seconds duration, which is the largest dataset to be used in studies.
- We have used a new class of optimal bi-orthogonal wavelet filters to decompose the signals into sub-bands.
- HP is used for feature extraction, which is a good discriminator for EEG signals.
- The proposed method achieved good accuracy and  $\kappa$  value utilizing only two unipolar EEG channels.
- Developed model is robust and generalizable as it performed well using on all six publically available databases.

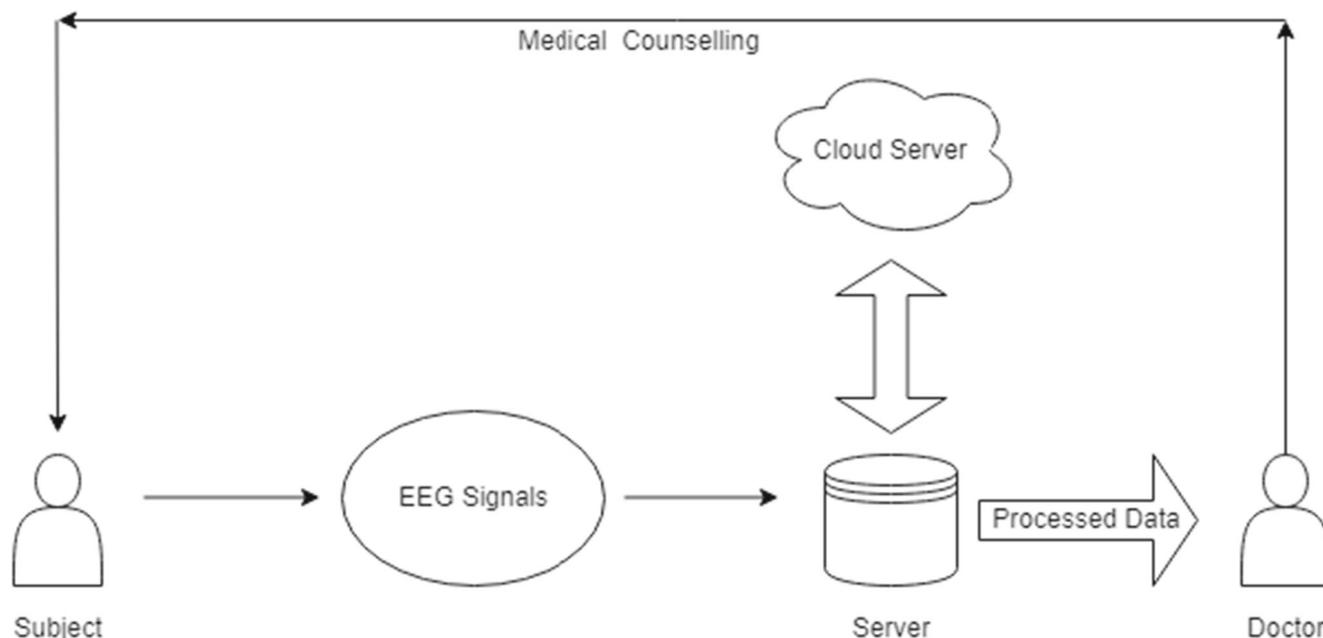
The limitation of the study are as follows:

- The datasets used are not balanced with respect to the number of epochs in each class, making it challenging for unbiased classification.
- As the datasets are unbalanced and has the least number of data for N1 sleep stage, it has the least accuracy. However, the performance accuracy of the N1 sleep stage is better than other reported state-of-the-art techniques.

The COVID-19 pandemic has paved the way for the Internet of things (IoT) in healthcare to obtain accurate diagnosis. The rise in demand for remote healthcare and the progress in cloud computing, machine learning, and biomedical sensors has made personal healthcare possible. Our proposed model can be used as an algorithm for cloud-based servers to detect sleep stages accurately and, the workflow

**Table 12** Comparison of the proposed method with state-of-the-art SSG studies developed using the WSC dataset

Database	Authors	Subjects	Channels	Features	Classifier	Accuracy (%)
CAP	Sharma et al. [13]	80	2 EEG (C4-A1 and F4-C4)	Norm	EBT	78.3
	Tripathi et al. [18]	25	4 EEG (P4-O2, C4-A1, F4-C4 and C4-P4)	Dispersion entropy and Bubble entropy	Hybrid learning	73.88
	Banluesombatkul et al. [25]	108	2 EEG (C3-A2/C3-P3, C4-A1)	Metasleep Learner	CNN	75.1
	Proposed Method	80	2 EEG (F4-C4 and C4-A1)	Hjorth Parameters	EBT	83.6
	Timplalexis et al. [19]	12	2 EEG (Fpz-Cz and Pz-Oz)	Mixture of time and frequency domain features	SVM	88.63
Sleep-EDF	Sharma et al. [15]	98	2 EEG (Fpz-Cz and Pz-Oz)	Fuzzy entropy	Gaussian SVM	89.8
	Mousavi et al. [20]	98	1 EEG (Fpz-Cz/Pz-Oz)	Time domain and frequency domain features	CNN-BIRNN	84.3
	Seo et al. [21]	42	1 EEG (Fpz-Cz)	IITNet	CRNN	83.9
	Supratak et al. [23]	20	2 EEG (Fpz-Cz and Pz-Cz)	Time-invariant features	CNN-BiLSTM	82.0
	Vilamala et al. [24]	20	2 EEG (Fpz-Cz and Pz-Oz)	VGG-FE and VGG-FT	CNN	81.3
	Michielli et al. [5]	10	1 EEG (Fpz-Cz)	mRMR	LSTM-RNN	83.6
	Yildirim et al. [28]	8	2 EEG (Fpz-Cz and Pz-Oz)	Raw-Signals	1D-CNN	90.8
	Proposed Method	98	2 EEG (Fpz-Cz and Pz-Oz)	Hjorth Parameters	EBT	90.6
	Banluesombatkul et al. [25]	110	1 EEG (C3-A2/C3-P3)	Metasleep Learner	CNN	75.2
	Tzamourta et al. [34]	100	6 EEG (F3-A2, C3-A2, O1-A2, F4-A1, C3-A1 and O2-A1)	Time and Frequency domain-based features	Random Forest	75.29
ISRUC	Ghimatgar et al. [29]	10	6 EEG (F3-A2, C3-A2, O1-A2, F4-A1, C3-A1 and O2-A1)	Time and Frequency domain-based features	HMM	79.64
	Proposed Method	100	6 EEG (F3-A2, C3-A2, O1-A2, F4-A1, and O2-A1)	Hjorth Parameters	EBT	77.4
	Rossow et al. [79]	18	EEG (C3-O1/C4-A1/O2-A1/C4-A1)	ARMA model coefficient	HMM	60.14
	Redmond et al. [80]	18	EEG and HRV (C3-O1/C4-A1/O2-A1/C4-A1)	RR interval features and ECG derived respiratory feature	Quadratic Discriminant	61
	Tripathy et al. [78]	18	EEG and HRV (C3-O1/C4-A1/O2-A1/C4-A1)	Dispersion entropy	DNN	73.70
SVUH	Proposed Method	18	EEG (C3-O1/C4-A1/O2-A1/C4-A1)	Hjorth Parameters	EBT	73.6
	Yuan et al. [44]	25	2 EEG (C3-A2 and C4-A1), 2 EOG, EMG and ECG	Hybrid CNN	SVM	74.2
	Cen et al. [81]	25	2 EEG (C3-A2 and C4-A1), 2 EOG, EMG and ECG	Deep Belief Network	HMM	69.78
	Proposed Method	25	2 EEG (C3-A2 and C4-A1)	Hjorth Parameters	EBT	75.4
WSC	Proposed Method	2431	2 EEG (C3-M2 and O1-M2)	Hjorth Parameters	EBT	83.2



**Fig. 5** Flow chart of an IoT-based automated sleep stage classification system

is shown in Fig. 5. First, the EEG signals can be collected from the subject using the smart wearable hardware. This data can be uploaded and processed at a local server, connected to a server hosted on the cloud. This processed data can be forwarded to medical professionals for examination, who will provide the medical assistance if needed.

## 6 Conclusion

The proposed study aims to implement an automatic SSG method using a novel class of stop-band energy localized biorthogonal filter bank and supervised machine learning algorithms. We have used a very large dataset comprising of 2431 PSG files and 21,65,205 epochs of 30 seconds duration each. With the help of a filter bank and five-level wavelet decomposition, six subbands are created from each EEG epoch. After the wavelet decomposition, three HP are computed from each subband, resulting in 18 features. These features are then fed to supervised machine learning algorithms to classify the sleep stages. The classification is done using a ten-fold cross-validation strategy. Only two unipolar EEG signals (O1\_M2 and C3\_M2) are considered to carry out this study compared to PSG-based system, which makes our approach convenient for patients, and a portable device can be developed with it. The database employed in our study is very large and diverse involving many channels and several subdatasets. We have obtained the best accuracy of 83.2% and  $\kappa$  value of 0.7345 using an ensemble bagged tress (EBT) classifier with our 12 considered data subsets. We also observed that our system

is either better or competitive with existing state-of-art systems when we tested with CAP, sleep-EDF, ISRUC, MIT-BIH, and St. Vincent's University databases.

The choice of computationally less intensive features, optimal filter bank, and only two EEG channels make our developed model suitable for real-time application. Hence it can be implemented in a portable EEG device for sleep monitoring while sleeping comfortably in our homes. In the future, we plan to carry out sleep disorder identification automatically using the same WSC database.

## References

1. Vyazovskiy VV (2015) Sleep, recovery, and metaregulation: explaining the benefits of sleep. *Nature and Science of Sleep* 7:171
2. Potkin KT, Bunney WE Jr (2012) Sleep improves memory: the effect of sleep on long term memory in early adolescence
3. Takahashi Y, Kipnis D, Daughaday W et al (1968) Growth hormone secretion during sleep. *J Clin Investig* 47(9):2079–2090
4. Loh HW, Ooi CP, Vicnesh J, Oh SL, Faust O, Gertych A, Acharya UR (2020) Automated detection of sleep stages using deep learning techniques: a systematic review of the last decade (2010–2020). *Appl Sci* 10(24):8963
5. Michielli N, Acharya UR, Molinari F (2019) Cascaded lstm recurrent neural network for automated sleep stage classification using single-channel eeg signals. *Computers in Biology and Medicine* 106:71–81
6. Moser D, Anderer P, Gruber G, Parapaties S, Loretz E, Boeck M, Kloesch G, Heller E, Schmidt A, Danker-Hopfe H et al (2009) Sleep classification according to aasm and rechtschaffen & kales: effects on sleep scoring parameters. *Sleep* 32(2):139–149
7. Graven S (2006) Sleep and brain development. *Clin Perinatol* 33(3):693–706

8. Amzica F, Steriade M (2002) The functional significance of k-complexes. *Sleep Med Rev* 6(2):139–149
9. De Gennaro L, Ferrara M (2003) Sleep spindles: an overview. *Sleep Med Rev* 7(5):423–440
10. Toogood AA, O'Neill PA, Shalet SM (1996) Beyond the somatopause: growth hormone deficiency in adults over the age of 60 years. *J Clin Endocrinol Metab* 81(2):460–465
11. Siegel JM (2017) Rapid eye movement sleep. *Principles and Pract Sleep Med* 6:78–95
12. Carskadon MA, Dement WC et al (2005) Normal human sleep: an overview. *Princ Pract Sleep Med* 4(1):13–23
13. Sharma M, Tiwari J, Acharya UR (2021) Automatic sleep-stage scoring in healthy and sleep disorder patients using optimal wavelet filter bank technique with eeg signals. *Int J Environ Public Health* 18(6):3087
14. Sharma M, Acharya UR (2021) Automated detection of schizophrenia using optimal wavelet-based  $l_1$  norm features extracted from single-channel eeg. *Cogn Neurodyn* 15(4):661–674
15. Sharma M, Patel S, Choudhary S, Acharya UR (2020) Automated detection of sleep stages using energy-localized orthogonal wavelet filter banks. *Arab J Sci Eng* 45(4):2531–2544
16. Loh HW, Ooi CP, Dhok SG, Sharma M, Bhurane AA, Acharya UR (2022) Automated detection of cyclic alternating pattern and classification of sleep stages using deep neural network. *Appl Intell* 52(3):2903–2917
17. Sharma M, Kumar K, Kumar P, Tan R-S, Acharya UR (2022) Pulse oximetry spo<sub>2</sub> signal for automated identification of sleep apnea: a review and future trends. *Physiological Measurement*
18. Tripathy RK, Ghosh SK, Gajbhiye P, Acharya UR (2020) Development of automated sleep stage classification system using multivariate projection-based fixed boundary empirical wavelet transform and entropy features extracted from multichannel eeg signals. *Entropy* 22(10):1141
19. Timplalexis C (2019) Classification of sleep stages using machine learning methods
20. Mousavi S, Afghah F, Acharya UR (2019) Sleepeegnet: Automated sleep stage scoring with sequence to sequence deep learning approach. *PloS One* 14(5):0216456
21. Seo H, Back S, Lee S, Park D, Kim T, Lee K (2020) Intra- and inter-epoch temporal context network (iitnet) using sub-epoch features for automatic sleep scoring on raw single-channel eeg. *Biomedical Signal Processing and Control* 61:102037
22. Zhang X, Xu M, Li Y, Su M, Xu Z, Wang C, Kang D, Li H, Mu X, Ding X et al (2020) Automated multi-model deep neural network for sleep stage scoring with unfiltered clinical data. *Sleep and Breathing* 24(2):581–590
23. Supratak A, Dong H, Wu C, Guo Y (2017) Deepsleepnet: a model for automatic sleep stage scoring based on raw single-channel eeg. *IEEE Trans Neural Syst Rehabil Eng* 25(11):1998–2008
24. Vilamala A, Madsen KH, Hansen LK (2017) 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, pp 1–6
25. Banluesombatkul N, Ouppaphan P, Leelaarporn P, Lakhan P, Chaitusaney B, Jaimcharyatam N, Chuangsuwanich E, Chen W, Phan H, Dilokthanakul N et al (2020) Metasleeplearner: a pilot study on fast adaptation of bio-signals-based sleep stage classifier to new individual subject using meta-learning. *IEEE J Biomed Health Inform* 25(6):1949–1963
26. Lesmana TF, Isa SM, Surantha N (2018) Sleep stage identification using the combination of elm and pso based on eeg signal and hrv. In: 2018 3rd international conference on computer and communication systems (ICCCS), IEEE, pp 258–262
27. Lee H, Choi J, Kim S, Jun SC, Lee B-G (2019) A compressive sensing-based automatic sleep-stage classification system with radial basis function neural network. *IEEE Access* 7:186499–186509
28. Yildirim O, Baloglu UB, Acharya UR (2019) A deep learning model for automated sleep stages classification using psg signals. *Int J Environ Res Public Health* 16(4):599
29. Ghimatgar H, Kazemi K, Helfroush MS, Aarabi A (2019) An automatic single-channel eeg-based sleep stage scoring method based on hidden markov model. *J Neurosci Methods* 324:108320
30. Wei R, Zhang X, Wang J, Dang X (2018) The research of sleep staging based on single-lead electrocardiogram and deep neural network. *Biomed Eng Lett* 8(1):87–93
31. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE (2000) Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation* 101(23):215–220
32. Terzano MG, Parrino L, Sherieri A, Chervin R, Chokroverty S, Guilleminault C, Hirshkowitz M, Mahowald M, Moldofsky H, Rosa A et al (2001) Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (cap) in human sleep. *Sleep Medicine* 2(6):537–554
33. Långkvist M, Karlsson L, Loutfi A (2012) Sleep stage classification using unsupervised feature learning. *Advances in Artificial Neural Systems* 2012
34. Tzimirta KD, Tsilimbaris A, Tzioukalia K, Tzallas AT, Tsipouras MG, Astrakas LG, Giannakeas N (2018) Eeg-based automatic sleep stage classification. *Biomed J* 1:6
35. Khalighi S, Sousa T, Santos JM, Nunes U (2016) Isruc-sleep: a comprehensive public dataset for sleep researchers. *Comput Methods Prog Biomed* 124:180–192
36. Kemp B, Zwirnerman A, Tuk B, Kamphuisen H, Oberyé J (2018) Sleep-edf database expanded. *Physionet org*
37. Ichimaru Y, Moody G (1999) Development of the polysomnographic database on cd-rom. *Psychiatry Clin Neurosci* 53(2):175–177
38. Willemsen T, Van Deun D, Verhaert V, Vandekerckhove M, Exadaktylos V, Verbraecken J, Van Huffel S, Haex B, Vander Sloten J (2013) An evaluation of cardiorespiratory and movement features with respect to sleep-stage classification. *IEEE J Biomed Health Inform* 18(2):661–669
39. Fonseca P, Long X, Radha M, Haakma R, Aarts RM, Rolink J (2015) Sleep stage classification with eeg and respiratory effort. *Physiological Measurement* 36(10):2027
40. Kloth G, Kemp B, Penzel T, Schlogl A, Rappelsberger P, Trenker E, Gruber G, Zeithofer J, Saletu B, Herrmann W et al (2001) The siesta project polygraphic and clinical database. *IEEE Eng Med Biol Mag* 20(3):51–57
41. Phan H, Andreotti F, Cooray N, Chén OY, De Vos M (2018) Joint classification and prediction cnn framework for automatic sleep stage classification. *IEEE Trans Biomed Eng* 66(5):1285–1296
42. O'reilly C, Gosselin N, Carrier J, Nielsen T (2014) Montreal archive of sleep studies: an open-access resource for instrument benchmarking and exploratory research. *J Sleep Res* 23(6):628–635
43. Shi J, Liu X, Li Y, Zhang Q, Li Y, Ying S (2015) Multi-channel eeg-based sleep stage classification with joint collaborative representation and multiple kernel learning. *Journal of Neuroscience Methods* 254:94–101
44. Yuan Y, Jia K, Ma F, Xun G, Wang Y, Su L, Zhang A (2019) A hybrid self-attention deep learning framework for multivariate sleep stage classification. *BMC Bioinformatics* 20(16):1–10
45. Radha M, Fonseca P, Moreau A, Ross M, Cerny A, Anderer P, Long X, Aarts RM (2019) Sleep stage classification from heart-

- rate variability using long short-term memory neural networks. *Scientific Reports* 9(1):1–11
46. Heneghan C (2011) St. Vincent's University Hospital/University College Dublin Sleep Apnea Database. *Physionet*
  47. Sharma M, Yadav A, Tiwari J, Karabatak M, Yildirim O, Acharya UR (2022) An automated wavelet-based sleep scoring model using eeg, emg, and eog signals with more than 8000 subjects. *Int J Environ Res Public Health* 19(12):7176
  48. Sharma M, Kumbhani D, Tiwari J, Kumar TS, Acharya UR (2022) Automated detection of obstructive sleep apnea in more than 8000 subjects using frequency optimized orthogonal wavelet filter bank with respiratory and oximetry signals. *Comput Biol Med* 144:105364
  49. Sharma M, Acharya UR (2019) A new method to identify coronary artery disease with eeg signals and time-frequency concentrated antisymmetric biorthogonal wavelet filter bank. *Pattern Recogn Lett* 125:235–240
  50. Sharma M, Patel V, Acharya UR (2021) Automated identification of insomnia using optimal bi-orthogonal wavelet transform technique with single-channel eeg signals. *Knowledge-based Systems*, pp 107078–107083. <https://doi.org/10.1016/j.knsys.2021.107078>
  51. Sharma M, Kolte R, Patwardhan P, Gadre V (2010) Time-frequency localization optimized biorthogonal wavelets. In: 2010 international conference on signal processing and communications (SPCOM), IEEE, pp 1–5
  52. Sharma M, Singh T, Bhati D, Gadre V (2014) Design of two-channel linear phase biorthogonal wavelet filter banks via convex optimization. In: 2014 international conference on signal processing and communications (SPCOM), IEEE, pp 1–6
  53. Sharma M, Dhere A, Pachori RB, Gadre VM (2017) Optimal duration-bandwidth localized antisymmetric biorthogonal wavelet filters. *Signal Process* 134:87–99
  54. Sharma M, Acharya UR (2018) Analysis of knee-joint vibroarthrographic signals using bandwidth-duration localized three-channel filter bank. *Comput Elect Eng* 72:191–202
  55. Sharma M, Tan R-S, Acharya UR (2019) Automated heartbeat classification and detection of arrhythmia using optimal orthogonal wavelet filters. *Informatics in Medicine Unlocked* 16:100221
  56. Sharma M, Achuth P, Deb D, Puthankattil SD, Acharya UR (2018) An automated diagnosis of depression using three-channel bandwidth-duration localized wavelet filter bank with eeg signals. *Cogn Syst Res* 52:508–520
  57. Phoong S-M, Kim CW, Vaidyanathan P, Ansari R (1995) A new class of two-channel biorthogonal filter banks and wavelet bases. *IEEE Trans Signal Process* 43(3):649–665
  58. Sharma M, Bapodara S, Tiwari J, Acharya UR (2022) Automated sleep apnea detection in pregnant women using wavelet-based features. *Informatics in Medicine Unlocked* 32:101026
  59. Sharma M, Darji J, Thakrar M, Acharya UR (2022) Automated identification of sleep disorders using wavelet-based features extracted from electrooculogram and electromyogram signals. *Computers in Biology and Medicine* 105224
  60. Sharma M, Patel S, Acharya UR (2021) Expert system for detection of congestive heart failure using optimal wavelet and heart rate variability signals for wireless cloud-based environment. *Expert Syst*, 12903
  61. Rajput JS, Sharma M, San Tan R, Acharya UR (2020) Automated detection of severity of hypertension eeg signals using an optimal bi-orthogonal wavelet filter bank. *Comput Biol Med* 123:103924
  62. Grant M, Boyd S (2014) CVX: Matlab Software For Disciplined Convex Programming, version 2.1. <http://cvxr.com/cvx>
  63. Grant M, Boyd S, Ye Y (2008) CVX: Matlab software for disciplined convex programming
  64. Upadhyay A, Sharma M, Pachori RB (2017) Determination of instantaneous fundamental frequency of speech signals using variational mode decomposition. *Comput Elect Eng* 62:630–647
  65. Hjorth B (1970) Eeg analysis based on time domain properties. *Electroencephalogr Clin Neurophysiol* 29(3):306–310
  66. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297
  67. Noble WS (2006) What is a support vector machine? *Nature Biotechnology* 24(12):1565–1567
  68. Quinlan JR (1986) Induction of decision trees. *Mach Learn* 1(1):81–106
  69. Bühlmann P. (2012) Bagging, boosting and ensemble methods. In: *Handbook of computational statistics*, Springer, pp 985–1022
  70. Rokach L, Maimon O (2005) Decision trees. In: *Data mining and knowledge discovery handbook*, Springer, pp 165–192
  71. Peterson LE (2009) K-nearest neighbor. *Scholarpedia* 4(2):1883
  72. Bayes T (1968) Naive bayes classifier. *Article Sources and Contributors*, pp 1–9
  73. Alickovic E, Subasi A (2018) Ensemble svm method for automatic sleep stage classification. *IEEE Trans Instrum Meas* 67(6):1258–1265
  74. Ferreira-Santos D, Amorim P, Martins TS, Monteiro-Soares M, Rodrigues PP et al (2022) Enabling early obstructive sleep apnea diagnosis with machine learning: Systematic review. *J Med Internet Res* 24(9):39452
  75. Faust O, Hagiwara Y, Hong TJ, Lih OS, Acharya UR (2018) Deep learning for healthcare applications based on physiological signals: a review. *Computer Methods and Programs in Biomedicine* 161:1–13
  76. Sahayadhas A, Sundaraj K, Murugappan M (2012) Detecting driver drowsiness based on sensors: a review. *Sensors* 12(12):16937–16953
  77. Sharma M, Goyal D, Achuth P, Acharya UR (2018) An accurate sleep stages classification system using a new class of optimally time-frequency localized three-band wavelet filter bank. *Comput Biol Med* 98:58–75
  78. Tripathy R, Acharya UR (2018) Use of features from rr-time series and eeg signals for automated classification of sleep stages in deep neural network framework. *Biocybernetics and Biomedical Engineering* 38(4):890–902
  79. Rossow AB, Salles EOT, Côco KF (2011) Automatic sleep staging using a single-channel eeg modeling by kalman filter and hmm. In: *ISSNIP biosignals and biorobotics conference 2011*, IEEE, pp 1–6
  80. Redmond S, Heneghan C (2003) Electrocardiogram-based automatic sleep staging in sleep disordered breathing. In: *Computers in cardiology, 2003*, IEEE, pp 609–612
  81. Cen L, Yu ZL, Tang Y, Shi W, Kluge T, Ser W (2017) Deep learning method for sleep stage classification. In: *International conference on neural information processing*, Springer, pp 796–802

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Affiliations

Manish Sharma<sup>1</sup>  · Paresh Makwana<sup>1</sup> · Rajesh Singh Chad<sup>1</sup> · U Rajendra Acharya<sup>2,3,4</sup>

Paresh Makwana  
paresh.makwana.18e@iitram.ac.in

Rajesh Singh Chad  
rajesh.chad.18e@iitram.ac.in

U Rajendra Acharya  
RajendraUdyavaraACHARYA@np.edu.sg

- <sup>1</sup> Department of Electrical and Computer Science Engineering, Institute of Infrastructure, Technology, Research and Management (IITRAM), Ahmedabad, 380026, India
- <sup>2</sup> School of Engineering, Ngee Ann Polytechnic, Singapore, 599489, Singapore
- <sup>3</sup> Department of Bioinformatics and Medical Engineering, Asia University, Taichung, 41354, Taiwan
- <sup>4</sup> Department of Biomedical Engineering, School of Science and Technology, SUSS University, Singapore, Singapore